CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
Telephone 31 - 20 592 9333, telex 12571 (mactr nl),
telefax 31 - 20 592 4199

CWI is the nationally funded Dutch institute for research in Mathematics and Computer Science.

One-dependent processes: two-block-factors and non-two-block-factors

V. de Valk

ISBN 90 6196 437 7
NUGI-code: 811

Printed in the Netherlands

One-Dependent Processes:
Two-Block-Factors and Non Two-Block-Factors

Vincent de Valk
Mathematical Institute
University of Groningen
P.O. Box 800
9700 AV Groningen
The Netherlands

September 8, 1993

Contents

The Six Articles

Classification Code (MSC 1991)
Primary 60G10, 28D05, 05B20, 54H20
Secondary 15A36, 15A45, 26D15, 28A75, 47A35, 52A20, 60J10

1 Introduction

This monograph consists of six articles, mainly on one-dependent processes, with connections to combinatorics, analysis, functional analysis, dynamical systems, matrix theory and variational problems. Before describing the articles in this introduction, a survey is given of the theory of m-dependence (a generalization of one-dependence) with its applications to renormalization theory and other fields of statistical physics (section 3). In section 4 we introduce $(m+1)$-block-factors of i.i.d. sequences as examples of m-dependent processes and we give a counterexample of a one-dependent process that is not a block-factor at all. In section 5 we prove that all one-dependent renewal processes are two-block-factors. In section 6 we consider one-dependent Markov processes. We show that a one-dependent Markov process with no more than 4 states is a two-block-factor of an i.i.d. sequence. The main part of this section is devoted to a counterexample of a one-dependent Markov process with 5 states that is not a two-block-factor. In section 7 we discuss the problem under which conditions a one-dependent process necessarily is a two-block-factor. m-dependent processes can be described in terms of Hilbert spaces, operators and vectors. These Hilbert space representations seem to be the right way to investigate m-dependent processes. Section 8 deals with two-correlations that play an important role in the problem which one-dependent processes are two-block-factors. In section 9 some applications are given of the results on two-correlations. The six articles are summarized in section 11 and in section 12 we give a list of open problems and conjectures, to which this monograph gives rise.

2 m-dependent processes

Discrete time stochastic processes $(X_N)_{N\in\mathbb{Z}}$ have been studied thoroughly by probabilists. An important class of these processes are the independent processes. The class of independent processes can be considered as a part of a wider class, such as the Markov processes. Another way of generalizing the notion of independence is by defining m-dependence. An independent process has the property that two events are independent whenever they are separated by a time-interval with positive length, and an *m-dependent process* has the property that two events are independent whenever they are separated by a time-interval with length more than m. To be more precise: at each (discrete) time t the future $(X_N)_{N\geq t+m}$ is independent of the past $(X_N)_{N<t}$. Although "almost everything" is known about Markov processes, not so much is known about m-dependent processes; the theory is young but growing. We give a survey in the next sections.

3 Renormalization theory and statistical physics

Many models in statistical physics have rescaling-properties for critical values (e.g. critical temperature) of their parameters, as is conjectured by physicists. This means that these models are invariant under rescaling-operations (as e.g. fractals). So several random fields in statistical physics (concerning e.g. magnetization, Ising model etc.) with the place as parameter should typically be one-dependent, in contrast to processes in many other applications of probability theory with the time as parameter that are typically Markovian. Nevertheless sometimes we will consider the parameter of a one-dependent process as the time and then we write about "past" and "future". We will show that one-dependent processes occur as limits of rescaling operations in renormalization theory (see e.g.[O'Br.]). Let the process $X^{(0)}$ be $0-1$ valued, let r be an integer greater than one, and let

$$\phi : \{0,1\}^r \to \{0,1\}$$

be a function. We define a new process $X^{(1)}$ by

$$X_i^{(1)} := \phi\left(X_{ri}^{(0)}, \ldots, X_{ri+r-1}^{(0)}\right), \quad i \in \mathbf{Z}.$$

We can iterate this procedure, obtaining a sequence of processes $X^{(N)}$. When we assume $X^{(0)}$ to be stationary, then it is trivial that $X^{(N)}$ is also stationary. Because $X_1^{(N)}$ depends on $X_{r^N}^{(0)}, \ldots.., X_{2r^N-1}^{(0)}$ and $X_{-1}^{(N)}$ depends on $X_{-r^N}^{(0)}, \ldots., X_{-1}^{(0)}$, it is easy to see that if $(X^{(N)})_{N=1}^{\infty}$ has a subsequence that converges (in distribution) to some limit, then this limit is one-dependent, assuming that $X^{(0)}$ satisfies the following mixing condition. O'Brien assumes that there exists a decreasing sequence $(g(K))_{K=0}^{\infty}$ converging to zero, such that

$$|P(A \cap B) - P(A) \cdot P(B)| \leq g(K)$$

for all events A depending on $\{\ldots, -3, -2, -1\}$ and all events B depending on $\{K, K+1, K+2, \ldots\}$. (P is the underlying probability measure corresponding to the process $X^{(0)}$.)

4 m-block-factors

In addition to being limits of rescaling operations, m-dependent processes can be obtained in a simpler way: as an $m+1$-block-factor of an i.i.d. sequence $(Y_N)_{N \in \mathbf{Z}}$. Let the process X be defined by

$$X_N := f(Y_N, Y_{N+1}, \ldots, Y_{N+m}) \quad (N \in \mathbf{Z})$$

for some function f. Obviously X is an m-dependent process. For an m-block-factor X it is no restriction to assume that the underlying sequence $(Y_N)_{N \in \mathbf{Z}}$ is identically uniformly distributed over the unit interval.
In this section we will show that not all m-dependent processes are $m+1$-block-factors by giving a counterexample of a one-dependent process that is not an $m+1$-block-factor for any $m \in \mathbf{N}$.

Theorem 4.1 (Burton, Goulet and Meester, see [B.G.M.], Theorem 1) *There exists a stationary, one-dependent process with 4 states that is not a K-block-factor of an i.i.d. sequence for any $K \in \mathbf{N}$.*

Proof.
The process has state space $\{0, 1, (2,0), (2,1)\}$. We start the construction with an i.i.d. sequence for $(Z_n)_{n \in \mathbf{Z}}$ such that $P[Z_1 = x] = 1/3$ for $x \in \{0,1,2\}$. Define the random number $\tau(n)$ by

$$\tau(n) := \max\{m \in \mathbf{Z} : Z_m = 2, \ m < n\}$$

for $n \in \mathbf{Z}$ and let $d(n) := n - \tau(n) - 1$ be the number of elements strictly between Z_n and the previous 2. Let $S := \{n^2 : n \in \mathbf{N}\}$. We define now

$$X_n := \begin{cases} Z_n, \text{ if } Z_n \in \{0,1\} \\ \\ (2, i_n), \text{ if } Z_n = 2 \end{cases}$$

where the second coordinate i_n is defined by

$$i_n := \begin{cases} \sum_{j=\tau(n)+1}^{n-1} Z_j \bmod 2 & \text{if } d(n) \notin S \\ \\ 1 + \sum_{j=\tau(n)+1}^{n-1} Z_j \bmod 2 & \text{if } d(n) \in S \end{cases}$$

We call $\sum_{j=\tau(n)+1}^{n-1} Z_j \bmod 2$ the parity of the elements (all zeroes and ones) between $\tau(n)$ and n.It is trivial that $(X_n)_{n\in\mathbf{Z}}$ is stationary, because $(Z_n)_{n\in\mathbf{Z}}$ is stationary. First we prove the one-dependence and then that the process is not a K-block-factor.

Claim 1.
$(X_n)_{n\in\mathbf{Z}}$ is one-dependent.

Proof of Claim 1.
Let $A := \{X_j = a_j,\ \ j = -1, \dots, -m\}$ and $B := \{X_j = b_j,\ \ j = 1, \dots, n\}$ be two events with positive probability, where $m, n \in \mathbf{N}$. We will prove $P(B|A) = P(B)$ what implies one-dependence.
If all b_j are 0 or 1, it is trivial that A and B are clearly independent by construction. So, assume that $b_j \notin \{0,1\}$ for some j. Let

$$\lambda := \min\{1 \leq j \leq n : b_j \notin \{0,1\}\}.$$

By construction only the second coordinate of b_λ can depend on A. To exploit this observation we define the events

$$\begin{aligned} D &:= \{X_1 + \cdots + X_{\lambda-1} = b_1 + \cdots + b_{\lambda-1} \bmod 2\} \\ B_\lambda &:= \{X_\lambda = b_\lambda\}. \end{aligned}$$

We have

$$\begin{aligned} P(B|A) &= P(B|A, X_0 \notin \{0,1\})P(X_0 \notin \{0,1\}|A) + \\ &+ P(B|A, X_0 \in \{0,1\})P(X_0 \in \{0,1\}|A) = \\ &= (1/3)^{n-1} P(B_\lambda|A, D, X_0 \notin \{0,1\})P(X_0 \notin \{0,1\}|A) + \\ &+ (1/3)^{n-1} P(B_\lambda|A, D, X_0 \in \{0,1\})P(X_0 \in \{0,1\}|A). \end{aligned}$$

We want to remove the A from the above formula's. Clearly by construction

$$P(X_0 \notin \{0,1\}|A) = P(X_0 \notin \{0,1\}),$$

$$P(X_0 \in \{0,1\}|A) = P(X_0 \in \{0,1\}) \text{ and}$$
$$P(B_\lambda|A, D, X_0 \notin \{0,1\}) = P(B_\lambda|D, X_0 \notin \{0,1\}).$$

Finally, the symmetry of even and odd parities implies that the probability of seeing a one (or a zero) as second coordinate of a given 2 given any event which does not specify all coordinates back to the previous 2 is 1/2, independent of the exact form of the event. Because $\{A, D, X_0 \in \{0,1\}\}$ and $\{D, X_0 \in \{0,1\}\}$ both fall in this category we have

$$P(B_\lambda|A, D, X_0 \in \{0,1\}) = P(B_\lambda|D, X_0 \in \{0,1\}),$$

which completes the proof of the Claim.

Claim 2.
$(X_n)_{n\in\mathbb{Z}}$ is not a K-block-factor of an i.i.d. sequence.

Proof of Claim 2.
Assume that the process is a K-block-factor for some $K \in \mathbb{N}$. Let

$$X_n = f(Y_n, \ldots, Y_{n+K-1})$$

for some measurable function f and some i.i.d. sequence Y_n. We define the code c of a sequence of symbols $y = (y_1, \ldots, y_m)(m \geq K)$ by

$$c(y) := f(y_1, \ldots, y_K)f(y_2, \ldots, y_{K+1}) \ldots f(y_{m-K+1}, \ldots, y_m)$$

We will write $[0]^n$for a sequence of n zeroes. Because $(n+1)^2 - n^2 \to \infty$ for $n \to \infty$ it is possible to choose $i, j, n \in \mathbb{N}$ such that

(i) $K < i < j < n^2$

(ii) $n^2 - i + l \notin S$ for all $l = 1, \ldots, K$

(iii) $j - l \notin S$ for all $l = 1, \ldots, K$

(iv) $j - i + l \notin S$ for all $l = -K+1, \ldots, K-1$.

Let $m := n^2 + K + 1$. We define for every m-tuple $i_1 < i_2 < \ldots < i_m$ the event

$$E(i_1, i_2, \ldots, i_m) := \left\{c(Y_{i_1}, Y_{i_2}, \ldots, Y_{i_m}) = (2,1)[0]^{n^2}(2,1)\right\}$$

From the construction of the process follows that $P[E(i_1, i_2, \ldots, i_m)] > 0$ for all $i_1 < i_2 < \ldots < i_m$. For every K-tuple $j_1 < \ldots < j_K$ we define the event

$$F(j_1, \ldots, j_K) := \{c(Y_{j_1}, \ldots, Y_{j_K}) \notin \{0,1\}\}.$$

The idea of the proof is to start with the event $E(1, 2, \ldots, m)$, then "pull this event apart" and insert a 2 in two different places and then show that this results in an impossible event having positive probability. We define

$$D_1 := E(1, 2, \ldots, i, i+K+1, \ldots, m+K) \cap F(i+1, i+2, \ldots, i+K).$$

In comparison with the event $E(1, 2, \ldots, m)$, $K-1$ zeroes in $X_1, \ldots, X_{n^2+2}$ are replaced by $2K-1$ new symbols, at least one of which is not in $\{0,1\}$ by construction.
On the event D_1 we have a.s.

$$X_1 X_2 \ldots X_{n^2+K+2} = (2,1)[0]^{i-K} \underbrace{* \cdots *}_{2K-1 \text{ times}} [0]^{n^2-i+1}(2,1),$$

where the stars are unspecified but at least one of them (the middle one) is not a zero or a one. We write $(2,*)$ for a symbol in this sequence that is not a zero or a one. We consider the rightmost $(2,*)$ among the stars, at a random position. Condition (ii) implies that the number of elements between this $(2,*)$ and the final $(2,1)$ is not a square. Hence, the parity of the stars between the rightmost $(2,*)$ and the final $(2,1)$ is necessarily 1 on the event D_1.
Consider the event D_2;

$$D_2 := E(1,\ 2, \ldots, j,\ j+K+1, \ldots, m+K) \cap F(j+1,\ j+2, \ldots, j+K).$$

Comparable with D_1 we have on D_2 a.s.

$$X_1 X_2 \ldots X_{n^2+K+2} = (2,1)[0]^{j-K} \underbrace{* \cdots *}_{2K-1 \text{ times}} [0]^{n^2-j+1}(2,1).$$

This time we consider the leftmost $(2,*)$ among the stars. Its (random) second coordinate is denoted by l. Condition (iii) implies that the parity of the stars to the left of this $(2,l)$ must be l.
We will now derive a contradiction by combining D_1 and D_2:

$$D_3 := E(1,2,\ldots,i,i+K+1,\ldots,j+K,j+2k+1,j+2K+2,\ldots,m+2K)$$
$$\cap F(i+1,i+2,\ldots,i+K) \cap F(j+K+1,j+K+2,\ldots,j+2K).$$

By construction we have obviously $P(D_3) > 0$. On D_3 we have a.s.

$$X_1 X_2 \ldots X_{n^2+2K+2} =$$
$$(2,1)[0]^{i-K} \underbrace{* \cdots * (2,*) * \cdots *}_{2K-1 \text{ elements}} [0]^{j-i-K+1} \underbrace{* \cdots * (2,l) * \cdots *}_{2K-1 \text{ elements}} [0]^{n^2-j+1}(2,1),$$

where the stars are not specified and l is random. Combining the observations of D_1 and D_2 above, we see that the parity of all elements between the designated $(2,*)$ and $(2,l)$ must be $1+l$. Condition (iv) implies that the number of elements between the designated $(2,*)$ and $(2,l)$ is not a square, hence $(2,l)$ has the wrong second coordinate and we conclude that $P(D_3) = 0$, which is the contradiction.

□

Remark. The above counterexample is slightly generalized in [B.G.M.] to a counterexample of a stationary, one-dependent process with finite energy which is not a K-block-factor of an i.i.d. sequence for any $K \in \mathbb{N}$.
A stationary process $(X_n)_{n\in\mathbb{Z}}$ with finite state space S satisfies the finite energy condition of Newman and Schulman if for any $s \in S$ and for any event A that is measurable with respect to the σ-field generated by $\{X_n : n \neq 0\}$ and with positive probability

$$P[X_0 = s \mid A] > 0$$

holds (see [N.S.]).

5 One-dependent renewal processes

In this section we prove the result by Aaronson, Gilat and Keane ([A.G.K.]) that every stationary one-dependent renewal process is a two-block-factor of an i.i.d. sequence. Let $(W_n)_{n=0}^{\infty}$ be a renewal process; i.e.

$$P\left\{\bigcap_{j=0}^{K}[W_{n_j}=1]\right\} = P\{W_{n_0}=1\}\prod_{j=1}^{K} u_{n_j-n_{j-1}}$$

for $0 \leq n_0 \leq n_1 \leq n_2 \leq \cdots \leq n_K$, where the sequence $(u_n)_{n=0}^{\infty}$

$$u_n := P[W_n = 1 \mid W_0 = 1] \qquad (n \geq 0]$$

is called the renewal sequence (see [C.]) of $(W_n)_{n=0}^{\infty}$. The sequence clearly satisfies $u_2 \geq u_1^2$ because

$$u_2 = P[W_2 = 1 \mid W_0 = 1] \geq P[W_2 = W_1 = 1 \mid W_0 = 1] = u_1^2.$$

We define

$$f_n := P[W_n = 1, W_K = 0, 1 \leq K \leq n-1 \mid W_0 = 1] \qquad (n \geq 1).$$

We have the renewal equation

$$u_n = \sum_{K=1}^{n} f_K u_{n-K} \qquad (n \geq 1).$$

It follows trivially from the definition of u_n that a stationary renewal process is one-dependent if and only if

$$u_n = P[W_0 = 1] =: b \qquad \text{for all } n \geq 2.$$

Further we define $a := u_1$. After these definitions we can prove

Theorem 5.1 ([A.G.K.], Theorem 1) *Any stationary, one-dependent renewal process is a two-block-factor of an i.i.d. sequence.*

Proof.
Let $(W_n)_{n=0}^{\infty}$ be a stationary, one-dependent renewal process with renewal sequence $(u_n)_{n=0}^{\infty}$ and with $(f_n)_{n=1}^{\infty}$, a, b as defined above.

Claim.
We claim that

$$0 \leq a \leq 1, \;\; a^2 \leq b \leq \frac{(1+a)^2}{4}. \tag{1}$$

Proof of the Claim.
We have $b = u_2 \geq u_1^2 = a^2$ as was shown above. A straightforward computation shows that

$$\frac{1}{(1-x)U(x)} = \sum_{n=0}^{\infty} c_n x^n \qquad (|x| < 1) \tag{2}$$

where $U(x) := \sum_{n=0}^{\infty} u_n x^n = 1 + ax + \frac{bx^2}{1-x}$ and $c_n := \sum_{K=n+1}^{\infty} f_K \geq 0$.
Writing

$$(1-x)U(x) = 1 - (1-a)x + (b-a)x^2 = (1 - r_+ x)(1 - r_- x)$$

we have that if $b > \frac{(1+a)^2}{4}$, then

$$r_\pm = re^{\pm i\theta} \text{ where } 0 < \theta < \pi, \ \ r > 0.$$

Expanding into partial fractions gives us

$$\frac{1}{(1-x)U(x)} = \frac{de^{i\delta_+}}{1 - re^{i\theta}x} + \frac{de^{i\delta_-}}{1 - re^{-i\theta}x} = \sum_{n=0}^{\infty} c_n x^n$$

where $d > 0$, $\delta_\pm \in (0, 2\pi)$, so we obtain

$$c_n = 2dr^n \cos(n\theta + \frac{1}{2}(\delta_+ - \delta_-)) \cos(\frac{1}{2}(\delta_+ + \delta_-))$$

which cannot be nonnegative for all $n \geq 1$. Hence $b \leq \frac{(1+a)^2}{4}$ which proves the Claim.

Let $(Y_n)_{n=0}^{\infty}$ be an i.i.d. sequence of random variables each uniformly distributed over the unit interval. Let A, B be two measurable subsets of the unit interval. It is easy to see that the two-block-factor

$$1_{[Y_n \in A, Y_{n+1} \in B]} \qquad (n \geq 0) \tag{3}$$

is a stationary, one-dependent renewal process. It is easily checked that for this renewal process

$$a = |B \cap A| \text{ and } b = |A| \cdot |B| = (a + |A \backslash B|)(a + |B \backslash A|),$$

where $|\cdot|$ denotes the Lebesgue measure. Using $1 + a \geq |A| + |B|$ it follows that we have once more (see (1)).

$$0 \leq a \leq 1, \ \ a^2 \leq b \leq \frac{(1+a)^2}{4}$$

One checks easily that this is the parametrization of the set of two-block-factors of the type as in (3). This proves the Theorem. □

6 One-dependent Markov processes

In the sixth article of this monograph (Proposition 7 [R.V.]) is proved that a stationary, one-dependent Markov process with only 2 states is an i.i.d. sequence. This does not hold any more for more than 2 states. It is easy to check that

$$\begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/6 & 1/6 & 2/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$$

is a transition matrix of a stationary, one-dependent Markov process with 3 states. Clearly it is not an independent process. However, under the symmetry condition $P^{(X_1,X_2)} = P^{(X_2,X_1)}$ in Proposition 10 [R.V.] is proved that a stationary, one-dependent Markov process is a i.i.d. sequence. Because every stationary, one-dependent Markov process with only 2 states satisfies this symmetry condition, Proposition 10 [R.V.] is a generalization of Proposition 7 [R.V.] mentioned above.
We give here a sketch of the rather long technical proof of a theorem by Aaronson, Gilat and Keane.

Theorem 6.1 ([A.G.K.], Corollary of Theorem 3) *Every stationary, one-dependent Markov chain with at most 4 states is a two-block-factor of an i.i.d. sequence.*

Sketch of proof.
Let P be the transition matrix and let S be the state space. The one-dependence implies that $P^2 = \Pi$ where Π is the matrix where each row is equal to the invariant measure π. We define the inner product

$$< x, y >:= \sum_{s \in S} \pi_s x_s y_s.$$

One can prove the existence of vectors x, y such that

$$p_{s,t} = \pi_t(1 + x_s y_t) \qquad \forall s, t \in S$$

and such that $\{x, y, \mathbf{1}\}$ is an orthogonal system, where $\mathbf{1}$ is the vector with all coordinates equal to 1.
Define $\alpha := \max_t x_t$, $\beta := \max_t -x_t$. Let $\Omega := \{-\beta, \alpha\}$ be a probability space with probabilities $p(-\beta) = \frac{\alpha}{\alpha+\beta}$, and $p(\alpha) = \frac{\beta}{\alpha+\beta}$. Define the random variables $a_s, b_s (s \in S)$ on Ω by

$$a_s(\omega) := \pi_s(1 + \omega y_s)$$
$$b_s(\omega) := 1 + \frac{\omega x_s}{\alpha\beta}.$$

We have $\mathsf{E}a_s = \pi_s$, $\mathsf{E}b_s = 1$, $\mathsf{E}b_s a_t = p_{s,t}$ for all $s, t \in S$ and $\sum_{s \in S} a_s(\omega) b_s(\omega^1) = 1 \ \forall \ \omega, \omega^1 \in \Omega$.
Define $\mathcal{X} := \Omega \times S^{\Omega \times \Omega}$ and let $\mathcal{U} = (U, \sigma)$ be a random variable (on some probability space) with values in $\mathcal{X}$ and distribution

$$P\left[A \times \bigcap_{K=1}^{n} [\sigma(\omega_K, \omega_K^1) = s_K]\right] = p(A) \cdot \prod_{K=1}^{n} a_{s_K}(\omega_K) b_{s_K}(\omega_K^1)$$

for all $s_1, \ldots, s_n \in S, A \subset \Omega, (\omega_1, \omega_1^1), \ldots, (\omega_n, \omega_n^1)$ different points in $\Omega \times \Omega$.
Let $\mathcal{U}_n = (U_n, \sigma_n)(n \in \mathbb{N})$ be i.i.d. random variables, each distributed as $\mathcal{U}$. Define $W_n := \sigma_n(U_n, U_{n+1})$. W_n is the desired two-block-factor. □

The number of 4 states in this Theorem is sharp, because there is ([A.G.K.], Theorem 4) an example of stationary, one-dependent Markov process with 5 states that is not a two-block-factor of an i.i.d. sequence.

Theorem 6.2 **([A.G.K.], Theorem 4)** *Let*

$$P := \begin{bmatrix} 2/5 & 1/5 & 0 & 1/10 & 3/10 \\ 1/5 & 2/5 & 1/10 & 0 & 3/10 \\ 2/5 & 0 & 1/10 & 3/10 & 1/5 \\ 0 & 2/5 & 3/10 & 1/10 & 1/5 \\ 0 & 0 & 1/2 & 1/2 & 0 \end{bmatrix}$$

be a transition matrix; its invariant measure is $\pi = (1/5, 1/5, 1/5, 1/5, 1/5)$. *The corresponding stationary Markov process is one-dependent, but is not a two-block-factor of an i.i.d. sequence.*

Sketch of proof.
The proof consists of 10 steps that we will sketch briefly (see Diagram 1). For a two-block-factor we can assume without loss of generality that the underlying i.i.d. sequence is distributed uniformly over the unit interval.
First we need some notation.
Let $\{A(s)\}_{s=1}^{5}$ be a measurable partition of the unit square $I \times I$, define $(s, t \in S = \{1, \ldots, 5\})$

$$\begin{aligned} p(s,t) &:= \int_0^1 \int_0^1 \int_0^1 1_{A(s)}(x,y) 1_{A(t)}(y,z) dx dy dz \qquad (4) \\ A_x &:= \{y \in I : (x,y) \in A\} \\ A^y &:= \{x \in I : (x,y) \in A\} \qquad (x, y \in I) \\ R(s) &:= \{y \in I : |A(s)^y| > 0\} \end{aligned}$$

We have $A(s) \subset I \times R(s) \bmod 0$ and $p(s,t) = \int_I |A(s)^y| \cdot |A(t)_y| dy$ hence

$$p(s,t) > 0 \Leftrightarrow |\{y \in R(s) : |A(t)_y| > 0\}| > 0 \qquad (5)$$

We denote both length and area by $|\cdot|$. We prove a slightly stronger statement, namely: There is no measurable partition $\{A(s)\}_{s=1}^{5}$ of the unit square such that

$$p(s,t) \begin{cases} = 0 \text{ for } (s,t) \in \{(1,3),(2,4),(3,2),(4,1),(5,1),(5,2),(5,5)\} \\ > 0 \text{ for } (s,t) \in \{(1,1),(1,2),(1,4),(2,3),(5,3),(5,4)\} \end{cases}$$

The proof is by contradiction; assume that such a partition exists, then

$$|A(s)| > 0 \text{ for all } s \in S.$$

Step 1. $R(5) \times I \subset A(3) \cup A(4) \bmod 0$. This follows directly from (5) and from $p(5,t) = 0$ for $t = 1, 2, 5$.

Step 2. $(R(5)^c \times I) \supset A(s) \bmod 0$ for $s = 1, 2$. This follows immediately from Step 1.

Step 3. $|(B \times R(5)^c) \cap A(s)| > 0$ for $s = 3, 4$ and for all $B \subset R(5)$ measurable with $|B| > 0$. If this is false e.g. for $s = 3$, then $(B \times R(5)^c) \subset A(s)$ for $s = 4$ (by Step 1) and for some $B \subset R(5)$, with $|B| > 0$. Using (4) and Step 2 this would lead to $p(4,1) > 0$. Analogously, if Step 3 does not hold for $s = 4$, this would imply $p(3,2) > 0$.

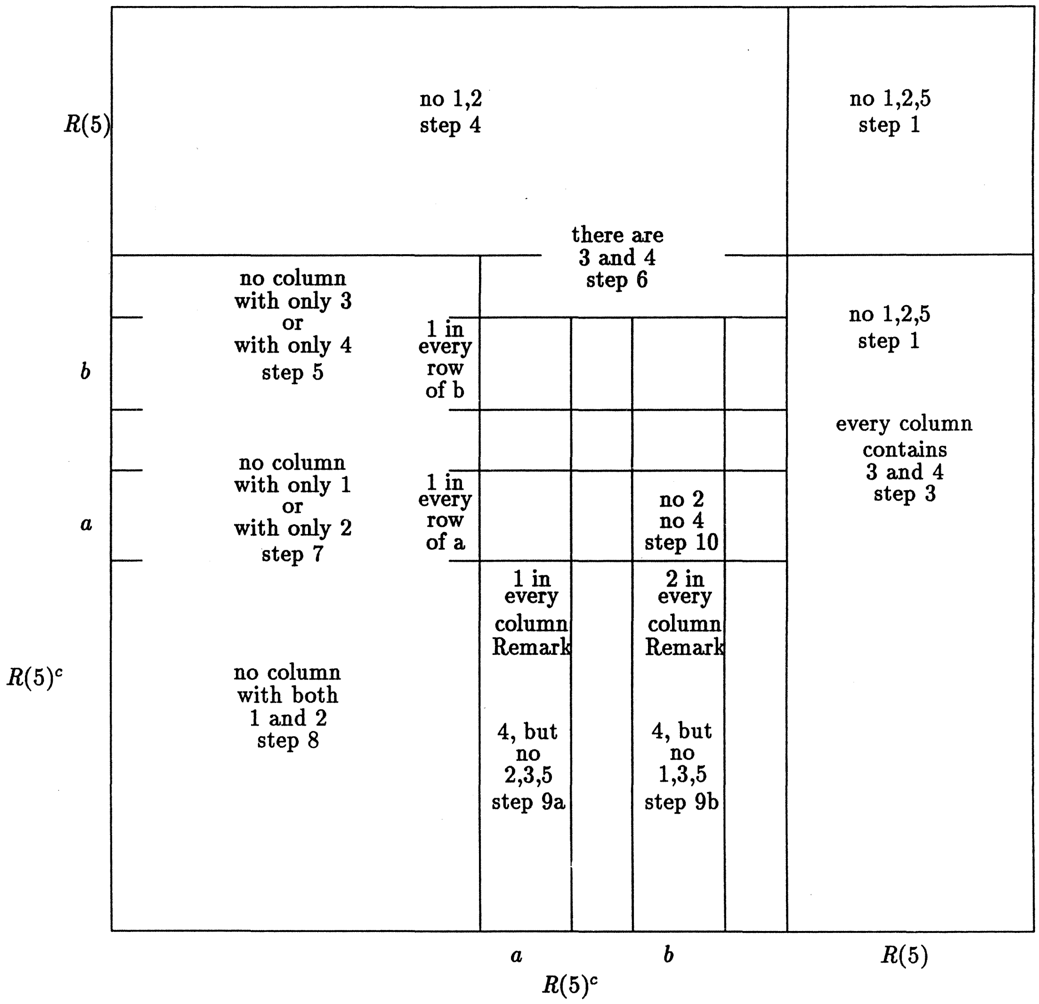

R(5)
no 1,2
step 4
no 1,2,5
step 1
there are
3 and 4
step 6
no column
with only 3
or
with only 4
step 5
b
1 in
every
row
of b
no 1,2,5
step 1
every column
contains
3 and 4
step 3
no column
with only 1
or
with only 2
step 7
a
1 in
every
row
of a
no 2
no 4
step 10
1 in
every
column
Remark
2 in
every
column
Remark
R(5)^c
no column
with both
1 and 2
step 8
4, but
no
2,3,5
step 9a
4, but
no
1,3,5
step 9b
a
b
R(5)
R(5)^c

Diagram 1

Step 4. $A(s) \subset R(5)^c \times R(5)^c$ for $s = 1,2$. If this is false for e.g. $A(1)$, then by Step 2 we would have $|(R(5)^c \times R(5)) \cap A(1)| > 0$. Using Step 3 this would lead to $p(1,4) > 0$. Analogously, if Step 4 is false for $A(2)$ then we would obtain $p(2,3) > 0$.

Step 5. $|(B \times R(5)^c) \backslash A(s)| > 0$ for $s = 3,4$ and for any measurable $B \subset R(5)^c, |B| > 0$. If this is false e.g. for $s = 3$, then $\exists B \subset R(5)^c$ such that $|B| > 0$ and $A(3) \supset B \times R(5)^c$ mod 0. Using Step 2 this leads to $p(3,2) > 0$. Analogously if Step 5 is false for $s = 4$, we would derive $p(4,1) > 0$.

Step 6. $|(R(5)^c \times I) \cap A(s)| > 0$ for $s = 3,4$. This follows (using Step 4) from $p(2,3) > 0$ and $p(1,4) > 0$.

Step 7. $|(B \times R(5)^c) \backslash A(t)| > 0$ for $t = 1,2$ and for any measurable $B \subset R(5)^c$, $|B| > 0$. If this is false for $t = 1$, then $\exists B \subset R(5)^c$ measurable, $|B| > 0$ such that $B \times R(5)^c \subset A(1)$ mod 0. Using Step 6 this would imply that $p(1,3) > 0$. Analogously if Step 7 is false for $t = 2$, we would derive $p(2,4) > 0$.

Step 8. $|A(1)_y| \cdot |A(2)_y| = 0$ a.e. on $R(5)^c$. If this is false, then $\exists B \subset R(5)^c$ measurable, $|B| > 0$, $\exists \varepsilon > 0$, such that $|A(t)_y| \geq \varepsilon \quad \forall y \in B$ for $t = 1,2$. By Step 1 $|(R(5) \times B) \cap A(s)| > 0$ for $s = 3$ or 4 (or both). This would lead to $p(3,2) > 0$ or $p(4,1) > 0$.

Remark. Because $p(1,1) > 0$ and $p(1,2) > 0$, Step 4 implies that $\exists a, b \subset R(1)$ measurable such that

$$|A(1)_y \cap R(5)^c| > 0 \text{ a.e. on } a$$

$$|A(2)_y \cap R(5)^c| > 0 \text{ a.e. on } b$$

Further note that $R(1) \subset R(5)^c$ by Step 4 and $|A(5)_y \cap R(5)^c| = 0$ a.e. because $A(5) \subset I \times R(5)$ mod 0.

Step 9a. $|A(4)_y \cap R(5)^c| > 0$, $|A(s)_y \cap R(5)^c| = 0$ $(s \neq 1,4)$ for a.e. $y \in a$. For a.e. $y \in a$ we have $|A(2)_y| = 0$ by Step 8 and $|A(3)_y| = 0$ because of $p(1,3) = 0$. Now $|A(5)_y \cap R(5)^c| = 0$ together with Step 7 implies $|A(4)_y \cap R(5)^c| > 0$, for a.e. $y \in a$.

Step 9b. $|A(4)_y \cap R(5)^c| > 0$, $|A(s)_y \cap R(5)^c| = 0$ $(s \neq 2,4)$ for a.e. $y \in b$. For a.e. $y \in b$ we have $|A(1)_y| = 0$ by Step 8 and $|A(3)_y| = 0$ because of $p(1,3) = 0$. Now $|A(5)_y \cap R(5)^c| = 0$ together with Step 7 implies $|A(4)_y \cap R(5)^c| > 0$, for a.e. $y \in b$.

Step 10. The contradiction. By Step 9b we have $b \times a \subset A(2) \cup A(4)$. But if $|(b \times a) \cap A(2)| > 0$, we would derive from Step 9a that $p(2,4) > 0$. Further if $|(b \times a) \cap A(4)| > 0$, we would derive from Step 9a that $p(4,1) > 0$. Once more a contradiction.

This completes the proof of the theorem.

$\square$

Remark. According to Matúš ([Ma.4]) this theorem also holds if we take as transition matrix

$$\frac{1-\varepsilon}{10}\begin{pmatrix} 4 & 2 & 0 & 1 & 3 \\ 2 & 4 & 1 & 0 & 3 \\ 4 & 0 & 1 & 3 & 2 \\ 0 & 4 & 3 & 1 & 2 \\ 0 & 0 & 5 & 5 & 0 \end{pmatrix} + \frac{\varepsilon}{5}\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

for $0 \leq \varepsilon < 10^{-6}$. This implies that there exist stationary one-dependent Markov processes with only positive transition probabilitites that are not two-block-factors of i.i.d. sequences.

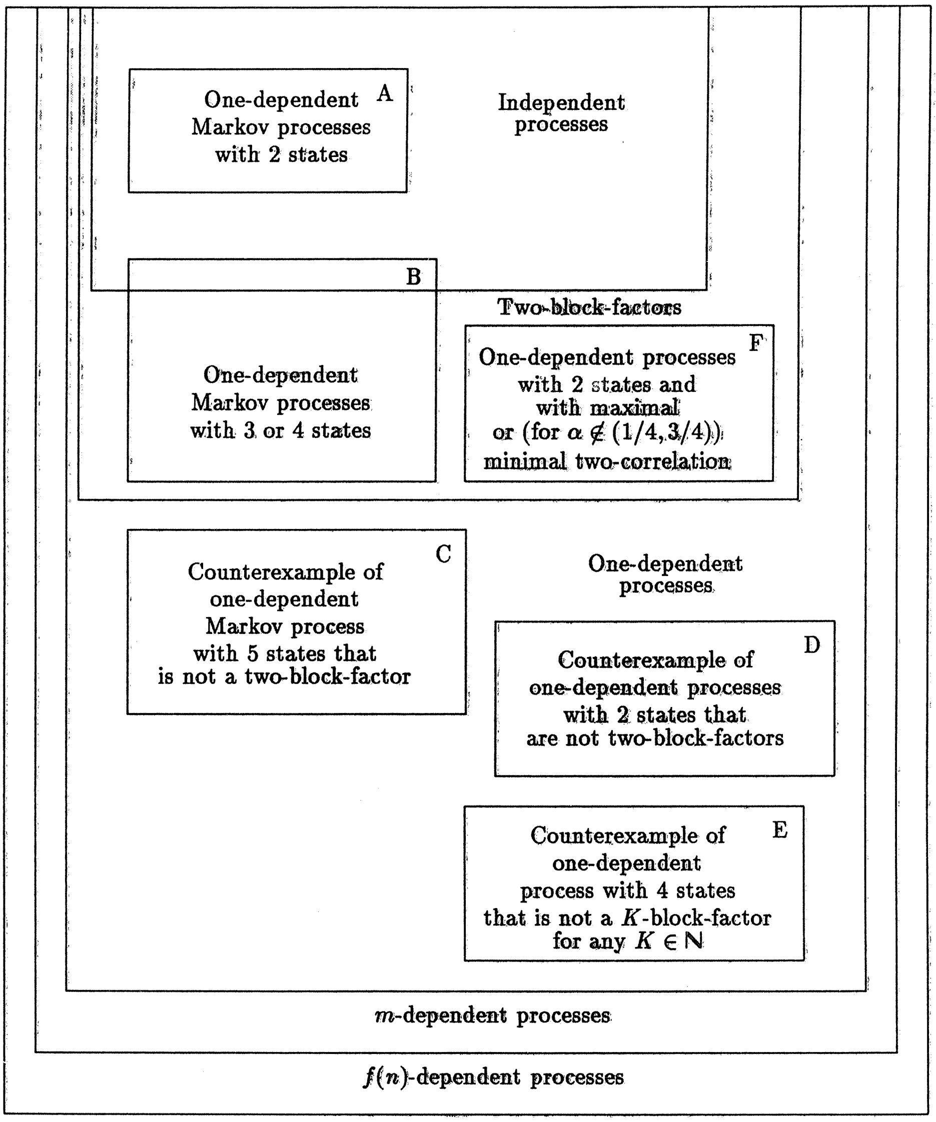
7 A conjecture

It is obvious to ask under which conditions m-dependent processes are $m + 1$-block-factors. If the m-dependent process is a Gaussian process, then it is necessarily an $m + 1$-block-factor, because there is a one to one correspondence between Gaussian stationary processes $(X_N)_{N\in\mathbb{Z}}$ and autocovariance functions R_N. Given such a process, there exists a positive definite function $R_N := \mathsf{E}(X_N X_0)$, and given a positive definite function R_N, there exists a unique Gaussian process with this autocovariance function. Now the notion of m-dependence means that $R_N = 0$ for $|N| > m$. These functions correspond to the set of $m + 1$-block-factors defined by

$$X_N := t_0 U_N + t_1 U_{N+1} + \dots + t_m U_{N+m}$$

where $(U_N)_{N\in\mathbb{Z}}$ is an i.i.d. sequence of Gaussian random variables. Although this was conjectured for quite a long time, a one-dependent process $(X_N)_{N\in\mathbb{Z}}$ is not necessarily a two-block-factor if $(X_N)_{N\in\mathbb{Z}}$ is not a Gaussian process. This has been stated yet by Ibragimov and Linnik ([Ibr.Li.]) in 1971, but unfortunately they did not give a counterexample to this conjecture. This conjecture appeared also in several other publications; [Be.], [G.H.2], [Ja.1-2] and [O'Ci.]. Several authors used this conjecture as a hypothesis. Janson ([Ja.2]) studied runs of ones in m-dependent processes. He proved his results only for $m + 1$-block-factors and he remarked that this is sufficient under this hypothesis. Later Van den Berg ([Be.]) and O'Cinneide ([O'Ci.]) also studied runs of ones, and they proved some of their results only for $m + 1$-block-factors. Götze and Hipp ([G.H.3]) and Heinrich ([He.5]) proved some of their local limit theorems and central limit theorems for m-dependent random fields only for block-factors. The results in the articles [Be.], [G.H.1-3], [He.1-6], [Ja.1-2] and [O'Ci] are essentially different from those of this monograph. In 1987 Aaronson and Gilat ([A.G.]) found a one-parameter-family of counterexamples. Later, in collaboration with Keane and De Valk ([A.G.K.V.], the second article of this monograph), they found a two-parameter-family. These counterexamples are all $0 - 1$ valued one-dependent processes where a run of three ones has probability zero.

In section 6 we showed a recent example (by Aaronson, Gilat and Keane, see [A.G.K.], 1992) of a one-dependent Markov chain (assuming only 5 values) that is not a two-block-factor. In section 4 we showed an even more recent example (by Burton, Goulet and Meester, see [B.G.M.]) of a 4-valued one-dependent process that is not a K-block-factor



A See [R.V.]
B See [A.G.K.]
C See [A.G.K.]
D See [A.G.K.V.]
E See [B.G.M.]
F See [G.K.V.]

Diagram 2

for any $K \in \mathbb{N}$. In the fifth article of this monograph the construction of the counterexamples from [A.G.K.V.] is generalized by representing one-dependent processes in terms of Hilbert spaces, vectors and bounded linear operators on Hilbert spaces. All m-dependent processes admit a Hilbert space representation. The dimension of the smallest Hilbert space that represents a process is a measure for the complexity of the structure of the process. The difference between two-block-factors and non-two-block-factors seems to be determined by the geometry of cones that are invariant under certain operators.

We summarize some facts in Diagram 2 (the definition of $f(N)$-dependence will be given in section 10).

8 Two-correlations and the conjecture

Although the conjecture does not hold generally, it is true under certain extremal conditions on $0-1$ valued one-dependent processes.

Fix an α in the unit interval and consider the subclass of $0-1$ valued one-dependent processes with probability of a one equal to α. In [G.K.V.] (the third article of this monograph) is proved that in this subclass the probability of a run of two ones (a *two-correlation*) has maximal value $\alpha^{\frac{3}{2}}$ (if $\alpha \geq \frac{1}{2}$) and $2\alpha - 1 + (1-\alpha)^{\frac{3}{2}}$ (if $\alpha \leq \frac{1}{2}$). This supremum is attained uniquely if α is not equal to $\frac{1}{2}$, and for $\alpha = \frac{1}{2}$ there exist exactly two processes with maximal two-correlation. The processes with maximal two-correlation are all two-block-factors. Further, a $0-1$ valued one-dependent process with minimal two-correlation (for fixed α) is necessarily a two-block-factor if $\alpha \notin (\frac{1}{4}, \frac{3}{4})$ ([G.K.V.]).

9 More two-correlations and applications

The maximal two-correlation of two-block-factors (translated to our terminology) was computed by Katz ([Ka.]) and later by Finke ([Fi.]), who interpreted Katz' mathematical objects as two-correlations in stochastic processes. The minimal two-correlation of two-block-factors is computed in [V.1] (the first article in this monograph). A rather sharp lowerbound $\frac{\alpha}{3}(4\alpha-1)$ for the minimal two-correlation of two-block-factors was computed by Matúš and Tuzar ([M.T.], see also [Tu.]) in a remarkable elementary way. Their lower bound is very close to the minimal two-correlation when α is close to $\frac{1}{2}$. These two-correlations have applications to matrix theory and graph theory, when we restrict our attention to $0-1$ valued one-dependent processes that are two-block-factors of an independent sequence of random variables, uniformly distributed over a finite number of values. The problem of the maximal or minimal value of a two-correlation in this discretized setting is equivalent to the problem of finding the maximal or minimal number of paths of length two in a directed graph (as was remarked in [Fi.]) with a fixed number of edges and vertices. This problem is also equivalent to finding the maximal or minimal value of $||M^2||$ over the class of $0-1$ valued $N \times N$ matrices M with K ones (for fixed N and K). This problem is solved in [V.3] (the fourth article of this monograph).

10 Other publications on m-dependence

Although the following articles consider different problems than those dealt within this monograph, they are mentioned to give a survey over the field of m-dependence. Hoeffding and Robbins ([Ho.Ro.]) have studied $f(N)$-dependent processes, i.e. processes $(X_N)_{N=1}^{\infty}$ such that

$$\{X_1, ..., X_{K_1}\} \quad \text{and} \quad \{X_{K_2}, ..., X_N\} \quad \text{are independent}$$

$$\text{whenever} \quad K_2 - K_1 > f(N), \quad \text{for some function} \quad f.$$

When f is constant, then we have m-dependence. They proved entral limit theorems for these processes.
There is a lot of literature on central limit theorems (and related limit theorems) for m-dependent processes and m-dependent random fields; by e.g. Diananda ([Di.1-3]), Götze and Hipp ([G.H.1-3]), Guyon and Richardson ([Gu.Ri.]), Heinrich ([He.1-6]), Petrov ([Pe.]), Prakasa Rao ([P.R.]), Shergin ([Sh.]), Takahata ([Ta.]) and Tikhomirov ([Ti.]). Haiman ([Ha.1-2]), Newell ([Ne.]) and Watson ([W.]) wrote about extreme value theory for m-dependent processes. Janson studied renewal theory ([Ja.1]) and runs ([Ja.2]) in m-dependent processes. Smorodinsky ([Sm.]) proved that stationary m-dependent processes of the same entropy are finitarily isomorphic.
Tsirelson ([Ts.]) wrote recently a paper on the connection between inequalities for quantum theory, for partition functions in statistical physics and for one-dependent processes (as in [A.G.K.V.] and [G.K.V.]).
Recently Matúš ([Ma.4]) proved that a stationary process $(X_N)_{N \in \mathbb{N}}$ is equal in distribution to a two-block-factor of an i.i.d. sequence if and only if there exists a jointly exchangeable and dissociated array $(Z_{N,M})_{N,M \in \mathbb{N}}$ such that its superdiagonal $(Z_{N,N+1})_{N \in \mathbb{N}}$ is equal in distribution to $(X_N)_{N \in \mathbb{N}}$. An array $(Z_{N,M})_{N,M \in \mathbb{N}}$ is called jointly exchangeable if its distribution is equal to the distribution of $(Z_{\pi(N),\pi(M)})_{N,M \in \mathbb{N}}$ for every permutation π which moves only a finite number of positive integers. $(Z_{N,M})_{N,M \in \mathbb{N}}$ is called dissociated if $(Z_{N,M})_{N,M<K}$ is independent of $(Z_{N,M})_{N,M \geq K}$ for every $K > 1$. As a consequence in [Ma.4] is proved that the class of two-block-factors is closed w.r.t. the weak topology, hence two-block-factors are not dense in the class of one-dependent processes.

11 Comment on the six articles

I. [V.1] "*The maximal and minimal* 2*-correlation of a class of* 1*-dependent* $0-1$ *valued processes*"

In this article we consider $0-1$ valued two-block-factors $(X_N)_{N \in \mathbb{Z}}$ of an independent sequence $(U_N)_{N \in \mathbb{Z}}$ of random variables that are uniformly distributed over the unit interval. Because such two-block-factors are completely determined by the indicator function of a subset A of the unit square, defining

$$X_N := 1_A(U_N, U_{N+1}),$$

these processes are also called *indicator processes.* The probability of a one is equal to the Lebesgue measure of A, and the probability of a run of two ones (a *two-correlation*) is equal to

$$I_A := \int_0^1 H_A(x) \cdot V_A(x)\,dx$$

where H_A and V_A are the horizontal and vertical sections of A. The computation of the least possible two-correlation (for fixed probability of a one) over the class of $0-1$ valued two-block-factors turns out to be a variational problem, equivalent to computing the minimal value of I_A for fixed Lebesgue measure of A. This problem gives rise to some questions (see also section 12), some of which are solved in [G.K.V.]. The articles [G.K.V.] and [V.3] can be considered as continuations of this article.

II. [A.G.K.V.] *"An algebraic construction of a class of one-dependent processes" (with J. Aaronson, D. Gilat and M.S. Keane)*

In this article a rather old conjecture is disproved. The authors construct in an algebraic way a continuum number of $0-1$ valued stationary one-dependent processes that are not two-block-factors of i.i.d. sequences, and in this way they disprove the conjecture that each one-dependent process is a two-block-factor. All these counterexamples have the property that a run of three ones has probability zero. The class of counterexamples is parametrized by α (the probability of a one) and β (the probability of a run of two ones; a two-correlation). These parameters (together with the fact that a run of three ones has probability zero and the property of one-dependence) uniquely determine the measure of all cylinder sets. To determine for which values of the parameters a process exists, it is enough to check whether the measures of all cylinder sets are non-negative. This turns out to be equivalent to the problem whether the orbit of $(1,1)$ under successive applications of certain mappings φ_0 and $\varphi_1 : \mathbb{R}^2 \to \mathbb{R}^2$ in any order always remains in the unit square. It is known for which values of α and β a two-block-factor exists (by methods as in [V.1]) and it turns out that there exists a two-parameter-family of counterexamples to the conjecture.

The construction of these counterexamples is generalized in [V.4], that is inspired by this article.

III. [G.K.V.] *"Extremal two-correlations of two-valued stationary one-dependent processes" (with A. Gandolfi and M.S. Keane)*

This article can be considered as a continuation of [V.1].

The authors compute the maximal value of a two-correlation (probability of a run of two ones) over the class of $0-1$ valued, stationary, one-dependent processes. This is a simplification and generalization of [Ka.], where the maximal two-correlation over the class of two-block-factors was computed. The authors prove that this supremum is uniquely attained when the fixed probability of a one is not equal to $\frac{1}{2}$, and that there exist exactly two processes with maximal two-correlation when the fixed probability of a one is equal to $\frac{1}{2}$. The processes with maximal two-correlation are all two-block-factors.

Further, the minimal two-correlation over the class of $0-1$ valued, stationary, one-dependent processes is computed in the case that the fixed probability of a one is $\leq \frac{1}{3}$ or $\geq \frac{2}{3}$. The computed lower bound is the same as the minimal two-correlation over the class of two-block-factors ([V.1]). In the case that the fixed probability of a one is $\leq \frac{1}{4}$ or $\geq \frac{3}{4}$ it is proved that the infimum over the class of one-dependent processes is uniquely attained, and the corresponding processes are all two-block-factors. The upper- and lower-bounds for the two-correlation are computed by showing that the measure of some cylinder sets becomes negative when we assume that the two-correlation has a value greater than the upper-bound c.q. smaller than the lower-bound. So the computation is probabilistic, in contrast to the analytic and combinatoric computation in [V.1].

IV. [V.3] *"A problem on $0-1$ matrices"*

In terms of matrices the maximal and minimal value of $\|M^2\|$ is computed over the class of $0-1$ valued $N \times N$ matrices M with K entries equal to one (for fixed N and K). In terms of one-dependent processes, the maximal and minimal value of the two-correlation over the class of $0-1$ valued two-block-factors of the N-shift (for fixed N and fixed probability of a one) is computed. This article can be considered as a discretized version of [V.1]. In terms of graphs, this corresponds to the maximal and minimal number of different paths of length two in a directed graph with N vertices and K edges (for fixed N and K). The solution is found by means of analysis and combinatorics.

V. [V.4] *"Hilbert space representations of m-dependent processes"*

This article can be considered as a continuation of [A.G.K.V.]. The construction in [A.G.K.V.] of one-dependent processes that are not two-block-factors, is generalized by a representation of one-dependent processes in terms of Hilbert spaces, vectors and bounded linear operators on Hilbert spaces. Moreover all m-dependent processes admit a representation.

If there is in the Hilbert space a closed convex cone that is invariant under certain operators and that is spanned by a finite number of linearly independent vectors, then the corresponding process is a two-block-factor of an independent process. Apparently the geometry of invariant cones determines the difference between two-block-factors and non-two-block-factors. The dimension of the smallest Hilbert space that represents a process is a measure for the complexity of the structure of the process. One-dependent processes, represented by a one-dimensional Hilbert space, are i.i.d. sequences. One-dependent processes, represented by a 2-dimensional Hilbert space, are two-block-factors. The counterexamples from [A.G.K.V.] fit with a 3-dimensional Hilbert space. If a two-valued one-dependent process has a cylinder with measure equal to zero, then this process can be represented by a Hilbert space with dimension smaller than or equal to the length of this cylinder. In the two-valued case a cylinder (with measure equal to zero) whose length is minimal and ≤ 7, is symmetric. We conjecture that all minimal zero-cylinders are symmetric and we give examples of minimal zero-cylinders.

VI. [R.V.] *" On regression representations of stochastic processes" (with L. Rüschendorf)*

In this article we construct almost sure nonlinear regression representations of general stochastic processes $(X_n)_{n\in\mathbb{N}}$. Given a process X we construct an i.i.d. sequence $(U_n)_{n\in\mathbb{N}}$ and a sequence of functions $(f_n)_{n\in\mathbb{N}}$ such that

(i) $X_n = f_n(X_1, \dots, X_{n-1}, U_n)$ a.s. for all $n \in \mathbb{N}$
and

(ii) U_n is independent of $(X_1, \dots, X_{n-1})$.

We call (i) the Markov Regression of X.

In this paper we also present the Standard Representation $X_n = g_n(U_1, \dots, U_n)$ of an arbitrary process by constructing functions $(g_n)_n$ and an i.i.d. sequence $(U_n)_n$ for a given process $(X_n)_n$. If X is an m-Markov process, then the Markov Regression reduces to $X_n = f_n(X_{n-m}, \dots, X_{n-1}, U_n)$. Assume that X is a generalized m-block-factor of U; i.e. $X_n = g_n(U_{n-m+1}, \dots, U_n)$. We can ask the question whether the Standard Representation of X gives us $(U_n)_n$ and $(g_n)_n$ in return. If this would always be the case, then we would have a method to check whether a process is an m-block-factor or not. Unfortunately we can only prove this for a special case; namely the monotone block-factors.

12 Open problems and conjectures

In this section we give a list of open problems and conjectures, to which this monograph gives rise.

On [V.1] and [G.K.V.].

1. Is the value of the minimal two-correlation (for fixed probability of a one) over the class of $0-1$ valued two-block-factors (as in [V.1]) equal to the value of the minimal two-correlation over the class of $0-1$ valued one-dependent processes? In [G.K.V.] this problem is solved in the case that the fixed probability of a one is $\leq \frac{1}{3}$ or $\geq \frac{2}{3}$. It seems that this problem becomes more and more complicated when the fixed probability of a one tends to $\frac{1}{2}$.

2. If the answer to question (1) is yes, are the one-dependent processes with minimal two-correlation all two-block-factors? In [G.K.V.] this problem is solved in the case that the fixed probability of a one is $\leq \frac{1}{4}$ or $\geq \frac{3}{4}$. Just as question (1), it seems that this problem becomes more and more complicated when the fixed probability of a one approaches $\frac{1}{2}$. In particular we do not know whether the minimal two-correlation is equal to $\frac{1}{6}$ when the fixed probability of a one is $\frac{1}{2}$ (question (1)), and if the answer to this question is yes, we do not know whether this minimum is uniquely attained in the following process $(X_N)_{N \in \mathbf{Z}}$ (question (2)). Let $(Y_N)_{N \in \mathbf{Z}}$ be an i.i.d. sequence of random variables, uniformly distributed over the unit interval. Let $X_N := 0$ if $Y_N < Y_{N+1}$ and $X_N := 1$ if $Y_N \geq Y_{N+1}$. This problem seems to be interesting in the theory of order-statistics.

3. Can the computation of the minimal two-correlation in [V.1] be simplified, just as the computation of the maximal two-correlation in [Ka.] is simplified (and generalized) in [G.K.V.]? The elementary computation of the lower bound in [M.T.] seems to be a first step in the direction of a simplified proof. The lower bound in [M.T.] is very close to the minimal two-correlation when the fixed probability of a one is close to $\frac{1}{2}$.

4. The computation in [V.1] is not probabilistic but analytic and combinatoric. Can the computation in [V.1] be "probabilized", just as [Ka.] is probabilized by [G.K.V.]?

5. What extremal conditions on N-correlations (the probability of a run of N ones) are needed to assure that two-valued m-dependent processes are always $m+1$-block-factors?

   On [A.G.K.V.].

6. The counterexamples of one-dependent processes that are not two-block-factors are constructed in an algebraic way. Can they be constructed in a probabilistic way, such that their structure becomes more natural and clear (can the counterexamples be probabilized)?

7. Are these counterexamples m-block-factors of i.i.d. sequences for some $m \geq 3$?

8. Do there exist two-valued counterexamples, not having the property that a run of three ones has probability zero, or even having the property that each cylinder set has positive measure?

9. For which values of the parameters α and β do there exist processes in the "unexplored area"? It seems that this problem becomes more and more complicated when (α, β) approaches $(\frac{1}{3}, \frac{1}{27})$.

10. Can the counterexamples be described as limits of a rescaling operation (see [O'Br.]) of a mixing process?

11. Are the counterexamples functions of Markov processes, or even functions of m-dependent Markov processes?

12. Do there exist m-dependent processes (for some $m \geq 2$) that are not $m+1$-block-factors, and that are not $m-1$-dependent?

    On [V.3].

13. Can the computation of $\mathrm{Max}(N,K)$ and $\mathrm{Min}(N,K)$ be more straightforward? Methods as used in the computation of the lower bound in [M.T.] might be of some help. There exist values of N and K such that $\mathrm{int}\{N^3 \cdot \mathrm{Max}(K/N^2)\} > \mathrm{Max}(N,K)$ and other values such that $1+\mathrm{int}\{N^3 \cdot \mathrm{Min}(K/N^2)\} < \mathrm{Min}(N,K)$ ($\mathrm{int}(x)$ is the integer part of x), and therefore it is not possible to prove the maximality or minimality of some matrix M by stating that I_M (an integer) is in this case the best integer approximation (the entier) to $N^3 \cdot \mathrm{Max}(K/N^2)$ (in the maximum case), c.q. the best integer approximation (one + the entier) to $N^3 \cdot \mathrm{Min}(K/N^2)$ (in the minimum case). Note that always: $N^3 \cdot \mathrm{Max}(K/N^2) \geq \mathrm{Max}(N,K)$ and $N^3 \cdot \mathrm{Min}(K/N^2) \leq \mathrm{Min}(N,K)$.

    On [V.4].

14. The essential difference between two-block-factors and one-dependent processes that are not two-block-factors seems to be determined by the geometry of the invariant cone. What are the crucial aspects of the geometry of the invariant cone that determine this difference?

15. Can a $0-1$ valued one-dependent process have no other minimal zero-cylinders than [101], [010], $[1^N]$, and $[0^N]$ ($N \in \mathbb{N}$)? The minimal dimensions are $2, 2, N$ and N respectively.

16. Do there exist for any $N \in \mathbb{N}(N \geq 3)$ a one-dependent process, that is not a two-block-factor, with minimal dimension equal to N, and without zero-cylinders?

17. Do there exist for any $N \in \mathbb{N}(N \geq 3)$ a one-dependent process, that is not a two-block-factor, with minimal dimension equal to N, and with a minimal zero-cylinder with length N?

18. Do there exist for any $N \in \mathbb{N}(N \geq 1)$ a two-block-factor with minimal dimension equal to N, and without zero-cylinders?

19. Do there exist for any $N \in \mathbb{N}(N \geq 1)$ a two-block-factor with minimal dimension equal to N, and with a minimal zero-cylinder with length N?

20. Are one-dependent processes always functions of Markov processes, or even functions of m-dependent Markov processes?

21. Do there exist one-dependent m-block-factors ($m \geq 3$) that can not be written as a two-block-factor?

22. Is a one-dependent process with an m-dimensional Hilbert space representation always an m-block-factor ($m \geq 3$)?

23. Under which conditions is a one-dependent Markov process necessarily a two-block-factor?

24. Are the two-block-factors extreme points of the set of one-dependent processes?

25. Do there exist two-dependent processes that are not two-block-factors of one-dependent processes?

    On [R.V.].

26. Is an m-dependent process $(X_N)_{N\in\mathbb{Z}}$ always a finitary-block-factor of an i.i.d. sequence $(Y_N)_{N\in\mathbb{Z}}$; i.e. $X_N = f_N(Y_N, ..., Y_{N+K_N})$ for some sequence of integers $(K_N)_{N\in\mathbb{Z}}$ and some sequence of functions $(f_N)_{N\in\mathbb{Z}}$?

27. Under which conditions on an m-block-factor does the Standard Representation construction return the m-block-factor representation?

28. How restrictive is the condition of monotonicity of a two-block-factor?

13 References

[A.G.] J. Aaronson and D.Gilat, On the structure of stationary one-dependent processes, preprint, School of Mathematical Sciences, Tel Aviv University, Israel, 1987.

[A.G.K.] J. Aaronson, D. Gilat and M.S. Keane, On the structure of 1-dependent Markov shifts, *J. Theoretical Probab.* **5**, 1992, 545–561.

[A.G.K.V.] J. Aaronson, D. Gilat, M.S. Keane and V. de Valk, An algebraic construction of a class of one-dependent processes, *Ann. Probab.* **17**, 1989, 128–143, (also in this monograph).

[Be.] J. van den Berg, On some results by S. Janson concerning runs in m-dependent sequences, preprint, 1986.

[B.G.M.] R.M. Burton, M. Goulet, and R.W.J. Meester, On one-dependent processes and K-block-factors, 1992, to appear in *Ann. Probab.*

[C.] K.L. Chung, Markov chains with stationary transition probabilities, 2nd edition, Springer, Berlin, Heidelberg, 1967.

[Di.1] P.H. Diananda, Some probability limit theorems with statistical applications, *Proc. Cambridge Phil. Soc.* **49**, 1953, 239–246.

[Di.2] P.H. Diananda, The central limit theorem for $m$dependent variables asymptotically stationary to second order, *Proc. Cambridge Phil. Soc.* **50**, 1954, 287–292.

[Di.3] P.H. Diananda, The central limit theorem for m-dependent variables, *Proc. Cambridge Phil. Soc.* **51**, 1955, 92–95.

[F.N.S.] M. Fannes, B. Nachtergaele and L. Slegers, Functions of Markov processes and algebraic measures, *Rev. Math. Phys.* **4**, 1992, 39–64.

[Fi.] L. Finke, Two maximization problems, paper submitted to Oregon State University in partial fulfillment of the requirements for the degree of Master of Arts, 1982.

[G.K.V.] A. Gandolfi, M.S. Keane and V. de Valk, Extremal two-correlations of two-valued stationary one-dependent processes, *Probab. Theory Related Fields* **80**, 1989, 475–480, (also in this monograph).

[G.H.1] F. Götze and C. Hipp, Asymptotic expansions for sums of weakly dependent random vectors, *Z. Wahrsch. Verw. Gebiete* **64**, 1983, 211–239.

[G.H.2] F. Götze and C.Hipp, Asymptotic expansions for potential functions of i.i.d. random fields, *Probab. Theory Related Fields* **82**, 1989, 349–370.

[G.H.3] F. Götze and C. Hipp, Local limit theorems for sums of finite range potentials of a Gibbsian random field, *Ann. Probab.* **18**, 1990, 810–828.

[G.] M. Goulet, One-dependence and K-block-factors, Ph.D. Thesis, Oregon State University.

[Gu.Ri.] X. Guyon and S. Richardson, Vitesse de convergence du théorème de la limite centrale pour des champs faiblement dépendants, *Z. Wahrsch. Verw. Gebiete* **66**, 1984, 297–314.

[Ha.1] M.G. Haiman, Valeus extrémales de suites stationnaires de variables aléatoires m-dépendantes, *Ann. Inst. H. Poincaré Sec.* **B 17**, 1981, 309–330.

[Ha.2] M.G. Haiman, Étude des extrêmes d'une suite stationnaire m-dépendante avec une application relative aux accroissements du processus de Wiener, *Ann. Inst. H. Poincaré Prob. Stat.* **23**, 1987, 425–458.

[Han] D.L. Hanson, On the representation problem for stationary stochastic processes with trivial tail field, *J. Math. Mech.* **12**, 1963, 293–301.

[H.L.P.] G.H. Hardy, J.E. Littlewood and G. Pólya, Inequalities, Cambridge University Press, 1934.

[He.1] L. Heinrich, A method for the derivation of limit theorems for random fields with finite range dependence, *Z. Wahrsch. Verw. Gebiete* **60**, 1982, 501–515.

[He.2] L. Heinrich, Non-uniform estimates and asymptotic expansions of the remainder in the central limit theorem for m-dependent random variables, *Math. Nachr.* **115**, 1984, 7–20.

[He.3] L. Heinrich, Stable limit theorems for sums of multiply indexed m-dependent random variables, *Math. Nachr.* **127**, 1986, 193–210.

[He.4] L. Heinrich, Asymptotic expansions in the central limit theorem for a special class of m-dependent random fields I, *Math. Nachr.* **134**, 1987, 83–106.

[He.5] L. Heinrich, Asymptotic expansions in the central limit theorem for a special class of m-dependent random fields II - Lattice case, *Math. Nachr.* **145**, 1990, 309–327.

[He.6] L. Heinrich, Non-uniform bounds for the error in the central limit theorem for random fields generated by functions of independent random variables, *Math. Nachr.* **145**, 1990, 345–364.

[Ho.Ro] W. Hoeffding and H. Robbins, The central limit theorem for dependent random variables, *Duke Math. J.* **15**, 1948, 773–780.

[Ibr.Li.] I.A. Ibragimov and Y.V. Linnik, Independent and stationary sequences of random variables, Wolters Noordhoff, Groningen, 1971.

[Ja.1] S. Janson, Renewal theory for m-dependent variables, *Ann. Probab.* **11**, 1983, 558–568.

[Ja.2] S. Janson, Runs in m-dependent sequences, *Ann. Probab.* **12**, 1984, 805–818.

[Ka.] M. Katz, Rearrangements of $(0-1)$matrices, *Israel J. Math* **9**, 1971, 53–72.

[Kh.1] A. Khintchine, Über eine Ungleichung, *Mat. Sb.* **39**, 1932, 35–39.

[Kh.2] A. Khintchine, Über ein metrisches Problem der additieven Zahlentheorie, *Mat Sb.* **40**, 1932, 180–189.

[Lo.] G.G. Lorentz, A problem of plane measure, *Amer. J. Math.* **71**, 1949, 417–426.

[Lux.] W.A.J. Luxemburg, On a inequality of A. Khintchine for zero-one matrices, *J. Comb. Th. A.* **12**, 1972, 289–296.

[Ma.1] F. Matúš, Sur les chemins de longeur deux et les produits d'un espace probabilisé, research report nr. 1586, ÚTIA ČSAV, Prague, 1989.

[Ma.2] F. Matúš, Inequalities concerning the demi-degrees and numbers of paths, research report nr. 1652, ÚTIA ČSAV, Prague, 1990.

[Ma.3] F. Matúš, Independence structures modelling stochastic and algebraic independences and connectedness, submitted to *Theoretical Comp. Sc.*

[Ma.4] F. Matúš, On two-block-factor sequences and one-dependence, submitted to *J. Theor. Probab.*

[M.T.] F. Matúš and A. Tuzar, Short proofs of Khintchine-type inequalities for zero-one matrices, *J. Comb. Th. series A.* **59**, 1992, 155–159.

[Ne.] G.F. Newell, Asymptotic extremes for m-dependent random variables, *Ann. Math. Stat.* **35**, 1964, 1322–1325.

[N.S.] C.M. Newman, and L.S. Schulman, Infinite clusters in percolation theory, *J. Stat. Phys.* **26**, 1981, 613–628.

[O'Br.] G.L. O'Brien, Scaling transformations for $\{0,1\}$-valued sequences, *Z. Wahrsch. Verw. Gebiete* **53**, 1980, 35–49.

[O'Ci.] C.A. O'Cinneide, Some properties of one-dependent sequences, preprint, 1987.

[Pe.] V.V. Petrov, On the central limit theorem for m-dependent variables, *Selected Transl. Math. Statist. and Prob.* **9**, 1970, 83–88.

[Pz.] D. Petz, A de Finetti-type Theorem with m-dependent states, *Probab. Theory Related Fields* **85**, 1990, 65–72.

[P.R.] B.L.S. Prakasa Rao, A non-uniform estimate of the rate of convergence in the central limit theorem for m-dependent random fields., *Z. Wahrsch. Verw. Gebiete* **58**, 1981, 247–256.

[R.R.] S.T. Rachev and L. Rüschendorf, A transformation property of minimal metrics, *Theor. Probab. Appl.* **35**, 1990, 131–137.

[Ro.1] M. Rosenblatt, Remarks on a multivariate transformation, *Ann. qMath. Stat.* **23**, 1953, 470–472.

[Ro.2] M. Rosenblatt, Stationary Markov chains and independent random variables, *J. Math. Mech.* **9**, 1960, 945–949.

[Ro.3] M. Rosenblatt, Addendum to "Stationary Markov chains and independent random variables", *J. Math. Mech.* **11**, 1962, 317.

[Rü.] L. Rüschendorf, Stochastically ordered distributions and monotonicity of the OC-function of sequential probability ratio tests, *Math. Operationforsch. Ser. Statist.* **12**, 1981, 327–338.

[R.V.] L. Rüschendorf and V. de Valk, On regression representations of stochastic processes, *Stochastic Processes and their Applications* **46**, 1993, 183–198. (also in this monograph).

[Sh.] V.V. Shergin, On the convergence rate in the central limit theorem for m-dependent random variables, *Theory of Probability and its Applications* **24**, 1979, 782–796.

[Sk.] A.V. Skorohod, On a representation of random variables, *Theory Probab. Appl.* **21**, 1976, 628–631.

[Sm.] M. Smorodinsky, Finitary isomorphism of m-dependent processes, *Contemporary Mathematics* **135**, 1992, 373–376.

[Ta.] H. Takahata, On the rates in the central limit theorem for weakly dependent random fields, *Z. Wahrsch. Verw. Gebiete* **64**, 1983, 445–456.

[Ti.] A.N. Tikhomirov, On the convergence rate in the central limit theorem for weakly dependent random variables, *Theory of Probability and its Applications* **25**, 1980, 790–809.

[Ts.] B.S. Tsirelson, A new framework for old Bell inequalities, preprint, Tel Aviv University, 1993.

[Tu.] A. Tuzar, Remark to a problem on 0–1 matrices, *Compositio Math.* **86**, 1993, 97–100.

[V.1] V. de Valk, The maximal and minimal 2-correlation of a class of 1-dependent $0-1$ valued processes, *Israel J. Math.* **62**, 1988, 181–205, (also in this monograph).

[V.2] V. de Valk, One-dependent processes, dissertation, Delft University of Technology, Delft University Press, 1988.

[V.3] V. de Valk, A problem on $0-1$ matrices, *Compositio Math.* **71**, 1989, 139–179, (also in this monograph).

[V.4] V. de Valk, Hilbert space representations of m-dependent processes, to appear in *Ann. Probab.*, (also in this monograph).

[V.5] V. de Valk, Two-fold-mixing implies N-fold-mixing for all N for processes with compact Hilbert space representation, preprint 637, State University of Utrecht, 1991.

[W.] G.S. Watson, Extreme values in samples from m-dependent stationary stochastic processes, *Ann. Math. Stat.* **25**, 1954, 798–800.

[Z.] A. Zygmund, Trigonometric series, 2nd edition, Cambridge University Press, Cambridge, U.K., 1959.

14 Summary

This monograph consists of six articles on one-dependent processes. Therefore, the subject is in the first place probability theory, although the methods and applications not only appear in probability theory, but also in statistical physics, analysis, functional analysis, dynamical systems, variational problems, matrix theory and combinatorics. One-dependent processes are stationary, discrete time processes $(X_N)_{N\in\mathbb{Z}}$ with the property that at each time t the future $(X_N)_{N>t}$ is independent of the past $(X_N)_{N<t}$. Such processes can be constructed as a two-block-factor of an i.i.d. sequence $(Y_N)_{N\in\mathbb{Z}}$ by defining $X_N := f(Y_N, Y_{N+1})$ for some function f. Although it was conjectured for quite a long time that each one-dependent process is a two-block-factor, in the second article of this monograph we construct a continuum number of counterexamples of $0-1$ valued one-dependent processes that are not two-block-factors. In section 6 of this Introduction we show a counterexample (by Aaronson, Gilat and Keane) of a one-dependent Markov process (assuming only 5 values) that is not a two-block-factor.

In the third article of this monograph is proved that under certain extremal conditions on the two-correlations (the probability of a run of two ones) a $0-1$ valued one-dependent process is a two-block-factor. The maximal value of a two-correlation over the class of $0-1$ valued one-dependent processes (for fixed probability of a one) is computed and it turns out that the processes where this maximum is attained, are all two-block-factors. If the fixed probability of a one is not equal to $\frac{1}{2}$, this maximum is uniquely attained and there exist exactly two processes with maximal two-correlation in the case that the fixed probability of a one equals $\frac{1}{2}$. Further partial results are proved on minimal two-correlations. The third article of this monograph is also a simplification and a generalization of [Ka.], where the maximal two-correlation over the class of $0-1$ valued two-block-factors is computed (for fixed probability of a one).

In the first article of this monograph the minimal two-correlation over the class of $0-1$ valued two-block-factors is computed (for fixed probability of a one). In the fourth article of this monograph the maximal and minimal value of $\|M^2\|$ is computed over the class of $0-1$ valued $N \times N$ matrices M with K ones (for fixed N and K). In terms of two-correlations this corresponds to the maximal and minimal value of the two-correlation over the class of $0-1$ valued two-block-factors of an i.i.d. sequence of random variables that are all uniformly distributed over N values (for fixed N and fixed probability of a one).

In the fifth article of this monograph the construction (in the second article) of counterexamples of one-dependent processes that are not two-block-factors is generalized by a representation in terms of Hilbert spaces, vectors and bounded linear operators on Hilbert spaces. All one-dependent processes admit a representation. The difference between two-block-factors and non-two-block-factors is determined by the geometry of a closed convex cone that is invariant under certain operators. The dimension of the smallest Hilbert space that represents a process is a measure for the complexity of the structure of the process.

In the sixth article of this monograph we construct for an arbitrary process $(X_n)_{n\in\mathbb{N}}$ a nonlinear autoregression representation $X_n = f_n(X_1, \ldots, X_{n-1}, U_n)$ and a representation $X_n = g_n(U_1, \ldots, U_n)$, where $(U_n)_{n\in\mathbb{N}}$is an i.i.d. sequence with the property that U_n and $(X_1, \ldots, X_{n-1})$ are independent.
For a special class of processes this provides a method to check whether a process is an m-block-factor of an i.i.d. process.

15 Errata

Article I, page 46, line 3 from below. "reduces by one." should be "reduces by one (or two in Case I $\cap$ Case II)."

Article I, page 53, line 12 "with $d_i = d$ for all $i \neq i_0$ (for some i_0)" should be "with $d_i = d$ for all $i \neq i_0$ (for some i_0) and $d_{i_0} \leq d$."

Article VI, page 171, lines 3 and 4 from below. interchange the formula's $X_n = f_n(X_1, \ldots, X_{n-1}, U_n)$ and $X_n = g_n(U_1, \ldots, U_n)$.

16 Acknowledgements

A considerable part of the investigation presented in this monograph was carried out as a part of the research program of the Netherlands Foundation for Mathematics (SMC) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO, now the Netherlands Organization for Scientific Research NWO).

After finishing his undergraduate studies (with emphasis on functional-analysis) at the University of Amsterdam the author was from September 1984 till September 1988 re-

searcher of ZWO on the research project "Coding problems in ergodic theory". This project was carried out at the Vakgroep Statistics, Stochastics and Operations Research (SSOR) of the Faculty of Technical Mathematics and Informatics of the Delft University of Technology. The author collaborated with M.S. Keane (project leader), A. Gandolfi (C.N.R., Italy), J. Aaronson and D. Gilat (both from Tel Aviv University, Israel) and L. Rüschendorf (Westfälische Wilhelms-Universität, Münster, Germany). On 14 June 1988 the author received his doctoral degree at the Delft University of Technology. At the thesis defence the committee consisted of M.S. Keane (thesis advisor), F.M. Dekking, F.A. Lootsma, C.L. Scheffer (all from Delft), R.M. Burton (Oregon State University, Corvallis, USA), D. van Dulst (University of Amsterdam), P.E. Greenwood (University of British Columbia, Vancouver, Canada) and M. Smorodinsky (Tel Aviv University, Israel). This monograph is the revised and extended version of the author's dissertation "One-dependent processes".

From September 1988 until March 1989 the author worked with a stipend from NWO at the School of Mathematical Sciences of Tel Aviv University in Israel. From March 1989 until September 1989 he worked (with the same NWO-stipend) at the Department of Mathematics of Oregon State University in Corvallis, U.S.A. From September 1989 till March 1991 he worked at the Mathematical Institute of the State University of Utrecht. Since March 1991 he works at the Mathematical Institute of the University of Groningen.

The author thanks all these people, universities and institutions for their collegiality, hospitality and support.
In addition to these people the author thanks H.C.P. Berbee (Center for Mathematics and Computer Science, Amsterdam), H.G. Dehling (University of Groningen), M. Goulet (Oregon State University, Corvallis, U.S.A.), S. Janson (Uppsala University, Sweden) and F. Matúš (Academy of Sciences of the Czech Republic, Prague) for many inspiring discussions.

Finally the author thanks the publishers of the journals (in which the 6 articles are published) for their permission to publish these articles in this monograph:
The Magnum Press, The Hebrew University, Jerusalem, Israel (holder of the copyright of the first article, published in the Israel Journal of Mathematics)
The Institute of Mathematical Statistics (holder of the copyright of the second and fifth article, published in The Annals of Probability)
Springer Verlag, Heidelberg, Germany (holder of the copyright of the third article, published in Probability Theory and Related Fields)
Kluwer Academic Publishers, Dordrecht (holder of the copyright of the fourth article, published in Compositio Mathematica) and
Elsevier Science Publishers B.V., Amsterdam (holder of the copyright of the sixth article, published in Stochastic Processes and their Applications).

Groningen, September 13, 1993. Vincent de Valk

ISRAEL JOURNAL OF MATHEMATICS, Vol. 62, No. 2, 1988

THE MAXIMAL AND MINIMAL 2-CORRELATION OF A CLASS OF 1-DEPENDENT 0–1 VALUED PROCESSES

BY
V. DE VALK†
Department of Technical Mathematics & Informatics, Delft University of Technology, P.O. Box 356, 2600 AJ Delft, The Netherlands

ABSTRACT
We compute the maximal and minimal value of $P[X_N = X_{N+1} = 1]$ for fixed $P[X_N = 1]$, where $(X_N)_{N \in \mathbf{Z}}$ is a 0–1 valued 1-dependent process obtained by a coding of an i.i.d.-sequence of uniformly [0,1] distributed random variables with a subset of the unit square.

1. Introduction

A stationary, 0–1 valued, stochastic process $(X_N)_{N \in \mathbf{Z}}$ is 1-*dependent* if

$$P[X_{-N} = i_{-N}, \ldots, X_{-1} = i_{-1}, X_1 = i_1, \ldots, X_N = i_N]$$
$$= P[X_{-N} = i_{-N}, \ldots, X_{-1} = i_{-1}] \cdot P[X_1 = i_1, \ldots, X_N = i_N]$$

for all $N \geqq 1$ and for all $i_{-N}, \ldots, i_{-1}, i_1, \ldots, i_N \in \{0, 1\}$.

For quite a long time it seemed to be folklore to conjecture that each 1-dependent process is an indicator process (we will define that), but recently Aaronson and Gilat ([AG]) found a counterexample of a 1-dependent process that is not an indicator process. A paper by Aaronson, Gilat, Keane and De Valk [AGKV] on a two-parameter family of such counterexamples has been written.

Let J be the unit interval, J^2 the unit square, let λ and μ be Lebesgue measure on J and J^2 resp. and let A be the collection of μ-measurable sets in J^2.

† This research was supported by the Netherlands Foundation for Mathematics (S.M.C.) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO).
Received November 5, 1986 and in revised form August 16, 1987

Let $(U_N)_{N\in\mathbf{Z}}$ be an i.i.d. sequence of random variables uniformly distributed over J. Define for each $A \in \mathcal{A}$ the corresponding *indicator process* $(X_N)_{N\in\mathbf{Z}}$:

$$X_N := \begin{cases} 0, & \text{if } (U_N, U_{N+1}) \notin A, \\ 1, & \text{if } (U_N, U_{N+1}) \in A. \end{cases}$$

It is easy to see that each indicator process is a 1-dependent process and that

$$P[X_N = 1] = \mu(A).$$

From now on we reserve α for the Lebesgue measure of A (thus $\alpha = \mu(A)$ is the probability of a one).

In 1971 Katz [Ka] computed (translated to our terminology) the maximal value of a 2-correlation $P[X_N = X_{N+1} = 1]$ over the class of indicator processes for fixed α.

Finke [F] (1982) was the first to interpret Katz's mathematical objects as correlations in stochastic processes.

Recently Gandolfi, Keane and De Valk [GKV] proved a more general result about the maximal value of a 2-correlation over the class of 1-dependent processes. They computed that the 2-correlation (for fixed probability of a one) has the same upper bound over the class of 1-dependent processes as over the class of indicator processes.

Further, they proved that there exists a unique 1-dependent process with this 2-correlation if the probability of a one is not $\frac{1}{2}$. If the probability of a one is $\frac{1}{2}$, there exist exactly two 1-dependent processes with this 2-correlation (and both are indicator processes). So, the conjecture mentioned in the beginning of this section does not hold in general, but is true for these extremal cases.

In this paper we will compute the minimal 2-correlation for all indicator processes. For $\alpha \notin (\frac{1}{3}, \frac{2}{3})$ we have been able to compute the minimal 2-correlation for 1-dependent processes, finding the same lower bound ([GKV]).

For $\alpha \notin (\frac{1}{4}, \frac{3}{4})$ we know that there exists a unique process with this 2-correlation ([GKV]).

2. Basic properties

For $A \in \mathcal{A}$ we define the *horizontal and vertical sections* H_A and V_A:

$$H_A(y) := \lambda\{x \in J : (x, y) \in A\}, \qquad y \in J,$$

$$V_A(x) := \lambda\{y \in J : (x, y) \in A\}, \qquad x \in J,$$

and we define I_A:

$$I_A := \int_0^1 H_A(x)V_A(x)d\lambda(x).$$

LEMMA 1. *The 2-correlation $P[X_N = X_{N+1} = 1]$ of an indicator process is equal to I_A.*

PROOF. Directly from the definitions,

$$\begin{aligned} P[X_N = X_{N+1} = 1] &= P[(U_N, U_{N+1}) \in A, (U_{N+1}, U_{N+2}) \in A] \\ &= \int_0^1 P[(U_N, U_{N+1}) \in A, (U_{N+1}, U_{N+2}) \in A \mid U_{N+1} = x] d\lambda(x) \\ &= \int_0^1 H_A(x)V_A(x)d\lambda(x) \\ &= I_A. \end{aligned}$$

□

We define the *maximal and minimal 2-correlations* of an indicator process by

$$\mathrm{Max}(\alpha) := \sup\{I_A : A \in \mathcal{A}, \mu(A) = \alpha\},$$

$$\mathrm{Min}(\alpha) := \inf\{I_A : A \in \mathcal{A}, \mu(A) = \alpha\}, \qquad \alpha \in J.$$

Before we describe the sets for which these extremal values are attained, we state three simple lemmas.

Let $A^c := J^2 \setminus A$ be the *complement* of A.

LEMMA 2 (Complement Lemma). *For $A \in \mathcal{A}$ with $\mu(A) = \alpha$ we have*

$$I_A = I_{A^c} + 2\alpha - 1$$

and therefore (for $\alpha \in J$)

$$\mathrm{Min}(\alpha) = \mathrm{Min}(1 - \alpha) + 2\alpha - 1 \quad \text{and} \quad \mathrm{Max}(\alpha) = \mathrm{Max}(1 - \alpha) + 2\alpha - 1.$$

PROOF. We have $H_{A^c}(x) = 1 - H_A(x)$ and $V_{A^c}(x) = 1 - V_A(x)$ which implies

$$\begin{aligned} I_{A^c} &= \int_0^1 (1 - H_A(x))(1 - V_A(x))d\lambda(x) \\ &= \int_0^1 \{1 - H_A(x) - V_A(x) + H_A(x)V_A(x)\}d\lambda(x) \\ &= 1 - 2\alpha + I_A. \end{aligned} \qquad \square$$

Note that the supremum (infimum) is attained in A for α iff the supremum (infimum) is attained in A^c for $1 - \alpha$, so that we may assume $\alpha \leqq \frac{1}{2}$.

We call the sets $\{(x, x) \in J^2 : x \in J\}$, $\{(x, 1 - x) \in J^2 : x \in J\}$ the *diagonal*, the *cross diagonal*, resp.

Let R_d, resp. R_c be reflection w.r.t. these diagonals. We call a transformation

$$(T \times T) : J^2 \to J^2$$

a *product isomorphism* if $T : J \to J$ is measurable, measure preserving and almost everywhere 1–1.

LEMMA 3 (Reflection and Invariance Lemma). *For $A \in \mathcal{A}$ and for a product isomorphism $T \times T$ we have*

$$I_A = I_{R_dA} = I_{R_cA} = I_{(T\times T)A}.$$

PROOF. We have $H_{R_dA} = V_A$, $H_{R_cA}(x) = V_A(1 - x)$ and $H_{(T\times T)A}(x) = H_A(T^{-1}x)$ (and similar formulas for V_A) which imply the statement. $\square$

We will identify two sets A and B if $\mu(A \triangle B) = 0$, and we introduce the habitual metric d:

$$d(A, B) := \mu(A \triangle B), \qquad A, B \in \mathcal{A}.$$

LEMMA 4 (Continuity Lemma). *For $A, B \in \mathcal{A}$ we have*

$$|I_A - I_B| \leqq 2\mu(A \triangle B)$$

and therefore (for $\alpha, \beta \in J$)

$$|\mathrm{Max}(\alpha) - \mathrm{Max}(\beta)| \leqq 2|\alpha - \beta| \quad \textit{and} \quad |\mathrm{Min}(\alpha) - \mathrm{Min}(\beta)| \leqq 2|\alpha - \beta|.$$

PROOF. The first inequality follows from

$$\begin{aligned} |I_A - I_B| &= \left| \int H_A(V_A - V_B) + V_B(H_A - H_B)d\lambda \right| \\ &\leqq \int |V_A - V_B| d\lambda + \int |H_A - H_B| d\lambda \\ &\leqq 2\mu(A \triangle B). \end{aligned}$$

The second inequality follows by choosing for $\alpha > \beta$ a set A with measure α such that I_A is close to $\mathrm{Max}(\alpha)$, and a subset B of A with measure β. Then $\mu(A \triangle B) = \alpha - \beta$, and application of the first inequality yields the second inequality.

The third inequality follows analogously. □

3. The sets where the maximal and minimal 2-correlations are attained

We define the following sets for $0 \leqq \alpha \leqq \frac{1}{2}$:

$$A_\alpha^{\max} := ([0, 1 - \sqrt{1-\alpha}] \times [0, 1]) \cup ([1 - \sqrt{1-\alpha}, 1] \times [0, 1 - \sqrt{1-\alpha}]).$$

For $\alpha < \frac{1}{2}$, let

$$s := \frac{1 + \sqrt{1 - 2\alpha\left(\frac{N+1}{N}\right)}}{N+1}$$

where

$$N := \mathrm{int}\left(\frac{1}{1-2\alpha}\right)$$

is such that

$$\frac{1}{2} - \frac{1}{2N} \leqq \alpha < \frac{1}{2} - \frac{1}{2(N+1)}.$$

Now let

$$A_\alpha^{\min} := \{(x, y) \in J^2 : y \leqq s \cdot \mathrm{int}(x/s)\}$$

or equivalently

$$A_\alpha^{\min} := \bigcup_{i=1}^{N-1} ([is, (i+1)s] \times [0, is]) \cup ([Ns, 1] \times [0, Ns]).$$

Finally we define

$$A_{1/2}^{\min} := \{(x, y) \in J^2 : y \leqq x\}.$$

We call $A_\alpha^{\min}$ a *staircase set*.

Straightforward computations show that both $A_\alpha^{\max}$ and $A_\alpha^{\min}$ have measure α (see Figs. 1–11).

We call a set A with $\mu(A) = \alpha < \frac{1}{2}$ a *disturbed staircase set* if A is not a staircase set and if there exists a set B such that (N and s as above)

$$A = B \cup \bigcup_{i=1}^{N-1} ([x_i, x_{i+1}] \times [0, x_i])$$

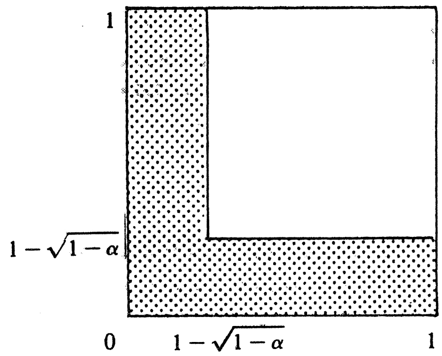


Fig. 1. $A_\alpha^{\max}$ $(0 \leqq \alpha \leqq \frac{1}{2})$.

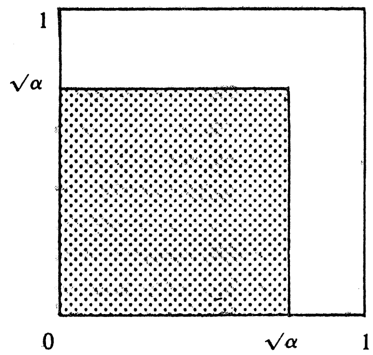


Fig. 2. $R_c(A_{1-\alpha}^{\max})^c$ $(\frac{1}{2} \leqq \alpha \leqq 1)$.

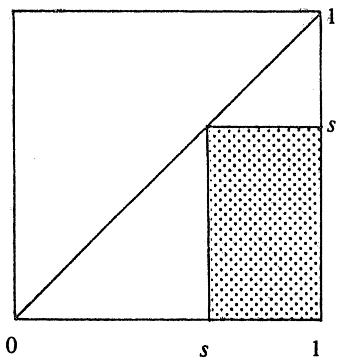


Fig. 3. $A_\alpha^{\min}$ $(0 \leqq \alpha < \frac{1}{4})$.

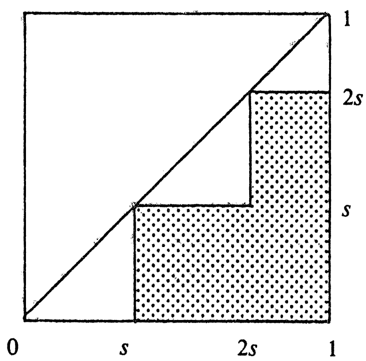


Fig. 4. $A_\alpha^{\min}$ $(\frac{1}{4} \leqq \alpha < \frac{1}{3})$.

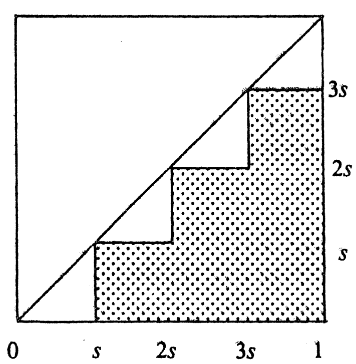


Fig. 5. $A_\alpha^{\min}$ $(\frac{1}{3} \leqq \alpha < \frac{3}{8})$.

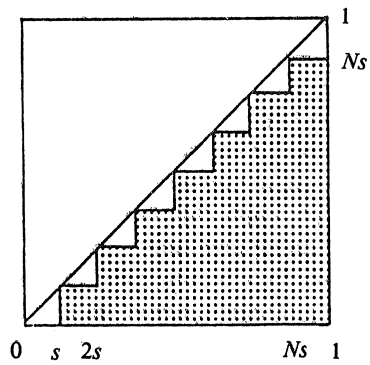


Fig. 6. $A_\alpha^{\min}$ $(\frac{1}{2} - 1/(2N) \leqq \alpha < \frac{1}{2} - 1/(2(N+1)))$.

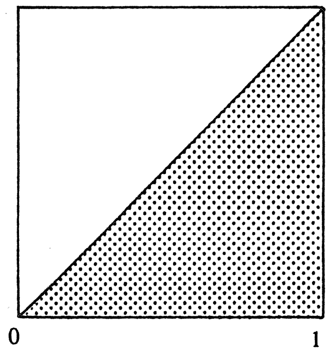


Fig. 7. $A_{1/2}^{\min}$.

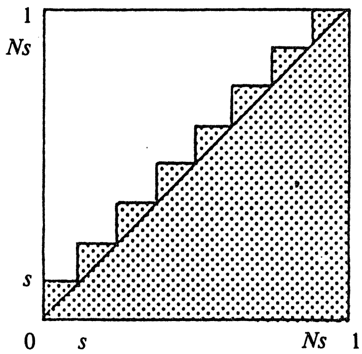


Fig. 8. $R_d(A_{1-\alpha}^{\min})^c\,(\frac{1}{2}+1/(2(N+1))<\alpha\leqq\frac{1}{2}+1/2N)$.

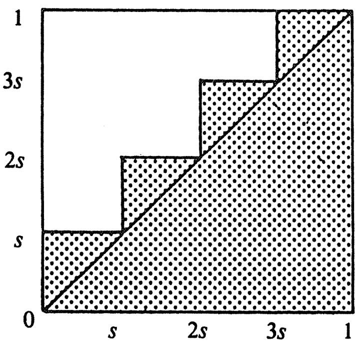


Fig. 9. $R_d(A_{1-\alpha}^{\min})^c\,(\frac{5}{8}<\alpha\leqq\frac{2}{3})$.

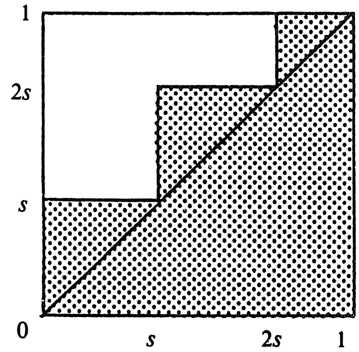


Fig. 10. $R_d(A_{1-\alpha}^{\min})^c\,(\frac{2}{3}<\alpha\leqq\frac{3}{4})$.

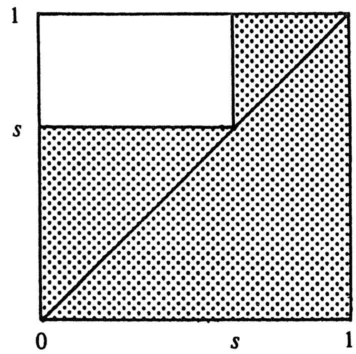


Fig. 11. $R_d(A_{1-\alpha}^{\min})^c\,(\frac{3}{4}<\alpha\leqq 1)$.

with $0(=x_0)<x_1<\cdots<x_{N-1}<x_N=1$, and for some $i_0\in\{0,1,\ldots,N-1\}$

$$x_{i+1}-x_i=\begin{cases}1-(N-1)s, & \text{if } i=i_0\\ s, & \text{if } i\neq i_0,\quad i\in\{0,1,\ldots,N-1\}\end{cases}$$

and, for some $0<\gamma<x_{i_0+1}-x_{i_0}$, B is a subset of

$$[x_{i_0}+\gamma,x_{i_0+1}]\times[x_{i_0},x_{i_0}+\gamma].$$

See Figs. 12–15.

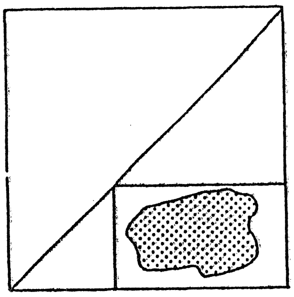


Fig. 12. Two disturbed staircase sets and the complements of the reflected (w.r.t. the diagonal) sets of two disturbed staircase sets.

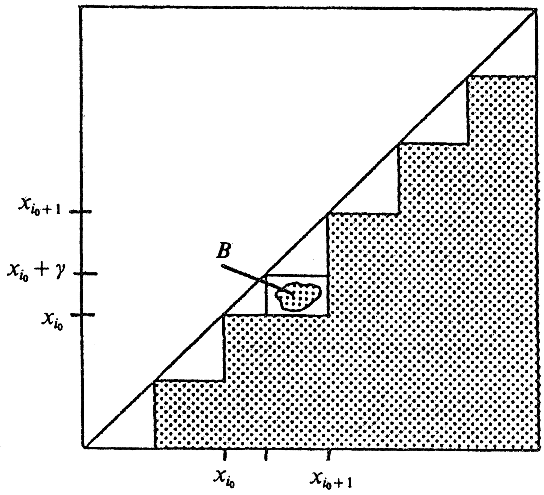


Fig. 13.

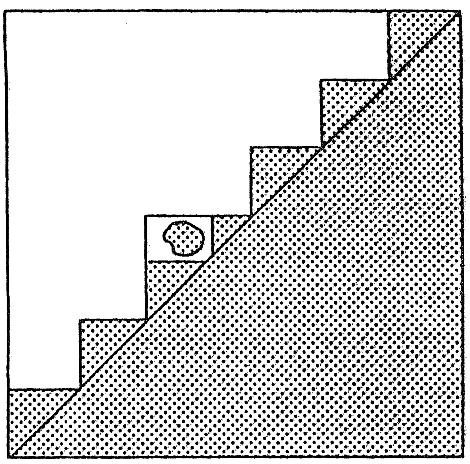


Fig. 14.

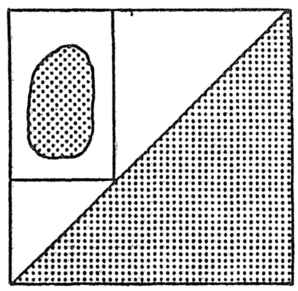


Fig. 15.

4. Results

THEOREM 1.

$$\operatorname{Max}(\alpha)=\begin{cases}2\alpha-1+(1-\alpha)^{3/2}, & 0\leqq\alpha\leqq\frac{1}{2}\\ \alpha^{3/2}, & \frac{1}{2}\leqq\alpha\leqq 1.\end{cases}$$

PROPOSITION 1. *This supremum is attained in the sets $A_\alpha^{\max}$ for $0\leqq\alpha\leqq\frac{1}{2}$ and in $(A_{1-\alpha}^{\max})^c$ for $\frac{1}{2}\leqq\alpha\leqq 1$.*

Conversely, each set A with measure α and $I_A=\operatorname{Max}(\alpha)$ is product isomorphic to one of the above-mentioned sets.

For the proof of Theorem 1 we refer to Katz [Ka] or Finke [F] or Gandolfi, Keane and de Valk [GKV].

Proposition 1 is proved in [GKV].

THEOREM 2.

$$\operatorname{Min}(\alpha)=\begin{cases}\dfrac{(N-1)N}{6(N+1)^2}(1-2\delta)(1+\delta)^2, & \text{if } 0\leqq\alpha<\frac{1}{2},\\ \frac{1}{6}, & \text{if } \alpha=\frac{1}{2},\\ 2\alpha-1+\operatorname{Min}(1-\alpha), & \text{if } \frac{1}{2}<\alpha\leqq 1,\end{cases}$$

with

$$N=\operatorname{int}\left(\frac{1}{1-2\alpha}\right)\quad \text{and}\quad \delta=\sqrt{1-2\alpha\left(\frac{N+1}{N}\right)}.$$

REMARK. For $\frac{1}{2}-1/(2N)\leqq\alpha<\frac{1}{2}-1/(2(N+1))$ we have $1/N\geqq\delta>0$, so $\delta\to 0$ if $\alpha\to\frac{1}{2}$. Note that if $1/(1-2\alpha)$ is an integer we have

$$\operatorname{Min}(\alpha)=\operatorname{Min}\left(\frac{1}{2}\pm\frac{1}{2N}\right)=\frac{(N\pm 1)(N\pm 2)}{6N^2}=\frac{\alpha(4\alpha-1)}{3}$$

and in these points the function Min has a left derivative which is smaller than the right derivative. For the function Max this phenomenon only occurs at $\alpha=\frac{1}{2}$. (See Fig. 16.)

Note further that $\operatorname{Min}(\alpha)\geqq\alpha(4\alpha-1)/3$ for all $\alpha\in J$.

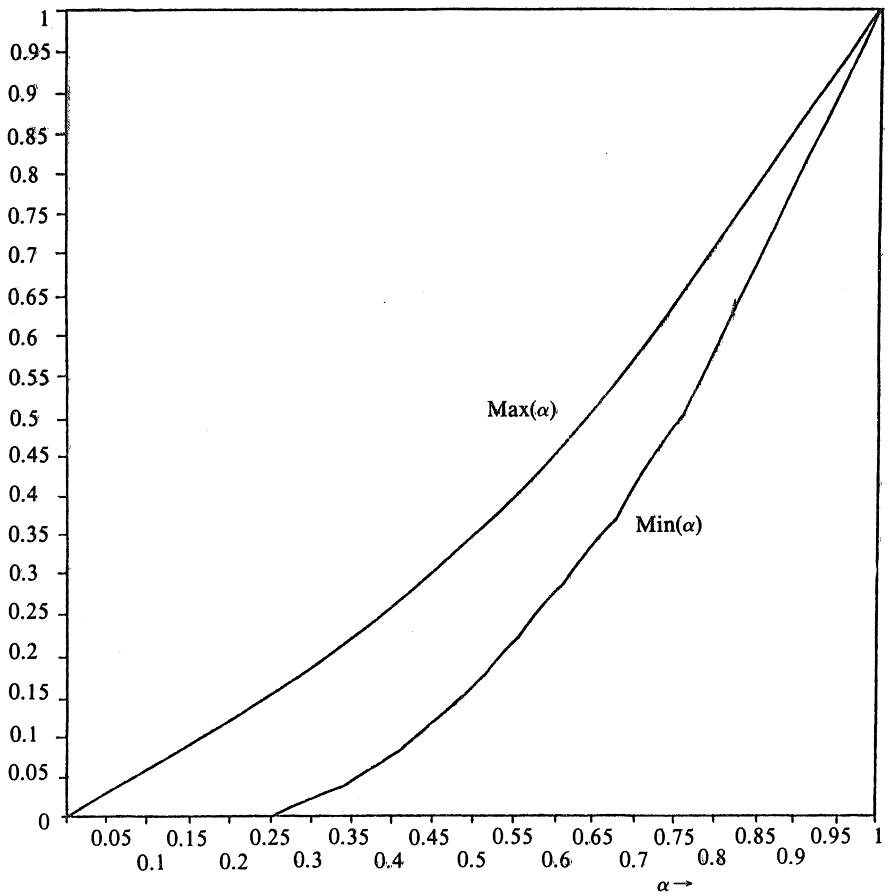


Fig. 16. The functions Max and Min.

PROPOSITION 2. *The infimum is attained in the staircase sets* $A_\alpha^{\min}$ *for* $0 \leqq \alpha \leqq \frac{1}{2}$, $(A_{1-\alpha}^{\min})^c$ *for* $\frac{1}{2} \leqq \alpha \leqq 1$, *and it is also attained in the disturbed staircase sets for* $\alpha < \frac{1}{2}$ *and in the complements of these for* $\alpha > \frac{1}{2}$.

Conversely, when $1/(1-2\alpha)$ *is an integer or* $\alpha = \frac{1}{2}$, *if the infimum is attained in some set* $A \in \mathcal{A}$ *with measure* α, *then A is product isomorphic to a staircase set* ($\alpha \leqq \frac{1}{2}$), *or to the complement of a staircase set* ($\alpha > \frac{1}{2}$).

When $\alpha \neq \frac{1}{2}$ *and* $1/(1-2\alpha)$ *is not an integer, if the infimum is attained in some set* $A \in \mathcal{A}$ *with measure* α, *then A is product isomorphic to a staircase set or to a disturbed staircase set* ($\alpha < \frac{1}{2}$) *or to the complement of one of these sets* ($\alpha > \frac{1}{2}$).

We prove Theorem 2 in Section 5 and we prove Proposition 2 in Section 6.

5. Proof of Theorem 2

Let $\alpha > 0$ be fixed. In six steps we will, by various rearrangement procedures, gradually diminish the size of the collection of sets A for which $I_A = \mathrm{Min}(\alpha)$, until we reach the staircase sets, so proving the statement of Theorem 2.

Step 1. Standardization

By the continuity lemma we may approximate a set A ($\mu(A) = \alpha$) by a finite union of squares of the form $[x, x + \delta) \times [y, y + \delta)$ with $x, y \in J$, where $\delta > 0$ is the reciprocal of an integer.

Then H_A and V_A are constant on intervals. We rearrange J with a transformation T (a permutation of intervals) such that $H_{(T \times T)A}$ is non-increasing (see Figs. 17 and 18). We use the notation $\tau := T \times T$.

The Invariance Lemma implies that $I_{\tau A} = I_A$.

We say that a set is in *standard form* if it is the set under (the graph of) a non-decreasing function. We will obtain from τA a set A' in standard form with $I_{A'} \leqq I_{\tau A}$. This is accomplished by moving squares horizontally to the right.

If τA is not in standard form, then there exist squares S_1 and S_2 such that

$$S_1 := [x_1, x_1 + \delta) \times [y, y + \delta) \text{ is a subset of } \tau A,$$

$$S_2 := [x_2, x_2 + \delta) \times [y, y + \delta) \text{ is disjoint with } \tau A,$$

for some $x_1 < x_2$. Define the set $\sigma\tau A$ (Fig. 19):

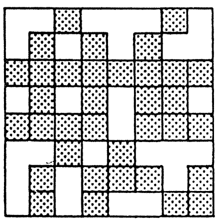


Fig. 17. A.

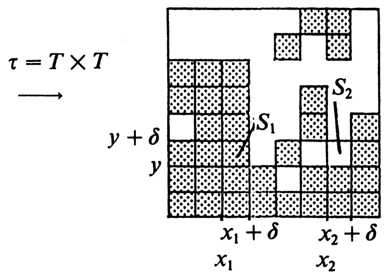


Fig. 18. τA.

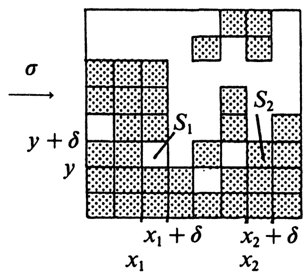


Fig. 19. $\sigma\tau A$.

$$\sigma\tau A := (\tau A \setminus S_1) \cup S_2;$$

then $\mu(\sigma\tau A) = \mu(\tau A)$ and we will prove that $I_{\sigma\tau A} \leqq I_{\tau A}$.

We have $H_{\sigma\tau A} = H_{\tau A} =: H$ and

$$V_{\sigma\tau A}(x) = \begin{cases} V_{\tau A}(x) - \delta, & \text{if } x \in [x_1, x_1 + \delta), \\ V_{\tau A}(x) + \delta, & \text{if } x \in [x_2, x_2 + \delta), \\ V_{\tau A}(x), & \text{else.} \end{cases}$$

Therefore

$$\begin{aligned} I_{\tau A} - I_{\sigma\tau A} &= \int_{x_1}^{x_1+\delta} \delta H(x)dx - \int_{x_2}^{x_2+\delta} \delta H(x)dx \\ &= \delta^2\{H(x_1) - H(x_2)\} \\ &\geqq 0. \end{aligned}$$

Note that we have equality iff H is constant on $[x_1, x_2 + \delta)$. The set A' (in standard form) is obtained from A by applying τ (once) and a finite number of shifts of the type σ. Using these facts we obtain the next claim, in which we introduce the notation f_A and A_f (to stress the correspondence between a non-decreasing function f and a set A in standard form that is the set under f).

CLAIM 1 (Standardization).

$$\mathrm{Min}(\alpha) = \inf\{I_A : \mu(A) = \alpha, A = A_f \in A \text{ in standard form}, \\ f_A \text{ finite valued}\} \qquad (\alpha \in J).$$

REMARK. From Helly's selection principle ([Luk] par. 3.5) it follows that the infimum is actually attained in some set in standard form.

Step 2. (*Under the diagonal*)

Because of the Complement Lemma we assume that $\alpha < \frac{1}{2}$. It is easy to see that $\mathrm{Min}(\alpha) = 0$ for $\alpha \leqq \frac{1}{4}$; take e.g. $A = [1 - \sqrt{\alpha}, 1] \times [0, \sqrt{\alpha}]$.

Therefore we assume further in this proof that $\frac{1}{4} < \alpha < \frac{1}{2}$.

Take a set A in standard form with measure α and such that f_A is finite valued. Assume that A does not lie under the diagonal (a set A lies *under the diagonal* if A is a subset of $A^{\min}_{1/2}$). We will transform A to a set lying under the diagonal such that I_A does not increase.

Let A be a union of $\delta \times \delta$ squares. We choose

$$S_1 := [x_1, x_1 + \delta) \times [y_1, y_1 + \delta) \text{ subset of } A$$

and

$$S_2 := [x_2, x_2 + \delta) \times [y_2, y_2 + \delta) \text{ disjoint with } A$$

such that S_1 lies above the diagonal and S_2 lies under the diagonal (by passing from δ to $\frac{1}{2}\delta$ we may assume that there exist such squares entirely above or under the diagonal), and such that

$$f_A(x_1-)\leqq y_1,\qquad f_A(x_2+\delta+)\geqq y_2+\delta$$

(these conditions guarantee that the transformed set will be in standard form). Let g be such that

$$A_g=(A_f\setminus S_1)\cup S_2.$$

We will prove that $I_{A_g}<I_{A_f}$.

We say that a rectangle $[x',x'')\times[y',y'')$ (disjoint with the diagonal and a subset of a set A in standard form) *interferes* with the horizontal sections $H_A(x)$ with $x'\leqq x<x''$ and with the vertical sections $V_A(y)$ with $y'\leqq y<y''$.

We introduce this definition because the removal of this rectangle from A decreases I_A by the amount (as follows from the computation in this step)

$$(y''-y')\cdot(x''-x')\cdot\left\{\frac{\int_{x'}^{x''}H_A(x)dx}{(x''-x')}+\frac{\int_{y'}^{y''}V_A(y)dy}{(y''-y')}\right\},$$

i.e., the change in I_A equals the area of the rectangle times the average value of the sections with which the rectangle interferes.

The intuitive idea behind the inequality $I_{A_g}<I_{A_f}$ is the fact that the square S_1 interferes with the sections marked with a $-$ sign and the square S_2 interferes with the sections marked with a $+$ sign (in Fig. 20); the first total is larger than the second. We have

$$\begin{aligned}I_{A_f}-I_{A_f\setminus S_1}&=\int_{x_1}^{x_1+\delta}H_A(x)\delta dx+\int_{y_1}^{y_1+\delta}V_A(y)\delta dy\\&\geqq\delta\cdot(1-x_1)\cdot\delta+\delta\cdot(y_1+\delta)\cdot\delta\end{aligned}$$

and analogously

$$\begin{aligned}I_{A_f\setminus S_1}-I_{A_g}&=-\int_{x_2}^{x_2+\delta}H_A(x)\delta dx-\int_{y_2}^{y_2+\delta}V_A(y)\delta dy\\&\geqq-\delta\cdot(1-x_2-\delta)\cdot\delta-\delta y_2\delta\end{aligned}$$

which implies

$$I_{A_f}-I_{A_g}\geqq\delta^2\{y_1-x_1+x_2-y_2+2\delta\}\geqq4\delta^3>0$$

since $x_1+\delta\leqq y_1$ and $x_2\geqq y_2+\delta$.

We write our conclusions in the next claim:

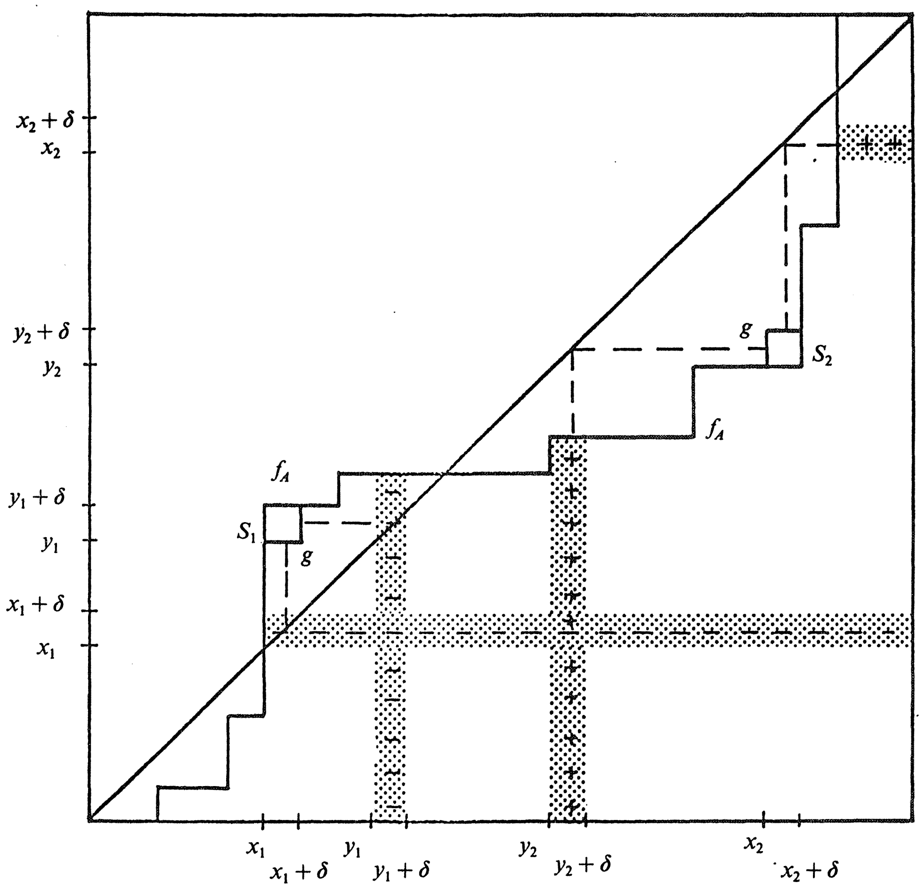


Fig. 20.

CLAIM 2 (Under the diagonal). For $\frac{1}{4} < \alpha < \frac{1}{2}$ we have

$$\begin{aligned}\mathrm{Min}(\alpha) = \inf\{I_A : A \in \mathcal{A}, \mu(A) = \alpha, &A \text{ in standard form,}\\ &A \text{ under the diagonal}, f_A \text{ finite valued}\}.\end{aligned}$$

LEMMA 5 (Windowing). *Let $f_A : J \to J$ be a non-decreasing function such that $f_A(a) = a, f_A(b) = b$ for some $0 \leqq a < b \leqq 1$. Define*

$$A^w := A \cap ([a, b] \times [a, b])$$

and let H^w and V^w be the corresponding sections on $[a, b]$ (Fig. 21),

$$H^w := H_A - (1 - b) \quad \textit{and} \quad V^w := V_A - a.$$

Let $\alpha^w := \mu(A^w)$ and

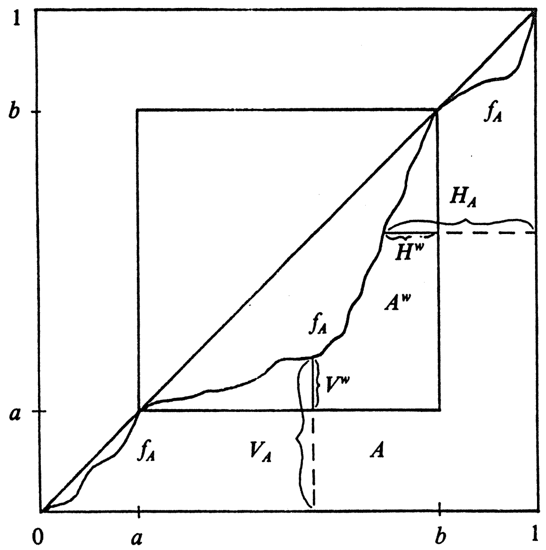


Fig. 21.

$$I_{A^w} = \int_a^b H^w(x)V^w(x)dx,$$

then

$$I_A = I_{A^w} + \int_{[0,a]\cup[b,1]} H_A(x)V_A(x)dx + (1-b+a)\alpha^w + (b-a)(1-b)a.$$

PROOF. We have

$$\begin{aligned} I_A - \int_{[0,a]\cup[b,1]} H_A(x)V_A(x)dx &= \int_a^b (H^w(x)+1-b)(V^w(x)+a)dx \\ &= \int_a^b H^w(x)V^w(x)dx + (1-b)\int_a^b V^w(x)dx \\ &\quad + a\int_a^b H^w(x)dx + (b-a)(1-b)a \\ &= I_{A^w} + (1-b+a)\alpha^w + (b-a)(1-b)a. \quad \square \end{aligned}$$

COROLLARY. *Assume f_A is as in this lemma, and we change f_A on (a,b)* area preservingly *to f_B (i.e. $\mu(A)=\mu(B)$). Then I_A will change by the same amount as I_{A^w}.*

Step 3. (*Moving to the diagonal*)

Let A be a set in standard form, lying under the diagonal, such that for some positive integer N

$$f_A = \sum_{i=1}^{N} y_i \cdot 1_{[x_i, x_{i+1})}$$

with $0 = x_0 < x_1 < \cdots < x_{N+1} = 1$.

Let $d_i := x_i - x_{i-1}$ and $c_i := y_i - y_{i-1}$ $(i = 1, \ldots, N+1)$. Assume that

$$U(f_A) := \operatorname{card}\{i : y_i < x_i\} > 0,$$

then we will prove the existence of a set B in standard form, lying under the diagonal, with the same measure as A, and with a finite valued function f_B such that $I_B \leqq I_A$ and $U(f_B) < U(f_A)$.

We first give an intuitive sketch of our procedure (cf. Figs. 22 and 23). Let i be the first index such that $y_i < x_i$. We will change f_A on $[x_{i-1}, x_{i+1})$. Because of the Windowing Lemma we may restrict our attention to the square $[x_{i-1}, 1] \times [x_{i-1}, 1]$.

We transform the rectangle $[x_i, x_{i+1}) \times [y_{i-1}, y_i)$ (with area $d_{i+1} \cdot c_i$) such that $U(f_A)$ reduces by one. We change it to a rectangle with height $c_i + c_{i+1}$. This is possible if (Case I, Figs. 22 and 23)

$$d_{i+1} \cdot c_i \leqq (x_{i+1} - y_{i+1}) \cdot (c_i + c_{i+1}).$$

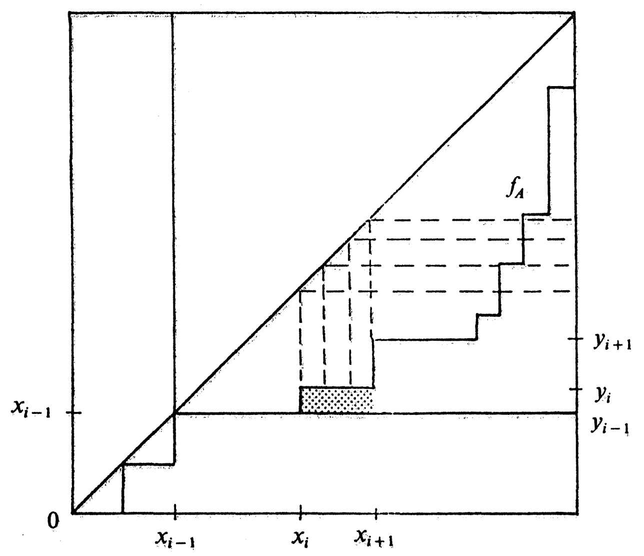


Fig. 22. Case I. Before the transformation.

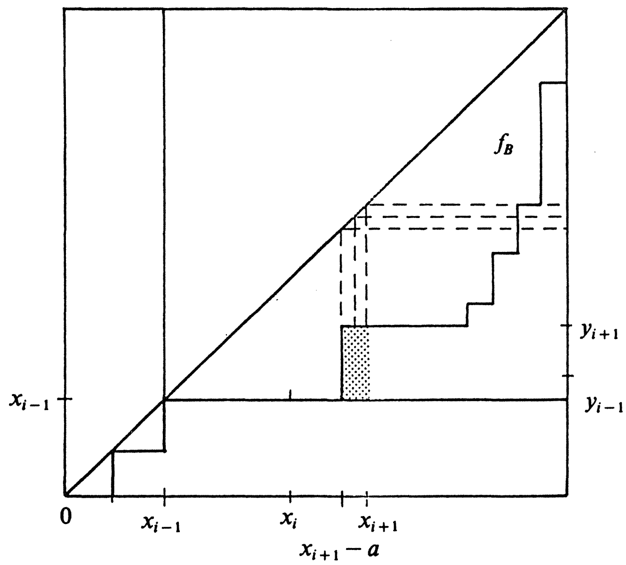


Fig. 23. Case I. After the transformation.

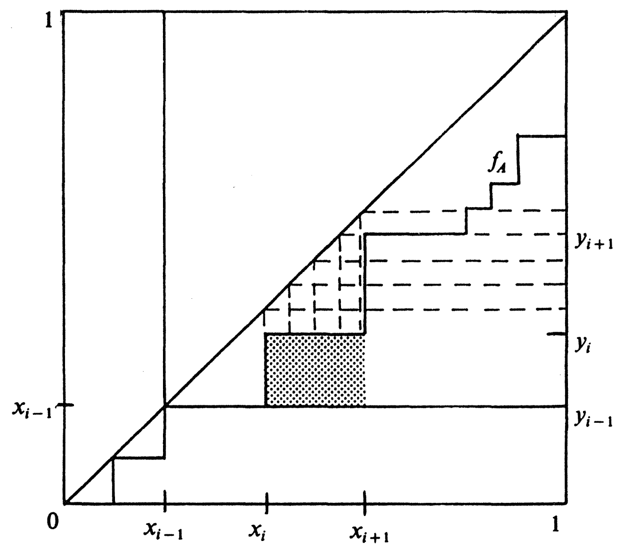


Fig. 24. Case II. Before the transformation.

Otherwise (Case II, Figs. 24 and 25) we transform it to a rectangle that has one corner on the diagonal and lies as far as possible to the right. The rectangle with area $d_{i+1}\cdot c_i$ interferes before the transformation with the set of horizontal sections $H_A(x)$, $x_i \leqq x < x_{i+1}$ and after the transformation it interferes with

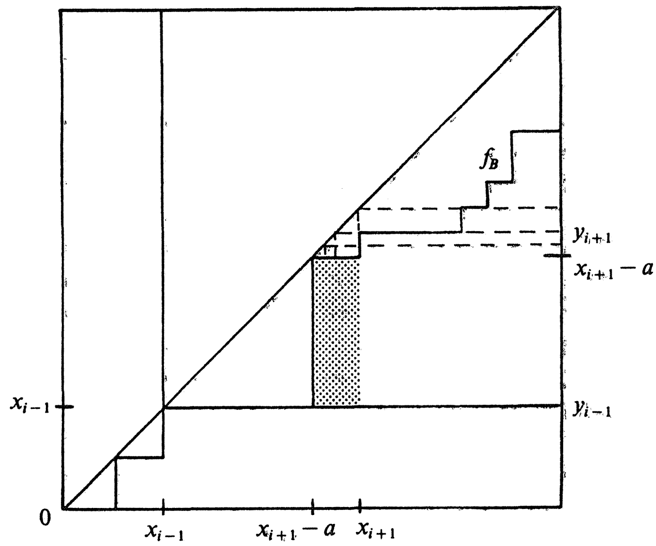


Fig. 25. Case II. After the transformation.

the subset of horizontal sections $H_A(x)$, $x_{i+1}-a \leqq x < x_{i+1}$ (in Case I, $a := d_{i+1}\cdot c_i/(c_i + c_{i+1})$; in Case II, choose $0 < a < d_{i+1}$ such that $a\cdot(d_i + d_{i+1} - a) = d_{i+1}\cdot c_i$).

Because the first set contains some large sections, which are not contained in the subset, this subset has a smaller average value (see definition of interference). This crucial observation implies that I_A will decrease. (Note that the vertical sections V^w with which the described rectangles interfere, have length zero.)

Case I. $(d_{i+1}\cdot c_i \leqq (x_{i+1} - y_{i+1})\cdot(c_i + c_{i+1}))$

Replacing $[x_i, x_{i+1})\times[y_{i-1}, y_i)$ by $[x_{i+1}-a, x_{i+1})\times[y_{i-1}, y_{i+1})$ we have

$$\begin{aligned} I_{A_f} - I_B &= \int_{x_i}^{x_{i+1}} H^w(x)c_i dx - \int_{x_{i+1}-a}^{x_{i+1}} H^w(x)(c_i + c_{i+1})dx \\ &= \int_{x_i}^{x_{i+1}-a} H^w(x)c_i dx - \int_{x_{i+1}-a}^{x_{i+1}} H^w(x)c_{i+1}dx \\ &\geqq (d_{i+1}-a)\cdot H^w(x_{i+1}-a)\cdot c_i - a\cdot H^w(x_{i+1}-a)\cdot c_{i+1} \\ &= 0. \end{aligned}$$

Case II. $(d_{i+1}c_i \geqq (x_{i+1} - y_{i+1})\cdot(c_i + c_{i+1}))$

Replacing $[x_i, x_{i+1})\times[y_{i-1}, y_i)$ by $[x_{i+1}-a, x_{i+1})\times[y_{i-1}, x_{i+1}-a)$ we have

$$
\begin{aligned}
I_{A_f} - I_B &= \int_{x_i}^{x_{i+1}} H^w(x)c_i dx - \int_{x_{i+1}-a}^{x_{i+1}} H^w(x)(x_{i+1} - a - y_{i-1})dx \\
&= \int_{x_i}^{x_{i+1}-a} H^w(x)c_i dx - \int_{x_{i+1}-a}^{x_{i+1}} H^w(x)(x_{i+1} - a - y_{i-1} - c_i)dx \\
&\geqq (d_{i+1} - a)\cdot H^w(x_{i+1} - a)\cdot c_i \\
&\quad - a \cdot H^w(x_{i+1} - a)\cdot(x_{i+1} - a - y_{i-1} - c_i) \\
&= 0.
\end{aligned}
$$

It is easy to see that we can reduce $U(f_A)$ to zero, while I_A does not increase, and we conclude

CLAIM 3 (Moving to the diagonal). For $\frac{1}{4} < \alpha < \frac{1}{2}$ we have

$$
\mathrm{Min}(\alpha) = \inf \left\{ I_A : A \in \mathcal{A}, \mu(A) = \alpha, \text{ for some } N \in \mathbf{N}, f_A = \sum_{i=1}^{N} x_i \cdot 1_{[x_i, x_{i+1})}, \right.
$$
$$
\left. 0 = x_0 < x_1 < \cdots < x_N < x_{N+1} = 1 \right\}.
$$

Step 4. (*Rearrangement*)

We will prove that we may assume that $(d_i)_{i=1}^{N+1}$ is a non-increasing sequence. Let A be a set as in Claim 3 and assume that for some $i \in \{1, \ldots, N\}$ we have

$$d_i < d_{i+1}.$$

We will change f_A on $[x_i, x_{i+1})$ area-preservingly such that for the new function g we have $d_i' > d_{i+1}'$ and $I_{A_f} = I_{A_g}$.

The intuitive idea behind this equality is the fact that both rectangles with area $d_{i+1} \cdot d_i$ interfere with horizontal sections of the same constant length (see Fig. 26). Because of the Windowing Lemma we may restrict our attention to

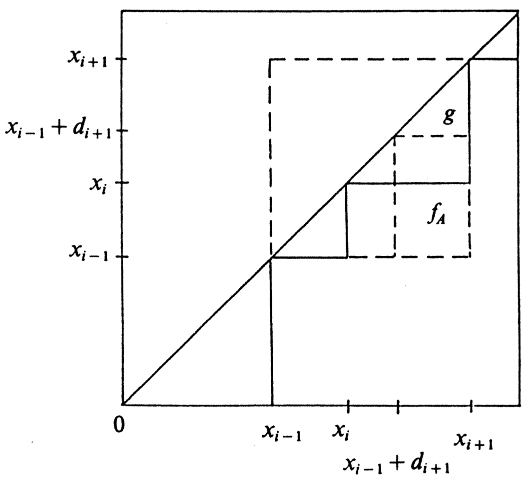


Fig. 26.

$[x_{i-1}, x_{i+1}) \times [x_{i-1}, x_{i+1})$. We replace $[x_{i-1}+d_i, x_{i+1}) \times [x_{i-1}, x_{i-1}+d_i)$ by $[x_{i-1}+d_{i+1}, x_{i+1}) \times [x_{i-1}, x_{i-1}+d_{i+1})$ and it is trivial that

$$I'_{A_f} = I'_{A_g} \qquad (\text{because } H^w = 0 \text{ or } V^w = 0).$$

We conclude.

Claim 4 (Rearrangement). For $\frac{1}{4} < \alpha < \frac{1}{2}$ we have

$$\begin{aligned} \mathrm{Min}(\alpha) = \inf \Big\{ I'_A : A \in \mathcal{A},\ \mu(A) = \alpha, \text{ for some } N \in \mathbf{N}, f_A = \sum_{i=1}^{N} x_i \cdot 1_{[x_i, x_{i+1})}, \\ 0 = x_0 < x_1 < \cdots < x_N < x_{N+1} = 1, \\ d_1 \geqq d_2 \geqq \cdots \geqq d_N \geqq d_{N+1} \Big\}. \end{aligned}$$

Step 5. (Equality of Differences)

Let A be as in Claim 4. We will prove that we may assume that

$$d_1 = d_2 = \cdots = d_N \geqq d_{N+1}.$$

Assume that for some $i \in \{1, \ldots, N-1\}$ we have

$$d_i > d_{i+1} \geqq d_{i+2}.$$

We will change f_A area-preservingly to g on $[x_{i-1}, x_{i+2})$ (cf. Fig. 27). Because

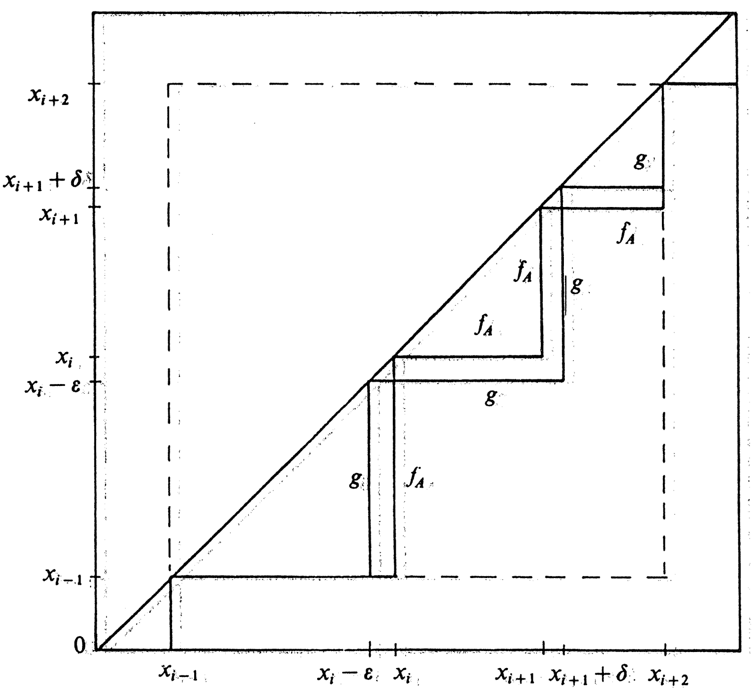


Fig. 27.

of the Windowing Lemma we may restrict our attention to $[x_{i-1}, x_{i+2}) \times [x_{i-1}, x_{i+2})$. We will obtain $I_{A_g} < I_{A_f}$.

Let $0 < \varepsilon < d_i - d_{i+1}$. Since $d_i > d_{i+1} \geqq d_{i+2}$ we can find $\delta > 0$ such that

$$(*) \qquad \varepsilon(d_i - \varepsilon) + \delta(d_{i+2} - \delta) = \varepsilon d_{i+1} + \delta d_{i+1} + \varepsilon\delta.$$

Define

$$d_i' := d_i - \varepsilon, \quad d_{i+1}' := d_{i+1} + \varepsilon + \delta, \quad d_{i+2}' := d_{i+2} - \delta,$$

and let g be the changed version of f (corresponding to d_i', d_{i+1}', d_{i+2}'), then

$$I_{A_f} - I_{A_g} = d_{i+1} \cdot d_i \cdot d_{i+2} - d_{i+1}' \cdot d_i' \cdot d_{i+2}' \qquad (\text{use } (*))$$

$$= \varepsilon(d_i - d_{i+2})(d_i - d_{i+1}) - \varepsilon^2(d_i - d_{i+2}) - \varepsilon\delta(d_{i+1} - d_{i+2} + \varepsilon + \delta)$$

and this is positive if ε is small enough ($\delta \to 0$ if $\varepsilon \to 0$). We conclude

CLAIM 5 (Equality of Differences). For $\frac{1}{4} < \alpha < \frac{1}{2}$ we have

$$\mathrm{Min}(\alpha) = \inf \left\{ I_A : A \in \mathcal{A}, \mu(A) = \alpha, \text{ for some } N \in \mathbf{N}, f_A = \sum_{i=1}^{N} x_i \cdot 1_{[x_i, x_{i+1})}, \right.$$
$$0 = x_0 < x_1 < \cdots < x_N < x_{N+1} = 1,$$
$$\left. d_1 = \cdots = d_N \geqq d_{N+1} \right\}.$$

Step 6. (*Conclusion*)

These computations will prove Theorem 2. Let A be as in Claim 5, and set $s := d_1 = \cdots = d_N$. We have $1/(N+1) \leqq s \leqq 1/N$ and (see e.g. Fig. 6)

$$\alpha = \mu(A) = \sum_{i=1}^{N-1} s \cdot is + (1 - Ns) \cdot Ns = Ns - \frac{N(N+1)s^2}{2},$$

this implies

$$s = \frac{1 + \sqrt{1 - 2\alpha\left(\frac{N+1}{N}\right)}}{N+1} \qquad \left(+ \text{ sign because } s \geqq \frac{1}{N+1} \right).$$

Further, for I_A we have

$$I_A = \sum_{i=1}^{N-1} s \cdot is(1 - is - s) = \frac{s^2 \cdot N(N-1)}{6} \cdot \{3 - 2(N+1)s\}$$

$$= \frac{(N-1)N}{6(N+1)^2} \cdot (1 - 2\delta)(1 + \delta)^2,$$

when we write

$$\delta := \sqrt{1 - 2\alpha\left(\frac{N+1}{N}\right)}.$$

This is the formula for $\mathrm{Min}(\alpha)$ in Theorem 2. We have computed $\mathrm{Min}(\alpha)$ for $\alpha < \frac{1}{2}$. The continuity of $\mathrm{Min}(\alpha)$ leads to $\mathrm{Min}(\frac{1}{2}) = \frac{1}{6}$ (use $N \to \infty$ and $\delta \to 0$ if $\alpha \to \frac{1}{2}$).

The Complement Lemma leads to the formula for $\alpha > \frac{1}{2}$. □

6. Proof of Proposition 2

In Step 6 of Section 5 we proved that the infimum is attained in the staircase sets. A straightforward computation shows that the infimum is also attained in the disturbed staircase sets. Observe that the subset B (in the definition of a disturbed staircase set in Section 3) interferes with sections of the same size as a rectangle.

Let $A \in \mathcal{A}$ be a set with measure $\alpha < \frac{1}{2}$, where the infimum is attained. We can generalize Steps 1–5 of the proof of Theorem 2 to A with $I_A = \mathrm{Min}(\alpha)$. Two integrable functions $f, g : [0, \infty) \to [0, \infty)$ are called *equimeasurable* (see [HLP] par. 10.12) if

$$\lambda\{x : f(x) \geqq y\} = \lambda\{x : g(x) \geqq y\} \qquad \text{for all } y > 0.$$

Let $f: [0, \infty) \to [0, \infty)$ be an integrable function. It is a well-known fact that there exists a non-increasing function g (the so-called *equimeasurable decreasing rearrangement* of f) such that f and g are equimeasurable.

Let H_A^* be the equimeasurable decreasing rearrangement of H_A. We define

$$A_1 := \{(x, y) \in J^2 : 1 - H_A^*(y) \leqq x\}.$$

Then A_1 is a set in standard form and this method of standardization generalizes Step 1. A simple approximation argument yields $I'_{A_1} \leqq I_A$, but $I_{A_1} = I_A$ because the infimum is attained in A.

If A_1 does not lie under the diagonal, then we can strictly reduce I_{A_1} (and obtain a set A_2) by moving a part of A_1 lying above the diagonal to a place under the diagonal, as a slight modification of Step 2 shows (consider the interference in Fig. 20). Therefore we may assume $A_1 = A_2$, and this set lies under the diagonal.

A modification of Step 3 (approximation by stepfunctions) transforms A_2 to a set of the type in Claim 3, such that $I_{A_3} \leqq I_{A_2}$ (but again $I_{A_3} = I_{A_2}$).

Application of Steps 4 and 5 (unmodified) leads to sets A_4 and A_5 of the type in Claim 4, Claim 5, resp. with $I_{A_3} = I_{A_4} = I_{A_5}$.

We consider A_5 and go backwards to determine what A can be. If A_4 is not of the type in Claim 5, then the computation in Step 5 would imply that $I_{A_5} < I_{A_4}$. So $A_4 = A_5$.

Because rearrangement does not change I_A, we conclude that A_3 is of the type in Claim 3 with $d_i = d$ for all $i \neq i_0$ (for some i_0). Because $I_{A_2} = I_{A_3}$, the set A_2 $(= A_1)$ is a staircase set or a disturbed staircase set in standard form (see the interference in Figs. 24 and 25). Note that the edgepoints (x_i, x_i) $(i \neq i_0 - 1)$ cannot be removed from the diagonal without changing the measure of the set. We consider the effect of moving some subset of A_1 horizontally to the left. If the new set is still a staircase set or a disturbed staircase set, then I_{A_1} will not change. But if the new set is no longer of this type, then I_{A_1} will change as a step-1-type computation shows (consider the interference with the horizontal sections). So before the process of moving rectangles to the right (as in Step 1), we already had a staircase set or a disturbed one. So A is product isomorphic to a set of this type. This proves Proposition 2 for the case $\alpha < \frac{1}{2}$.

The case $\alpha = \frac{1}{2}$ can be proved analogously. The case $\alpha > \frac{1}{2}$ follows from the case $\alpha < \frac{1}{2}$ (use the Complement Lemma). □

7. Remarks

(1) Katz proved a kind of symmetrization theorem ([Ka], Th. 3, p. 66) for the maximum case; for each set A, which is the set under (the graph of) a non-increasing function f_A (standard form in maximum case) and which is not symmetric (w.r.t. the diagonal), there exists a symmetric set A^{SYM} (in standard form with the same measure as A) such that

$$I_{A^{\mathrm{SYM}}} > I_A .$$

A^{SYM} is obtained in the following way. Let

$$x_0 := \sup\{x \in J : f_A(x) > x\},$$

let $f_{R_d(A)}$ be the function corresponding to $R_d(A)$, let

$$g := \tfrac{1}{2}(f_A + f_{R_d(A)}) \cdot 1_{[0,x_0]},$$

let C_g be the set under g, define A^{SYM} as

$$A^{\mathrm{SYM}} := C_g \cup R_d(C_g).$$

(See Fig. 28.) This symmetrization method does not work in the minimum case; i.e., given a set A (set under a non-decreasing function) we can construct a set A'^{SYM} (symmetric w.r.t. the cross diagonal) in an analogous way, but we will not always have $I_{A'^{\mathrm{SYM}}} \leqq I_A$, as the next counterexample shows.

Let $f = \frac{2}{5} \cdot 1_{[2/5,4/5]} + \frac{4}{5} \cdot 1_{[4/5,1)}$, then $\alpha = \frac{8}{25}$ and $I_{A_f} = 0.032 = \mathrm{Min}(\alpha)$, but $I_{A'^{\mathrm{SYM}}} = 0.036$. The infimum is attained in A, but not in A'^{SYM}, which does not touch the diagonal in each step (terminology from Step 3).

(2) Extension of the problem from J^2 to $\mathbf{R}^2$ is not possible. Given a set $A \subset \mathbf{R}^2$ with measure α we can define H_A and V_A in the usual way, but the problem is that

$$\int_{-\infty}^{+\infty} H_A(x) V_A(x) dx$$

can diverge. So the supremum is infinite. Further, the infimum is zero (take e.g. $A \subset (0, \infty) \times (-\infty, 0)$).

(3) In the minimum case there exists a continuous (w.r.t. d) mapping

$$F : J \rightarrow \mathcal{A}$$

such that

$$I_{F(\alpha)} = \mathrm{Min}(\alpha) \qquad \text{for } \alpha \in J.$$

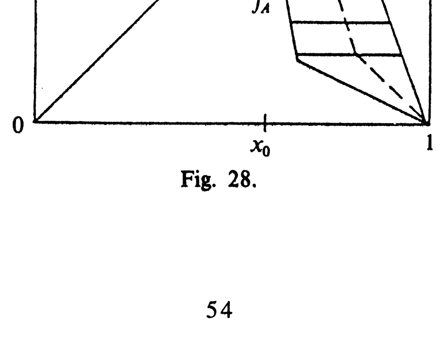


Fig. 28.

In the maximum case such a mapping is discontinuous in $\alpha = \frac{1}{2}$. For $\alpha < \frac{1}{2}$ we have (independent of the choice of the set A)

$$\mathrm{Range}(H_A) = \{1 - \sqrt{1-\alpha},\, 1\}$$

and for $\alpha > \frac{1}{2}$

$$\mathrm{Range}(H_A) = \{\sqrt{\alpha},\, 0\}.$$

If F could be chosen to be continuous, then the range of H_A would depend continuously on α.

In other words: in the maximum case the sections of $A_{1/2}^{\max}$ and $(A_{1/2}^{\max})^c$ have different ranges and in the minimum case the sections of $A_{1/2}^{\min}$ and $(A_{1/2}^{\min})^c$ have the same range.

Acknowledgements

The author thanks his colleagues F. M. Dekking, A. Gandolfi, M. S. Keane and A. Sieders of the Delft University of Technology and H. C. P. Berbee of the Centre for Mathematics and Computer Science in Amsterdam for many helpful discussions.

References

[AG] J. Aaronson and D. Gilat, *On the structure of stationary one dependent processes*, preprint, Tel Aviv University (1987).

[AGKV] J. Aaronson, D. Gilat, M. S. Keane and V. de Valk, *An algebraic construction of a class of one-dependent processes*, Ann. Probab., submitted.

[F] L. Finke, *Two maximization problems*, paper submitted to Oregon State University in partial fulfillment of the requirements for the degree of Master of Arts (1982).

[GKV] A. Gandolfi, M. S. Keane and V. de Valk, *Extremal two-correlations of two-valued stationary one-dependent processes*, Probability Theory and Related Fields, submitted.

[HLP] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, Cambridge University Press, 1934.

[Ka] M. Katz, *Rearrangements of* (0–1) *matrices*, Isr. J. Math. **9** (1971), 53–72.

[Kh] A. Khintchine, *Uber eine Ungleichung*, Mat. Sb. **39** (1932), 35–39.

[Lo] G. G. Lorentz, *A problem of plane measure*, Am. J. Math. **71** (1949), 417–426.

[Luk] E. Lukacs, *Characteristic functions*, Griffin, London, 1960.

[Lux] W. A. J. Luxemburg, *On an inequality of A. Khintchine for zero-one matrices*, J. Comb. Theory **12** (1972), 289–296.

The Annals of Probability
1989, Vol. 17, No. 1, 128–143

AN ALGEBRAIC CONSTRUCTION OF A CLASS OF ONE-DEPENDENT PROCESSES

BY JON AARONSON, DAVID GILAT, MICHAEL KEANE AND VINCENT DE VALK[1]

Tel Aviv University, Tel Aviv University, Delft University of Technology and Delft University of Technology

A special class of stationary one-dependent two-valued stochastic processes is defined. We associate to each member of this class two parameter values, whereby different members receive different parameter values. For any given values of the parameters, we show how to determine whether:

1. a process exists having the given parameter values, and if so,
2. this process can be obtained as a two-block factor from an independent process.

This determines a two-parameter subfamily of the class of stationary one-dependent two-valued stochastic processes which are not two-block factors of independent processes.

Introduction. A discrete time stochastic process $X = (X_n)$ is *one-dependent* if at any given time n, its past $(X_k)_{k<n}$ is independent of its future $(X_k)_{k>n}$. In contrast to the Markovian concept, a weakening of independence which has been investigated thoroughly, no knowledge of the present value X_n is assumed. One-dependent processes arise naturally as limits of rescaling operations in renormalization theory (see, e.g., O'Brien [8]). In an analogous manner m-dependence ($m \geq 1$) can be defined, considering the present to be given by m successive observations. The works [2], [4]–[7] and [10] deal with various aspects of m-dependent processes.

Examples of m-dependent processes are given by so-called $(m + 1)$-*block factors*: Let $Y = (Y_n)$ be an independent process and f a function of $m + 1$ variables. If we define

$$X_n = f(Y_n, \ldots, Y_{n+m}),$$

then the $(m + 1)$-block factor $X = (X_n)$ is an m-dependent process.

In this article we restrict our attention to one-dependent processes X which are *stationary* and *assume two values only*, denoted in the following by 0 and 1. It is not difficult to see that if X is a *two-block factor*, then it may be assumed that the underlying independent sequence Y is identically distributed with the

Received October 1987; revised May 1988.
[1]The last author was supported by the Netherlands Foundation for Mathematics (SMC) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO).
AMS 1980 *subject classifications.* 60G10, 28D05, 54H20.
Key words and phrases. Stationary process, one-dependence, m-dependence, block factors, cylinder functions, dynamical systems.

uniform distribution on the unit interval as the common distribution, and that f can be identified with the subset A of the unit square on which it assumes one of the values, say 1. Hence the distribution of a two-block factor is completely described by a measurable subset A of the unit square, which we call an *indicator* of the two-block factor. Of course, different A's may give rise to two-block factors having the same distributions.

It is natural to ask ([6], [7], [9]) whether all one-dependent processes arise as two-block factors. Under certain extremal conditions, this is true ([3]). However, in the following we produce a two-parameter family of stationary 0-1-valued one-dependent processes which are *not* two-block factors. This extends a one-parameter family of such examples recently obtained by two of us [1] based on unpublished results of the other two of us.

The plan of the article is as follows. In Section 1 we show that every one-dependent process can be parametrized by the collection of probabilities it associates to runs of 1's. Here we define *cylinder functions* for arbitrary parameter values and note that a one-dependent process exists if and only if the corresponding one-dependent cylinder function assumes only nonnegative values.

In general, it seems to be difficult to decide whether a given set of parameter values yields a positive cylinder function and thus a process. However, if we restrict our attention to a class of cylinder functions which we call *special* (for lack of a better name), defined by requiring that three or more 1's in a row have probability 0, then an effective algorithm can be given to decide whether a one-dependent process, with prescribed values of the probabilities α of a single 1 and β of two successive 1's, exists. In Section 2 we present the basis for this algorithm.

Section 3 contains a classification of those pairs (α, β) corresponding to special two-block factors. This section is essentially independent of the other results.

In Section 4 we continue the development of our algorithm, which has the following form. Two mappings ϕ_0 and ϕ_1, depending on α and β, are defined on $\mathbb{R}^2$, and a special one-dependent process exists for (α, β) if and only if the orbit of $(1,1)$ under successive applications of ϕ_0 and ϕ_1 in any order always remains in the unit square. Section 4 is devoted to dynamical properties of the more complicated mapping ϕ_0.

Theorem 5 of Section 5 contains the final form of our algorithm, and the remainder of this section is devoted to the determination of those (α, β) giving rise to special one-dependent processes. Although we have an effective decision procedure for any given pair (α, β), the time needed for decision grows as (α, β) approaches $(\frac{1}{3}, \frac{1}{27})$ and no closed form expression for the admissible set of parameters in a neighborhood of this point has been found. Away from this point, things become easier, and several results are given. For example, if $0 \le \alpha \le \frac{1}{4}$, then a special one-dependent process exists for every $0 \le \beta \le \frac{1}{4}\alpha$ (and no other β), whereas a two-block factor requires (for $0 \le \alpha \le \frac{2}{9}$)

$$0 \le \beta \le \tfrac{1}{8}(1 + \sqrt{1 - 4\alpha})\alpha.$$

The sum of our investigations is recorded in Figure 2 of Section 5.

It is the opinion of the authors that this paper raises more questions than it resolves. We mention two such questions. First of all, our methods are algebraic in nature and seem to give no probabilistic mechanism to produce the processes which we have discovered. In particular, we have not been able to determine if they are m-block factors for some $m \geq 3$. Second, our methods for studying ϕ_0 and ϕ_1 are at best amateuristic, and a more canonical approach is desirable.

1. Cylinder functions. Let W be the set of all finite sequences of 0's and 1's. An element of W is called a *word*. The empty word will be denoted by e and the word consisting of n 1's by 1^n. If $w_1, \ldots, w_n \in W$, then $w = w_1 \cdots w_n \in W$ is the concatenation of the words $w_1, \ldots, w_n$, and the w_i are *subwords* of w.

DEFINITION. A (normalized) *cylinder function* is a mapping

$$\mu\colon W \to \mathbb{R}$$

such that

(i) $$\mu(e) = 1,$$

(ii) $$\mu(w) = \mu(0w) + \mu(1w), \qquad w \in W,$$

(iii) $$\mu(w) = \mu(w0) + \mu(w1), \qquad w \in W.$$

The cylinder function μ is *positive* if

$$\mu(w) \geq 0, \qquad w \in W,$$

and *one-dependent* if

$$\mu(v)\mu(w) = \mu(v0w) + \mu(v1w), \qquad v, w \in W.$$

By elementary measure theory, the set of positive cylinder functions is in one-to-one correspondence with the set of distributions of stationary 0–1-valued discrete time stochastic processes, $\mu(w)$ being the probability of "seeing" the word w. Moreover, such a process is one-dependent if and only if its corresponding cylinder function is one-dependent.

THEOREM 1. *Let $\gamma = (\gamma_1, \gamma_2, \ldots)$ be any sequence of real numbers. Then there exists a unique one-dependent cylinder function μ_γ such that*

$$\mu_\gamma(1^n) = \gamma_n, \qquad n \geq 1.$$

PROOF. In the proof of this theorem and the next theorem, we denote the number of zeroes in a word w by $n_0(w)$. Set $\gamma_0 = 1$. The requirement, together with (i) of the definition of a cylinder function, defines $\mu_\gamma(w)$ for all $w \in W$ with $n_0(w) = 0$. We now proceed by induction on $n_0(w)$, as follows. If $w \in W$ with $n_0(w) > 0$, then clearly

$$w = 1^n 0 v$$

for some $n \geq 0$ and $v \in W$, and

$$n_0(v) = n_0(w) - 1.$$

One-dependence now dictates that

$$\mu_\gamma(1^n)\mu_\gamma(v) = \mu_\gamma(w) + \mu_\gamma(1^{n+1}v),$$

and since $n_0(1^{n+1}v) = n_0(v) < n_0(w)$, the formula

$$\mu_\gamma(w) = \gamma_n\mu_\gamma(v) - \mu_\gamma(1^{n+1}v)$$

defines μ_γ inductively on all of W. Straightforward induction arguments now show that μ_γ is a one-dependent cylinder function, whose uniqueness is obvious from the inductive definition. □

Theorem 2. *If for some $m \geq 1$ we have*

$$\gamma_m = \gamma_{m+1} = \cdots = 0,$$

and if 1^m is a subword of $w \in W$, then

$$\mu_\gamma(w) = 0.$$

Proof. The hypothesis states that $\mu_\gamma(w) = 0$ if $n_0(w) = 0$ and if 1^m is a subword of w. Now proceed by induction: If $n_0(w) > 0$, write as above

$$w = 1^n 0 v,$$

with

$$\mu_\gamma(w) = \gamma_n\mu_\gamma(v) - \mu_\gamma(1^{n+1}v).$$

If 1^m is a subword of w, then either $n \geq m$ and $\gamma_n = 0$ or $n < m$ and 1^m is a subword of v. In both cases, 1^m is a subword of $1^{n+1}v$, and hence $\mu_\gamma(w) = 0$ by induction. □

In the sequel we restrict our attention exclusively to one-dependent cylinder functions $\mu = \mu_\gamma$ for which $\gamma_3 = \gamma_4 = \cdots = 0$. For the sake of brevity (and in want of a more suitable name), such μ are called *special*. By Theorem 2, if μ is special and if 111 is a subword of w, then $\mu(w) = 0$. Hence positive special cylinder functions correspond bijectively to stationary 0–1-valued one-dependent processes for which the probability of three 1's in a row is 0; we refer to these as *special processes*.

Remark 1. Suppose that μ is a one-dependent cylinder function such that $\mu(w) = 0$ whenever 101 is a subword of w. Set $\alpha = \mu(1)$ and $\beta = \mu(11)$. Then

$$\mu(11111) = \mu(11)\cdot\mu(11) - \mu(11011) = \beta^2,$$

but also

$$\begin{aligned}\mu(11111) &= \mu(111)\cdot\mu(1) - \mu(11101)\\ &= \mu(1)(\mu(1)\cdot\mu(1) - \mu(101))\\ &= \mu(1)\cdot\mu(1)\cdot\mu(1) = \alpha^3.\end{aligned}$$

Hence $\beta^2 = \alpha^3$. This remark is intended to persuade the reader to examine the induction arguments of the above proofs carefully.

REMARK 2. Theorem 1 can be viewed as a parametrization result for one-dependent cylinder functions with parameter γ: Each cylinder function yields a parameter, different cylinder functions possess different parameters and γ is the parameter of a process if and only if μ_γ is positive. In the sequel, we set

$$\gamma_1 = \alpha, \qquad \gamma_2 = \beta, \qquad \gamma_3 = \gamma_4 = \cdots = 0$$

and discuss the *admissible pairs* (α, β) yielding special processes.

2. Positivity of special cylinder functions. In this section we derive a necessary and sufficient condition for the positivity of the special cylinder function defined by

$$\mu(1) = \alpha, \qquad \mu(11) = \beta, \qquad \mu(1^n) = 0, \qquad n \geq 3.$$

By Theorem 2, we need only examine words not having 111 as a subword. Let V be the set of all such words and denote by V_n those words of V having exactly n 0's. Then

$$V_0 = \{e, 1, 11\},$$

and if we define the set of words

$$U = \{0, 10, 110\},$$

then for each $n \geq 0$ the set of words V_n can be identified with

$$U^n \times V_0.$$

That is, each $v \in V_n$ has a unique representation

$$v = u_n u_{n-1} \cdots u_1 v_0,$$

with $v_0 \in V_0$ and $u_k \in U$, $1 \leq k \leq n$.

We now describe an algorithm for calculating the values of $\mu(v)$, $v \in V$. For each $v \in V$, define the column vector $\mathbf{v} \in \mathbb{R}^3$ by

$$\mathbf{v} = \begin{pmatrix} x(v) \\ y(v) \\ z(v) \end{pmatrix}$$

with

$$x(v) = \mu(0v), \qquad y(v) = \mu(10v), \qquad z(v) = \mu(110v).$$

Also set

$$\mathbf{f} = \begin{pmatrix} \mu(e) \\ \mu(1) \\ \mu(11) \end{pmatrix} = \begin{pmatrix} 1 \\ \alpha \\ \beta \end{pmatrix}.$$

Finally, define the 3×3 matrices

$$M_0 = \begin{pmatrix} 1 & -1 & 0 \\ \alpha & 0 & -1 \\ \beta & 0 & 0 \end{pmatrix}, \qquad M_{10} = \begin{pmatrix} 0 & 1 & -1 \\ 0 & \alpha & 0 \\ 0 & \beta & 0 \end{pmatrix}, \qquad M_{110} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & \alpha \\ 0 & 0 & \beta \end{pmatrix}$$

indexed by elements of U.

THEOREM 3. *If $v \in V_n$, then*

$$\mathbf{v} = M_{u_n} M_{u_{n-1}} \cdots M_{u_1} M_{v_0 0} \mathbf{f}.$$

PROOF. The case $n = 0$ is easily checked from the definitions. Now use induction on n, together with the following formulas:

$$x(v) = \mu(0v) = \mu(v) - \mu(1v) = \begin{cases} \mu(v) - \mu(10v'), & \text{if } v = 0v', \\ \mu(v) - \mu(110v'), & \text{if } v = 10v', \\ \mu(v), & \text{if } v = 110v', \end{cases}$$

$$y(v) = \mu(10v) = \alpha\mu(v) - \mu(11v) = \begin{cases} \alpha\mu(v) - \mu(110v'), & \text{if } v = 0v', \\ \alpha\mu(v), & \text{if } v = (1)10v', \end{cases}$$

$$z(v) = \mu(110v) = \beta\mu(v) - \mu(111v) = \beta\mu(v).$$

The formula for $x(v)$ shows that the first rows of the matrices M are correct, and those for $y(v)$ and $z(v)$ verify the second and third row, respectively. □

COROLLARY. *For $(x, y) \in \mathbb{R}^2 \setminus \{(x, y)\colon xy = 0\}$ set*

$$\phi_0(x, y) = \left(1 - \frac{\alpha y}{x}, 1 - \frac{\beta}{\alpha x}\right),$$

$$\phi_1(x, y) = \left(1 - \frac{\beta}{\alpha y}, 1\right).$$

Then the pair (α, β) is admissible if and only if either $\alpha = \beta = 0$ or $0 < \alpha \le 1$, $0 \le \beta \le \alpha$, and all iterates of the point $(1, 1)$ under successive applications of ϕ_0 and ϕ_1 in any order remain in the unit square $S = \{(x, y)\colon 0 < x \le 1, 0 < y \le 1\}$.

PROOF. Theorem 3 yields all values $\mu(v)$, $v \in V$, as iterates of $\mathbf{f}$ under the three M-matrices. In testing positivity we can disregard M_{110} since it brings us back to a multiple of $\mathbf{f}$. Next, reduce the dimension by normalizing such that the third coordinate is always equal to β, i.e., set

$$\Phi_0(x, y, \beta) = \frac{1}{x} M_0 \begin{pmatrix} x \\ y \\ \beta \end{pmatrix}, \qquad \Phi_1(x, y, \beta) = \frac{1}{y} M_{10} \begin{pmatrix} x \\ y \\ \beta \end{pmatrix}$$

and then drop β to obtain

$$\Phi_0(x, y) = \left(1 - \frac{y}{x}, \alpha - \frac{\beta}{x}\right),$$

$$\Phi_1(x, y) = \left(1 - \frac{\beta}{y}, \alpha\right),$$

with initial value $(1, \alpha)$. Clearly $\alpha = \mu(1)$ must lie in the unit interval, and $0 \le \beta = \mu(11) \le \alpha$ is also necessary. The case $\alpha = \beta = 0$ yields the special process which is given by all 0's, and if $\alpha > 0$, then we can replace y by αy, which results in the given ϕ_0 and ϕ_1, with initial value $(1,1)$. Noting now that if (x, y), $\phi_0(x, y)$, $\phi_1(x, y)$ have positive coordinates, then $\phi_0(x, y)$ and $\phi_1(x, y)$ cannot have a coordinate greater than 1 and that $x = 0$ or $y = 0$ leads to a negative coordinate, we see that the proof is finished. □

3. Determination of the parameter set corresponding to two-block factors. Let μ_A be the cylinder function corresponding to a two-block factor with indicator A, such that $\mu_A(111) = 0$. In this section we determine the range of possible values for $\alpha = \mu_A(1)$ and $\beta = \mu_A(11)$. By the definition, we have for any $n \geq 1$,

$$\mu_A(1^n) = \int_0^1 \cdots \int_0^1 1_A(x_0, x_1) \cdots 1_A(x_{n-1}, x_n)\, dx_0 \cdots dx_n.$$

Moreover, if T: $[0,1] \to [0,1]$ preserves Lebesgue measure, then A and $(T \times T)^{-1}(A)$ give rise to the same process.

Examples of sets A for which $\mu_A(111) = 0$ can be obtained in the following manner. Let $a, b \in [0,1]$ with $a \leq b$ and define

$$F(a, b) = ([a, b) \times [0, a)) \cup ([b, 1] \times [0, b)).$$

If $A \subseteq F(a, b)$ and if $(x_0, x_1) \in A$, $(x_1, x_2) \in A$, then clearly $x_1 < b$ and hence $x_2 < a$, so that no choice of x_3 permits $(x_2, x_3) \in A$. That is,

$$A \subseteq F(a, b) \Rightarrow \mu_A(111) = 0.$$

The following lemma shows that, up to a measure preserving transformation T, the reverse implication is valid.

Lemma. *If $\mu_A(111) = 0$, then there exists a transformation T: $[0,1] \to [0,1]$ preserving Lebesgue measure and $a, b \in [0,1]$ with $a \leq b$ such that*

$$(T \times T)^{-1} A \subseteq F(a, b)$$

modulo Lebesgue measure on the unit square.

Proof. Define

$$A_2 = \left\{x_2 \in [0,1]: \int_0^1 1_A(x_2, x_3)\, dx_3 > 0\right\},$$

$$A_1 = \left\{x_1 \in [0,1]: \int_{A_2} 1_A(x_1, x_2)\, dx_2 > 0\right\},$$

$$A_0 = \left\{x_0 \in [0,1]: \int_{A_1} 1_A(x_0, x_1)\, dx_1 > 0\right\}.$$

Then $A_2 \supseteq A_1 \supseteq A_0$, and the formula

$$0 = \mu_A(111) = \int_{A_0}\left(\int_{A_1} 1_A(x_0, x_1)\left(\int_{A_2} 1_A(x_1, x_2)\left(\int_0^1 1_A(x_2, x_3)\, dx_3\right) dx_2\right) dx_1\right) dx_0$$

allows us to conclude that the Lebesgue measure of A_0 is 0. Choosing

$$a = 1 - \text{Lebesgue measure}(A_2),$$
$$b = 1 - \text{Lebesgue measure}(A_1)$$

and $\dot{T}$ measure preserving with

$$T((a,1]) = A_2,$$
$$T((b,1]) = A_1$$

yields the desired result. □

In accordance with our previous usage, a set A such that $\mu_A(111) = 0$ will be called *special*. In order to calculate α and β, note that the first formula of this section for $n = 1$ and $n = 2$ reduces to

$$\alpha = \mu_A(1) = \text{Lebesgue measure}(A)$$

and

$$\beta = \mu_A(11) = \int_0^1 H_A(x) V_A(x)\, dx,$$

where $H_A(x)$ and $V_A(x)$ denote the Lebesgue measures of the horizontal and vertical sections of A at x, respectively. In particular, if $A \subseteq F(a, b)$, the part of A lying in the lower right rectangle $[b,1] \times [0, a)$ does not contribute to $\beta = \mu_A(11)$. A simple but tedious calculation (which we omit) now shows that for fixed α, the minimal value of β occurs when $A = F(a, b)$ for suitable a and b, and the maximal value of β (for $0 \le \alpha \le 2/9$) occurs when

$$A = G(a, b) := F(a, b) \setminus ([b,1] \times [0, a)),$$

again for suitable a and b. Further reduction eventually produces

THEOREM 4. *Let μ be a cylinder function with $\alpha = \mu(1)$, $\beta = \mu(11)$ and $0 = \mu(1^n)$ for $n \ge 3$. Then μ is the cylinder function of a two-block factor if and only if*

(i) $0 \le \alpha \le \frac{1}{3}$ *and*
(ii) $m(\alpha) \le \beta \le M(\alpha)$, *where*

$$m(\alpha) = \begin{cases} 0, & 0 \le \alpha \le \frac{1}{4}, \\ \frac{1}{3}\alpha - \frac{2}{27}\left\{1 + (1 - 3\alpha)^{3/2}\right\}, & \frac{1}{4} \le \alpha \le \frac{1}{3} \end{cases}$$

and

$$M(\alpha) = \begin{cases} \frac{1}{8}(1 + \sqrt{1 - 4\alpha})\alpha, & 0 \le \alpha \le \frac{2}{9}, \\ \frac{1}{27}, & \frac{2}{9} \le \alpha \le \frac{1}{3}. \end{cases}$$

For related results and similar calculation we refer to de Valk [2]. In the next sections we shall need the following observation.

LEMMA. *If* $A = F(a, b)$ *and* $\alpha = \mu_A(1)$, $\beta = \mu_A(11)$, *then the equation*

$$x^3 - x^2 + \alpha x - \beta = 0$$

has the three real roots $r_1 = a$, $r_2 = b - a$ *and* $r_3 = 1 - b$.

PROOF. One easily calculates

$$\alpha = a(b - a) + (1 - b)b = r_1 r_2 + r_2 r_3 + r_1 r_3$$

and

$$\beta = a(1 - b)(b - a) = r_1 r_2 r_3. \qquad \square$$

4. A study of ϕ_0. Before using the corollary of Section 2 to determine admissible pairs (α, β), we investigate the mapping ϕ_0. Recall that for fixed $0 < \alpha \leq 1$ and $0 \leq \beta \leq \alpha$,

$$\phi_0(x, y) = \left(1 - \frac{\alpha y}{x}, 1 - \frac{\beta}{\alpha x}\right).$$

4.1. *Fixed points.* These are given by solutions to the equations

$$x = 1 - \frac{\alpha y}{x}$$

and

$$y = 1 - \frac{\beta}{\alpha x};$$

eliminating y results in

$$\rho(x) := x^3 - x^2 + \alpha x - \beta = 0.$$

This equation can have either one real root and two complex roots, or three real roots. As $\rho(0) = -\beta \leq 0$ and $\rho(1) = \alpha - \beta \geq 0$, one root must lie in the unit interval. The sum of the roots is 1, so that if the other two are also real, they also lie in the unit interval, because they have the same sign. If we denote these roots by r_1, r_2, r_3 and set $a = r_1$, $b = r_1 + r_2$, then it follows from the lemma at the end of Section 3 that the cylinder function μ corresponding to the pair (α, β) is equal to μ_A, with $A = F(a, b)$. Hence we have proved the

PROPOSITION. *If* $x^3 - x^2 + \alpha x - \beta = 0$ *has three real roots and if* $0 < \alpha \leq 1$, $0 \leq \beta \leq \alpha$, *then the pair* (α, β) *is admissible and corresponds to a two-block factor with indicator* $A = F(a, b)$ *for suitable* a *and* b.

Having discovered the situation for three real roots, we now restrict our attention to those (α, β) for which

$$x^3 - x^2 + \alpha x - \beta = 0$$

has only one real root $x_0 \in [0,1]$. If now $x_0 = 0$, then we have $\beta = 0$ and $\alpha > \frac{1}{4}$, and a simple application of the corollary of Section 2 shows that (α, β) cannot be admissible. Hence we may also assume that $\beta > 0$ and $x_0 > 0$. Now set

$$y_0 = 1 - \frac{\beta}{\alpha x_0}.$$

By the foregoing, $\phi_0(x_0, y_0) = (x_0, y_0)$.

4.2. *Regions of increase and decrease.* Recall that

$$S = \{(x, y) \colon 0 < x \le 1, 0 < y \le 1\}$$

and define

$$\begin{aligned}
X_+ &= \left\{(x, y) \in S \colon 1 - \frac{\alpha y}{x} \ge x\right\},\\
X_- &= \left\{(x, y) \in S \colon 1 - \frac{\alpha y}{x} \le x\right\},\\
Y_+ &= \left\{(x, y) \in S \colon 1 - \frac{\beta}{\alpha x} \ge y\right\},\\
Y_- &= \left\{(x, y) \in S \colon 1 - \frac{\beta}{\alpha x} \le y\right\},\\
\mathrm{I} &= X_- \cap Y_-,\\
\mathrm{II} &= X_+ \cap Y_-,\\
\mathrm{III} &= X_+ \cap Y_+,\\
\mathrm{IV} &= X_- \cap Y_+,
\end{aligned}$$

thus dividing S into four regions whose boundaries are segments of the parabola

$$P \colon y = \frac{1}{\alpha} x(1 - x)$$

and/or the hyperbola

$$H \colon y = 1 - \frac{\beta}{\alpha x}.$$

Figure 1 has two parts, according to whether $\alpha \le \frac{1}{4}$ or $\alpha > \frac{1}{4}$.
By definition:

(i) If $(x, y) \in \mathrm{I}$, then $\phi_0(x, y)$ is to the left and below (x, y).
(ii) If $(x, y) \in \mathrm{II}$, then $\phi_0(x, y)$ is to the right and below (x, y).
(iii) If $(x, y) \in \mathrm{III}$, then $\phi_0(x, y)$ is to the right and above (x, y).
(iv) If $(x, y) \in \mathrm{IV}$, then $\phi_0(x, y)$ is to the left and above (x, y).

4.3. *Line segments.* It is trivial to check that if L is a straight line segment in S, then $\phi_0(L)$ is a straight line segment.

4.4. *Image of P.* It is trivial to check that $\phi_0(P) \subseteq H$.

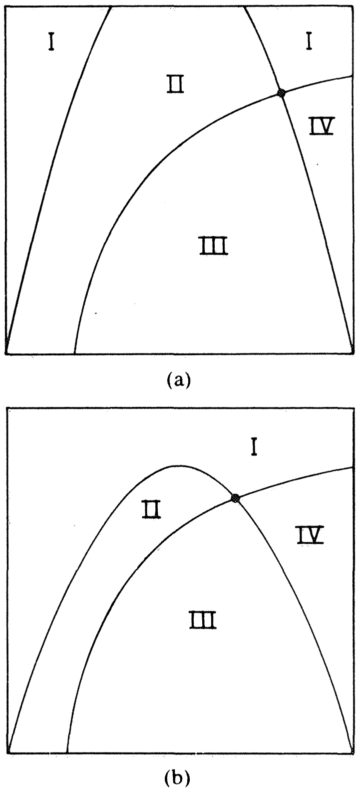


FIG. 1. (a) $\alpha \le \frac{1}{4}$; (b) $\alpha > \frac{1}{4}$ [$\bullet = (x_0, y_0)$].

4.5. *Images of regions.* It follows from Sections 4.3, 4.4 and the definitions that

$$\phi_0(\mathrm{I}) \cap S \subseteq \mathrm{I} \cup \mathrm{II},$$
$$\phi_0(\mathrm{II}) \cap S \subseteq \mathrm{III},$$
$$\phi_0(\mathrm{III}) \cap S \subseteq \mathrm{III} \cup \mathrm{IV},$$
$$\phi_0(\mathrm{IV}) \cap S \subseteq \mathrm{I}.$$

4.6. *Entering region* II. We now show that our hypothesis of one real root (= one point of intersection of P and H) implies that for each $(x, y) \in \mathrm{I}$, there exists n such that

$$\phi_0^{(n)}(x, y) \notin \mathrm{I},$$

i.e., either

$$\phi_0^{(n)}(x, y) \in \mathrm{II}$$

or it leaves S. Assume the contrary. Then Sections 4.2 and 4.3 imply that some line through (x_0, y_0) must be taken into itself by ϕ_0. If L is such a line then either

(i) L is vertical or
(ii) L is tangent to P at (x_0, y_0) or
(iii) L intersects P (not necessarily in S) at some point $(x_1, y_1) \neq (x_0, y_0)$.

Now (i) is impossible because $1 - (\alpha y)/x_0$ cannot be equal to x_0 for more than one value of y and (ii) implies (by Section 4.4) that P and H are tangent at (x_0, y_0), which says that x_0 is a root of multiplicity three of $\rho(x) = 0$ and is excluded by hypothesis. But (iii) also is impossible, since ϕ_0 maps $(x_1, y_1) \in P$ to a point $(x_1, y_1') \in H$ (by Section 4.4) with $y_1' \neq y_1$.

4.7. *Invariant polygons.* Let $0 < t < 1$ and set $x_1 = t,\ y_1 = 1$. Suppose that we successively apply ϕ_0 to (x_1, y_1), obtaining a sequence (x_n, y_n) which remains in S. Then by Sections 4.6 and 4.2 there is an $n \geq 1$ such that

$$(x_k, y_k) \in \mathrm{I} \text{ for } 1 \leq k < n,$$

$$(x_n, y_n) \in \mathrm{II}$$

and

$$(x_{n+1}, y_{n+1}) \in \mathrm{III}.$$

We now claim that the points

$$(1,1), (x_1, y_1), \ldots, (x_{n+1}, y_{n+1}), (1, y_{n+1})$$

are the vertices of a convex polygon $C(t)$, and that $\phi_0(C(t)) \subseteq C(t)$. By the properties in Section 4.2, connecting the given points in the given order forms a nonself-intersecting polygon, and the inclusion is obvious if one notes that $\phi_0(1,1)$ lies on the line segment joining $(1,1)$ and (x_2, y_2) and that $\phi_0(1, y_{n+1})$ lies on the line segment joining $\phi_0(1,1)$ and $(1, 1 - \beta/\alpha) = \phi_0(1,0)$. The convexity of $C(t)$ is also easy to show, but we omit the calculation as it is not needed in the sequel.

5. Determination of admissibility. Now we can use the results of the previous section, together with the corollary of Section 2, to determine the admissibility of a given pair (α, β). Suppose first that (α, β) is admissible; if C denotes the convex hull of the orbit closure of $(1,1)$ under ϕ_0 and ϕ_1, then $\phi_0(C) \subseteq C$, $\phi_1(C) \subseteq C$ and $C \subseteq S$. Now set

$$y^* = \min\{y\colon (x, y) \in C\},$$

$$t^* = 1 - \frac{\beta}{\alpha y^*},$$

$$L^* = \{(x,1)\colon t^* \leq x \leq 1\}.$$

Then $\phi_1(x, y^*) = (t^*, 1)$ implies that $L^* \subseteq C$. If we set

$$t = \min\{x\colon (x,1) \in C\},$$

then $t \leq t^*$ and the ϕ_0 invariant polygon $C(t)$ of the previous section is also contained in C and hence ϕ_1-invariant. Conversely, if for some $0 < t < 1$ the polygon $C(t)$ is also ϕ_1-invariant, then clearly (α, β) is admissible, since the orbit of $(1,1)$ is contained in $C(t)$. We have shown

THEOREM 5. *The pair (α, β) is admissible if and only if*

(i) $0 \leq \alpha \leq 1$, $0 \leq \beta \leq \alpha$ *and either*
(ii) *the equation*

$$x^3 - x^2 + \alpha x - \beta = 0$$

has three (not necessarily distinct) real roots or
(ii′) *the equation*

$$x^3 - x^2 + \alpha x - \beta = 0$$

has exactly one real root, and there exists $t \in (0,1)$ such that $C(t)$ is well defined [i.e., the ϕ_0-orbit of $(1,1)$ enters region III *without previously leaving S] and such that*

$$1 - \frac{\beta}{\alpha y^*} \geq t,$$

where

$$y^* = \min\{y\colon (x, y) \in C(t)\}.$$

A computer program has been written which decides, within the limits of machine accuracy, whether for given (α, β) the conditions of the above theorem are verified or not, and a copy is available on request. Moreover, we have the following rigorous results concerning admissibility.

1. If (α, β) is admissible, then $0 \leq \alpha \leq \frac{1}{2}$ and $0 \leq \beta \leq \alpha/4$.
2. If $0 \leq \alpha \leq \frac{1}{4}$ and $0 \leq \beta \leq \alpha/4$, then (α, β) is admissible.
3. If $\frac{1}{4} < \alpha < \frac{1}{2}$ and $2\alpha^{3/2} - \alpha \leq \beta \leq \frac{1}{2}(\alpha - \alpha^{3/2})$, then (α, β) is admissible.
4. In the following ranges, (α, β) is not admissible:
   (i) $\frac{1}{4} \leq \alpha \leq \frac{1}{3}$ and $27\beta < 9\alpha - 2(1 - 3\alpha)^{3/2} - 2$,
   (ii) $\frac{1}{3} \leq \alpha \leq \frac{1}{2}$ and $27\beta < 9\alpha - 2$ and
   (iii) $\frac{1}{4} < \alpha \leq \frac{1}{2}$ and $\beta > \frac{1}{2}(\alpha - \alpha^{3/2})$.

These results, together with the two-block factor region, are summarized in Figure 2.

Finally, we sketch our proofs of results 1–4.

1. If $\alpha > \frac{1}{2}$, then either the x-coordinate of $\phi_0\phi_1(1,1)$,

$$\frac{1 - \alpha - \beta/\alpha}{1 - \beta/\alpha},$$

is negative, or if this is nonnegative, the x-coordinate of $\phi_0^2\phi_1(1,1)$,

$$\frac{(1 - 2\alpha)(1 - \beta/\alpha)}{1 - \alpha - \beta/\alpha},$$

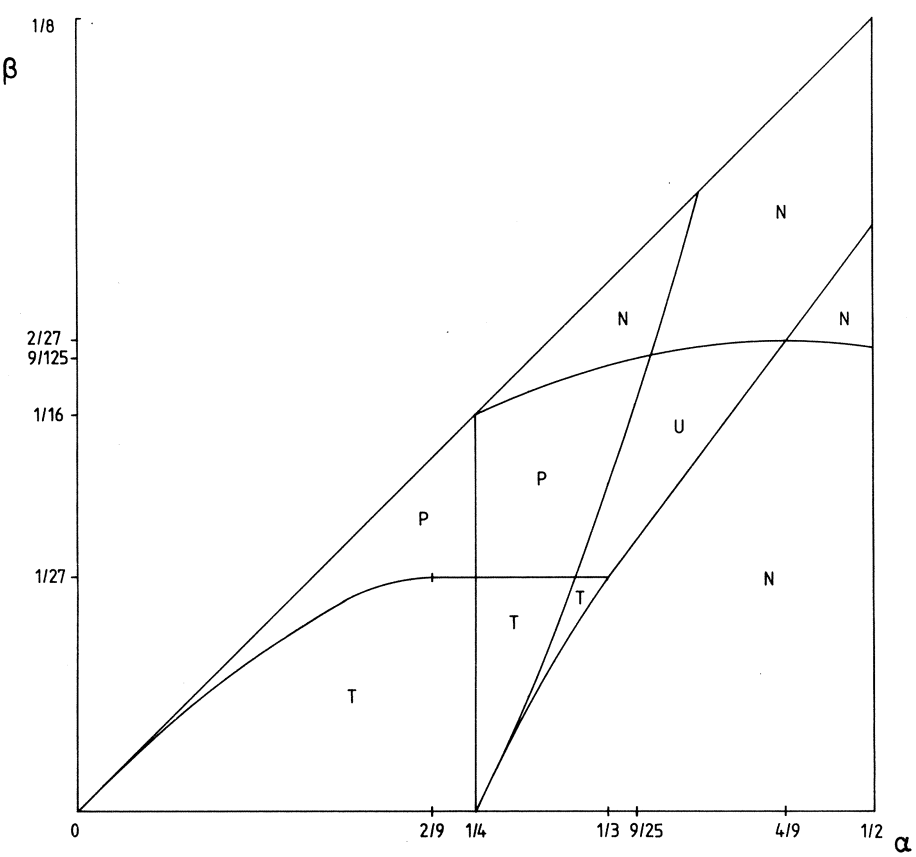


FIG. 2. *T, one-dependent processes which are two-block factors, P, one-dependent processes which are not two-block factors, N, no one-dependent processes, U, unexplored.*

is negative. If $\beta > \alpha/4$, then $(\phi_1\phi_0)^n(1,1)$ becomes negative in its x-coordinate for some n, since

$$(\phi_1\phi_0)^n(1,1) = \left(g^{2n}(1),1\right)$$

with

$$g(t) = 1 - \frac{\beta}{\alpha t},$$

and $g^{2n}(1)$ is eventually negative iff $t = g(t)$ has no real root, leading to $\beta > \alpha/4$.

2. This is the simplest polygon case, corresponding to $\alpha \leq \frac{1}{4}$ in Figure 1. Here $(\frac{1}{2},1)$ belongs to region II, so

$$\phi_0\left(\frac{1}{2},1\right) = \left(1 - 2\alpha, 1 - \frac{2\beta}{\alpha}\right) \in \mathrm{III}.$$

The quadrilateral with vertices

$$(1,1), \qquad \left(\frac{1}{2},1\right), \qquad \phi_0\left(\frac{1}{2},1\right), \qquad \left(1,1-\frac{2\beta}{\alpha}\right)$$

has lowest y-coordinate

$$y^* = 1 - \frac{2\beta}{\alpha}$$

with

$$1 - \frac{\beta}{\alpha y^*} \geq \frac{1}{2}$$

and is thus invariant under ϕ_0 and ϕ_1.

3. This is the next polygon case. For $t \in (0,1)$, $C(t)$ is a pentagon [i.e., $\phi_0(t,1) \in$ II] if t satisfies

$$\alpha t^2 - (\alpha + \beta)t + \alpha^2 \leq 0,$$

and ϕ_1-invariance holds if

$$(\alpha - \beta)t^2 - (\alpha^2 + \alpha - 2\beta)t + \alpha(\alpha - \beta) \leq 0.$$

Discriminant calculation and elementary considerations lead to the bounds given in result 3.

4(i) and 4(ii). Here one can show directly that $\mu(0^n) = z_n$ is negative for some n. By one-dependence one easily derives the recurrence

$$z_n = z_{n-1} - \alpha z_{n-2} + \beta z_{n-3},$$

whose characteristic equation is

$$\rho(x) = x^3 - x^2 + \alpha x - \beta = 0.$$

In the ranges indicated, there is one real root and two complex roots whose real part is larger than the real root, and it follows that z_n becomes negative.

4(iii). Here we have (similar result to 1)

$$(\phi_1\phi_0^2)^n(1,1) = (g^n(1),1)$$

with

$$g(t) = \frac{(\alpha - 2\beta)t - \alpha(\alpha - \beta)}{(\alpha - \beta)t - \alpha^2}.$$

Hence if $g(t) = t$ has no real root, then $g^n(1)$ becomes negative for some n. A discriminant calculation leads to the given bound.

Acknowledgments. This paper is a result of the close and ongoing collaboration between members of the School of Mathematical Sciences of Tel Aviv University and the Vakgroep Statistiek, Stochastiek en Operationele Analyse of the Delft University of Technology. We would like to thank both institutions warmly for their continued and substantial support of this collaboration.

REFERENCES

[1] AARONSON, J. and GILAT, D. (1987). On the structure of stationary one-dependent processes. School of Mathematical Sciences, Tel Aviv Univ., Israel.

[2] DE VALK, V. (1988). The maximal and minimal 2-correlation of a class of 1-dependent 0–1 valued processes. *Israel J. Math.* **62** 181–205.

[3] DE VALK, V. (1988). One-dependent processes. Dissertation, Delft University Press, Delft.

[4] HAIMAN, M. G. (1981). Valeurs extrémales de suites stationnaires de variables aléatoires m-dépendantes. *Ann. Inst. H. Poincaré Sect. B (N.S.)* **17** 309–330.

[5] HOEFFDING, W. and ROBBINS, H. (1948). The central limit theorem for dependent random variables. *Duke Math. J.* **15** 773–780.

[6] IBRAGIMOV, I. A. and LINNIK, Y. V. (1971). *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff, Groningen.

[7] JANSON, S. (1984). Runs in m-dependent sequences. *Ann. Probab.* **12** 805–818.

[8] O'BRIEN, G. L. (1980). Scaling transformations for $\{0,1\}$-valued sequences. *Z. Wahrsch. verw. Gebiete* **53** 35–49.

[9] O'CINNEIDE, C. A. (1987). Some properties of one-dependent sequences. Unpublished.

[10] VAN DEN BERG, J. (1986). On some results by S. Janson concerning runs in m-dependent sequences. Unpublished.

JON AARONSON AND DAVID GILAT
SCHOOL OF MATHEMATICAL SCIENCES
TEL AVIV UNIVERSITY
RAMAT AVIV
69978 TEL AVIV
ISRAEL

MICHAEL KEANE AND VINCENT DE VALK
DEPARTMENT OF MATHEMATICS AND INFORMATICS
DELFT UNIVERSITY OF TECHNOLOGY
JULIANALAAN 132
2628 BL DELFT
THE NETHERLANDS

Probab. Th. Rel. Fields 80, 475–480 (1989)

Probability Theory and Related Fields

Extremal Two-Correlations of Two-Valued Stationary One-Dependent Processes

A. Gandolfi*, M. Keane, and V. de Valk**

Delft University of Technology, Faculty of Technical Mathematics and Informatics, Julianalaan 132. 2628 BL Delft, The Netherlands

Summary. The maximal value of the two-correlation for two-valued stationary one-dependent processes with fixed probability α of a single symbol is determined. We show that the process attaining this bound is unique except when $\alpha = 1/2$, when there are exactly two different processes. The analogous problem for minimal two-correlation is discussed, and partial results are obtained.

Introduction

Let μ be the distribution on sequence space of a discrete time, stationary, 0-1-valued one-dependent stochastic process. Suppose that the probability of a one is α. Then, generalizing and simplifying a result of M. Katz ([3], see also L. Finke [2]), we show that the probability of two ones in a row is at most

$$\alpha^{3/2} \quad \text{if} \quad 1/2 \leqq \alpha \leqq 1$$

and

$$(2\alpha - 1) + (1-\alpha)^{3/2} \quad \text{if} \quad 0 \leqq \alpha \leqq 1/2.$$

Moreover, if equality holds we show that there is a unique process with this two-correlation for $\alpha \neq 1/2$, and exactly two processes when $\alpha = 1/2$. These extremal processes are identified as two-block factors of two-state Bernoulli processes.

In the second section we discuss the minimal possible two-correlations for one-dependent processes. For $0 \leqq \alpha \leqq 1/4$ the corresponding results holds trivially and for $1/4 < \alpha \leqq 1/3$ we can produce a bound which is attained, but we do not know whether uniqueness holds at the bound. We conjecture a value for the lower bound for all α, and also that uniqueness holds at this value.

1. Maximal Two-Correlations of Two-Valued Stationary One-Dependent Processes

The distribution of a stationary 0-1-valued stochastic process is given by a shift-invariant probability measure on the space X of doubly infinite sequences

* Supported by C.N.R., Italy

** Supported by Z.W.O., The Netherlands

of zeroes and ones. As we shall only be interested in distributional properties of such objects, such a measure will be called a *process*. If w is a finite sequence of zeroes and ones, then the probability of "seeing" w starting from a given time t does not depend upon t. We denote this probability by $[w]$, suppressing the measure.

Definition 1. A stationary 0-1-valued stochastic process is said to be *one-dependent* if for any finite sequences u and v,

$$[u*v]=[u]\cdot[v],$$

where $[u*v]$ is defined by

$$[u*v]=[u0v]+[u1v].$$

Example. Fix $0\leqq\alpha\leqq 1$. Define the mapping

$$\varphi\colon X\to X$$

by setting

$$\varphi(x)_t := x_t\, x_{t+1}$$

for $x=(x_t)\in X$. Let μ_α be the image under φ of the product (Bernoulli) measure on X which assigns the probability $\sqrt{\alpha}$ to the symbol one in each coordinate. Clearly μ_α is one-dependent, with

$$[1]=\alpha$$

and

$$[11]=\alpha^{3/2}.$$

Theorem 1. *Let μ be one-dependent with $[1]=\alpha$ and $1/2<\alpha\leqq 1$. Then*

$$[11]\leqq\alpha^{3/2}.$$

Moreover, if $[11]=\alpha^{3/2}$ then $\mu=\mu_\alpha$.

Proof. Set $\alpha=[1]$ and $\beta=[11]$. By one-dependence and linearity we have

$$\begin{aligned}
[11010]&=[11]\cdot[10]-[11110]\\
&=\beta(\alpha-\beta)-[111]\cdot[0]+[11100]\\
&=\beta(\alpha-\beta)-([1]\cdot[1]-[101])(1-\alpha)+[1]\cdot[100]-[10100]\\
&=\beta(\alpha-\beta)-\alpha^2(1-\alpha)+[101](1-\alpha)+\alpha([10]-[101])-[10100]\\
&=\beta(\alpha-\beta)-\alpha^2(1-\alpha)+\alpha(\alpha-\beta)+(1-2\alpha)[101]-[10100]\\
&=+\alpha^3-\beta^2+(1-2\alpha)[101]-[10100].
\end{aligned}$$

From $[11010]\geqq 0$ and $\alpha>1/2$ we conclude that

$$\beta^2\leqq\alpha^3,$$

which proves the first assertion. Moreover, if $\beta^2=\alpha^3$ then it follows that $[101]=0$. This implies that there is at most one one-dependent process with $[1]=\alpha$ and $[11]=\beta$, since the knowledge of the measure of *one* cylinder set of length

n for each $n \geqq 1$ clearly fixes the measure of each cylinder set by a simple calculation using one-dependence, and $[101]=0$ implies $[w]=0$ for any w containing 101. But the measure μ_α of the example satisfies $[101]=0$, so that we must have $\mu=\mu_\alpha$.

Definition 2. Denote by $\tilde{\mu}_\alpha$ the measure on X which is the image of μ_α under the map from X to X which interchanges all zeroes and ones.

Theorem 2. *Let μ be one-dependent with $[1]=\alpha$.*
Then

$$[11] \leqq \begin{cases} \alpha^{3/2} & \text{if } 1/2 \leqq \alpha \leqq 1 \\ 2\alpha-1+(1-\alpha)^{3/2} & \text{if } 0 \leqq \alpha \leqq 1/2. \end{cases}$$

Moreover, if equality holds, then

$$\mu = \begin{cases} \mu_\alpha & \text{if } 1/2 < \alpha \leqq 1 \\ \tilde{\mu}_\alpha & \text{if } 0 \leqq \alpha < 1/2 \end{cases}$$
$$\text{either } \mu_{1/2} \quad \text{if } \alpha = 1/2$$
$$\text{or } \tilde{\mu}_{1/2}$$

Proof. If $\alpha > 1/2$, then this is just Theorem 1, and for $\alpha < 1/2$ the statement follows by interchanging zeroes and ones, since then $[1] > 1/2$ and

$$[00] = -1 + 2 \cdot [0] + [11].$$

For the case $\alpha = 1/2$, return to the calculation of Theorem 1, which shows that the inequality holds, and also that if equality holds, then

$$[10100] = 0.$$

But then

$$[1010] \cdot [00] = [1010 * 00] = [1010000] + [1010100] = 0.$$

which shows that $[1010]=0$, since $[00]=\sqrt{2}/4$. Similarly,

$$[101] \cdot [010] = [101 * 010] = [1010010] + [1011010] = 0,$$

so that either $[101]=0$ and $\mu=\mu_{1/2}$ or $[010]=0$ and $\mu=\tilde{\mu}_{1/2}$.

Remarks. 1. In particular, Theorem 2 applies to those one-dependent processes which are two-block factors. That is, the results of [2] and [3] are corollaries of Theorem 2, which is both more general (see [1]) and easier to prove.

2. A solution for the discrete version of the question raised in [3] is contained in [5]. Unfortunately, the above method does not seem to be applicable.

2. Minimal Two-Correlations of Two-Valued Stationary One Dependent Processes

We are not able to prove as much as in section one, although we suspect that similar results are valid. Our notation is the same as in the first paragraph.

Case 1. Suppose μ is one-dependent with $\alpha=[1]$ and $0\leqq\alpha\leqq 1/4$. Then clearly

$$[11]\geqq 0$$

and zero is the best lower bound. The map $\psi\colon X\to X$ with

$$\psi(x)_t := x_t(1-x_{t+1})$$

carries the Bernoulli measure with parameter γ(=probability of one) to a one-dependent measure with $[11]=0$ and

$$[1]=\gamma(1-\gamma);$$

for $0\leqq\alpha\leqq 1/4$ we can choose γ such that $\alpha=\gamma(1-\gamma)$.

Case 2. Let μ be one-dependent with $\alpha=[1]$ and $1/4<\alpha\leqq 1/3$. Then we can show that

$$[11]\geqq\frac{(1-2\sqrt{1-3\alpha})(1+\sqrt{1-3\alpha})^2}{27},$$

and exhibit a measure μ with equality, but we do not know whether this measure is unique.

In general, we suspect that if $N=\left[\dfrac{1}{1-2\alpha}\right]$ (in Case 1, $N=1$ and in Case 2, $N=2$), then

$$[11]\geqq\frac{N(N-1)}{6(N+1)^2}(1-2\delta)(1+\delta)^2,$$

with

$$\delta=\sqrt{1-2\alpha\left(\frac{N+1}{N}\right)}.$$

Particularly intriguing is the Case infinity, when $\alpha=1/2$; here we conjecture that

$$[11]\geqq 1/6$$

with uniqueness at equality. The article [4] shows that the bounds given above are attained and unique in the class of two-block factors.

We now prove the result stated above in Case 2. The proof will be divided into two parts.

Part 1. Assume that μ is one-dependent with

$$1/4 < [1] = \alpha \leqq 1/3,$$
$$[11] = \beta < \frac{(1-2\sqrt{1-3\alpha})(1+\sqrt{1-3\alpha})^2}{27}$$

and

$$[111] = 0.$$

(The assumption $[111]=0$ will be removed in Part 2.)

Setting $f(n) := [0^n]$, we shall derive a recurrence relation for $f(n)$ and show that under the above conditions, there exists n such that $f(n) < 0$, yielding a contradiction. We have for $n \geqq 4$

$$\begin{aligned} f(n) = [0^n] &= [0^{n-1}] - [0^{n-1}1] \\ &= f(n-1) - [0^{n-2}]\cdot[1] + [0^{n-2}11] \\ &= f(n-1) - \alpha f(n-2) + [0^{n-3}]\cdot[11]. \end{aligned}$$

Since $[0^{n-3}111]=0$ by assumption. Hence

$$f(n) = f(n-1) - \alpha f(n-2) + \beta f(n-3).$$

A simple calculation now shows that the characteristic polynomial

$$P(x) = x^3 - x^2 + \alpha x - \beta$$

for $f(n)$ has one real root λ_1 and two complex roots λ_2 and $\lambda_3 = \overline{\lambda_2}$, and that

$$|\lambda_2| > \lambda_1,$$

for a given bound on β. This implies that for some n, $f(n) < 0$, since $f(n)$ is a linear combination of the λ_i^n, $1 \leqq i \leqq 3$, with non-zero coefficients.

Part 2. The assumptions are as in Part 1 except that $[111] > 0$. Now let

$$g(n) := [0^n].$$

We claim that for each $n \geqq 4$,

$$g(n) = f(n) - \sum_{k=0}^{n-4} f(n-k-4)\cdot[0^k 1^3] - [0^{n-3}1^3]. \qquad (*)$$

where $f(n)$ is as in part 1. For small n, we have

$$\begin{aligned} g(1) &= f(1) \\ g(2) &= f(2) \\ g(3) &= f(3) - [1^3], \end{aligned}$$

and $(*)$ follows easily by induction for $n \geqq 4$. Now Part 1 implies that there is a first n for which $f(n) < 0$, so that for this n, we also have $g(n) < 0$ by the above. This concludes the proof of Case 2.

References

1. Aaronson, J., Gilat, D., Keane, M.S., De Valk, V.: An algebraic construction of a class of one-dependent processes. Ann. Probab. (in press)
2. Finke, L.: Two maximization problems. A paper submitted to Oregon State University in partial fulfillment of the requirements for the degree of Master of Arts, 1982
3. Katz, M.: Rearrangements of (0-1) matrices. Israel J. Math. **9**, 53–72, (1971)
4. De Valk, V.: The maximal and minimal 2-correlation of a class of 1-dependent 0-1 valued processes. Israel J. Math. (in press)
5. De Valk, V.: A problem on 0-1 matrices. Compositio Mathematica (in press)

Received January 13, 1988; in revised form June 6, 1988

Compositio Mathematica **71**: 139–179, 1989.

A problem on 0–1 matrices

V. DE VALK
Delft University of Technology, Faculty of Technical Mathematics and Informatics, Julianalaan 132, 2628 BL Delft, The Netherlands.

Received 5 August 1988; accepted in revised form 2 December 1988

Key words: norms of matrices of zeroes and ones, one-dependent 0–1 valued stationary processes, two-block factors

Abstract. We compute the maximal and the minimal value of $\|M^2\|$ over the class of 0–1 valued N × N matrices M with K entries equal to one for fixed K and N, where $\|\cdot\|$ denotes the sum of the entries. This result has applications to graph theory and probability theory.

1. Introduction

1.0. A despotic problem

A country has 38 airports. Between these airports exist 639 direct flights. The despot of this country wants to get more control over the population by diminishing the interlocal traffic. Because of the public opinion in the rest of the world, he can not change the number of airports or the number of direct flights.

How should the despot distribute the 639 direct flights over the (ordered) pairs of airports, such that the number of different flights with one transit is minimized?

This problem can be solved by applying Theorem 2 of this paper. The minimal number of flights with one transit is 6239.

1.1. The matrix problem

Let $\|M\|$ denote the sum of the absolute values of the entries of a matrix M. Let $\mathscr{M}_{N,K}$ be the set of 0–1 valued $N \times N$ matrices with $\|M\| = K$.

In this paper we compute the maximal and minimal value of $\|M^2\|$ over $\mathscr{M}_{N,K}$ for fixed N and $K(0 \leqslant K \leqslant N^2)$. So we are looking for

$$\max(N, K) := \max\{\|M^2\| : M \in \mathscr{M}_{N,K}\}$$

AMS 1980 classification
primary 05B20
secondary 60G10, 28D05, 15A36, 15A45, 26D15, 28A75

and

$$\min(N, K) := \min\{\|M^2\| : M \in \mathscr{M}_{N,K}\}.$$

We give an application of this problem to graph theory and to stochastic processes.

1.2. The problem in terms of graphs

Let G be a directed graph consisting of N vertices and K edges. Solving the matrix problem is equivalent (as [F.] remarks) to solving the problem of finding for fixed N and K the maximal and minimal number of paths of length two, i.e. pairs of edges $a = (v, v')$, $b = (v', v'')$.

1.3. The problem in terms of two-correlations of stochastic processes

Let $(Y_n)_{n \in Z}$ be an i.i.d. sequence of random variables. A *two-block factor* $(X_n)_n$ of this sequence is defined by

$$X_n := f(Y_n, Y_{n+1})$$

for some function f.

The process $(X_n)_n$ has the property of *one-dependence*, i.e. for each integer time t the future $(X_n)_{n>t}$ is independent of the past $(X_n)_{n<t}$, as is easily checked. [A.G.] and [A.G.K.V.] have shown that not all one-dependent two-state processes are two-block factors (this was conjectured for several years).

We return to our matrices by restricting our attention to two-block factors of an i.i.d. sequence $(D_n)_n$, each D_n uniformly distributed over a finite set $\{1, \dots, N\}$. A matrix $M \in \mathscr{M}_{N,K}$ yields a two-block factor as follows

$$X_n := M_{D_n, D_{n+1}}.$$

Define $H_j := \Sigma_{i=1}^N M_{ij}$ and $V_j := \Sigma_{i=1}^N M_{ji} (i, j = 1, \dots, N)$. We have

$$P[X_n = 1] = K/N^2$$

and for the two-correlation $P[X_n = X_{n+1} = 1]$ we have

$$N^3 \cdot P[X_n = X_{n+1} = 1] = \sum_{i=1}^N \sum_{j=1}^N (M^2)_{i,j} = \sum_{i=1}^N \sum_{j=1}^N \left(\sum_{t=1}^N M_{it} M_{tj} \right)$$
$$= \sum_{i=1}^N H_i V_i =: I_M.$$

We conclude that the matrix problem above is equivalent to the problem of computing the maximal and minimal two-correlation, for fixed probability of a one, over the class of two-block factors of i.i.d. sequences $(D_n)_n$, where D_n is uniformly distributed over $\{1,\dots,N\}$.

Let $(Y_n)_n$ be an i.i.d. sequence, each Y_n uniformly distributed over the unit interval. Given a Lebesgue-measurable set A in the unit square we construct a two-block factor (the corresponding *indicator process*) $(X_n)_n$ by taking f equal to the indicator function of A (see [V.] for more details).

Let max(α) and min(α) be the maximal, minimal resp., two-correlation over the class of indicator processes for fixed probability α of a one. An approximation argument (approximation of the uniform distribution by discrete distributions) shows that the connection between max(α) and max(N, K), min(α) and min(N, K) resp., is

$$\max(\alpha) = \operatorname*{Sup}_{N,K} \left\{\frac{\max(N,K)}{N^3} : \alpha \geqslant K/N^2\right\}$$

and

$$\min(\alpha) = \operatorname*{Inf}_{N,K} \left\{\frac{\min(N,K)}{.N^3} : \alpha \leqslant K/N^2\right\}.$$

The discretization of the variational problems max(α) and min(α) was the motivation for this research.

We associate to a matrix $M \in \mathcal{M}_{N,K}$ a subset A_M of $[0, N] \times [0, N]$ by setting

$$A_M := \bigcup_{\{(i,j): M_{i,j}=1\}} \langle i-1, i] x \langle j-1, j].$$

We remark that the class of two-block factors of an i.i.d. sequence $(D_n)_n$ (each D_n uniformly distributed over $\{1,\dots,N\}$) is a subclass of the class of indicator processes, by taking $A = (1/N)A_M$ for the associated matrix $M \in \mathcal{M}_{N,K}$.

1.4. Previous results

For the class of two-block factors the problem of the maximal two-correlation (max(α)) was solved in [Ka.] and [F.] and the problem of the minimal two-correlation min(α)) was solved in [V.] (α denotes the fixed probability of a one). The results are

$$\max(\alpha) = \begin{cases} 2\alpha - 1 + (1-\alpha)^{3/2}, & 0 \leqslant \alpha \leqslant \frac{1}{2} \\ \alpha^{3/2}, & \frac{1}{2} \leqslant \alpha \leqslant 1 \end{cases}$$

and

$$\min(\alpha) = \begin{cases} \dfrac{m(m-1)}{6(m+1)^2}(1-2\delta)(1+\delta)^2, & 0 \leqslant \alpha < \frac{1}{2} \\ \frac{1}{6}, & \alpha = \frac{1}{2} \\ 2\alpha - 1 + \min(1-\alpha), & \frac{1}{2} < \alpha \leqslant 1 \end{cases}$$

with $m := \operatorname{int}(1/(1-2\alpha))$ and $\delta := \sqrt{1 - 2\alpha((m+1)/m)}$. (Here $\operatorname{int}(x)$ is the integer part of x).

The upper bound $\max(\alpha)$ also holds for the wider class of one-dependent processes. For $\alpha \neq \frac{1}{2}$ there is a unique one-dependent process with two-correlation $\max(\alpha)$, and for $\alpha = \frac{1}{2}$ there are exactly two such processes. These processes are all two-block factors, determined by the sets

$$A = [0, 1 - \sqrt{1-\alpha}] \times [0, 1] \cup [1 - \sqrt{1-\alpha}, 1] \times [0, 1 - \sqrt{1-\alpha}]$$
$$\text{for } 0 \leqslant \alpha \leqslant \tfrac{1}{2},$$

and

$$A = [0, \sqrt{\alpha}] \times [0, \sqrt{\alpha}] \quad \text{for } \tfrac{1}{2} \leqslant \alpha \leqslant 1.$$

For proofs see [G.K.V.].

1.5. Introductory remarks

Let l be the $N \times N$ matrix with all entries equal to one.

The following lemma shows that we may restrict our attention to the case $K \leqslant \frac{1}{2}N^2$ and that the maximum (c.q. minimum) is attained in M (for K) iff it is attained in $l - M$ (for $N^2 - K$).

We will use this observation in Theorem 2.

COMPLEMENT LEMMA. *For a matrix $M \in \mathcal{M}_{N,K}$ we have*

$$I_{l-M} = I_M - 2NK + N^3.$$

We omit the straightforward proof. (see also the Complement Lemma in [V.])

REFLECTION LEMMA. *Let $M \in \mathcal{M}_{N,K}$. Let $M', M'' \in \mathcal{M}_{N,K}$ be the matrices obtained by reflecting M with respect to the diagonal, the cross-diagonal resp., i.e., $M'_{i,j} = M_{j,i}$ and $M''_{i,j} = M_{N+1-j,N+1-i}$. Then $I_{M'} = I_{M''} = I_M$.*

We omit the straightforward proof.

2. The results

THEOREM 1 (Maximum). *Let $\mathcal{M}_{N,K}$ be the class of* 0–1 *valued $N \times N$ matrices with K entries equal to one. Then $I_M = \|M^2\|$ attains its maximal value* $\max(N, K)$ *over $\mathcal{M}_{N,K}$ in (at least) one of the types I, II, III and IV.*

THEOREM 2 (Minimum). *Let $\mathcal{M}_{N,K}$ be the class of* 0–1 *valued $N \times N$ matrices with K entries equal to one. The following table gives the possible types where $I_M = \|M^2\|$ can attain its minimal value* $\min(N, K)$ *over $\mathcal{M}_{N,K}$ for the corresponding ranges of K.*

Range of K	*Type*
(a) $0 \leqslant K \leqslant \frac{1}{4}N^2$	V
(b) $\frac{1}{4}N^2 < K < \frac{1}{2}N(N-1)$	VI, VII or VIII
(c) $\frac{1}{2}N(N-1) \leqslant K \leqslant \frac{1}{2}N(N+1)$	IX
(d) $\frac{1}{2}N(N+1) < K < \frac{3}{4}N^2$	complement of VI, VII or VIII
(e) $\frac{3}{4}N^2 \leqslant K \leqslant N^2$	complement of V

In each matrix of these types $I_M = \min(N, K)$ and for each pair (N, K) there exists a matrix of these types. In case (c) there exists a unique matrix of the described type. In cases (b) and (d) there exists exactly one or exactly two matrices of the corresponding types.

The solution to the despotic problem is found by computing the corresponding parameters of the type VI, VII and VIII. It turns out that only type VIII is suitable for the despotic problem. We shall give the solution in the Appendix.

The types of matrices where I_M attains its maximal and minimal value

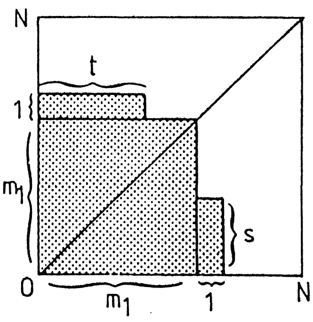


Figure I

Type I: (Maximum)

$$0 \leqslant s \leqslant t \leqslant m_1 \leqslant N,$$
$$t - s \leqslant 1,$$
$$K = m_1^2 + s + t,$$
$$I_M = m_1^3 + s(m_1 + 1) + tm_1 + st.$$

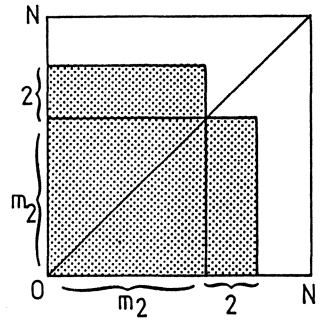


Figure II

Type II: (Maximum)

$$K = (m_2 + 2)^2 - 4,$$
$$I_M = m_2(m_2^2 + 6m_2 + 4).$$

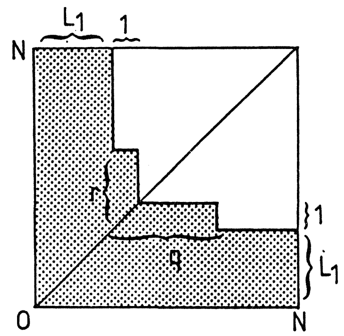


Figure III

Type III: (Maximum)

This type is the complement of type I reflected in the diagonal through $(0, N)$.

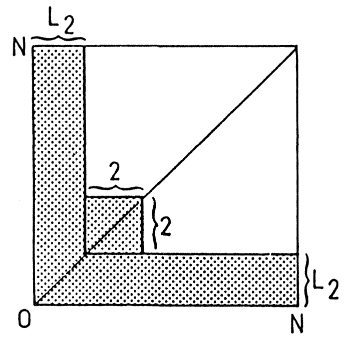


Figure IV

Type V: (Minimum)

This type is the complement of type II reflected in the diagonal through $(0, N)$.

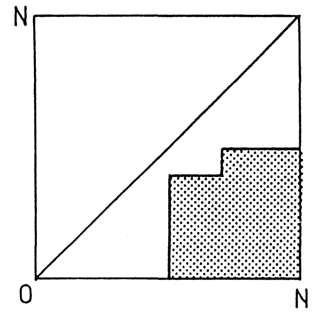


Figure V

Type V: (Minimum)

$$A_M \subset [\mathrm{int}(\tfrac{1}{2}N), N] \times [0, \mathrm{int}(\tfrac{1}{2}N)],$$
$$I_M = \min(N, K) = 0.$$
A_M is as in Figure V.

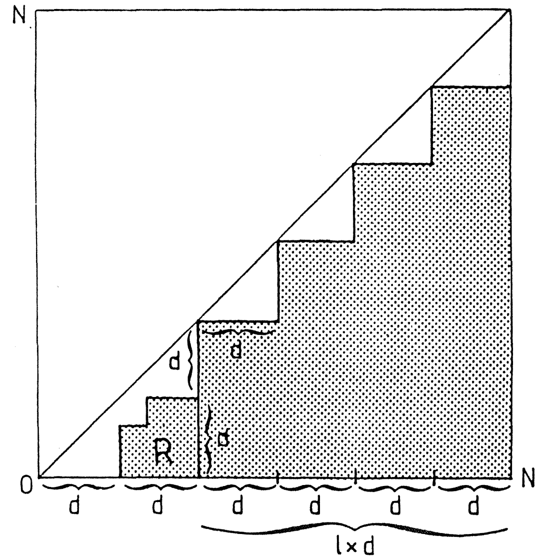


Figure VI

Type VI: (Minimum)

$\exists d, 1 \leqslant d < N$,

d divides N such that

$\frac{1}{2}N(N-d) - K \in \{0, 1\}$.

A_M is as in Figure VI.

ℓ and R are defined by $N = (\ell + 2)d$ and $N^2 - 2K = (\ell + 4)d^2 - 2R$. This implies $R \in \{d^2 - 1, d^2\}$.
There are R ones within the $d \times d$ square with corners at $(d, 0)$ and $(2d, d)$.
Further $V_{jd+i} = jd$ for $2 \leqslant j \leqslant \ell + 1$ and $1 \leqslant i \leqslant d$.

$$I_M = \min(N, K) = R\ell d + \frac{d^3}{6}\ell(\ell - 1)(\ell + 4).$$

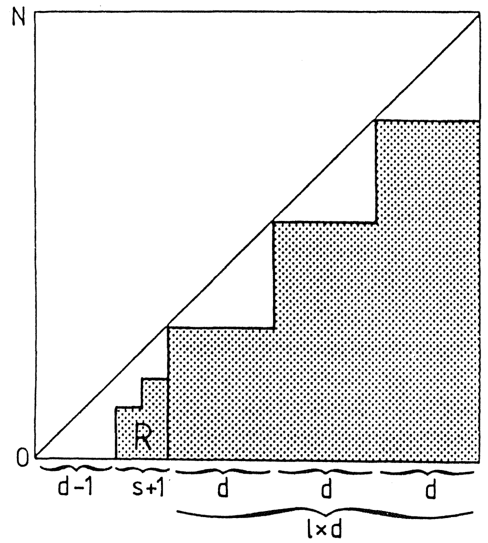


Figure VII

Type VII: (Minimum)

$\exists d, 1 \leqslant d < N$,

$\exists s, 0 \leqslant s \leqslant d - 1$,

such that

$N = (\ell + 1)d + s$

for some integer ℓ and

$\frac{1}{2}N(N-d) + \frac{1}{2}s(d-s) - K - 1 \in \{-d + s, \ldots, \min(d - s, sd - 2)\}$.

A_M is as in Figure VII.

R is defined by $N^2 - 2K = \ell d^2 + (d+s)^2 - 2R$.
This implies $1 \leqslant R$ and $(s-1)(d+1) \leqslant R \leqslant (s+1)(d-1)$.
There are R ones within the $(s+1) \times (d-1)$ rectangle with corners at $(d-1, 0)$ and $(d+s, d-1)$.
Further $V_{s+jd+i} = s + jd$ for $1 \leqslant j \leqslant \ell$ and $1 \leqslant i \leqslant d$.

$$I_M = \min(N, K) = R\ell d + \frac{d^2}{6}\ell(\ell-1)\{(\ell+1)d + 3s\}.$$

Type VIII: (Minimum)

$\exists d, 1 \leqslant d < N,$

$\exists s, 1 \leqslant s \leqslant d,$

$\exists p, q \geqslant 1,$

such that

$$N = (p+1)d + q(d+1) + s$$

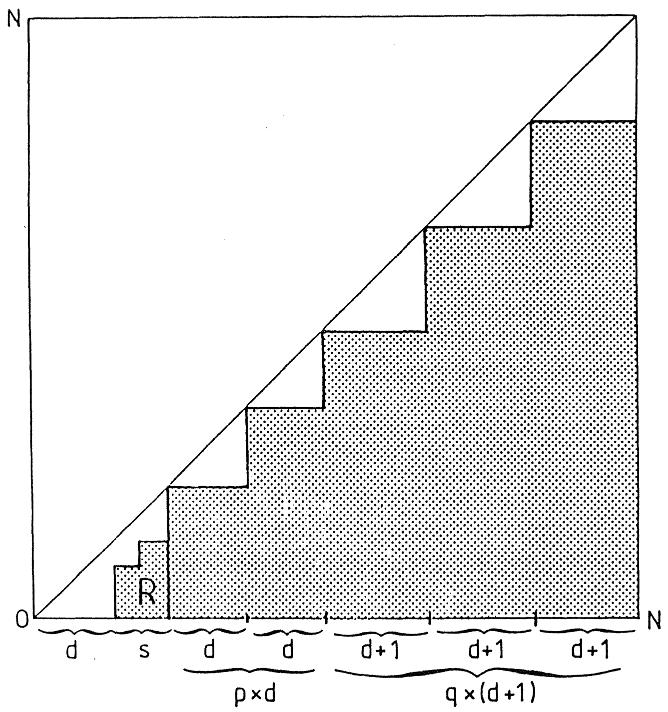


Figure VIII

and

$$\tfrac{1}{2}N(N-d)+\tfrac{1}{2}d(s-q)-\tfrac{1}{2}s^2-\tfrac{1}{2}q-K-1\in\{-1,\ldots,\min(d-s,sd-2)\}.$$

A_M is as in Figure VIII.

R is defined by $N^2-2K=q(d+1)^2+pd^2+(d+s)^2-2R$.
This implies $1\leqslant R$ and $(s-1)(d+1)\leqslant R\leqslant sd$.
There are R ones within the $s\times d$ rectangle with corners at $(d,0)$ and $(d+s,d)$.
Further $V_{s+jd+i}=s+jd$ for $1\leqslant j\leqslant p$ and $1\leqslant i\leqslant d$, and $V_{s+(p+1)d+j(d+1)+i}=s+(p+1)d+j(d+1)$ for $0\leqslant j\leqslant q-1$ and $1\leqslant i\leqslant d+1$.

$$I_M=\min(N,K)=R\{pd+q(d+1)\}+\frac{d^2}{6}p(p+1)\{(p-1)d+3q(d+1)-3s\}+$$

$$+sdp(pd+q(d+1))+\frac{(d+1)^3}{6}q(q-1)(q+1)+$$

$$+\frac{(d+1)^2}{2}q(q-1)(pd+s-1).$$

Type IX: (Minimum)

$$V_i=i\quad\text{for}\quad 1\leqslant i\leqslant K-\tfrac{1}{2}N(N-1)$$

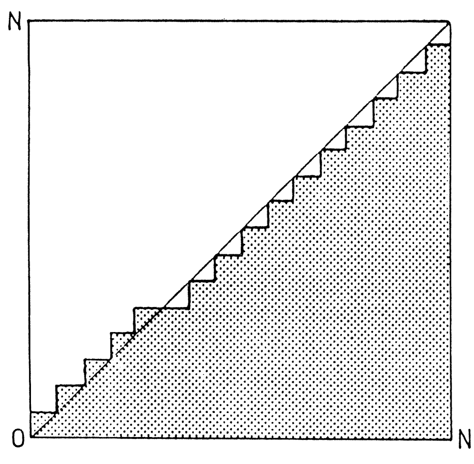


Figure IX

and

$$V_i = i - 1 \quad \text{for} \quad K - \tfrac{1}{2}N(N-1) < i \leqslant N.$$

A_M is as in Figure IX.

$$I_M = \min(N, K) = NK - \tfrac{1}{3}N(N-1)(N+1).$$

3. Proof of Theorem 1

Fix natural numbers N and K.
In several steps we will show that solutions to the problem belong to smaller and smaller subclasses of the class $\mathscr{M}_{N,K}$. To facilitate reading we refer to the appendix for technical details.

3.1. PERMUTATION LEMMA. *Let $M \in \mathscr{M}_{N,K}$, let T be a permutation of $\{1, \ldots, N\}$. Then I_M is invariant under $T \times T$.*

We omit the easy proof.

Step 1. Permutation. By taking T such that $\{H_{T_i}\}_{i=1}^N$ is a non-increasing sequence, we may assume that M is such that $(H_i)_{i=1}^N$ is non-increasing.

3.2. STANDARDIZATION LEMMA. *Let $M \in \mathscr{M}_{N,K}$ be a matrix such that $(H_i)_{i=1}^N$ is non-increasing. Then there exists a matrix $M' \in \mathscr{M}_{N,K}$ in* standard form, *i.e.,*

$$M'_{i_0,j_0} = 1 \Rightarrow \mathrm{M}'_{i,j} = 1 \quad \textit{for all } i \leqslant i_0, j \leqslant j_0,$$

such that $I_{M'} \geqslant I_M$.

Proof. Let M be a matrix, M not in standard form, such that the horizontal sections are non-increasing. Then there exist indices $i_1 < i_2, j$ such that

$$M_{i_1,j} = 0, \quad M_{i_2,j} = 1.$$

Let M' be the matrix obtained by interchanging this 0 and 1. We claim that $I_{M'} \geqslant I_M$. We have

$$\begin{aligned} I_{M'} - I_M &= H_{i_1}(V_{i_1} + 1) + H_{i_2}(V_{i_2} - 1) - H_{i_1}V_{i_1} - H_{i_2}V_{i_2} \\ &= H_{i_1} - H_{i_2} \geqslant 0. \end{aligned}$$

By repeating this argument (moving squares horizontally to the left) we obtain a matrix in standard form, while I_M does not decrease. $\square$

Step 2. Standardization. We conclude that we may assume that $M \in \mathscr{M}_{N,K}$ is in standard form.

3.3. SYMMETRIZATION LEMMA. *Let $M \in \mathscr{M}_{N,K}$ be a matrix in standard form. Then there exists a matrix $M' \in \mathscr{M}_{N,K}$ in standard form that is symmetric or nearly-symmetric, i.e.,*

$$M'_{i,j} = M'_{j,i} \quad \textit{for all } (i,j) \quad \textit{except one pair} \quad (i,j),$$

such that $I_{M'} \geqslant I_M$.

Proof. Assume that M is not of this kind. Then there exit a, b, c, d such that $M_{a,b} = M_{c,d} = 1$ and $M_{b,a} = M_{d,c} = 0$. Let M' be the matrix obtained by interchanging $M_{c,d}$ and $M_{b,a}$. We claim that $I_{M'} > I_M$. (See Appendix 1.) □

Step 3. Symmetrization. We conclude that we may assume that $M \in \mathscr{M}_{N,K}$ is in standard form and symmetric or nearly-symmetric.

With a matrix M in standard form we associate a left-continuous function $f_M\colon [0,N] \to [0,N]$ given by

$$f_M(x) = V_i \quad \text{for } x \in \langle i-1, i].$$

This implies that

$$A_M = \{(x,y) \in [0,N] \times [0,N] : y \leqslant f_M(x)\}.$$

Assume $f_M(a) \geqslant d, f_M(b) \geqslant c, f_M(c) \geqslant b, f_M(d) \geqslant a, b \leqslant c$.

Let H^w and V^w be the sections corresponding to the set

$$A^w := A_M \cap (\langle a,b] \times \langle c,d] \cup \langle c,d] \times \langle a,b]).$$

So, $H^w = H - c$ on $\langle a,b]$, $H^w = H - a$ on $\langle c,d]$ and $H^w = 0$ else, the same holds for V^w.

Let $I_{M^w} := \sum_i H_i^w V_i^w$.

3.4. WINDOWING LEMMA. *When we rearrange ones (preserving K that is the total number of ones) within $\langle a,b] \times \langle c,d] \cup \langle c,d] \times \langle a,b]$ (obtaining M') then*

$$I_{M'} - I_M = I_{M'^w} - I_{M^w}.$$

Conclusion. So, when we compute the influence of this rearrangement on I_M, we can pass over from H and V to H^w and V^w. (*Proof:* see Appendix 2.)

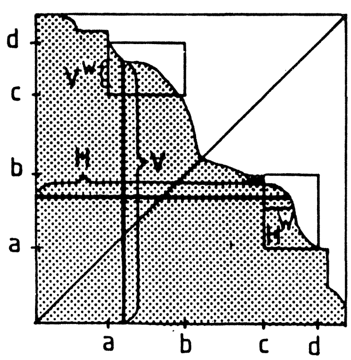


Figure X

3.5. LOCAL REFLECTION LEMMA. *Assume that:*

$$f_M(a) \geqslant d, f_M(b) \geqslant c, f_M(c) \geqslant b, f_M(d) \geqslant a, b \leqslant c$$

and also that $b - a = d - c$.

When we (obtaining M′) reflect $A_M \cap \langle a, b] \times \langle c, d]$ with respect to the line $y = x + c - a$ and $A_M \cap \langle c, d] \times \langle a, b]$ with respect to the line $y = x + a - c$, then I_M is invariant. (See fig. XI). (*Proof*: see Appendix 3).

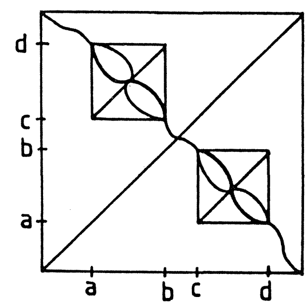


Figure XI

3.6. CONTRIBUTION OF A SQUARE LEMMA. *Let* $M \in \mathcal{M}_{N,K}$ *be in standard form and symmetric or nearly-symmetric. Let* $(a, b)(a, b \in \{1, \ldots, N\})$ *be a* corner point *of* M, *i.e.,* $f_M(a) = b$ *and* $f_M(a+1) < b$ *or* $a = N$. *Let* M' *be the matrix obtained from* M *by removing* $(a, b)(M'_{i,j} = M_{i,j} - \delta_{a,i} \cdot \delta_{b,j})$. *Then*

$$I_M - I_{M'} = \begin{cases} a + b & \text{if } a \neq b, \quad M_{b,a} = 1 \\ a + b - 2 & \text{if } a \neq b, \quad M_{b,a} = 0 \\ a + a - 1 & \text{if } a = b. \end{cases}$$

Proof. See Appendix 4. □

From now on all rearrangements of ones in M will be done such that M remains in standard form and (nearly-)symmetric. This means that a rearrangement of ones within $\langle a, b] \times \langle c, d]$ (above the diagonal) is attended with a (in some sense reflected rearrangement within $\langle c, d] \times \langle a, b]$ (under the diagonal).

This will not lead to confusion.

Spreading out

We will consider quasi-blocks and we will decrease the number of these quasi-blocks and so we will diminish the class of matrices.
Let

$$f_M = \sum_{i=1}^{m} y_i \cdot 1_{\langle x_{i-1}, x_i]}$$

be the function associated with M as defined in step 3 ($0 = x_0 < x_1 < \cdots < x_m = N$). We call a rectangle $\langle x_{k-1}, x_k] \times \langle y_{k+1}, y_k]$ a *block* if it is disjoint with the diagonal. Note that the points (x_k, y_k) are corner points.

We call a set $\langle x_{k-1}, x_k] \times \langle y_{k+2}, y_k] \cup \langle x_k, x_{k+1}] \times \langle y_{k+2}, y_{k+1}]$ (disjoint with the diagonal) a *quasi-block* if $y_k - y_{k+1} = 1$ or $x_{k+1} - x_k = 1$. We call in these cases $x_k - x_{k-1}$ c.q. $y_{k+1} - y_{k+2}$ the *remainder* of the quasi-block.

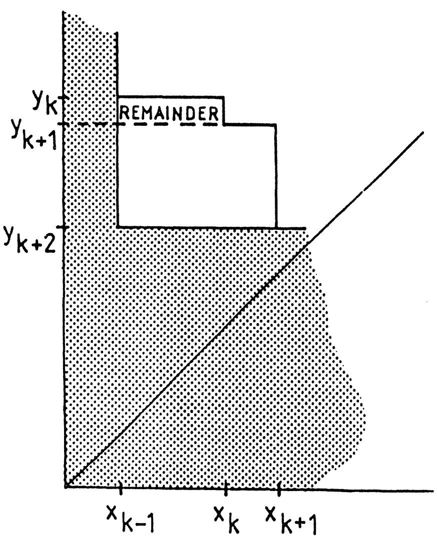


Figure XII. A quasi-block with $y_k - y_{k+1} = 1$.

We consider blocks as special quasi-blocks (with remainder equal to zero).
We shall spread out a quasi-block along the longest segment ($\langle x_{K-1}, x_{K+2}]$ or $\langle y_{K+2}, y_{K-1}]$), using the Local Reflection Lemma and the Contribution of a Square Lemma.

3.7. SPREADING-OUT LEMMA. *Let* $f_M = \Sigma_{i=1}^{m} y_i \cdot 1_{\langle x_{i-1} x_i]}$, *assume that*

$$\langle x_{K-1}, x_K] \times \langle y_{K+2}, y_K] \cup \langle x_K, x_{K+1}] \times \langle y_{K+2}, y_{K+1}]$$

is a quasi-block. Assume $y_{K+2} \geqslant x_{K+2}$. *Then this quasi-block can be replaced by a quasi-block of the type*

$$\langle x_{K-1}, t] \times \langle y_{K+2}, r+1] \cup \langle t, x_{K+2}] \times \langle y_{K+2}, r]$$

or by a quasi-block of the type

$$\langle x_{K-1}, r] \times \langle y_{K+2}, y_{K-1}] \cup \langle r, r+1,] \times \langle y_{K+2}, t]$$

such that I_M *does not decrease.*

Proof. See Appendix 5. □

3.8. TWO QUASI-BLOCKS LEMMA. *Let* $f_M = \Sigma_{i=1}^{m} y_i \cdot 1_{\langle x_{i-1}, x_i]}$, *assume that*

$$\langle x_{K-1}, x_K] \times \langle y_{K+2}, y_K] \cup \langle x_K, x_{K+1}] \times \langle y_{K+2}, y_{K+1}]$$

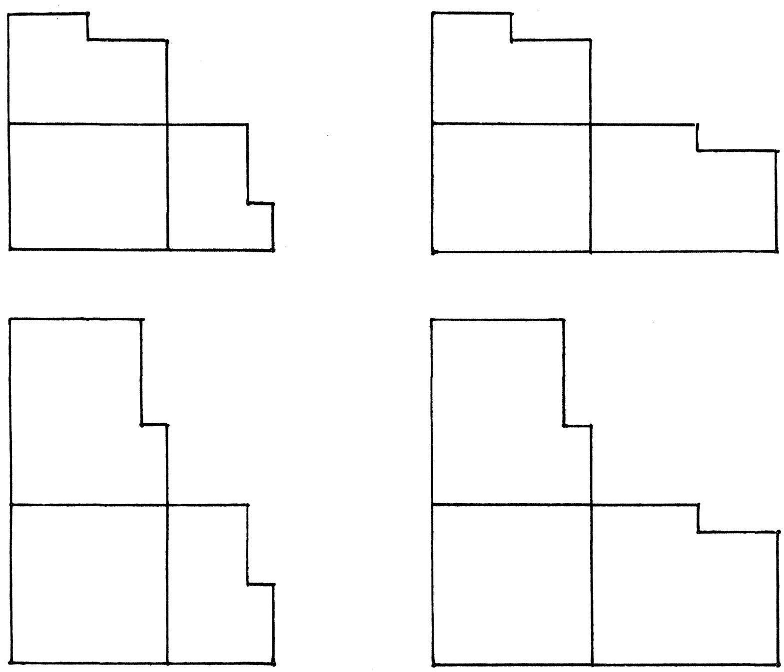


Figure XII *a–d* (4 cases)

and

$$\langle x_{K+1}, x_{K+2}] \times \langle y_{K+4}, y_{K+2}] \cup \langle x_{K+2}, x_{K+3}] \times \langle y_{K+4}, y_{K+3}]$$

are quasi-blocks. Assume $y_{K+4} \geqslant x_{K+4}$. Then these two quasi-blocks can be joined to one quasi-block, preserving standard form and (near-)symmetry, such that I_M does not decrease.

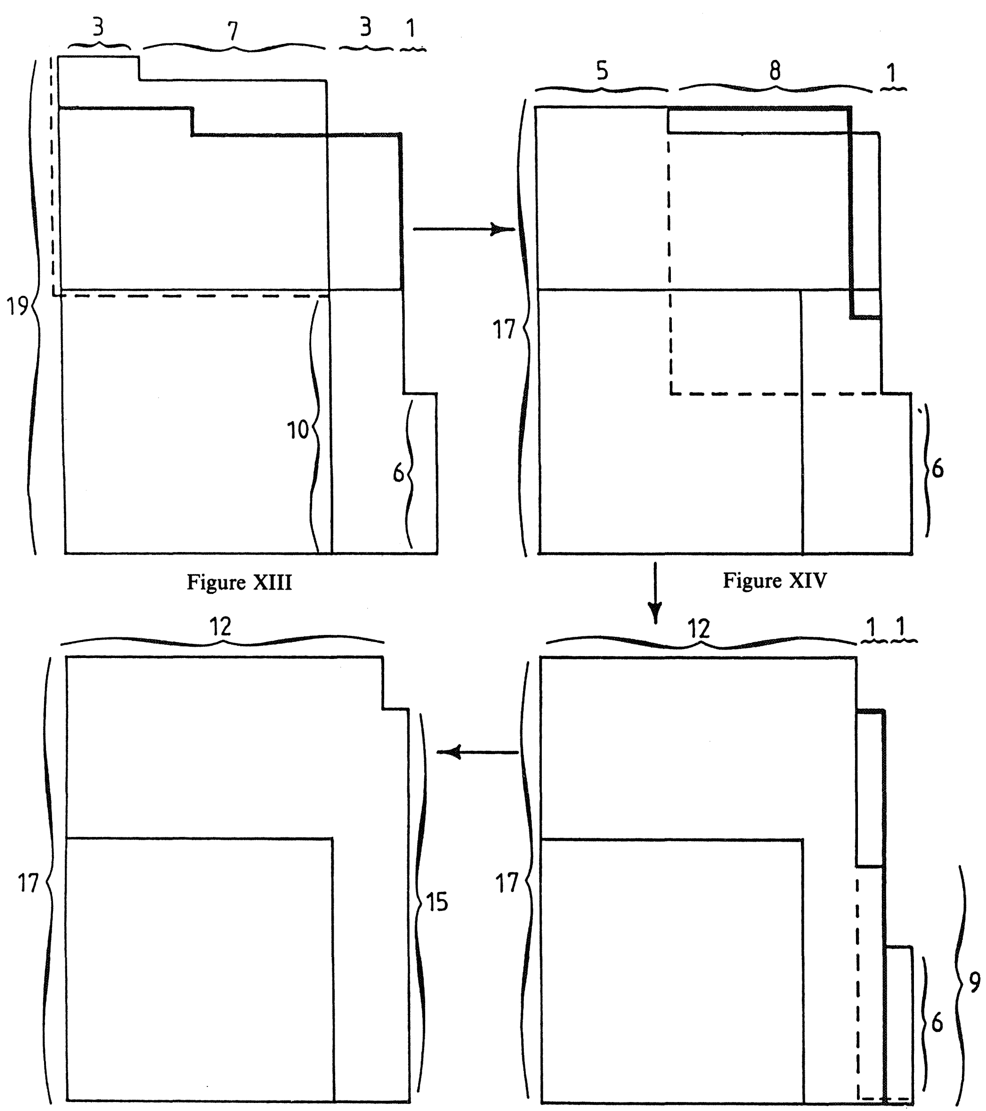


Figure XIII

Figure XIV

Figure XVI (EXAMPLE) Figure XV

We use the Spreading-out Lemma and the Local Reflection Lemma. There are 4 cases (see Fig. XII), depending whether the two remainders are horizontal or vertical strips.

To avoid a long and detailed list of cases and subcases, we restrict ourselves to the case of a quasi-block with horizontal remainder with at its right lower side a quasi-block with vertical remainder.

The given example is typical for this case. Just as in Appendix 5 we may assume that the remainder of a quasi-block is a horizontal strip if $x_{K+1} - x_{K-1} \geqslant y_K - y_{K+2}$ and a vertical strip if $x_{K+1} - x_{K-1} < y_K - y_{K+2}$.

First we spread out horizontally the left upper quasi-block (Fig. XIII). Then, by a reflection, we obtain one quasi-block consisting of one strip and a remainder (Fig. XIV). We spread out this quasi-block and we are finished (Fig. XV).

We spread out the quasi-blocks marked with $\llcorner$ – –. (thin lines before the transformation, thick lines after the transformation) (Computation: see Appendix 6).

COROLLARY. *Let $M \in \mathscr{M}_{N,K}$ be a matrix in standard form and (nearly-)symmetric. Then there exists a matrix $M' \in \mathscr{M}_{N,K}$ of type A or B such that $I_{M'} \geqslant I_M$.*

Proof. Apply Lemma 3.8 iteratively. □

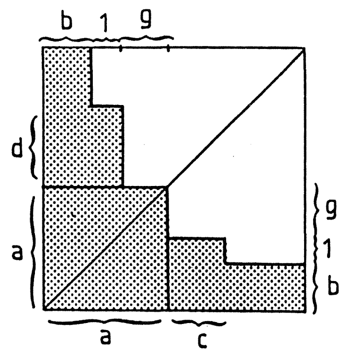


Figure XVIIC

Type A:

$|c - d| \leqslant 1.$

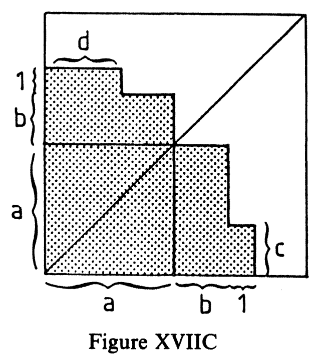


Figure XVIIC

Type B:

$|c - d| \leqslant 1.$

3.9. *Last Step.* To complete the Proof of Theorem 1 we will reduce this class of matrices to the types I, II, III, IV. (see Appendix 7). □

REMARK. If $K = m^2 > \frac{1}{2}N^2$ for some integer m, then the maximal value of I_M is attained when we take A_M equal to $m \times m$ square of ones, and if $K = N^2 - m^2 < \frac{1}{2}N^2$ for some integer m, then we obtain the maximal value of I_M by taking the complement of a $m \times m$ square. This directly follows from the fact that in these cases I_M assumes the value $N^3 \cdot \max(\alpha)$ (with $\alpha = K/N^2$).

In other cases I_M is strictly less than $N^3 \cdot \max(\alpha)$.

Generally, if $\alpha = K/N^2 > \frac{1}{2}$ the maximal value of I_M is attained in type I or II, and if $\alpha < \frac{1}{2}$ in type III or IV, because in these types $(1/N)A_M$ is an approximation of the corresponding $\sqrt{\alpha} \times \sqrt{\alpha}$ square (the solution of the continuous version for $\alpha > \frac{1}{2}$) c.q. the complement of a $\sqrt{1-\alpha} \times \sqrt{1-\alpha}$ square (the solution of the continuous version for $\alpha < \frac{1}{2}$). However, for $\alpha \approx \frac{1}{2}$ this can be different, as the following example shows. (See also the table at the end of this paper, before the appendix.)

EXAMPLE. Take $N = 10$ and $K = 49$, then $\alpha = 0.49 < \frac{1}{2}$. The maximal value of I_M is attained in type I (see Figure 1) where $I_M = 7 \times 7 \times 7 = 343$, and not in type III (see Figure 2) where $I_M = 339$.

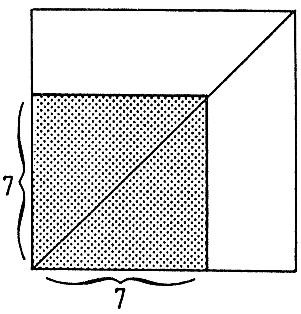


Figure 1

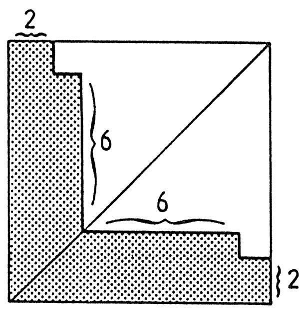


Figure 2

EXAMPLE. We show the existence of three sequences $(N_1)_{i=1}^{\infty}, (K_i)_{i=1}^{\infty}, (v_i)_{i=1}^{\infty}$ (each tending to infinity) such that

(1) $K_i = \frac{1}{2}N_i^2 - v_i$ and
(2) I_M attains its maximal value $\max(N_i, K_i)$ in type I and not in type III or IV.

From the theory of continued fractions follows the existence of increasing integer sequences $(p_i)_{i=1}^{\infty}, (q_i)_{i=1}^{\infty}$ such that

$$0 < \frac{1}{\sqrt{2}} - \frac{p_i}{q_i} < \frac{1}{q_i^2}, \qquad \text{and all } q_i \text{ are odd.}$$

This implies

(3) $0 < \frac{1}{2}q_1^2 - p_1^2 < \sqrt{2}$.

Now (3) implies $\frac{1}{2}q_i^2 - p_i^2 = \frac{1}{2}$.

We define

$$N_i := q_i + 3 \quad \text{and} \quad K_i := (p_i + 2)^2.$$

Defining v_i by

$$K_i = \tfrac{1}{2}N_i^2 - v_i$$

we have

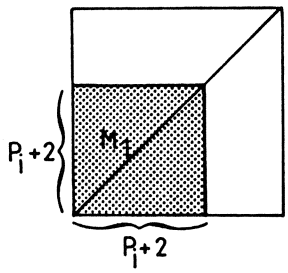
$$v_i = 3q_i - 4p_i + 1$$

So we have

$$(3\sqrt{2} - 4)p_i < v_i < (3\sqrt{2} - 4)p_i + 6.$$

So, for p_i sufficiently large we have

$$v_i < \tfrac{1}{2}p_i.$$

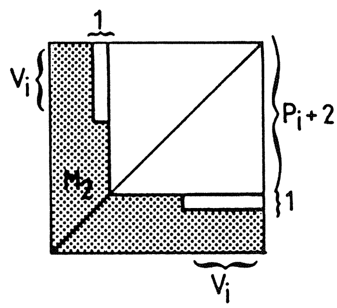


Taking M_1 of type I ($a(p_i + 2) \times (p_i + 2)$ square of ones) we have

$$I_{M_1} = (p_i + 2)^3.$$

Taking M_2 of type III (see figure) we have (by the Complement Lemma)

$$\begin{aligned} I_{M_2} &= v_i(p_i + 3)^2 + (p_i + 2 - v_i) \\ &\quad \times (p_i + 2)^2 + v_i^2 + 2(q_i + 3) \\ &\quad \times (p_i + 2)^2 - (q_i + 3)^3 \\ &= (p_i + 2)^3 - v_i(2p_i - q_i) < I_{M_1}. \end{aligned}$$



This proves the statement of the example.

4. Proof of Theorem 2

Fix natural numbers N and K. The case $K \leqslant \frac{1}{4}N^2$ is trivial because $I_M = 0$. In several steps we will show that solutions to the problem exist in smaller and smaller subclasses of $\mathscr{M}_{N,K}$. After the third step we will discriminate the cases $\frac{1}{4}N^2 < K < \frac{1}{2}N(N-1)$ and $\frac{1}{2}N(N-1) \leqslant K \leqslant \frac{1}{2}N^2$. To facilitate reading we refer to the Appendix for technical details.

4.1. STEP 1. PERMUTATION. Using Lemma 3.1 (Permutation Lemma) we may assume that (H_i) is a non-increasing sequence.

We define a new standard form; M is in *standard form* when

$$M_{i,j} = 1 \quad \text{if} \quad M_{i_0,j_0} = 1 \quad \text{for} \quad i \geqslant i_0, j \leqslant j_0.$$

4.2. STANDARDIZATION LEMMA. *Let $M \in \mathscr{M}_{N,K}$ be a matrix such that $(H_i)_{i=1}^N$ is a non-increasing sequence. Then there exists a matrix $M' \in \mathscr{M}_{N,K}$ in standard form such that $I_{M'} \leqslant I_M$.*

Proof. Analogous to the proof of Lemma 3.2 (Standardization Lemma). □

Step 2. Standardization. We conclude that we may assume that $M \in \mathscr{M}_{N,K}$ is in standard form.

We associate with the matrix M in standard form a right-continuous function $f_M: [0, N] \to [0, N]$ given by

$$f_M(x) = V_i \quad \text{for} \quad x \in [i-1, i).$$

We redefine A_M equal to

$$A_M := \bigcup_{\{(i,j): M_{i,j} = 1\}} [i-1, i\rangle \times \langle j-1, j].$$

This implies that $A_M = \{(x, y) \in [0, N] \times [0, N]: y \leqslant f_M(x)\}$. Except the right-continuity, f_M is the same as in the proof of Theorem 1. We call (a, b) a *corner point of M if $f_M(a-1) = b$* and $f_M(a-2) < b$.

4.3. UNDER THE DIAGONAL LEMMA. *Let $M \in \mathscr{M}_{N,K}$ be a matrix in standard form. If $\frac{1}{4}N^2 < K < \frac{1}{2}N(N-1)$, then there exists a matrix $M' \in \mathscr{M}_{N,K}$ such that $I_{M'} \leqslant I_M$ and M' lies under the diagonal, i.e.,*

$$M'_{i,j} = 0 \quad \text{if} \quad j \geqslant i.$$

If $\frac{1}{2}N(N-1) \leqslant K \leqslant \frac{1}{2}N^2$, then there exists a matrix $M' \in \mathscr{M}_{N,K}$ such that

$I_{M'} \leqslant I_M$ and

$$M'_{i,j} = 1 \quad \text{if} \quad j < i.$$

Proof. Assume that M has some corner point (i,j) not lying unde the diagonal ($j \geqslant i$), and assume that there exists a point $(i_1 j_1)$ under the diagonal ($i_1 \geqslant j_1 + 1$) such that $M_{i_1,j_1} = 0$ and $f_M(i_1) \geqslant j_1 + 1$.

We move the one from (i,j) to (i_1,j_1) and we consider two cases. In both cases I_M will decrease (see Appendix 8).

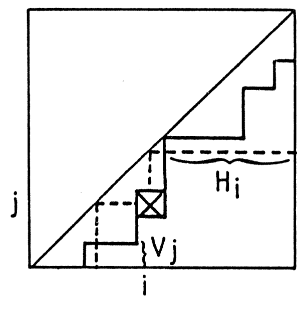
Step 3. *Under the diagonal.* We conclude that we may assume that in the case $\frac{1}{4}N^2 < K < \frac{1}{2}N(N-1)$, M lies under the diagonal and that in the case $\frac{1}{2}N(N-1) \leqslant K \leqslant \frac{1}{2}N^2, M_{i,j} = 1$ if $j < i$.

Now we consider the CASE $\frac{1}{4}N^2 < K < \frac{1}{2}N(N-1)$.

4.4. REMARK: Changing of I_M by a corner point. We consider the influence on I_M of removing a square from a corner point (i,j) of M to obtain a matrix M'.

We have

$$I_M - I_{M'} = H_i V_i + H_j V_j - H_i(V_i - 1) - (H_j - 1)V_j = H_i + V_j.$$



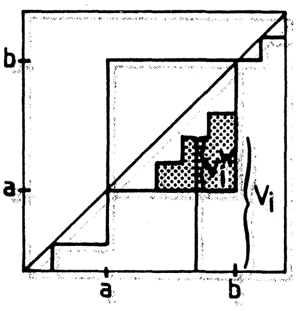
We say that the corner point (i,j) changes I_M by the sections H_i and V_j.

4.5. WINDOWING LEMMA. *Assume $f_M(a) = a, f_M(b) = b$ for some $0 \leqslant a < b \leqslant N$.*

Let H^w and V^w be the sections of $A_M \cap [a, b] \times [a, b]$, and let

$$I_{M^w} := \sum_i H_i^w V_i^w.$$

When we rearrange ones (preserving K that is the total number of ones) within



$[a, b] \times [a, b]$ *(obtaining M') we have*

$$I_M - I_{M'} = I_{M^w} - I_{M'^w}$$

The proof is analogous to the proof of Lemma 3.4. So, when we rearrange within $[a, b] \times [a, b]$, we can compute the influence on I_M by passing over from H_i and V_i to H_i^w and V_i^w.

4.6. LOCAL REFLECTION LEMMA. *Assume again $f_M(a) = a, f_M(b) = b$ $(a < b)$. Then I_M is invariant under reflecting $A_M \cap [a, b] \times [a, b]$ with respect to the line $y = -x + a + b$.*

We leave the straightforward proof to the reader (use the Windowing Lemma).

We say that a corner (i, j) lies *strictly under the diagonal* resp. *on the diagonal* if

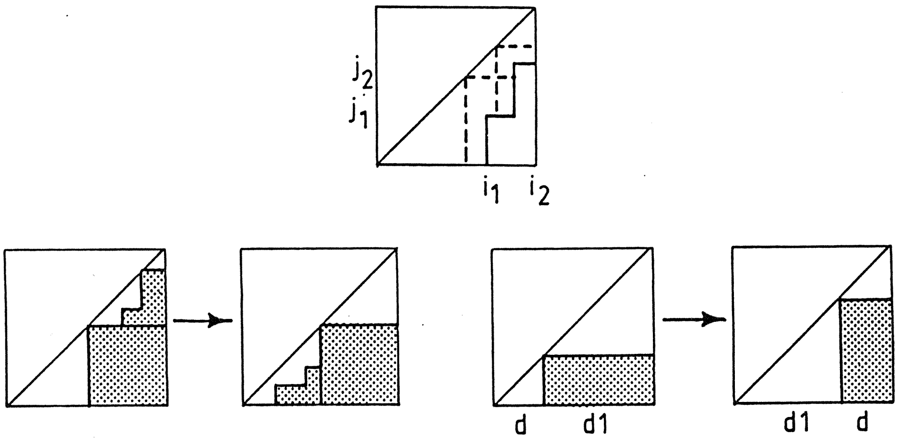
$$i \geqslant j + 2 \text{ resp. } i = j + 1.$$

4.7. MOVING TO THE DIAGONAL LEMMA. *Let $M \in \mathcal{M}_{N,K}$ (for $\frac{1}{4}N^2 < K < \frac{1}{2}N(N-1)$) be a matrix in standard form, lying under the diagonal. Then there exists a matrix $M' \in \mathcal{M}_{N,K}$ in standard form, lying under the diagonal, such that $I_{M'} \leqslant I_M$, and such that M' has at most two corners (i_1, j_1) and (i_2, j_2) $(i_1 < i_2$ and $j_2 < i_1)$ strictly under the diagonal.*

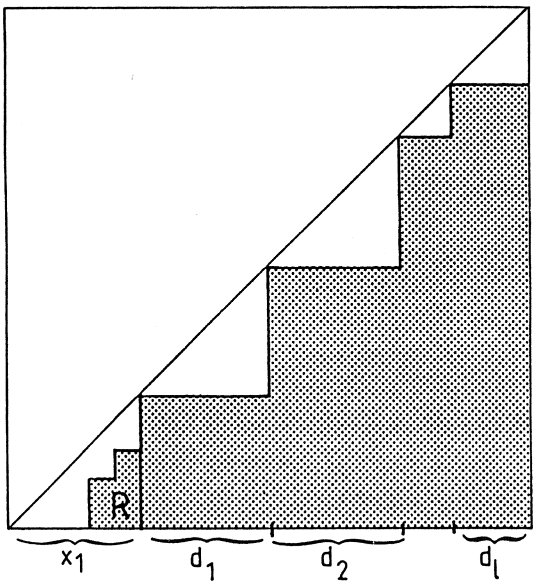
Proof. See Appendix 9. □

Note that $j_2 < i_1$ means that the corner point (i_1, j_1) changes I_M by horizontal sections at a higher level than j_2 and that (i_2, j_2) changes I_M by vertical sections lying more to the left than i_1 (see picture below).

Step 4. Moving to the diagonal. We conclude that we may assume that in the case $\frac{1}{4}N^2 < K < \frac{1}{2}N(N-1)$, M has at most two corner points $(i_1, j_1), (i_2, j_2)$ $(i_1 < i_2$ and $j_2 < i_1)$ lying strictly under the diagonal.



Using some local reflections (shown above), that leave I_M invariant, we can assume that M is the following type:



$0 < x_1 < \cdots < x_{\ell+1} = N, d_i = x_{i+1} - x_i.$

$V_j = x_i \quad \text{if} \quad x_i < j \leqslant x_{i+1}$

and there are R ones $(1 \leqslant R \leqslant \frac{1}{4}x_1^2)$ within some rectangle which is a subset of

$\{(i,j): 1 \leqslant i \leqslant x_1, j < i\}$.

$$N = x_1 + \sum_{i=1}^{\ell} d_i,$$

$$N^2 - 2K = \sum_{i=1}^{\ell} d_i^2 + x_1^2 - 2R.$$

We call R the *remainder*.

Note that we can interchange the d_i's by the Local Reflection Lemma.

4.8. LESS INEQUALITY BETWEEN DIFFERENCES: With various rearrangements we will prove that we may assume that

(8a) $x_1 \geqslant d_i \quad \forall i$
(8b) $x_1 \leqslant 2d_i \quad \forall i,$
(8c) $R \geqslant (s-1)(d+1) \quad (d := \min\{d_i : i \geqslant 1\}, \qquad s := x_1 - d),$
(8d) $d \leqslant d_i \leqslant d+1 \quad \forall i,$
(8e) $R \leqslant sd \quad$ if $\quad d_i = d+1 \quad$ for some i,
(8f) $R \leqslant (s+1)(d-1) \quad$ or $\quad R = d^2$.

See Appendix 10.

We have now reached the class of matrices of the types VI, VII, VIII.

We will prove that in each of these types I_M attains its minimal value $\min(N, K)$, and that for each pair (N, K) there exist at most two matrices of these types.

Our method is a lexicographical ordering $\ll$ on the class of matrices of the types VI, VII, VIII. We will prove that if $M_1 \ll M_2$ then $K_1 < K_2$ or $K_1 = K_2$ and $I_{M_1} = I_{M_2}$. Further we prove that if $M_1 \ll M_2 \ll M_3$ then $K_1 < K_3$. These facts imply the theorem for the case $\frac{1}{4}N^2 < K < \frac{1}{2}N(N-1)$ (See Appendix 11).

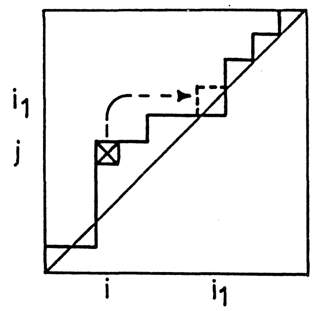
CASE $\frac{1}{2}N(N-1) \leqslant K \leqslant \frac{1}{2}N^2$.

Assume that M is not of type IX. Then, by Step 3, we can move a one from a corner point (i,j) above the diagonal $(j > i)$ to a place (i_1, i_1) at the diagonal. We obtain a matrix M'.

This transformation yields

$$\begin{aligned} I_M - I_{M'} &= H_i V_i + H_j V_j + H_{i_1} V_{i_1} - H_i(V_i - 1) - (H_j - 1)V_j - \\ &- (H_{i_1} + 1)(V_{i_1} + 1) = H_i + V_j - H_{i_1} - V_{i_1} - 1 \geqslant (N - i + 1) + \\ &+ j - (i_1 - 1) - (N - i_1) - 1 = j - i + 1 \geqslant 2. \end{aligned}$$

These last considerations prove Theorem 2. $\square$



REMARK. The two-correlation in the discrete case attains the infimum of the continuous case (min(α)) only in the following cases:

- case $\alpha \leqslant \frac{1}{4}$ (type V),
- case $\frac{1}{4} < \alpha < \frac{1}{2} \cdot \frac{N-1}{N}$ and $R = d^2$ (type VI),
- case $\frac{1}{4} < \alpha < \frac{1}{2} \cdot \frac{N-1}{N}$ and $R = ds$ (type VII),

and of course (by the Complement Lemma) in the complements of these configurations. In the other cases the (discrete) two-correlation will be strictly greater than min(α).

EXAMPLE

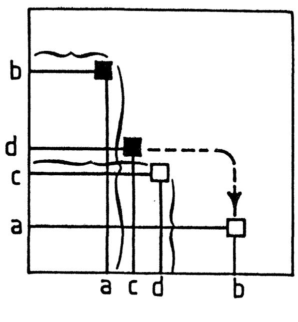
We give in a table the solutions of the minimality and maximality problem for $N = 10$ and $26 \leqslant K \leqslant 55$.

MINIMUM						*MAXIMUM*	
K	Type	d	p	R	min(10, K)	Type	max(10, K)
26	VII	5	1	1	5	III	142
27	VII	5	1	2	10	III	148
28	VII	5	1	3	15	III	156
29	VII	5	1	4	20	III	163
29	VII	4	1	5	20		
30	VII	4	1	6	24	III	172
31	VII	4	1	7	28	III	180
32	VII	4	1	8	32	III	190
33	VII	4	1	9	36	III	199
34	VII	3	2	1	42	III	210
35	VII	3	2	2	48	III	220

MINIMUM						*MAXIMUM*	
K	Type	d	p	R	min(10, K)	Type	max(10, K)
36	VII	3	2	3	54	III	232
37	VII	3	2	4	60	III	237
38	VIII	2	2	1	67	III	242
39	VIII	2	2	2	74	III	249
39	VI	2	3	3	74		
40	VI	2	3	4	80	IV	256
41	VII	2	4	1	88	III	263
42	VIII	1	2	1	96	III	270
43	VIII	1	4	1	104	III	279
44	VIII	1	6	1	112	III	287
45	IX				120	III	297
46	IX				130	III	306
46						I	306
47	IX				140	III	317
47						I	317
48	IX				150	I	330
49	IX				160	I	343
50	IX				170	III	350
50						I	350
51	IX				180	III	363
52	IX				190	III	370
53	IX				200	III	377
53						I	377
54	IX				210	III	386
54						I	386
55	IX				220	I	397

Appendix 1. (3.3. Symmetrization Lemma, Theorem 1)

We consider two cases:

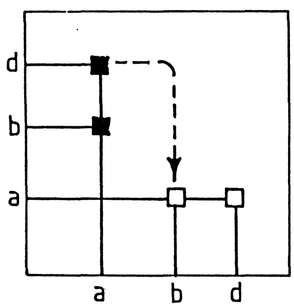


Case 1. a, b, c, d are all different. By permuting a, b, c, d it is no restriction to assume that

$$V_a + H_b \geqslant V_c + H_d.$$

We have:

$$\begin{aligned} I_{M'} - I_M &= (H_a + 1)V_a + H_b(V_b + 1) + H_c(V_c - 1) + \\ &\quad + (H_d - 1)V_d - H_aV_a - H_bV_b - H_cV_c - H_dV_d \\ &= V_a + H_b - H_c - V_d \geqslant V_a + H_b - (V_c - 1) - \\ &\quad - (H_d - 1) \geqslant 2. \end{aligned}$$



Case 2. $a = c$ (case $b = d$ analogously). We have:

$$\begin{aligned} I_{M'} - I_M &= (H_a + 1)(V_a - 1) + H_b(V_b + 1) + \\ &\quad + (H_d - 1)V_d - H_aV_a - H_bV_b - H_dV_d \\ &= V_a - H_a - 1 + H_b - V_d \geqslant \\ &\geqslant d - (b - 1) - 1 + a - (a - 1) \geqslant 2. \end{aligned}$$

Clearly this symmetrization can be done such that standard form is preserved.

Appendix 2. (3.4. Windowing Lemma, Theorem 1)

Proof.

$$\begin{aligned} I_{M'} - I_M &= \sum_{\substack{i=a+1,\dots,b \\ i=c+1,\dots,d}} (H_i'V_i' - H_iV_i) \\ &= \sum_{i=a+1,\dots,b} \{(H_i'^w + c)(V_i'^w + c) - (H_i^w + c)(V_i^w + c)\} + \\ &\quad + \sum_{i=c+1,\dots,d} \{(H_i'^w + a)(V_i'^w + a) - (H_i^w + a)(V_i^w + a)\} \\ &= \sum_{\substack{i=a+1,\dots,b \\ i=c+1,\dots,d}} (H_i'^w V_i'^w - H_i^w V_i^w) + c \cdot \sum_{i=a+1,\dots,b} (H_i'^w + V_i'^w - H_i^w - V_i^w) + \\ &\quad + a \cdot \sum_{i=c+1,\dots,d} (H_i'^w + V_i'^w - H_i^w - V_i^w) = I_{M'^w} - I_{M^w}. \end{aligned}$$

The last equality holds because the rearrangement preserves the number of ones ($=K$).

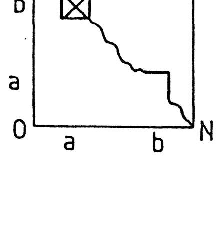
Appendix 3. (3.5. Local Reflection Lemma, Theorem 1).

Using the windowing principle we have

$$\begin{aligned} I_{M'} - I_M &= \sum_{i=a+1,\dots,b} (H_i'^w V_i'^w - H_i^w V_i^w) + \sum_{i=c+1,\dots,d} (H_i'^w V_i'^w - H_i^w V_i^w) = \\ &= \sum_{i=c+1,\dots,d} V_i^w H_i^w - \sum_{i=a+1,\dots,b} H_i^w V_i^w + \sum_{i=a+1,\dots,b} V_i^w H_i^w - \sum_{i=c+1,\dots,d} H_i^w V_i^w = 0. \end{aligned}$$

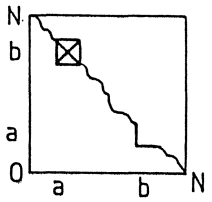
Appendix 4. (3.6. Contribution of a Square Lemma, Theorem 1)

We consider three cases in the (nearly-)symmetric situation:



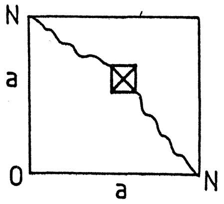
Case 1. $a \neq b, M_{b,a} = 1$.

$$\begin{aligned} I_M - I_{M'} &= b^2 + a^2 - \{b(b - 1) + a(a - 1)\} \\ &= a + b. \end{aligned}$$



Case 2. $a \neq b, M_{b,a} = 0$.

$$I_M - I_{M'} = b(b-1) + a(a-1) - \{(b-1)^2 + (a-1)^2\}$$
$$= a + b - 2.$$



Case 3. $a = b$.

$$I_M - I_{M'} = a^2 - (a-1)^2 = a + a - 1.$$

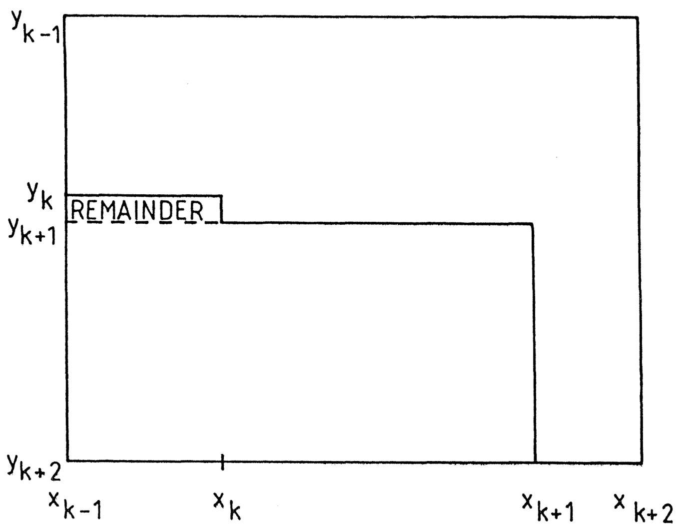
Note that removing both (a, b) and (b, a) leads to a decreasing of I_M by $2(a + b - 1)$. So, the average decreasing of I_M per square is the sum of the coordinates minus 1, just as in case 3.

Appendix 5. (3.7. Spreading-out Lemma, Theorem 1)

First we rearrange M such that the quasi-block lays with its longest side along the longest segment, i.e. if $x_{k+2} - x_{k-1} > y_{k-1} - y_{k+2}$ and $y_k - y_{k+2} > x_{k+1} - x_{k-1}$ or if $x_{k+2} - x_{k-1} < y_{k-1} - y_{k+2}$ and $y_k - y_{k+2} < x_{k+1} - x_{k-1}$, then we reflect the quasi-block with respect to the line $y = x - x_{k-1} + y_{k+2}$.

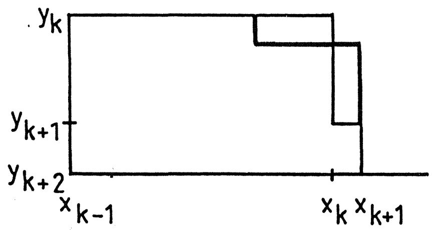
By the Local Reflection Lemma I_M is then invariant.

We consider the case $x_{k+2} - x_{k-1} \geqslant y_{k-1} - y_{k+2}$ (the other case goes analogously). We spread out the quasi-block from $\langle x_{k-1}, x_{k+1}]$ over $\langle x_{k-1}, x_{k+1} + 1]$.



If $x_{k+1} - x_k = 1$, we add the ones from $\langle x_{k+1} - y_k + y_{k+1}, x_k] \times \langle y_k - 1, y_k]$ to the remainder

$\langle x_k, x_{k+1}] \times \langle y_{k+1}, y_k - 1]$. So we obtain $y_k - y_{k+1} = 1$, while I_M is invariant by the Local Reflection Lemma.

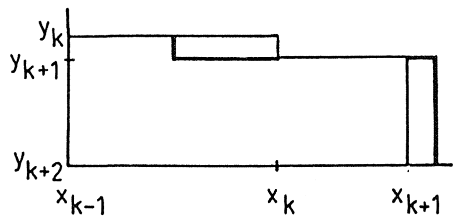


So we may assume that

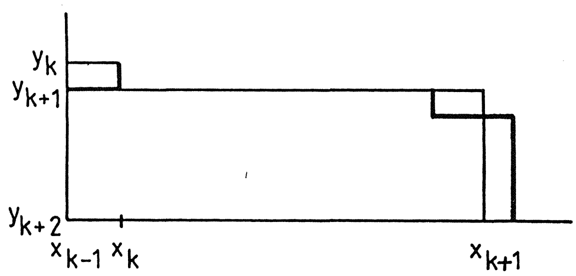
$$y_k - y_{k+1} = 1.$$

Now we have two cases:

Case 1. If the remainder $x_k - x_{k-1}$ is larger than (or equal to) $y_{k+1} - y_{k+2}$, then we spread out by moving $\langle x_k - y_{k+1} + y_{k+2}, x_k] \times \langle y_{k+1}, y_k]$ to $\langle x_{k+1}, x_{k+1} + 1] \times \langle y_{k+2}, y_{k+1}]$.



Case 2. If $x_k - x_{k-1} < y_{k+1} - y_{k+2}$, then we move $\langle x_{k-1}, x_k] \times \langle y_{k+1}, y_k]$ and $\langle x_{k+1} - y_{k+1} + 1 + y_{k+2} + x_k - x_{k-1}, x_{k+1}] \times \langle y_{k+1} - 1, y_{k+1}]$ to $\langle x_{k+1}, x_{k+1} + 1] \times \langle y_{k+2}, y_{k+1} - 1]$.



Considering the contributions of the various squares it is easy to see that in both cases I_M does not decrease.

Iterating this procedure we obtain a quasi-block of the form $\langle x_{k-1}, t] \times \langle y_{k+2}, r+1] \cup \langle t, x_{k+2}] \times \langle y_{k+2}, r]$ (for some r and t). In the case $x_{k+2} - x_{k-1} < y_{k+1} - y_{k+2}$ we spread out the quasi-block vertically from $\langle y_{k+2}, y_k]$ over $\langle y_{k+2}, y_{k-1}]$.

Appendix 6. (3.8. Two Quasi-Blocks Lemma, Theorem 1)

In the symmetric case I_{M^w} is in these four cases

$$3\cdot 19^2+7\cdot 18^2+3\cdot 10^2+1\cdot 6^2+6\cdot 14^2+4\cdot 13^2+8\cdot 10^2+3^2=6348,$$
$$5\cdot 17^2+8\cdot 16^2+1\cdot 6^2+6\cdot 14^2+10\cdot 13^2+1\cdot 5^2=6420,$$
$$12\cdot 17^2+1\cdot 9^2+1\cdot 6^2+6\cdot 14^2+3\cdot 13^2+8\cdot 12^2=6420, \quad \text{and}$$
$$12\cdot 17^2+1\cdot 15^2+15\cdot 13^2+2\cdot 12^2=6516.$$

So first I_M increases by 72, then I_M is constant, and finally I_M increases by 96.

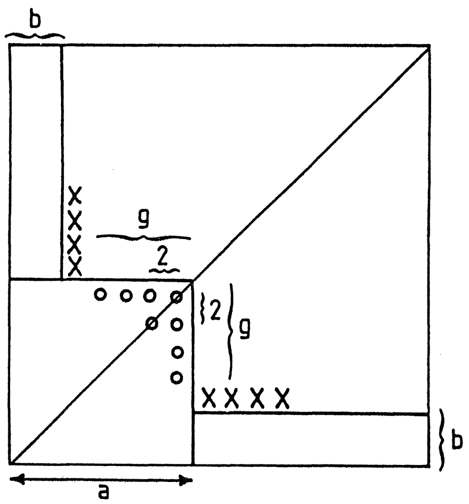
Appendix 7. (3.9. Last Step, Theorem 1)

Assume that M is of type A or B, but not of type I, II, III, IV.
We consider several cases and subcases.

Type A

Case 1. $a \leqslant N/2$ and $d = c = 0$.
We move $2g$ ones from $\langle a-g,a] \times \langle a-1,a] \cup \langle a-2,a] \times \langle a-2,a-1] \cup \langle a-1,a] \times \langle a-g,a-2]$ to $\langle b,b+1] \times \langle a,a+g] \cup \langle a,a+g] \times \langle b,b+1]$.

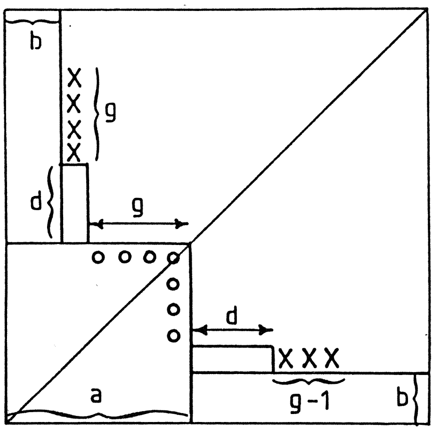
CASE 1.



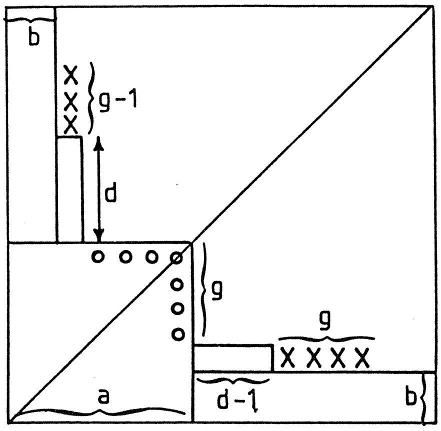
With the principle of the contribution of a square it is easy to see that I_M increases.

Case 2. $a \leqslant N/2$ and $(d > 0$ or $c > 0)$.
We move ones from the ath row and the ath column to the $b+1$th row and the $b+1$th column and (when there is no place enough in the $b+1$th row and the $b+1$th column) also to the $b+2$th row and the $b+2$th column.

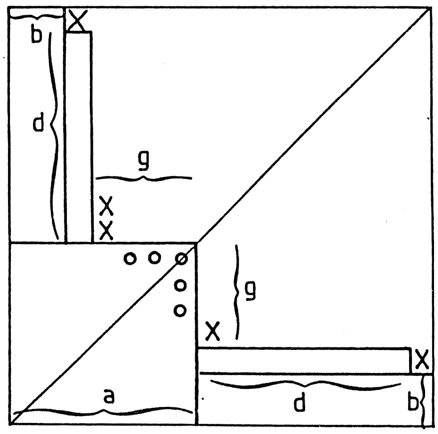
In detail we have 4 subcases (whether or not the matrix is symmetric or nearly-symmetric and whether the $b+2$th row and column are needed).



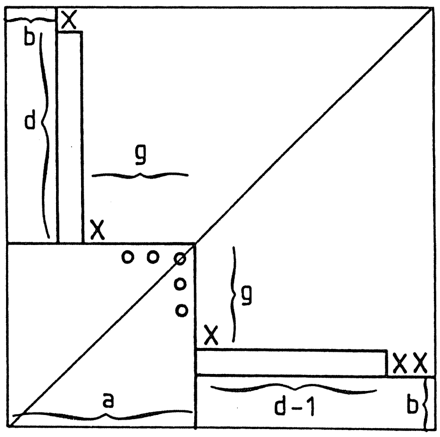
Subcase 2–1. M is symmetric and
$a + d + g \leqslant N$.



Subcase 2–2. M is nearly-symmetric and
$a + d + g - 1 \leqslant N$.



Subcase 2–3. M is symmetric and
$a + d + g > N$.

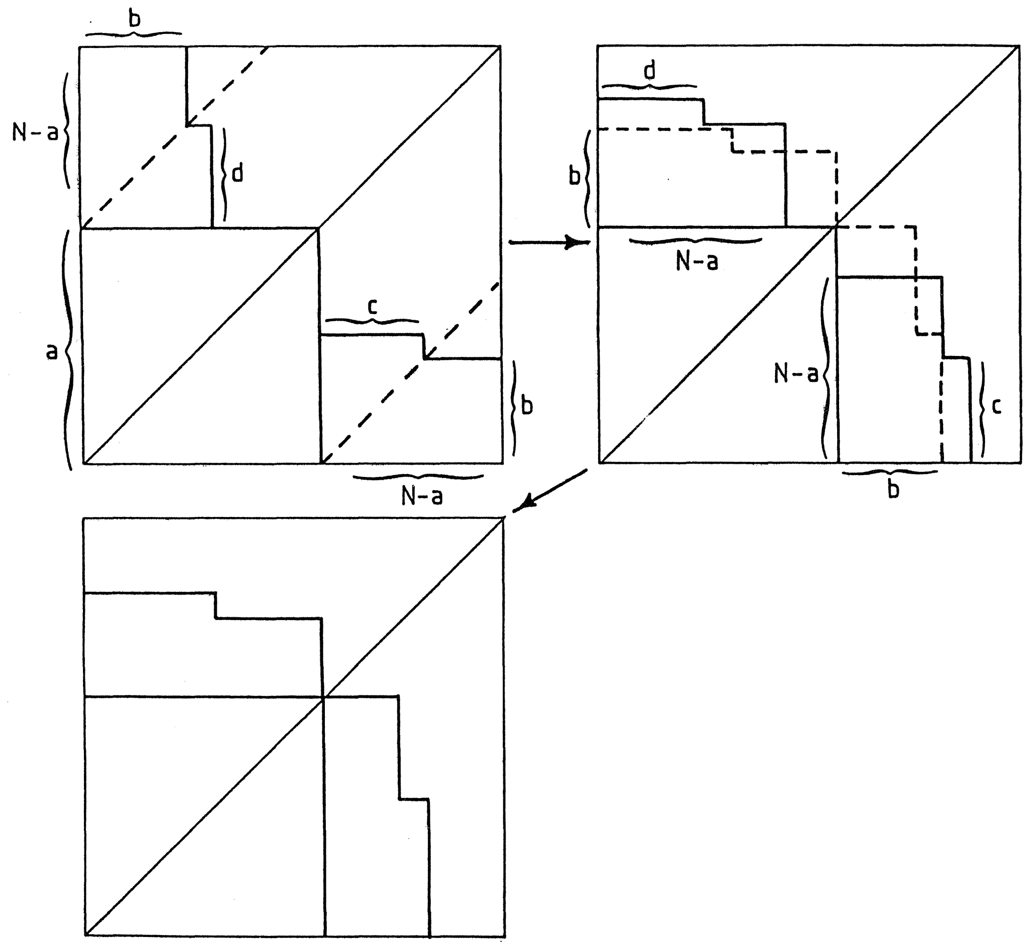


Subcase 2–4. M is nearly-symmetric and
$a + d + g - 1 > N$.

The squares marked with ○ are moved to squares marked with ×.

From the principle of the contribution of a square follows that I_M increases under these transformations.

Case 3. $a > N/2$ and $N - a \geqslant b + 1$.



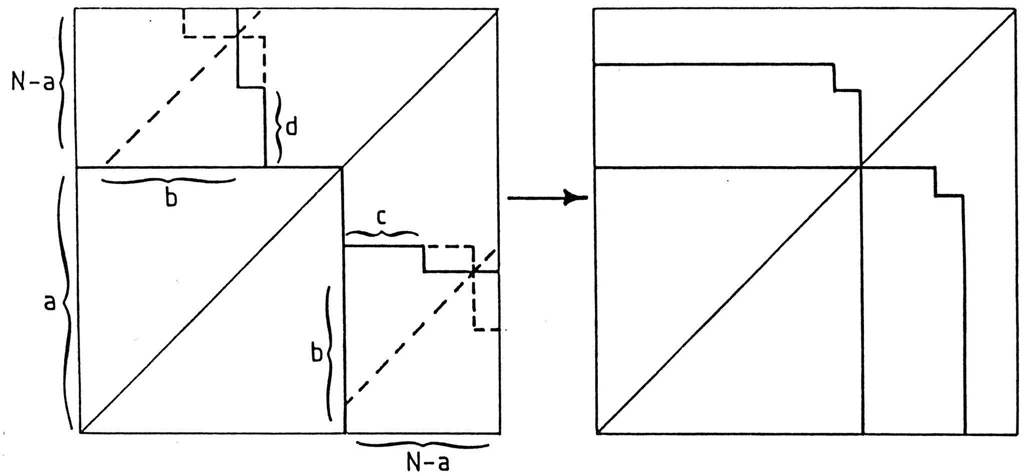
First we reflect the quasi-blocks with size $(N-a)b+d$, $(N-a)b+c$ resp. with respect to the lines $y=x+a$, $y=x-a$ resp. Then we spread out the quasi-blocks horizontally over $[0,a]$, vertically over $[0,a]$ resp.

Now we consider the complement of the obtained set A_M. This complement is of the type as in case 1 or 2. By the Complement Lemma I_{l-M} increases when I_M does. So, with the methods of cases 1 and 2 (applied to $l-M$) we can transform M to another matrix and in this way we prove that the maximum was not attained in M.

Case 4. $a > N/2$ and $N-a < b+1$.

First we reflect two quasi-blocks of size $(N-a)(N-a-1)+d$, $(N-a)(N-a-1)+c$ resp. with respect to the lines $y=x-(b+1)+N, y=x-N+b+1$ resp., then we spread out the two quasi-blocks of size $(N-a-1)(b+1)+b+1-(N-a-d)$, $(N-a-1)(b+1)+b+1-(N-a-c)$ resp., horizontally over $[0,a]$, vertically over $[0,a]$ resp.

Now we consider its complement and by an argument as in case 3 we are finished.



Type B

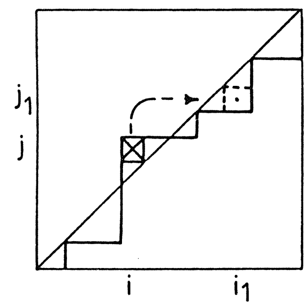
The complement of type B is of type A.

Again by the Complement Lemma we consider its complement and we deal with it as described above.

Now we have reached the matrices of the types I, II, III, IV and so we have proved Theorem 1. □

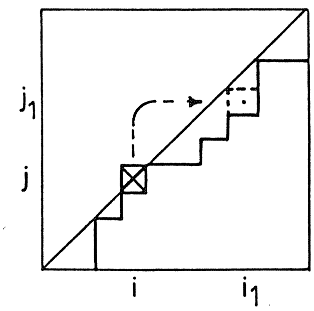
Appendix 8. (4.3. Under the diagonal Lemma, Theorem 2)

Case 1. $j > i$.



$$\begin{aligned}
I_M - I_{M'} &= H_iV_i + H_jV_j + H_{i_1}V_{i_1} + H_{j_1}V_{j_1} \\
&\quad - H_i(V_i - 1) - (H_j - 1)V_j - H_{i_1}(V_{i_1} + 1) \\
&\quad - (H_{j_1} + 1)V_{j_1} \\
&= H_i + V_j - H_{i_1} - V_{j_1} \geqslant (N - i + 1) + \\
&\quad + j - (N - i_1) - (j_1 - 1) \\
&= j - i + i_1 - j_1 + 2 \geqslant 3.
\end{aligned}$$

Case 2. $j = i$.



$$\begin{aligned}
I_M - I_{M'} &= H_iV_i + H_{i_1}V_{i_1} + H_{j_1}V_{j_1} \\
&\quad - (H_i - 1)(V_i - 1) - H_{i_1}(V_{i_1} + 1) \\
&\quad - (H_{j_1} + 1)V_{j_1} \\
&= H_i + V_j - 1 - H_{i_1} - V_{j_1} \geqslant \\
&\quad \geqslant (N - i + 1) + i - 1 - (N - i_1) - (j_1 - 1) \\
&= i_1 - j_1 + 1 \geqslant 2.
\end{aligned}$$

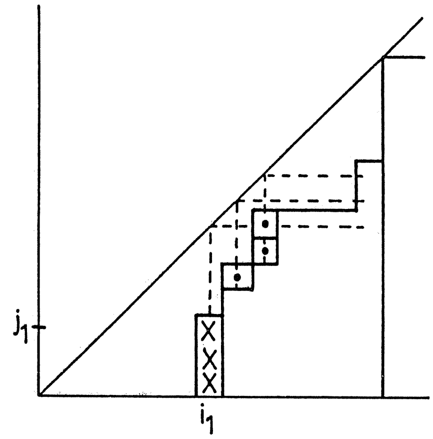
The conclusion follows directly.

Appendix 9. (4.7. Moving to the diagonal Lemma, Theorem 2).

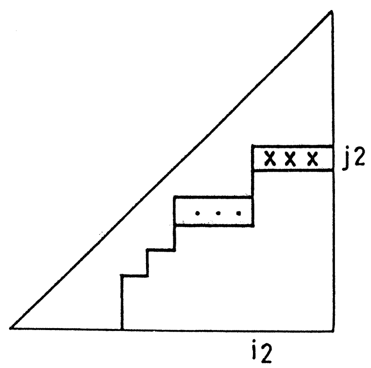
Assume that M is not of this kind. Take the first corner strictly under the diagonal (i_1, j_1) and the last one (i_2, j_2) (so i_1 minimal and i_2 maximal).

We first prove that we may assume that if (i_1, j_1) is a corner strictly under the diagonal then $(i_1 + 1, i_1 - 1)$ or $(i_1 + 1, i_1)$ is a corner.

If this is not the case, then we can move ones from the i_1th column to the $j_1 + 1$th row (or, when $M_{i_1+1, j_1+1} = 1$ to a row at higher level) and so on, until the second corner (i, j) strictly under the diagonal (with $i > i_1$ minimal) has the desired form. When the i_1th column is exhausted, we continue with moving ones from the $i_1 + 1$th column etc. We make the crucial observation that after the moving of ones to columns to the right these ones changes I_M by horizontal sections on a higher level; thus with smaller sections. So I_M does not increase.



Analogously we can assume that if (i_2, j_2) is the last corner strictly under the diagonal (with i_2 maximal) then $(j_2 + 1, j_2 - 1)$ or $(j_2, j_2 - 1)$ is a corner point.



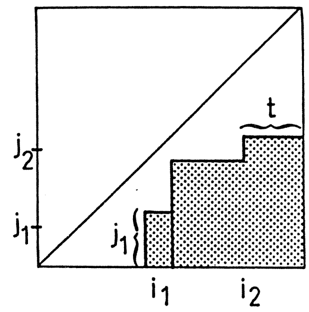
We now consider two cases:

Case 1. $i_1 = j_2$.

We use the Windowing Lemma. There are j_1 ones in the i_1th column. Assume there are t ones in the j_2th row (windowed).

Moving a one from (i_1, j_1) to $(i_2 - 1, j_2)$ yields

$$I_M - I_{M'} = j_1 t - (j_1 - 1)(t + 1) = t - j + 1$$



and moving a one from (i_2, j_2) to $(i_1, j_1 + 1)$ yields

$$I_M - I_{M'} = j_1 t - (j_1 + 1)(t - 1) = j - t + 1.$$

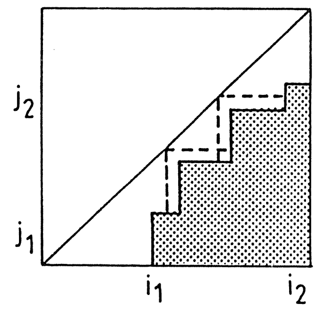
At least one of these transformations decreases I_M, so the minimum was not attained.

Case 2. $i_1 < j_2$.
Assume there are t ones in the j_2th row. Moving a one from (i_1, j_1) to $(i_2 - 1, j_2)$ now yields

$$I_M - I_{M'} = H^w_{i_1} - V^w_{j_2},$$

and moving a one from (i_2, j_2) to $(i_1, j_1 + 1)$ yields

$$I_M - I_{M'} = V^w_{j_2} - H^w_{i_1}.$$



If $H^w_{i_1} \neq V^w_{j_2}$, then clearly the minimum was not attained.

If $H^w_{i_1} = V^w_{j_2}$, then we can move ones from the i_1th column to the j_2th row (while I_M is invariant) until the column is exhausted or the row is full (i.e. the diagonal is reached); in both cases we have one corner less lying strictly under the diagonal.

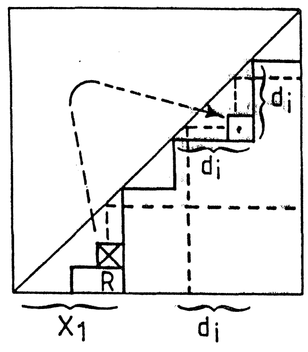
We conclude that we may assume that M has at most 2 corners lying strictly under the diagonal, and that in this case the first corner changes I_M by horizontal sections at a higher level than the second corner.

Appendix 10. (4.8. Less inequality between differences, Theorem 2).

8a. $x_1 \geqslant d_i \quad \forall i.$

First we prove that we may assume that $x_1 \geqslant d_i$ for all i.

Assume that $x_1 < d_i$ for some i, then we can move a one from R to $(x_{i+1}, x_i + 1)$ while I_M decreases (consider the changing of I_M) by $(N - x_1) - (N - d_i) = d_i - x_1 > 0$, so the minimum was not attained.



8b. $x_1 \leqslant 2d_i \quad \forall i.$

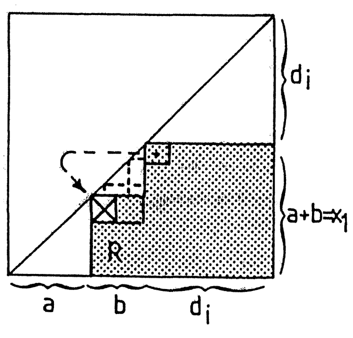
We now give an upper bound for x_1.

Assume that the R ones lie in a $a \times b$ rectangle

$\{(i, j): a + 1 \leqslant i \leqslant a + b, 1 \leqslant j \leqslant a\}.$

Because of local reflection it is no restriction to assume that $a \geqslant b$.

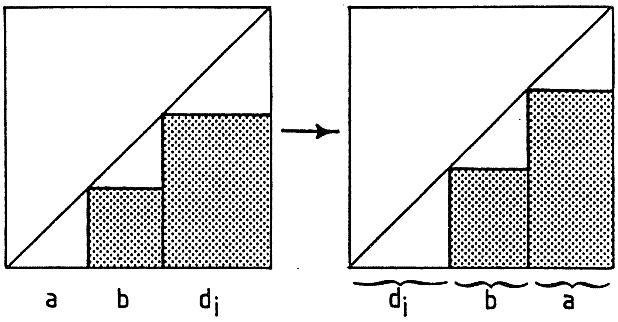
After a rearrangement we have the situation as in the picture.



First assume $R < a \cdot b$ and $b \geqslant 2$ (the case $b = 1$ is left as an exercise to the reader, use 8c). Then it is possible to move a one from $(x_1 + 1, x_1)$ to $(a + 1, a)$. This gives a decreasing of I_M by $a - d_i$ (consider the changing of I_M).

So we can assume that $d_i \geqslant a \geqslant b$, which implies $x_1 = a + b \leqslant 2d_i$.

When $R = a \cdot b$ we can reach the situation $x_1 \leqslant 2d_i$ by local reflecting

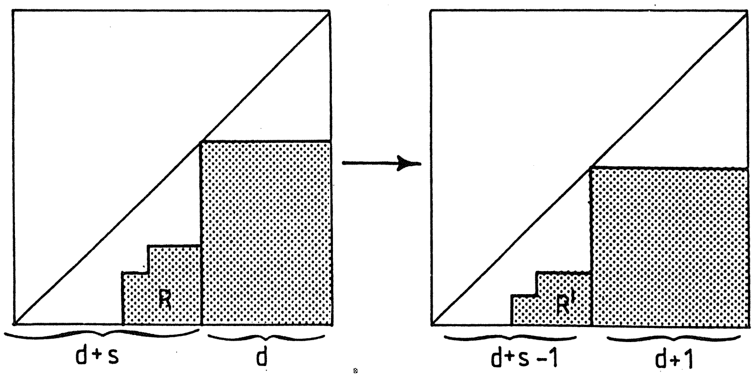


8c. $R \geqslant (s-1)(d+1)$.

In this substep we give a lower bound for R. Because

$$s - 1 + \frac{(d+s-1)^2}{4} - \frac{(d-s+1)^2}{4} = (s-1)(d+1)$$

we may assume that $R - s + 1 < (d+s-1)^2/4$ (otherwise the statement is trivially true).



After local reflections we obtain the situation as in the first above picture and by the Windowing Lemma we restrict our attention to the $R + (d+s)d$ ones in that picture. Because

$$R + (d+s)d = (R - s + 1) + (d + s - 1)(d + 1)$$

we can transform the matrix and obtain the second picture with $R' = R - s + 1$. This is possible because $R - s + 1 < (d + s - 1)^2/4$.

Note that $R_1 > 0$, else $I_{M'}$ is trivially smaller than I_M. We have

$$\begin{aligned} I_{M'} - I_M &= (R - s + 1)(d + 1) - Rd \\ &= R - (s-1)(d+1) \end{aligned}$$

and the statement follows.

8d. $d \leqslant d_i \leqslant d + 1 \quad \forall i$.

We will prove that the d_i's cannot differ more than 1 from each other. Assume that $d_i = d + v$ for some i and some $v \geqslant 2$.

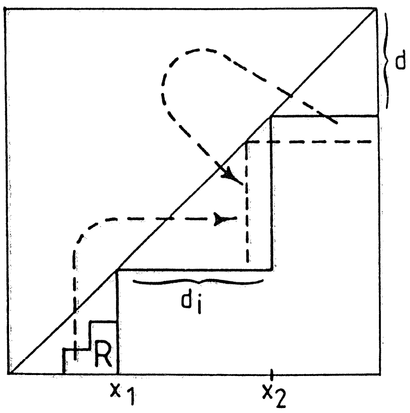
This implies $s = x_1 - d \geqslant d_i - d = v$ and $R \geqslant (s-1)(d+1) \geqslant v - 1$.

After the usual local reflections and rearrangements and windowing we obtain the next situation (see figure below).

We decrease d_i by one, and we add that to d, in practice this means the following. We move d ones from the x_2th row and $v - 1$ ones from R to the x_2th column and we obtain a matrix with $R' = R - v + 1$. We have

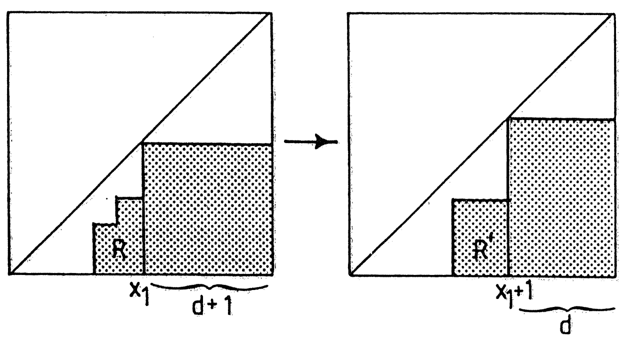
$$\begin{aligned} I_M - I_{M'} &= R(d + d_i) + x_1 d_i d - (R - v + 1)(d + d_i) - x_1(d_i - 1)(d + 1) \\ &= (v - 1)(d + d_i - x_1) > 0. \end{aligned}$$

So the minmum was not attained, and the statement is proved.



8e. $R \leqslant sd$ *if* $d_i = d + 1$ *for some* i.

This time we transform as follows:



From

$$R + x_1(d + 1) = R' + (x_1 + 1)d$$

follows

$$R' = R + s.$$

The second configuration is possible because

$$\frac{(x_1+1)^2}{4} = \frac{x_1^2}{4} + \frac{x_1}{2} + \frac{1}{4} > R + s = R'.$$

We have

$$I_M - I_{M'} = R(d+1) - (R+s)d = R - sd$$

and the minimality of I_M implies

$$R \leqslant sd.$$

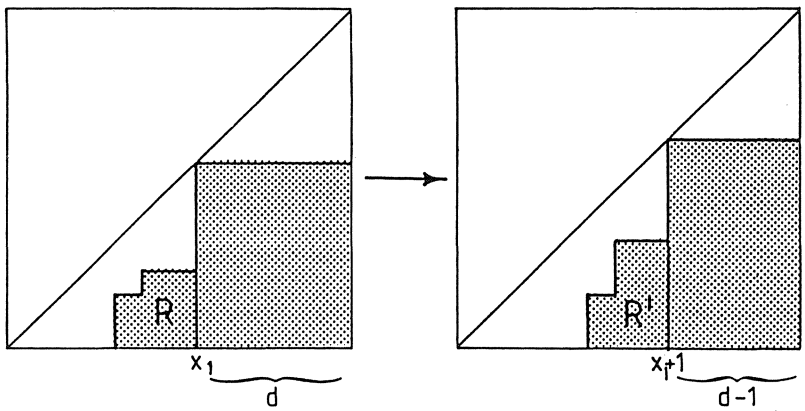
8f. $R \leqslant (s+1)(d-1)$ *or* $R = d^2$.

To prove this we consider some cases:

Case 1. $s = d$. Because $R \leqslant \frac{1}{4}x_1^2 = d^2$ and $(s+1)(d-1) = d^2 - 1$ in this case, the statement is clearly true.

Case 2. $s = d - 1$. Because $R \leqslant (x_1+1)/2 \cdot (x_1-1)/2 = (s+1)(d-1)$ in this case it is trivially true.

Case 3. $s \leqslant d - 2$.



This last time we transform the first configuration with $R + x_1 d$ ones to the second with $R' + (x_1+1)(d-1)$ ones ($R' = R + s + 1$). This last configuration is possible because

$$\frac{(x_1+1)^2}{4} = \frac{x_1^2}{4} + \frac{x_1}{2} + \frac{1}{4} \geqslant R + \frac{2s+2}{2} + \frac{1}{4} > R + s + 1.$$

We have

$$I_{M'} - I_M = (R+s+1)(d-1) - Rd = (s+1)(d-1) - R$$

and the statement follows from the minimality of M.

Appendix 11. (4.9, 4.10, 4.11, Theorem 2)

4.9. Representation by a triple

We represent a $N \times N$ matrix M of types VI, VII and VIII by a triple

$$(d, p, R).$$

If M is of type VI or VII we define $p := \ell$. We extend the parameter s to type VI, where we define $s := d$ and we extend the parameter q to types VI and VII, where we define $q := 0$.

We prove that (for fixed N) there corresponds at most one matrix M of the types VI, VII, VIII to a triple (d, p, R).

LEMMA. *Let N, d, p, R be integers. Then there exists at most one $N \times N$ matrix M of types VI, VII, VIII with the triple (d, p, R).*

Proof. We have

$$s + q(d + 1) = N - (p + 1)d$$

with $0 \leqslant s \leqslant d$. This implies

$$q = \operatorname{int}\left(\frac{N - (p + 1)d}{d + 1}\right)$$

and

$$s = N - (p + 1)d - q(d + 1).$$

Further, K follows now from

$$N^2 - 2K = pd^2 + q(d + 1)^2 + (d + s)^2 - 2R. \qquad \square$$

REMARK. The solution of the problem of the despot is represented by the triple (4, 4, 11). The other parameters are $N = 38$, $K = 639$, $q = 3$, $s = 3$.

Let $\mathscr{M}^*_{N,K} \subset \mathscr{M}_{N,K}$ be the subclass of matrices M with $M_{i,i} = 0$ for all i. It is more realistic to consider this problem over $\mathscr{M}^*_{N,K}$ instead of $\mathscr{M}_{N,K}$.

Theorem 2 shows that this makes no difference for these values of the parameters.

4.10. Ordering on the triples

Let M_1, M_2 be $N \times N$ matrices of types VI, VII or VIII with triples (d_1, p_1, R_1) and (d_2, p_2, R_2). We write

$$M_1 \ll M_2$$

if ($d_1 = d_2$ and $p_1 = p_2$ and $R_1 < R_2$) or if ($d_1 = d_2$ and $p_1 < p_2$) or if $d_1 > d_2$.

We call M_2 the *successor* of M_1 if

$$M_1 \ll M_2$$

and if there exists no matrix M_3 such that

$$M_1 \ll M_3 \ll M_2.$$

4.11. The successor

LEMMA. *Let M be a $N \times N$ matrix of types VI, VII or VIII with K ones and with triple (d, p, R). The successor M_2 of M is (distinguishing 16 cases) listed below. Let K_2 be the number of ones of the successor.*

No.	*Type M*	*Description case*	*Triple successor*	*Type successor*	K_2
1	VI	$R < d^2$	$(d, p, R+1)$	VI	$K+1$
2	VII	$R < (s+1)(d-1)$	$(d, p, R+1)$	VII	$K+1$
3	VIII	$R < s \cdot d$	$(d, p, R+1)$	VIII	$K+1$
4	VI	$R = d^2$	$(d, p+1, 1)$	VII	$K+1$
5	VII	$R = (s+1)(d-1), p = 1, s = d-3$	$(d-1, 1, d^2-2d)$	VI	K
6	VII	$R = (s+1)(d-1), p = 1, s = d-2$	$(d-1, 2, 1)$	VII	$K+1$
7	VII	$R = (s+1)(d-1), p = 1, s = d-1$	$(d-1, 1, 1)$	VIII	$K+1$
8	VII	$R = (s+1)(d-1), p \geqslant 2, s = d-2$	$(d-1, 2, 1)$	VIII	$K+1$
9	VII	$R = (s+1)(d-1), p \geqslant 2, s = d-1$	$(d-1, 1, 1)$	VIII	$K+1$
10	VII	$R = (s+1)(d-1), p = 1, s \leqslant d-4$	$(d-1, 1, (s+1)d)$	VII	K
11	VII	$R = (s+1)(d-1), p \geqslant 2, s \leqslant d-3$	$(d-1, 1, (s+1)d)$	VIII	K
12	VIII	$R = sd, q = 1, s = d-1$	$(d, p+1, d^2-1)$	VI	K
13	VIII	$R = sd, q = 1, s = d$	$(d, p+2, 1)$	VII	$K+1$
14	VIII	$R = sd, q \geqslant 2, s = d$	$(d, p+2, 1)$	VIII	$K+1$
15	VIII	$R = sd, q = 1, s \leqslant d-2$	$(d, p+1, s(d+1))$	VII	K
16	VIII	$R = sd, q \geqslant 2, s \leqslant d-1$	$(d, p+1, s(d+1))$	VIII	K

We leave the proof as an exercise to the reader.

In the cases 5, 10, 11, 12, 15 and 16 we have $K_2 = K$. Some easy calculations show that in these cases we also have $I_{M_2} = I_M$. It is also easy to verify that if $M_1 \ll M_2 \ll M_3$ then $K_1 < K_3$.

These facts prove our next lemma.

LEMMA. *Let M_1, M_2, M_3 be $N \times N$ matrices of types VI, VII or VIII, with K_1, K_2, K_3 resp. ones.*
If $M_1 \ll M_2$ then $K_2 = K_1 + 1$ or $K_2 = K_1$ and $I_{M_2} = I_{M_1}$.
If $M_1 \ll M_2 \ll M_3$ then $K_3 > K_1$.

Now the theorem follows in the case $\frac{1}{4}N^2 < K < \frac{1}{2}N(N-1)$.

Acknowledgements

This research was supported by the Netherlands Foundation for Mathematics (SMC) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO).

References

[A.G.] J. Aaronson and D. Gilat, On the structure of stationary one dependent processes, School of Mathematical Sciences, Tel Aviv Univ., Israel, (1987).

[A.G.K.V.] J. Aaronson, D. Gilat, M.S. Keane and V. de Valk, An algebraic construction of a class of one-dependent processes, *Annals of Probability* **17** (1989) 128–143.

[F.] L. Finke, Two maximization problems, a paper submitted to Oregon State Univ. in partial fulfillment of the requirements for the degree of Master of Arts, 1982.

[G.K.V.] A. Gandolfi, M.S. Keane and V. de Valk, Extremal two-correlations of two-valued stationary one-dependent processes, accepted by *Prob. Theory and Related Fields* (1988).

[H.L.P.] G.H. Hardy, J.E. Littlewood and G. Pólya, *Inequalities*, Cambridge Univ. Press (1934).

[Ka.] M. Katz, Rearrangements of (0, 1) matrices, *Israel Journ. of Mathematics*, 9 (1971) 53–72.

[Kh.] A. Khintchine, Über eine Ungleichung, *Mat. Sb.* 39 (1932) 35–39.

[Lo.] G.G. Lorentz, A problem of plane measure, *Amer. Journ. Math.* 71 (1949) 417–426.

[Lu.] W.A.J. Luxemburg, On an inequality of A. Khintchine for zero-one matrices, *Journ. of Combinatorial Theory* 12 (1972) 289–296.

[V.] V. de Valk, The maximal and minimal 2-correlation of a class of 1-dependent 0–1 valued processes, *Israel Journ. of Math.* 62 (1988) 181–205.

Hilbert Space Representations of m-Dependent Processes

Vincent de Valk
Mathematical Institute, University of Groningen
P.O.Box 800
9700 AV Groningen
The Netherlands

Abstract

A representation of one-dependent processes is given in terms of Hilbert spaces, vectors and bounded linear operators on Hilbert spaces. This generalizes a construction of one-dependent processes that are not two-block-factors. We show that all one-dependent processes admit a representation. We prove that if there is in the Hilbert space a closed convex cone that is invariant under certain operators and that is spanned by a finite number of linearly independent vectors, then the corresponding process is a two-block-factor of an independent process.
Apparently the difference between two-block-factors and non-two-block-factors is determined by the geometry of invariant cones. The dimension of the smallest Hilbert space that represents a process is a measure for the complexity of the structure of the process.
For two-valued one-dependent processes we prove that if there is a cylinder with measure equal to zero, then this process can be represented by a Hilbert space with dimension smaller than or equal to the length of this cylinder. In the two-valued case we show that a cylinder (with measure equal to zero) whose length is minimal and ≤ 7, is symmetric, and we give some examples of cylinders with measure equal to zero.
We generalize the concept of Hilbert space representation to m-dependent processes and it turns out that all m-dependent processes admit a representation. Several theorems are generalized to m-dependent processes.

Keywords: one-dependence, block-factors, Hilbert space representations, stationary process, m-dependence, dynamical systems, zero-cylinders, invariant cones.

AMS classification (MSC 1991). 60 G 10; 28 D 05; 54 H 20; 52 A 20.

1 Introduction

In this paper we consider ***one-dependent processes***, which are discrete time stationary stochastic processes $(X_N)_{N\in\mathbf{Z}}$ with the property that for any given time t the past $(X_N)_{N<t}$ is independent of the future $(X_N)_{N>t}$.

Just like Markov processes, one-dependent processes are a weakening of independence, but in contrast to these we assume no knowledge about the present value X_t. Although Markov processes have been investigated thoroughly for a long time the theory of one-dependence is still young but growing.

This paper is the first that uses Hilbert space techniques to investigate one-dependent processes. The concept of Hilbert space representations was initiated by Mike Keane.

One-dependent processes arise in renormalization theory as limits of rescaling operations (see [O'Br.]). In statistical physics many models have rescaling-properties for critical values (e.g. critical temperature) of their parameters (as is conjectured by physicists). This means that the model is invariant under rescaling operations (as e.g. fractals). Such random fields should therefore typically be one-dependent.

The notion of one-dependence can be generalized to m-***dependence*** ($m \in \mathbf{N}$); which means that for any given time t $(X_N)_{N<t}$ and $(X_N)_{N\geq t+m}$ are independent.

Examples of m-dependent processes are $m+1$-***block-factors*** ; let $(Y_N)_{N\in\mathbf{Z}}$ be an i.i.d. sequence and f a function of $m+1$ variables. If we define

$$X_N := f(Y_N, \ldots, Y_{N+m})$$

then the $m+1$-block-factor $(X_N)_{N\in\mathbf{Z}}$ is an m-dependent process, as follows immediately from the definition. It is easily checked that for $m+1$-block-factors it is no restriction to assume that the underlying sequence $(Y_N)_{N\in\mathbf{Z}}$ is identically distributed with the uniform distribution over the unit interval.

Although for quite a time probabilists conjectured ([Be.], [G.H.1], [Ibr.Li.], [Ja.1-2], [O'Ci.]) that all m-dependent processes are $m+1$-block-factors, in [A.G.K.V.] a two-parameter family is shown of counterexamples of one-dependent processes (assuming only two values) that are not two-block-factors. Recently Jon Aaronson, David Gilat and Mike Keane found an example of a five-state one-dependent Markov chain that is not a two-block-factor (a paper is in preparation). More recently Burton, Goulet and Meester found a counter example of a four-state one-dependent process that is not an m-block-factor for any $m \in \mathbf{N}$ (a paper is in preparation). Several authors ([Be.], [G.H.2], [He.2], [Ja.2], [O'Ci.]) used this conjecture as hypothesis and therefore some of their results on m-dependence are only valid for $m+1$-block-factors.

In this article we generalize the construction of the counterexamples from [A.G.K.V.] by representing one-dependent processes in terms of Hilbert spaces, vectors and bounded linear operators on Hilbert spaces.

A crucial difference between the operators in Hilbert space representations (HSR)

and operators in quantum probability is that the HSR operators are defined on the whole space and are in general not self-adjoint and not even normal, while the quantum probability operators are defined on a subspace and are self adjoint.
These Hilbert space representations can supply new tools to investigate the structure of one-dependent processes and especially the essential difference between two-block-factors and non-two-block-factors. The dimension of the smallest Hilbert space that represents a process is a measure of the complexity of the structure of the process.
One-dependent processes, represented by a one-dimensional Hilbert space, are i.i.d. sequences. One-dependent processes, represented by a 2-dimensional Hilbert space, are two-block-factors. The counterexamples from [A.G.K.V.] fit with a 3-dimensional Hilbert space.
The plan of this article is as follows.
In section 2 we describe the Hilbert space representation and we show that it actually represents a consistent probability measure that is one-dependent.
In section 3 we show that each one-dependent process (Theorem 3.2) admits a Hilbert space representation. We give some examples.
In section 4 we introduce closed convex cones that are invariant under certain operators. We prove that if there is an invariant cone that is spanned by a finite number of linearly independent vectors, then the one-dependent process is a two-block-factor (Theorem 4.4). This implies that one-dependent processes with a two-dimensional Hilbert space representation are two-block-factors (Theorem 4.3). It seems that the difference between two-block-factors and non-two-block-factors is determined by the geometry of invariant cones.
In section 5 we consider cylinders with measure equal to zero. Zero-cylinders play an important role in one-dependent processes. Extremal values of so called two-correlations are attained in processes with zero-cylinders ([G.K.V.], [V.1]) and the basis of the construction of the counterexamples in [A.G.K.V.] is the fact that [111] is a zero-cylinder. It turns out that if a two-valued one-dependent process has a cylinder with measure equal to zero; i.e. $P[X_1 = i_1, \ldots, X_N = i_N] = 0$ for some $i_1, \ldots, i_N$, then this one-dependent process can be represented by a Hilbert space with dimension smaller than or equal to N (Theorem 5.1). Further we give some examples of zero-cylinders in the two-valued case. We prove in the two-valued case that a zero-cylinder whose length is minimal and ≤ 7 is symmetric (Theorem 5.2). Finally we prove that [1001] can not appear as zero-cylinder with minimal length (Theorem 5.4). We conjecture that all zero-cylinders with minimal length are symmetric in the two-valued case. Actually we conjecture that only runs of ones, runs of zero's, [101] and [010] can appear as zero-cylinder with minimal length of a $0-1$ valued one-dependent process.
In section 6 we generalize the concept of Hilbert space representation to m-dependent processes and we prove that all m-dependent processes admit a representation (Theorem 6.2).

Several theorems on one-dependent processes are generalized to m-dependent processes.
In section 7 we give a contribution to the perpetuation of mathematics by a list of conjectures and open problems.

2 The Representation.

In this section we describe the Hilbert space representation and we show that it actually gives rise to a consistent probability measure that is one-dependent.
Let H be a real Hilbert space, let $K \geq 2$ be an integer, let $A_1, \dots, A_K : H \to H$ be linear, continuous operators, let $x, y \in H$ be two fixed vectors with $< x; y > = 1$.
We assume that

$$(A_1 + \dots + A_K)h = < h; x > y \text{ for all } h \in H$$

(so $A_1 + \dots + A_K$ has rank one).
Further we assume that

$$< A_{i_1} \dots A_{i_N} y; x > \geq 0$$

for all $N \in \mathbf{N}$ and for all $i_1, \dots, i_N \in \{1, \dots, K\}$.
We call $(H, x, y, A_1, \dots, A_K)$ a ***Hilbert space representation*** (HSR) of the one-dependent process $(X_N)_{N \in \mathbf{Z}}$ (with state space $\{1, \dots, K\}$) that is defined by

$$P\left[X_1 = i_1, \dots, X_N = i_N\right] := < A_{i_1} \dots A_{i_N} y; x >$$

(for $N \in \mathbf{N}$ and $i_1, \dots, i_N \in \{1, \dots, K\}$).
First we have to check that the innerproduct defines consistently a probability measure on $\{1, \dots, K\}^{\mathbf{Z}}$. We have (using the definitions)

$$\sum_{i_N=1}^{K} P\left[X_1 = i_1, \dots, X_N = i_N\right] = \sum_{i_N=1}^{K} < A_{i_1} \dots A_{i_N} y; x > =$$

$$= < A_{i_1} \dots A_{i_{N-1}} (A_1 + \dots + A_K) y; x > =$$

$$= < A_{i_1} \dots A_{i_{N-1}} < y; x > y; x > =$$

$$= < A_{i_1} \dots A_{i_{N-1}} y; x > = P\left[X_1 = i_1, \dots, X_{N-1} = i_{N-1}\right].$$

and

$$\sum_{i_1=1}^{K} P\left[X_1 = i_1, \dots, X_N = i_N\right] = \sum_{i_1=1}^{K} < A_{i_1} \dots A_{i_N} y; x > =$$

$$= < (A_1 + \dots + A_K) A_{i_2} \dots A_{i_N} y; x > =$$

$$= << A_{i_2} \dots A_{i_N} y; x > y; x > =$$

$$= < A_{i_2} \dots A_{i_N} y; x > = P\left[X_2 = i_2, \dots, X_N = i_N\right].$$

We see that

$$\sum_{i=1}^{K} P[X_1 = i] = < (A_1 + \ldots + A_K)y; x > =$$

$$= < < y; x > y; x > = < y; x >< y; x > = 1$$

and we conclude that the innerproduct (which was required to be non-negative) consistently defines a probability measure.
From

$$\sum_{i=1}^{K} P[X_1 = i_1, \ldots, X_{N-1} = i_{N-1}, X_N = i, X_{N+1} = i_{N+1}, \ldots, X_{N+M} = i_{N+M}] =$$

$$= < A_{i_1} \ldots A_{i_{N-1}} (A_1 + \ldots + A_K) A_{i_{N+1}} \ldots A_{i_{N+M}} y; x > =$$

$$= < A_{i_1} \ldots A_{i_{N-1}} < A_{i_{N+1}} \ldots A_{i_{N+M}} y; x > y; x > =$$

$$= < A_{i_1} \ldots A_{i_{N-1}} y; x >< A_{i_{N+1}} \ldots A_{i_{N+M}} y; x > =$$

$$= P[X_1 = i_1, \ldots, X_{N-1} = i_{N-1}] \cdot P[X_{N+1} = i_{N+1}, \ldots, X_{N+M} = i_{N+M}]$$

we conclude that $(X_N)_{N \in \mathbf{Z}}$ is a one-dependent process. □

3 Examples of Hilbert Space Representations.

In this section we show that every one-dependent process admits a Hilbert space representation and we give some examples of representations. First we need a technical theorem.

Theorem 3.1 *Let $(X_N)_{N\in\mathbf{Z}}$ be a one-dependent process over $\{1,\ldots,K\}^{\mathbf{Z}}$. Let H_0 be a Hilbert space, let $x \in H_0$ be a vector with $< x;x >= 1$, let $A_1,\ldots,A_K : H_0 \to H_0$ be linear, continuous operators such that $(A_1+\ldots+A_K)x = x$.*
Assume that

$$< A_{i_1}\ldots A_{i_N}x;x > = P[X_1 = i_1,\ldots,X_N = i_N]$$

for all $N \in \mathbf{N}$ and all $i_1,\ldots,i_N \in \{1,\ldots,K\}$.
Then there exists a closed separable subspace $H \subset H_0$ with $x \in H$, such that $(H,x,x,PA_1,\ldots,PA_K)$ is a HSR of $(X_N)_{N\in\mathbf{Z}}$, where $P : H_0 \to H$ is the orthogonal projection from H_0 on H.

Proof. We define the collection $\mathcal{H}$ of those closed subspaces H of H_0 with the properties that $x \in H$ and that for the orthogonal projection $P : H_0 \to H$ holds

$$< PA_{i_1}\ldots PA_{i_N}x;x > = P[X_1 = i_1,\ldots,X_N = i_N]$$

for all $N \in \mathbf{N}$ and all $i_1,\ldots,i_N \in \{1,\ldots,K\}$. We define a partial ordering on $\mathcal{H}$ by

$$H_1 \leq H_2 \text{ if } H_1 \supset H_2.$$

Note that $\mathcal{H} \neq \emptyset$ because $H_0 \in \mathcal{H}$. □

Claim 1. We claim that every totally ordered subset of $\mathcal{H}$ has an upper bound.
Proof of Claim 1. Let $\mathcal{H}_1 = \{H_\theta : \theta \in \Theta\}$ be a totally ordered subset of $\mathcal{H}$. Define $H_1 := \bigcap_{\theta\in\Theta} H_\theta$. We will show that H_1 is an upper bound of $\mathcal{H}_1$. First we prove the following claim.
Claim 2. $H_1 \in \mathcal{H}$.
Proof of Claim 2.
Because $H_1 \subset H_\theta$ for all θ, we have $H_1^\perp \supset H_\theta^\perp$ for all θ. So $H_1^\perp \supset \bigcup_\theta (H_\theta^\perp)$, and $H_1^\perp \supset \overline{\bigcup_\theta (H_\theta^\perp)}$. Assume that there exists a $h \in H_1^\perp$ such that $h \in \left(\overline{\bigcup_\theta (H_\theta^\perp)}\right)^\perp$. Then $h \in (H_\theta^\perp)^\perp = H_\theta$ for all θ, so $h \in \bigcap_\theta H_\theta = H_1$. But $h \in H_1^\perp$ and $h \in H_1$ implies $h = 0$. We conclude that $H_1^\perp = \overline{\bigcup_\theta (H_\theta^\perp)}$.
Let $P_1 : H_0 \to H_1$ and $P_\theta : H_0 \to H_\theta$ $(\theta \in \Theta)$ be the orthogonal projections.
Let $z \in H_1^\perp$. For any $\varepsilon > 0$ we can approximate z by a vector $h \in \bigcup_\theta (H_\theta^\perp)$ such

that $\|z-h\| < \varepsilon$. So $h \in H_{\theta_0}^{\perp}$ for some θ_0. For $H_\theta \geq H_{\theta_0}$ we have $P_\theta h \in H_\theta \cap H_{\theta_0}^{\perp} \subset H_\theta \cap H_\theta^{\perp} = \{0\}$. Therefore

$$\|P_\theta z\| = \|P_\theta z - P_\theta h\| \leq \|P_\theta\| \cdot \|z-h\| < \varepsilon$$

if $H_\theta \geq H_{\theta_0}$.
Now let $y \in H_0$. Take $z \in H_1^{\perp}$ and $w \in H_1$ such that $y = z + w$. Let $\varepsilon > 0$ be given. Take θ_0 as above. We have for $H_\theta \geq H_{\theta_0}$ that

$$\begin{aligned} \|(P_\theta - P_1)y\| &= \|P_\theta(z+w) - P_1(z+w)\| = \\ &= \|P_\theta z + w - 0 - w\| = \|P_\theta z\| < \varepsilon. \end{aligned}$$

We conclude that $P_\theta y \xrightarrow{\theta} P_1 y$ for all $y \in H_0$.
This implies that (for all i_1)

$$P[X_1 = i_1] = < P_\theta A_{i_1} x; x > \xrightarrow{\theta} < P_1 A_{i_1} x; x > .$$

Because

$$\|P_\theta A_{i_1} P_\theta A_{i_2} x - P_1 A_{i_1} P_1 A_{i_2} x\| =$$

$$= \|P_\theta A_{i_1}(P_\theta A_{i_2} x - P_1 A_{i_2} x) + (P_\theta - P_1)(A_{i_1} P_1 A_{i_2} x)\| \leq$$

$$\leq \|P_\theta\| \cdot \|A_{i_1}\| \cdot \|(P_\theta - P_1)(A_{i_2} x)\| + \|(P_\theta - P_1)(A_{i_1} P_1 A_{i_2} x)\|$$

and $\|P_\theta\| = 1$, we derive that (for all i_1, i_2)

$$P[X_1 = i_1, X_2 = i_2] = < P_\theta A_{i_1} P_\theta A_{i_2} x; x > \xrightarrow{\theta} < P_1 A_{i_1} P_1 A_{i_2} x; x > .$$

By induction (on N) we derive that

$$P[X_1 = i_1, \ldots, X_N = i_N] = < P_\theta A_{i_1} \ldots P_\theta A_{i_N} x; x > \xrightarrow{\theta} < P_1 A_{i_1} \ldots P_1 A_{i_N} x; x >$$

(for all $N \in \mathbf{N}$ and all $i_1, \ldots, i_N \in \{1, \ldots, K\}$).
Because $x \in H_\theta$ for all $\theta \in \Theta$, we have $x \in H_1$. We conclude that H_1 satisfies the 2 conditions in the definition of $\mathcal{H}$. Thus $H_1 \in \mathcal{H}$. This proves Claim 2.
Because $H_1 \subset H_\theta$ for all θ, H_1 is an upperbound of $\mathcal{H}_1$. This proves Claim 1.
Now we have proved (Claim 1) that every totally ordered subset of $\mathcal{H}$ has an upper bound, we can apply Zorn's Lemma that implies the existence of a maximal element. Let H be a maximal element in $\mathcal{H}$. Let $P : H_0 \to H$ be the orthogonal projection on H.
Claim 3. We claim that $(H, x, x, PA_1, \ldots, PA_K)$ is a HSR of $(X_N)_{N \in \mathbf{Z}}$.
Proof of Claim 3.
Consider the restricted operators $PA_1 \mid_H, \ldots, PA_K \mid_H$ from H to H. Let B_i :

$H \to H$ be the adjoints of these restricted operators ($i = 1, \dots, K$). We define the separable subspace

$$H_B := \overline{sp}\{B_{j_1} \dots B_{j_m} x : m \geq 0, j_1, \dots, j_m \in \{1, \dots, K\}\}.$$

To prove Claim 3 we first have to prove the following Claim.
Claim 4. We claim that $H = H_B$.
Assume that $H \overset{\supset}{\neq} H_B$. Apparently $B_i H_B \subset H_B$ for all i. Consider the restricted operators $B_i\,|_{H_B}$ ($i = 1, \dots, K$).
Let $C_i : H_B \to H_B$ be the adjoints of these restricted operators ($i = 1, \dots, K$). Now we will show that $H_B \in \mathcal{H}$ and that $H_B > H$, what contradicts the maximality of H.
Let $P_B : H_0 \to H_B$ be the orthogonal projection, let $y, z \in H_B$, then

$$< P_B A_i y; z > = < P_B P A_i y; z > =$$

$$= < P A_i y; P_B^* z > = < (P A_i\,|_H) y; P_B z > =$$

$$= < (P A_i\,|_H) y; z > = < y; B_i z > =$$

$$= < y; (B_i\,|_{H_B}) z > = < C_i y; z > .$$

This implies that $P_B A_i = C_i$ for all $i = 1, \dots, K$.
Further we have (for all N and for all $i_1, \dots, i_N$)

$$P[X_1 = i_1, \dots, X_N = i_N] = < P A_{i_1} \dots P A_{i_N} x; x > =$$

$$= < (P A_{i_1}\,|_H) \dots (P A_{i_N}\,|_H) x; x > = < x; B_{i_N} \dots B_{i_1} x > =$$

$$= < x; (B_{i_N}\,|_{H_B}) \dots (B_{i_1}\,|_{H_B}) x > = < C_{i_1} \dots C_{i_N} x; x > .$$

Together with $x \in H_B$ (by definition of H_B) this implies that $H_B \in \mathcal{H}$. Because we assumed $H_B \overset{\subset}{\neq} H$, we have $H_B > H$, what contradicts the maximality of H. We conclude that $H = H_B$. This proves Claim 4. To prove Claim 3 we have to show that

$$(P A_1 + \dots + P A_K) h = < h; x > x$$

for all $h \in H$.
This is equivalent to

$$< (P A_1 + \dots + P A_K) h; g > = < h; x > < x; g > \qquad (*)$$

for all $g, h \in H$.
Because

$$H = \overline{sp}\,\{B_{j_1} \dots B_{j_m} x : m \geq 0, j_1, \dots, j_m \in \{1, \dots, K\}\}\,,$$

and

$$H = \overline{sp}\{PA_{i_1} \dots PA_{i_N} x : N \geq 0, i_1, \dots, i_N \in \{1, \dots, K\}\}$$

(if the right hand side is a proper subspace of H, then this would contradict the maximality of H) and because (*) is a linear equation in h and g, it is sufficient to check (*) for $h = PA_{i_1} \dots PA_{i_N} x$ and $g = B_{j_1} \dots B_{j_m} x$ (for all $N, m \in \mathbf{N}, i_1, \dots, i_N, j_1, \dots, j_m \in \{1, \dots, K\}$).
For this h and g we have

$$< (PA_1 + \dots + PA_K)h;\ g >=$$

$$= < (PA_1 + \dots + PA_K)\ PA_{i_1} \dots PA_{i_N} x;\ B_{j_1} \dots B_{j_m} x > =$$

$$= < PA_{j_m} \dots PA_{j_1}\ (PA_1 + \dots + PA_K)\ PA_{i_1} \dots PA_{i_N} x;\ x > =$$

$$= \sum_{i=1}^{K} P[X_{-m} = j_m, \dots, X_{-1} = j_1,\ X_0 = i,\ X_1 = i_1, \dots, X_N = i_N] =$$

$$= P[X_{-m} = j_m, \dots, X_{-1} = j_1] \cdot P[X_1 = i_1, \dots, X_N = i_N] =$$

$$= < PA_{j_m} \dots PA_{j_1} x;\ x > < PA_{i_1} \dots PA_{i_N} x;\ x > =$$

$$= < x; B_{j_1} \dots B_{j_m} x > < PA_{i_1} \dots PA_{i_N} x;\ x > =$$

$$= < x; g > < h; x > .$$

This proves (*) and the proof of Claim 3 is finished. Claim 3 implies the theorem. □

Remark. We restricted ourselves in Theorem 3.1 to $H \subset H_0$ because in general (*) does not hold for all $h, g \in H_0$ (as is easy to see in the proof of Theorem 3.2, where we apply Theorem 3.1).
Now we can prove the main theorem of this section.

Theorem 3.2 *Let $(X_N)_{N \in \mathbf{Z}}$ be a K-valued (for some $K \in \mathbf{N}$) one-dependent process.*
Then there exists a HSR of $(X_N)_{N \in \mathbf{Z}}$.

Proof. Let $(X_N)_{N \in \mathbf{Z}}$ be a one-dependent process over $\{1, \dots, K\}^{\mathbf{Z}}$. $(X_N)_{N \in \mathbf{Z}}$ induces a probability measure P on $\{1, \dots, K\}^{\mathbf{N}}$. We define the Hilbert space $H_0 := L^2(P)$. Let $I \in H_0$ be the function that is identically one. We have $< I; I > = 1$.
We define the operators $A_1, \dots, A_K : H_0 \to H_0$ by $(A_i h)(w_1, w_2, w_3, \dots) :=$

$I_i(w_1)h(w_2, w_3, \ldots)$ for $h \in H_0$, where $I_i(w) := \begin{cases} 1 \text{ if } w = i \\ 0 \text{ if } w \neq i \end{cases}$.
Apparently $A_1, \ldots, A_K$ are linear and continuous and they satisfy the equation

$$< A_{i_1} \ldots A_{i_N} I; I > =$$

$$= \int I_{i_i}(w_1) I_{i_2}(w_2) \ldots I_{i_N}(w_N)\, d\, P(w) =$$

$$= P[X_1 = i_1, \ldots, X_N = i_N]$$

for all $N \in \mathbf{N}$ and all $i_1, \ldots, i_N \in \{1, \ldots, K\}$. Further $(A_1 + \ldots + A_K)I = I$ holds. Theorem 3.1 now implies the existence of a HSR of $(X_N)_{N \in \mathbf{Z}}$. □ □

The Hilbert space representation of a one-dependent process is not unique. In the Theorems 3.3, 3.4 and 3.5 we give some examples of HSR's.

Theorem 3.3 *Let $(X_N)_{N \in \mathbf{Z}}$ be a one-dependent process over $\{0,1\}^{\mathbf{Z}}$. Then there exists a HSR of $(X_N)_{N \in \mathbf{Z}}$ with Hilbert space ℓ^2.*

Proof. In [A.G.K.V.] (Theorem 1) it is proved that the distribution of a $0-1$ valued one-dependent process is uniquely determined by its values

$$[1^N] := P\left[X_1 = \ldots = X_N = 1\right] \qquad (N \in \mathbf{N}).$$

Let $H := \ell^2, y := \begin{pmatrix} 1 \\ [1] \\ [11] \\ \vdots \\ [1^N] \\ \vdots \end{pmatrix}, x := \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \end{pmatrix},$

$$A_0 = \begin{pmatrix} 1 & -1 & 0 & 0 & \ldots & & \\ [1] & 0 & -1 & 0 & \ldots & & \\ [11] & 0 & 0 & -1 & \vdots & & \\ \vdots & & & & & & \\ [1^N] & 0 & 0 & \ldots & \ldots & \ddots & -1\ldots \\ \vdots & & & & & & \end{pmatrix}, A_1 = \begin{pmatrix} 0 & 1 & 0 & 0 & \ldots & & \\ 0 & 0 & 1 & 0 & \ldots & & \\ 0 & 0 & 0 & 1 & \ldots & & \\ \vdots & & & & & & \\ 0 & 0 & 0 & 0 & \ldots & & 1\ldots \\ \vdots & & & & & & \end{pmatrix}.$$

Because

$$\left[1^{N+M+1}\right] \leq P\left[X_1 = \ldots = X_N = 1, X_{N+2} = \ldots = X_{N+M+1} = 1\right] =$$

$$= P\left[X_1 = \ldots = X_N = 1\right] P\left[X_{N+2} = \ldots = X_{N+M+1} = 1\right] = \left[1^N\right] \cdot \left[1^M\right]$$

it is easy to see that actually $x, y \in \ell^2$ and that A_0 and A_1 are continuous operators on ℓ^2.
It is trivial that $(A_0 + A_1)h = < h; x > y$ holds for all $h \in \ell^2$ and that $< x; y >= 1$.
From $< A_1^N y; x >= < \begin{pmatrix} [1^N] \\ [1^{N+1}] \\ \vdots \end{pmatrix}; \begin{pmatrix} 1 \\ 0 \\ \vdots \end{pmatrix} > = [1^N]$ $(\forall N \in \mathbf{N})$ and Theorem 1 of [A.G.K.V.] we conclude that (ℓ^2, x, y, A_0, A_1) is a HSR of $(X_N)_{N \in \mathbf{Z}}$. □ □

Remark. The "special" processes in [A.G.K.V.] are represented by $H = \mathbf{R}^3$, $y = \begin{pmatrix} 1 \\ \alpha \\ \beta \end{pmatrix}$, $x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $A_0 = \begin{pmatrix} 1 & -1 & 0 \\ \alpha & 0 & -1 \\ \beta & 0 & 0 \end{pmatrix}$, $A_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. The two-parameter family of counterexamples of one-dependent processes that are not two-block-factors corresponds with HSR's of this type.

Theorem 3.4 *Let $(X_N)_{N \in \mathbf{Z}}$ be a K-valued (for some $K \in \mathbf{N}$) two-block-factor of an i.i.d. sequence.*
Then there exists a HSR of $(X_N)_{N \in \mathbf{Z}}$ with Hilbert space $L^2[0,1]$.

Proof. Let $X_N = f(Y_N, Y_{N+1})$ for some function f and some i.i.d. sequence $(Y_N)_{N \in \mathbf{Z}}$ of random variables that are uniformly distributed over the unit interval. We define the sets $V_i (i = 1, \ldots, K)$ in the unit square;

$$V_i := \{(t, s) : f(t, s) = i\}$$

Let $H = L^2[0, 1]$, let the operators A_i be defined by

$$(A_i g)(t) := \int_0^1 I_{V_i}(t, s)\, g(s)\, ds \qquad (i = 1, \ldots, K)$$

where I_{V_i} is the indicatorfunction of V_i. Let $\mathbf{I} \in H$ be the function that is identically one.
It is an easy exercise to prove that $(H, \mathbf{I}, \mathbf{I}, A_1, \ldots, A_K)$ is a HSR of $(X_N)_{N \in \mathbf{Z}}$ □ □

We generalize this construction of a HSR to one-dependent m-block-factors (for any $m \in \mathbf{N}$). Generally an m-block-factor is $(m-1)$-dependent, but for special choices of the function f the m-block-factor $X_N = f(Y_N, \ldots, Y_{N+m-1})$ can be one-dependent. It is an open problem whether there exist one-dependent m-block-factors ($m \geq 3$) that can not be written as a two-block-factor.

Theorem 3.5 *Let $(X_N)_{N \in \mathbf{Z}}$ be a K-valued one-dependent m-block-factor of an i.i.d. sequence (for some $K, m \in \mathbf{N}$). Then there exists a HSR of $(X_N)_{N \in \mathbf{Z}}$ with as Hilbert space a subspace of $L^2([0,1]^{m-1})$.*

Proof. Let $X_N = f(Y_N, \ldots, Y_{N+m-1})$ for some function f of m variables and $(Y_N)_{N \in \mathbf{Z}}$ an i.i.d. sequence of random variables that are uniformly distributed over the unit interval. We define the sets V_i in the m-dimensional unit cube $[0,1]^m$;

$$V_i := \{(y_1, \ldots, y_m) : f(y_1, \ldots, y_m) = i\} \quad (i = 1, \ldots, K).$$

Let $H_0 := L^2([0,1]^{m-1})$.
We define the operators $A_i : H_0 \to H_0$ by

$$(A_i h)(y_1, \ldots, y_{m-1}) := \int_0^1 I_{V_i}(y_1, \ldots, y_m)\, h(y_2, \ldots, y_m)\, dy_m, \ (i = 1, \ldots, K)$$

where I_{V_i} is the indicator function of the set V_i $(i = 1, \ldots, K)$. Let I be the function on $[0,1]^{m-1}$ that is identically one. $(A_1 + \ldots + A_K)I = I$ holds. We have (as is easily checked) that

$$(A_{i_1} \ldots A_{i_N} I)\,(y_1, \ldots, y_{m-1}) =$$

$$= \int_0^1 I_{V_{i_1}}(y_1, \ldots, y_m)(A_{i_2} \ldots A_{i_N} I)(y_2, \ldots, y_m)\, dy_m =$$

$$= \int_0^1 I_{V_{i_1}}(y_1, \ldots, y_m) \int_0^1 I_{V_{i_2}}(y_2, \ldots, y_{m+1})(A_{i_3} \ldots A_{i_N} I)(y_3, \ldots, y_{m+1})\, dy_{m+1}\, dy_m = \ldots$$

$$= \int_0^1 \ldots \int_0^1 I_{V_{i_1}}(y_1, \ldots, y_m) I_{V_{i_2}}(y_2, \ldots, y_{m+1}) \ldots I_{V_{i_N}}(y_N, \ldots, y_{N+m-1})\, dy_{N+m-1} \ldots dy_m$$

and so we have

$$< A_{i_1} \ldots A_{i_N} I; I > =$$

$$= \int_0^1 \ldots \int_0^1 I_{V_{i_1}}(y_1, \ldots, y_m) I_{V_{i_2}}(y_2, \ldots, y_{m+1}) \ldots I_{V_{i_N}}(y_N, \ldots, y_{N+m-1})\, dy_1 \ldots dy_{N+m-1} =$$

$$P[X_1 = i_1, \ldots, X_N = i_N].$$

Now Theorem 3.1 implies the existence of a HSR of $(X_N)_{N \in \mathbf{Z}}$. □ □

The reversed process of a one-dependent process is also one-dependent. The following theorem gives a HSR.

Theorem 3.6 *Let $(H, x, y, A_1, \ldots, A_K)$ be a HSR of a one-dependent process $(X_N)_{N \in \mathbf{Z}}$. Let $(Y_N)_{N \in \mathbf{Z}}$ be the reversed process; i.e. $Y_N := X_{-N}$ $(N \in \mathbf{Z})$. Then $(H, y, x, A_1^*, \ldots, A_K^*)$ is a HSR of $(Y_N)_{N \in \mathbf{Z}}$.*

Proof. $< (A_1^* + \ldots + A_K^*)h;\ g > =$

$= < h;\ (A_1 + \ldots + A_K)g > = < h;\ < g; x > \ y > =$

$= < h; y > < g; x > = < < h; y > \ x; g >$

for all $h, g \in H$. This implies that

$(A_1^* + \ldots + A_K^*)h = < h; y > \ x \ \ \forall h \in H.$

Further

$< A_{i_1}^* \ldots A_{i_N}^* x; y > =$

$= < x; A_{i_N} \ldots A_{i_1} y > = P[X_1 = i_N, \ldots, X_N = i_1] =$

$= P[Y_1 = i_1, \ldots, Y_N = i_N].$

□ □

4 Finite Dimension and Invariant Cones.

In this section we prove that a HSR with 2-dimensional Hilbert space corresponds with a two-block-factor. Further we show that if there is an invariant (under $A_1, \ldots, A_K$) cone spanned by a finite number of linearly independent vectors, then the HSR corresponds with a two-block-factor. The first theorem is just a special case of the other one.
We need a technical theorem to show that it is no restriction to assume that the vectors x and y are equal.

Theorem 4.1 *Let $(H, x, y, A_1, \ldots, A_K)$ be a HSR of a one-dependent process $(X_N)_{N \in \mathbf{Z}}$. Then there exists a vector $x_0 \in H$ and there exist operators $B_1, \ldots, B_K : H \to H$ such that $(H, x_0, x_0, B_1, \ldots, B_K)$ is a HSR of $(X_N)_{N \in \mathbf{Z}}$.*

Proof. Case 1. If x and y are linearly dependent, then it is easy to see that $\left(H, \frac{x}{\|x\|}, \frac{x}{\|x\|}, A_1, \ldots, A_K\right)$ is a HSR of $(X_N)_{N \in \mathbf{Z}}$.
Case 2. If x and y are linearly independent, then we consider the 2-dimensional subspace H_0 that is spanned by x and y;

$$H_0 := sp\{x, y\}$$

and its orthogonal complement $H_0^{\perp}$;

$$H_0^{\perp} := \{h \in H :< h; x > = < h; y > = 0\}$$

Take some orthonormal basis of H_0, and assume that $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ with respect to this basis.
We have $1 = < x; y > = x_1 y_1 + x_2 y_2$.
Let $\lambda \in \mathbf{R}, \lambda \neq 0$. We define the linear operator $V : H \to H$ by

$$V \mid_{H_0} = \begin{pmatrix} y_1 & -\lambda x_2 \\ y_2 & \lambda x_1 \end{pmatrix}$$

and $V \mid_{H_0^{\perp}} =$ identity.
It is easy to see that V is invertible and

$$V^{-1} \mid_{H_0} = \frac{1}{\lambda} \begin{pmatrix} \lambda x_1 & \lambda x_2 \\ -y_2 & y_1 \end{pmatrix}$$

and $V^{-1} \mid_{H_0^{\perp}} =$ identity.
We claim that

$$(H, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, V^{-1} A_1 V, \ldots, V^{-1} A_K V)$$

is a HSR of $(X_N)_{N\in\mathbf{Z}}$. It is clear that $< \begin{pmatrix} 1 \\ 0 \end{pmatrix} ; \begin{pmatrix} 1 \\ 0 \end{pmatrix} > = 1$. Further, let $h \in H$. We have

$$(V^{-1}A_1V + \ldots + V^{-1}A_KV)h = V^{-1}(A_1 + \ldots + A_K)Vh =$$

$$= V^{-1} < Vh; x > y = < Vh; x > V^{-1}y = < Vh; x > \begin{pmatrix} 1 \\ 0 \end{pmatrix} =$$

$$= < h; V^*x > \begin{pmatrix} 1 \\ 0 \end{pmatrix} = < h; \begin{pmatrix} 1 \\ 0 \end{pmatrix} > \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

and

$$< V^{-1}A_{i_1}V \ldots V^{-1}A_{i_N}V \begin{pmatrix} 1 \\ 0 \end{pmatrix} ; \begin{pmatrix} 1 \\ 0 \end{pmatrix} > =$$

$$= < V^{-1}A_{i_1} \ldots A_{i_N}y; \begin{pmatrix} 1 \\ 0 \end{pmatrix} > = < A_{i_1} \ldots A_{i_N}y; (V^{-1})^* \begin{pmatrix} 1 \\ 0 \end{pmatrix} > =$$

$$= < A_{i_1} \ldots A_{i_N}y; x >,$$

which proves Theorem 4.1. □ □

Remark. The fact that any orthonormal basis of H_0 and any $\lambda \neq 0$ can be chosen in the proof of Theorem 4.1 shows the non-uniqueness of the Hilbert Space Representations.
In Theorem 4.3 we need the following Lemma.

Lemma 4.2 . *Let $(H, x, x, A_1, \ldots, A_K)$ be a HSR of a one-dependent process $(X_N)_{N\in\mathbf{Z}}$. Let*

$$T := \overline{co}\left\{\alpha A_{i_1} \ldots A_{i_N}x : \alpha \geq 0, N \in \mathbf{N}, i_1, \ldots, i_N \in \{1, \ldots, K\}\right\}.$$

If $\exists v \in T, v \neq 0$ with $< v, x > = 0$, then $(X_N)_{N\in\mathbf{Z}}$ has a HSR with Hilbert space

$$H_0 = \{v \in T :< v; x > = 0\}^\perp \overset{\subset}{\neq} H.$$

Proof.
Let $V := \overline{sp}\{v \in T :< v; x > = 0\}$, then $H_0 = V^\perp$. Note that $x \in H_0$. Let P be the orthogonal projection on H_0. We show that

$$(H_0, x, x, PA_1, \ldots, PA_K)$$

is a HSR of $(X_N)_{N\in\mathbf{Z}}$.

Let $v \in T$ with $< v; x > = 0$. Because $A_i T \subset T$ we have $< A_i v; x > \geq 0$ for all $i = 1, \dots, K$. Thus

$$0 \leq \sum_{i=1}^{K} < A_i v; x > = < (A_1 + \dots + A_K) v; x > =$$

$$= << v; x > x; x > = < v; x > = 0,$$

which implies that $< A_i v; x >= 0$ for all $i = 1, \dots, K$, and all $v \in V$. Hence $A_i V \subset V$ for all $i = 1, \dots, K$.
If $h \in H$, then $h - Ph \in V$, so $< A_{i_1} \dots A_{i_m}(h - Ph); x > = 0$ for all $m \in \mathbf{N}$ and all $i_1, \dots, i_m \in \{1, \dots, K\}$, and hence $< A_{i_1} \dots A_{i_m} Ph; x > = < A_{i_1} \dots A_{i_m} h; x >$.
Now we have ($h \in H_0$)

$$(PA_1 + \dots + PA_K)h = P(A_1 + \dots + A_K)h =$$

$$= P < h; x > x = < h; x > Px = < h; x > x,$$

and

$$< PA_{i_1} PA_{i_2} \dots PA_{i_N} x; x > = < A_{i_1} PA_{i_2} \dots PA_{i_N} x; P^* x > =$$

$$= < A_{i_1} A_{i_2} PA_{i_3} \dots PA_{i_N} x; x > = \dots =< A_{i_1} \dots A_{i_N} x; x > .$$

which proves our lemma. □

Now we consider the case that the Hilbert space has dimension one or two.

Theorem 4.3 *Let $(H, x, y, A_1, \dots, A_K)$ be a HSR of a one-dependent process $(X_N)_{N \in \mathbf{Z}}$.*
(a) If $\dim(H) = 1$, then $(X_N)_{N \in \mathbf{Z}}$ is an i.i.d. sequence.
(b) If $\dim(H) = 2$, then $(X_N)_{N \in \mathbf{Z}}$ is a two-block-factor of an i.i.d. sequence.

Proof. (a) If $\dim H = 1$, then $A_i = (a_i) \quad (i = 1, \dots, K)$. We have

$$P[X_1 = i_1, \dots, X_N = i_N] = < A_{i_1} \dots A_{i_N} y; x > =$$

$$= a_{i_1} \dots a_{i_N} < y; x > = a_{i_1} \dots a_{i_N} = P[X_1 = i_1] \dots P[X_N = i_N] . \square$$

(b) Theorem 4.1 implies that we may assume that $x = y$.
If $\dim H = 2$, then we consider the closed convex cone spanned by the orbit of x under the operators $A_1, \dots, A_K$;

$$T := \overline{co} \{ \alpha A_{i_1} \dots A_{i_N} x : \alpha \geq 0, N \in \mathbf{N}, i_1, \dots, i_N \in \{1, \dots, K\} \} .$$

Note that $x \in T$, and that $A_i T \subset T \quad \forall i = 1, \dots, K$.

We choose an orthonormal basis of $\mathbf{R}^2$ such that $x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

The lemma implies that there exist vectors $v = \begin{pmatrix} 1 \\ v_2 \end{pmatrix}$, $w = \begin{pmatrix} 1 \\ w_2 \end{pmatrix}$ such that $v_2 - w_2 > 0$ and

$$T = \overline{co}\{\alpha v, \alpha w : \alpha \geq 0\}.$$

Let $A_i = \begin{pmatrix} a^i_{11} & a^i_{12} \\ a^i_{21} & a^1_{22} \end{pmatrix}$ be the matrix of A_i with respect to the basis $\{v, w\}$ $(i = 1, \dots, K)$. Because $A_i v, A_i w \in \overline{co}\{\alpha v, \alpha w : \alpha \geq 0\}$ it follows that $a^i_{j_1 j_2} \geq 0$ $\forall i = 1, \dots, K$, $\forall j_1, j_2 \in \{1, 2\}$.
With respect to the standard basis we have

$$(A_1 + \ldots + A_K)v = < \begin{pmatrix} 1 \\ v_2 \end{pmatrix}; \begin{pmatrix} 1 \\ 0 \end{pmatrix} > \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = x$$

On the other hand we have

$$(A_1 + \ldots + A_K)v = \sum_{i=1}^{K} (a^i_{11} v + a^i_{21} w);$$

hence

$$1 = < x; x > = < \sum_{i=1}^{K} (a^i_{11} v + a^i_{21} w); x > =$$

$$= 1 = \sum_{i=1}^{K} (a^i_{11} + a^i_{21}). \tag{1}$$

Analogously (considering $(A_1 + \ldots + A_K)w$) we find that

$$\sum_{i=1}^{K} (a^i_{12} + a^i_{22}) = 1. \tag{2}$$

Further we have

$$0 = < x; \begin{pmatrix} 0 \\ 1 \end{pmatrix} > = < \sum_{i=1}^{K} (a^i_{11} v + a^i_{21} w); \begin{pmatrix} 0 \\ 1 \end{pmatrix} > =$$

$$= 0 = \sum_{i=1}^{K} (a^i_{11} v_2 + a^i_{21} w_2). \tag{3}$$

Analogously

$$\sum_{i=1}^{K}(a_{12}^{i}v_2 + a_{22}^{i}w_2) = 0. \tag{4}$$

(1), (2), (3), (4) imply that $\sum_{i=1}^{K} a_{11}^{i} = \sum_{i=1}^{K} a_{12}^{i}$ and $\sum_{i=1}^{K} a_{21}^{i} = \sum_{i=1}^{K} a_{22}^{i}$.

Let us define the matrix $S = \begin{pmatrix} 1 & 1 \\ v_2 & w_2 \end{pmatrix}$ then

$$S^{-1} = \begin{pmatrix} \sum_{i=1}^{K} a_{11}^{i} & \frac{1}{v_2-w_2} \\ \sum_{i=1}^{K} a_{21}^{i} & \frac{-1}{v_2-w_2} \end{pmatrix}$$

as is easily checked.

We note that A_i has matrix

$$S\begin{pmatrix} a_{11}^{i} & a_{12}^{i} \\ a_{21}^{i} & a_{22}^{i} \end{pmatrix} S^{-1}$$

with respect to the standard basis. So we have (with respect to the standard basis)

$$< A_{i_1} \dots A_{i_N}\begin{pmatrix} 1 \\ 0 \end{pmatrix}; \begin{pmatrix} 1 \\ 0 \end{pmatrix} > =$$

$$= < S\begin{pmatrix} a_{11}^{i_1} & a_{12}^{i_1} \\ a_{21}^{i_1} & a_{22}^{i_1} \end{pmatrix} S^{-1} \dots S\begin{pmatrix} a_{11}^{i_N} & a_{12}^{i_N} \\ a_{21}^{i_N} & a_{22}^{i_N} \end{pmatrix} S^{-1}\begin{pmatrix} 1 \\ 0 \end{pmatrix}; \begin{pmatrix} 1 \\ 0 \end{pmatrix} > =$$

$$= < \begin{pmatrix} a_{11}^{i_1} & a_{12}^{i_1} \\ a_{21}^{i_1} & a_{22}^{i_1} \end{pmatrix} \dots \begin{pmatrix} a_{11}^{i_N} & a_{12}^{i_N} \\ a_{21}^{i_N} & a_{22}^{i_N} \end{pmatrix} S^{-1}\begin{pmatrix} 1 \\ 0 \end{pmatrix}; S^{*}\begin{pmatrix} 1 \\ 0 \end{pmatrix} > =$$

$$= < \begin{pmatrix} a_{11}^{i_1} & a_{12}^{i_1} \\ a_{21}^{i_1} & a_{22}^{i_1} \end{pmatrix} \dots \begin{pmatrix} a_{11}^{i_N} & a_{12}^{i_N} \\ a_{21}^{i_N} & a_{22}^{i_N} \end{pmatrix} \begin{pmatrix} \sum_{i=1}^{K} a_{11}^{i} \\ \sum_{i=1}^{K} a_{21}^{i} \end{pmatrix}; \begin{pmatrix} 1 \\ 1 \end{pmatrix} > .$$

By induction (on $N \in \mathbf{N}$) it now follows that $(X_N)_{N\in\mathbf{Z}}$ is a two-block-factor of an i.i.d. sequence $(Y_N)_{N\in\mathbf{Z}}$ of random variables that are uniformly distributed over the unit interval. We have $X_N = f(Y_N, Y_{N+1})$ with

$$f(t,s) = i \qquad \Leftrightarrow$$

$$(t,s) \in [a_{11}^{1} + \dots + a_{11}^{i-1}; a_{11}^{1} + \dots + a_{11}^{i}) \times [0; a_{11}^{1} + \dots + a_{11}^{K})$$

or

$$(t,s) \in [a_{11}^1 + \ldots + a_{11}^K + a_{21}^1 + \ldots + a_{21}^{i-1}; a_{11}^1 + \ldots + a_{11}^K + a_{21}^1 + \ldots + a_{21}^i)$$
$$\times [0; a_{11}^1 + \ldots + a_{11}^K)$$

or

$$(t,s) \in [a_{12}^1 + \ldots + a_{12}^{i-1}; a_{12}^1 + \ldots + a_{12}^i) \times [a_{11}^1 + \ldots + a_{11}^K; 1)$$

or

$$(t,s) \in [a_{12}^1 + \ldots + a_{12}^K + a_{22}^1 + \ldots + a_{22}^{i-1}; a_{12}^1 + \ldots + a_{12}^K + a_{22}^1 + \ldots + a_{22}^i)$$
$$\times [a_{11}^1 + \ldots + a_{11}^K; 1)$$

This corresponds with the values of f shown in Figure 1. □
□

We generalize Theorem 4.3 to the case of more dimensions when there exists an invariant cone spanned by a finite number of linearly independent vectors.

Theorem 4.4 *Let* $(\mathbf{R}^N, x_0, x_0, A_1, \ldots, A_K)$ *be a HSR of a one-dependent process* $(X_N)_{N\in\mathbf{Z}}$.
Assume that there exist N linearly independent vectors $v_1, \ldots, v_N \in \mathbf{R}^N$ with $< v_i; x_0 >> 0$ for all i and such that the cone

$$T := \{\alpha_1 v_1 + \ldots + \alpha_N v_N : \alpha_1 \geq 0, \ldots, \alpha_N \geq 0\}$$

is invariant; i.e.

$$A_i T \subset T \text{ for all } i = 1, \ldots, K.$$

Assume further that $x_0 \in T$.
Then $(X_N)_{N\in\mathbf{Z}}$ is a two-block-factor of an i.i.d. sequence.

Proof. Let $A_{i_0}^v = \left(a_{i,j}^{i_0}\right)_{i,j=1}^N$ be the matrix of A_{i_0} with respect to $\{v_1, \ldots, v_N\}$. Because $A_{i_0} T \subset T$, we have $A_{i_0} v_j = \sum_{i=1}^N a_{ij}^{i_0} v_i \in T$ (for all i_0, j). This implies that

$$a_{i,j}^{i_0} \geq 0 \text{ for all } i_0, i, j.$$

Let S be the matrix of $\{v_1, \ldots, v_N\}$ with respect to the standard basis $\{x_0 = e_1, \ldots, e_N\}$, so

$$S = (v_{ij})_{i,j=1}^N; \text{ i.e. } v_j = \sum_{i=1}^N v_{ij} e_i \; \forall j.$$

Let $R = S^{-1}$ be the matrix of coordinates of the standard basis with respect to $\{v_1, \ldots, v_N\}$, so

$$R = (t_{ij})_{i,j=1}^{N} \ ; \ \ e_j = \sum_{i=1}^{N} t_{ij} v_i \ \forall j.$$

Because $e_1 = x_0 = \sum\limits_{i=1}^{N} t_{i1} v_i \in T$ we have

$$t_{i1} \geq 0 \text{ for all } i.$$

Because $< v_i; x_0 >> 0$ we can assume by multiplying the v_i that

$$v_{1i} = < v_i; x_0 > = 1 \text{ for all } i.$$

We have for all j

$$(A_1 + \ldots + A_K) v_j = \sum_{i_0=1}^{K} \sum_{i=1}^{N} a_{ij}^{i_0} v_i$$

and $(A_1 + \ldots + A_K) v_j = < v_j; x_0 > x_0 = x_0$.
This implies that for all j:

$$1 = < x_0; x_0 > = \sum_{i_0=1}^{K} \sum_{i=1}^{N} a_{ij}^{i_0} < v_i; x_0 >= \sum_{i_0=1}^{K} \sum_{i=1}^{N} a_{ij}^{i_0}.$$

Because apparently

$$x_0 = \sum_{i_0=1}^{K} \sum_{i=1}^{N} a_{ij}^{i_0} v_i = \sum_{i=1}^{N} t_{i1} v_i$$

we have that $t_{i1} = \sum\limits_{i_0=1}^{K} a_{ij}^{i_0}$ for all j and i (we make the crucial observation that this sum is independent of j).
Because A_{i_0} has matrix representation $SA_{i_0}^{v} R$ with respect to the standardbasis, we have

$$P\left[X_1 = i_1, \ldots, X_m = i_m\right] =$$

$$= < SA_{i_1}^{v} R \ldots SA_{i_m}^{v} R e_1; e_1 > =$$

$$= < SA_{i_1}^{v} \ldots A_{i_m}^{v} R e_1; e_1 > =$$

$$= < A_{i_1}^{v} \ldots A_{i_m}^{v} R e_1; S^* e_1 > =$$

$$= < A^v_{i_1} \dots A^v_{i_m} \begin{pmatrix} a^1_{11} + \dots + a^K_{11} \\ \vdots \\ a^1_{N1} + \dots + a^K_{N1} \end{pmatrix} ; \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} > .$$

By induction (on m) it is easy to show (just as in the proof of Theorem 4.3) that this corresponds with the two-block-factor shown in Figure 2 (where N is replaced by d). □ □

Remark. In section 3 we described the HSR of a class of counter-examples of one-dependent processes that are not two-block-factors. Their Hilbert space is 3-dimensional. Theorem 4.3 states that a 2-dimensional HSR is always a two-block-factor. From Theorem 4.4 follows that the crucial difference between 2 and 3 dimensions is apparently the geometry of cones. A closed convex cone in 2 dimensions is spanned by the convex hull of 2 linearly independent vectors. In 3 dimensions closed convex cones exist that are not spanned by the convex hull of 3 vectors, but of more than 3 vectors (a finite or even infinite number). Note that these vectors are the extreme points of a convex set. It seems that the difference between two-block-factors and non-two-block-factors is determined by the geometry of the invariant cone. We generalize Theorem 4.3(a) by showing that a one-dependent process is an i.i.d.sequence if the operators commute.

Theorem 4.5 *Let $(H, x, y, A_1, \dots, A_K)$ be a HSR of a one-dependent process $(X_N)_{N \in \mathbb{Z}}$. If the operators $A_1, \dots, A_K$ commute (i.e. $A_iA_j = A_jA_i$ for all i, j), then $(X_N)_{N\in\mathbb{Z}}$ is an i.i.d. sequence.*

Proof. We have

$$P[X_1 = i_1, \dots, X_N = i_N, X_{N+1} = j_1, \dots, X_{N+m} = j_m] =$$

$$= \sum_{i=1}^{K} P[X_1 = i_1, \dots, X_N = i_N, X_{N+1} = j_1, \dots, X_{N+m} = j_m, X_{N+m+1} = i] =$$

$$= \sum_{i=1}^{K} < A_{i_1} \dots A_{i_N} A_{j_1} \dots A_{j_m} A_i x; x > =$$

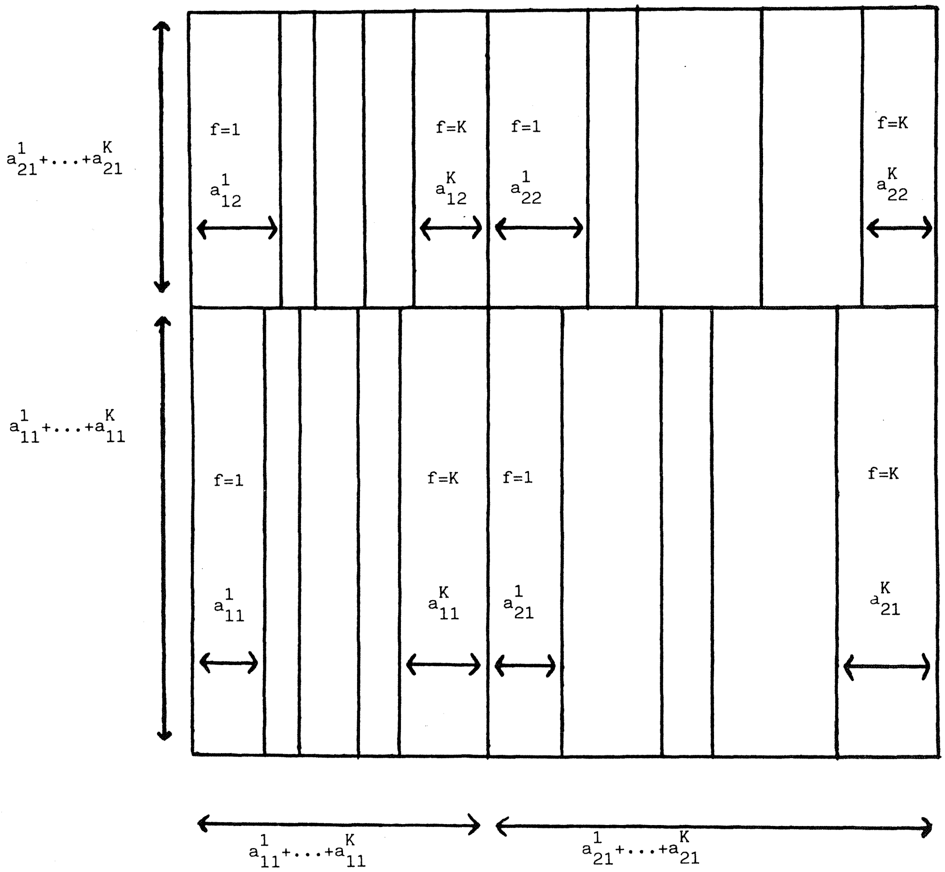
$$= \sum_{i=1}^{K} < A_{i_1} \dots A_{i_N} A_i A_{j_1} \dots A_{j_m} x; x > =$$

$$= \sum_{i=1}^{K} P[X_1 = i_1, \dots, X_N = i_N, X_{N+1} = i, X_{N+2} = j_1, \dots, X_{N+m+1} = j_m] =$$

$$= P[X_1 = i_1, \dots, X_N = i_N] \cdot P[X_{N+2} = j_1, \dots, X_{N+m+1} = j_m]$$

and the theorem follows. □ □

Remark. We conclude from Theorem 4.5 that an exchangeable one-dependent process is an i.i.d. sequence.



$a^1_{21}+...+a^K_{21}$
f=1
a^1_{12}
f=K
a^K_{12}
f=1
a^1_{22}
f=K
a^K_{22}
$a^1_{11}+...+a^K_{11}$
f=1
a^1_{11}
f=K
a^K_{11}
f=1
a^1_{21}
f=K
a^K_{21}
$a^1_{11}+...+a^K_{11}$
$a^1_{21}+...+a^K_{21}$

Figure 1.

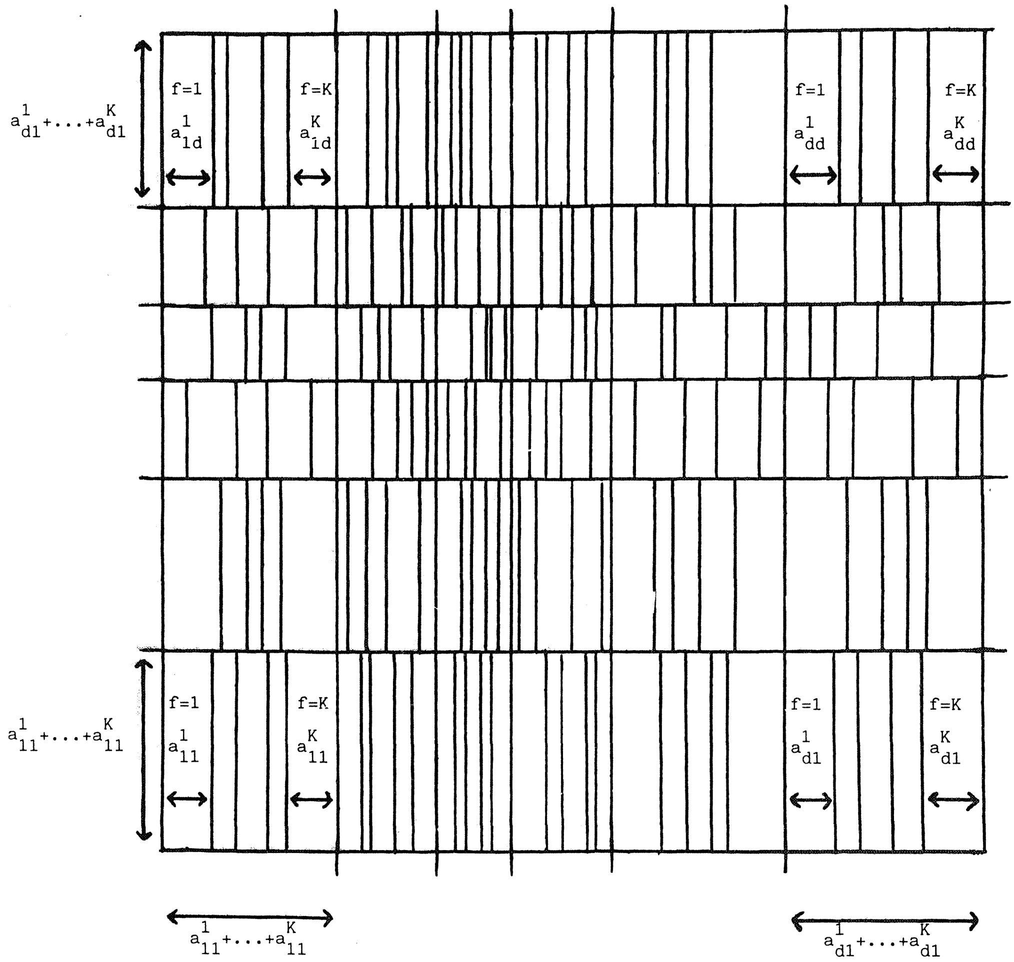

f=1
a^1_{1d}
f=K
a^K_{1d}
f=1
a^1_{dd}
f=K
a^K_{dd}
$a^1_{d1}+\dots+a^K_{d1}$
f=1
a^1_{11}
f=K
a^K_{11}
f=1
a^1_{d1}
f=K
a^K_{d1}
$a^1_{11}+\dots+a^K_{11}$
$a^1_{11}+\dots+a^K_{11}$
$a^1_{d1}+\dots+a^K_{d1}$

Figure 2.

5 Minimal Zero-Cylinders and Minimal Dimension.

Let $(X_N)_{N\in Z}$ be a one-dependent process over $\{1,\ldots,K\}^{\mathbf{Z}}$. We call the cylinder $[i_1,\ldots,i_N]$ a *minimal zero-cylinder* if $P[X_1=i_1,\ldots,X_N=i_N]=0$ and if $P[X_1=j_1,\ldots,X_m=j_m]>0$ for all $m<N$ and for all $j_1,\ldots,j_m\in\{1,\ldots,K\}$. We call N the length of the minimal zero-cylinder.
Let $(H,x,y,A_1,\ldots,A_K)$ be a HSR of a one-dependent process $(X_N)_{N\in\mathbf{Z}}$.
We call $\dim(H)$ the *minimal dimension* of $(X_N)_{N\in\mathbf{Z}}$ if for all HSR $(H',x',y',A_1',\ldots,A_K')$ of $(X_N)_{N\in\mathbf{Z}}$ holds $\dim(H')\geq\dim(H)$.
We show for two-valued one-dependent processes that if there is a zero-cylinder then the length of the minimal zero-cylinder is greater than or equal to the minimal dimension.

Theorem 5.1 *Let $(X_N)_{N\in\mathbf{Z}}$ be a one-dependent process over $\{0,1\}^{\mathbf{Z}}$. Assume that $[\varepsilon_1\ldots\varepsilon_{N_0}]$ is a minimal zero-cylinder with length N_0. Then there exists a HSR (H,x,x,A_0,A_1) of $(X_N)_{N\in\mathbf{Z}}$ with $\dim(H)\leq N_0$.*

Proof. Let (H,x,x,A_0,A_1) be a HSR of $(X_N)_{N\in\mathbf{Z}}$.
Just as in the proof of Theorem 4.3 we define the invariant cone T;

$$T:=\overline{co}\left\{\alpha A_{i_1}\ldots A_{i_N}x;\alpha\geq 0,N\in\mathbf{N},i_1,\ldots,i_N\in\{0,1\}\right\}.$$

Because $[\varepsilon_1\ldots\varepsilon_{N_0}]$ is a zero-cylinder we have

$$<A_{\varepsilon_1}\ldots A_{\varepsilon_{N_0}}x;x>=0$$

and

$$<A_{\varepsilon_1}\ldots A_{\varepsilon_{N_0}}A_{i_1}\ldots A_{i_N}x;x>=0$$

for all $N\in\mathbf{N}$ and all $i_1,\ldots,i_N\in\{0,1\}$.
Lemma 4.2 implies that we may assume that if $v\in T,<v;x>=0$ then $v=0$ (by passing over to subspace of H).
So we conclude that

$$(A_{\varepsilon_1}\ldots A_{\varepsilon_{N_0}})x=0$$

and

$$(A_{\varepsilon_1}\ldots A_{\varepsilon_{N_0}})(A_{i_1}\ldots A_{i_N}x)=0$$

for all $N\in\mathbf{N}$ and for all $i_1,\ldots,i_N\in\{0,1\}$.
This implies that

$$(A_{\varepsilon_1}\ldots A_{\varepsilon_{N_0}})(T)=\{0\}.$$

We can assume that $H = \overline{sp}\ (T)$, because apparently $\overline{sp}\ (T)$ is a closed subspace that is invariant under A_0 and A_1. By restricting A_0 and A_1 to $\overline{sp}\ (T)$ we have apparently a HSR of $(X_N)_{N\in\mathbf{Z}}$.
Therefore we can assume that

$$(A_{\varepsilon_1} \dots A_{\varepsilon_{N_0}})(H) = \{0\},$$

so $A_{\varepsilon_1} \dots A_{\varepsilon_{N_0}} = 0$.
Now we need a lemma: □

Lemma. *Let (H, x, x, A_0, A_1) be a HSR of a one-dependent process. Let $(\varepsilon_1, \dots, \varepsilon_N) \in \{0,1\}^N$, for some $n \in \mathbf{N}$. Then*

$$(*) \qquad A_0^N x \in \overline{sp}\ \{x, A_0 x, A_0^2 x, \dots, A_0^{N-1} x, A_{\varepsilon_1} \dots A_{\varepsilon_N} x\}.$$

Proof of the Lemma.
We prove this by induction on N.
For $N = 1$ we have to prove that

$$A_0 x \in \overline{sp}\ \{x, A_{\varepsilon_1} x\},$$

which is trivial if $\varepsilon_1 = 0$. If $\varepsilon_1 = 1$, we see that $A_0 x = x - A_1 x$.
Assume that (*) holds for N. We will prove (*) for $N + 1$.
If $\varepsilon_1 = 0$, then

$$\begin{array}{rcl} A_0^{N+1} x & \in & A_0(\overline{sp}\ \{x, A_0 x, \dots, A_0^{N-1} x, A_{\varepsilon_2} \dots A_{\varepsilon_{N+1}} x\}) = \\ & = & \overline{sp}\ \{A_0 x, A_0^2 x, \dots, A_0^N x, A_0 A_{\varepsilon_2} \dots A_{\varepsilon_{N+1}} x\} \subset \\ & \subset & \overline{sp}\{x, A_0 x, \dots, A_0^N x, A_{\varepsilon_1} \dots A_{\varepsilon_{N+1}} x\}. \end{array}$$

If $\varepsilon_1 = 1$, then

$$\begin{array}{rcl} A_0^{N+1} x & \in & \overline{sp}\ \{A_0 x, A_0^2 x, \dots, A_0^N x, A_0 A_{\varepsilon_2} \dots A_{\varepsilon_{N+1}} x\} = \\ & = & \overline{sp}\ \{A_0 x, \dots, A_0^N x, < A_{\varepsilon_2} \dots A_{\varepsilon_{N+1}} x; x > x - A_{\varepsilon_1} A_{\varepsilon_2} \dots A_{\varepsilon_{N+1}} x\} \subset \\ & \subset & \overline{sp}\ \{x, A_0 x, \dots, A_0^N x, A_{\varepsilon_1} \dots A_{\varepsilon_{N+1}} x\} \end{array}$$

This proves the Lemma. □
We define

$$H_0 := \overline{sp}\ \{x, A_0 x, A_0^2 x, \dots, A_0^{N_0 - 1} x\}.$$

The above Lemma implies that

$$A_0^{N_0} x \in H_0.$$

Apparently H_0 is invariant under A_0. Because $A_1 v = < v; x > x - A_0 v$ it follows that H_0 is also invariant under A_1.
By restricting A_0 and A_1 to H_0 we have a HSR of $(X_N)_{N\in\mathbf{Z}}$. Because

$$\dim(H_0) \leq N_0,$$

Theorem 5.1 follows. □

Remark. Analogously to the Lemma we have

$$A_1^N x \in \overline{sp}\,\{x, A_1x, A_1^2x, \ldots, A_1^{N-1}x, A_{\varepsilon_1}\ldots A_{\varepsilon_N}x\}$$

and we obtain that

$$H_1 := \overline{sp}\,\{x, A_1x, A_1^2x, \ldots, A_1^{N_0-1}x\}$$

is invariant under A_0 and A_1.
Therefore we can assume that

$$H = H_0 = H_1.$$

Examples. We know that the following minimal zero-cylinders of two-block-factors exist: [101], [010], $[1^N] := \underbrace{[11\ldots1]}_{N\text{ times}}$, $[0^N] := \underbrace{[00\ldots0]}_{N\text{ times}}$.
Let $A^{(\alpha)}_{[101]} := [0, \sqrt{\alpha}\,] \times [0, \sqrt{\alpha}\,]$ and

$$A^{(s)}_{[1^{N+1}]} := \bigcup_{i=1}^{N-1} ([is, (i+1)s] \times [0, is]) \cup ([Ns, 1] \times [0, Ns])$$

for $\frac{1}{N+1} \leq s < \frac{1}{N}$.
By taking f equal to the indicator-function of these sets, it is easy to check that $X_N := f(Y_N, Y_{N+1})$ defines a two-block-factor (of the i.i.d. sequence $(Y_N)_{N\in\mathbf{Z}}$, each Y_N uniformly distributed over the unit inverval) with the corresponding minimal zero-cylinders [101] and $[1^N]$. By replacing f by $1-f$ we obtain [010] and $[0^N]$ as minimal zero-cylinder.
In [V.1] is proved that in these two-block-factors the two-correlation is extremal; i.e. the probability of a run of two ones is extremal over the class of two-block-factors with fixed probability of a one.
In [G.K.V.] it is shown that the maximal two-correlation over the class of one-dependent processes is uniquely attained in two-block-factors corresponding with the set $A^{(\alpha)}_{[101]}$ and its complement in the unit square.
If α is the fixed probability of a one and $0 \leq \alpha \leq 1/4$ or $1/4 < \alpha \leq 1/3$ then the minimal two-correlation over the class of one-dependent processes is attained in the two-block-factors corresponding with $A^{(\alpha)}_{[11]}$ and $A^{(\alpha)}_{[111]}$.
The counterexamples in [A.G.K.V.] of one-dependent processes that are not two-block-factors have minimal zero-cylinder [111].
We show that minimal zero-cylinders with length ≤ 7 are symmetric (in the $0-1$ valued case).

Theorem 5.2 *Let $(X_N)_{N\in\mathbf{Z}}$ be a one-dependent process over $\{0,1\}^{\mathbf{Z}}$. Assume that $[\varepsilon_1 \ldots \varepsilon_{N_0}]$ is a minimal zero-cylinder with length N_0.*

1. If $N_0 \geq 2$ then $\varepsilon_1 = \varepsilon_{N_0}$

2. If $N_0 \geq 4$ then $\varepsilon_2 = \varepsilon_{N_0-1}$

3. If $N_0 \geq 6$ then $\varepsilon_3 = \varepsilon_{N_0-2}$

4. If $N_0 \leq 7$ then $\varepsilon_i = \varepsilon_{N_0+1-i}$ for all i

Proof. Assume that $N_0 \geq 2$ and that $\varepsilon_1 \neq \varepsilon_{N_0}$. Then

$$P[X_1 = \varepsilon_1, X_2 = \varepsilon_2, \ldots, X_{N_0-1} = \varepsilon_{N_0-1}]\, P[X_{N_0+1} = \varepsilon_2, \ldots, X_{2N_0-1} = \varepsilon_{N_0}] =$$

$$= P[X_1 = \varepsilon_1, X_2 = \varepsilon_2, \ldots, X_{N_0-1} = \varepsilon_{N_0-1}, X_{N_0} = \varepsilon_{N_0}, X_{N_0+1} = \varepsilon_2, \ldots, X_{2N_0-1} = \varepsilon_{N_0}]$$

$$+P[X_1 = \varepsilon_1, X_2 = \varepsilon_2, \ldots, X_{N_0-1} = \varepsilon_{N_0-1}, X_{N_0} = \varepsilon_1, X_{N_0+1} = \varepsilon_2, \ldots, X_{2N_0-1} = \varepsilon_{N_0}]$$

$$= 0 + 0 = 0$$

which implies that there is a zero-cylinder with length $N_0 - 1$. Contradiction. This proves (1).
Assume that $N_0 \geq 4$. We have $\varepsilon_1 = \varepsilon_{N_0}$. Assume that $\varepsilon_2 \neq \varepsilon_{N_0-1}$, then if $\varepsilon_1 \neq \varepsilon_2$ we have

$$P[X_1 = \varepsilon_1, X_2 = \varepsilon_2, \ldots, X_{N_0-1} = \varepsilon_{N_0-1}]\, P[X_{N_0+1} = \varepsilon_3, \ldots, X_{2N_0-2} = \varepsilon_{N_0}] =$$

$$= P[X_1 = \varepsilon_1, X_2 = \varepsilon_2, \ldots, X_{N_0-1} = \varepsilon_{N_0-1}, X_{N_0} = \varepsilon_{N_0}, X_{N_0+1} = \varepsilon_3, \ldots, X_{2N_0-2} = \varepsilon_{N_0}]$$

$$+P[X_1 = \varepsilon_1, X_2 = \varepsilon_2, \ldots, X_{N_0-1} = \varepsilon_{N_0-1}, X_{N_0} = 1 - \varepsilon_{N_0}, X_{N_0+1} = \varepsilon_3, \ldots, X_{2N_0-2} = \varepsilon_{N_0}]$$

$$= 0 + P[X_1 = \varepsilon_1, X_2 = \varepsilon_2, \ldots, X_{N_0-1} = \varepsilon_1, X_{N_0} = \varepsilon_2, X_{N_0+1} = \varepsilon_3, \ldots, X_{2N_0-2} = \varepsilon_{N_0}]$$

$$= 0 + 0 = 0$$

which implies that there is a zero-cylinder with length $N_0 - 1$ or $N_0 - 2$. If $\varepsilon_1 = \varepsilon_2$ then

$$P[X_1 = \varepsilon_1, X_2 = \varepsilon_2, \ldots, X_{N_0-2} = \varepsilon_{N_0-2}]\, P[X_{N_0} = \varepsilon_2, \ldots, X_{2N_0-2} = \varepsilon_{N_0}] =$$

$$= P[X_1 = \varepsilon_1, X_2 = \varepsilon_2, \ldots, X_{N_0-2} = \varepsilon_{N_0-2}, X_{N_0-1} = \varepsilon_{N_0-1}, X_{N_0} = \varepsilon_2, \ldots, X_{2N_0-2} = \varepsilon_{N_0}]$$

$$+P[X_1 = \varepsilon_1, X_2 = \varepsilon_2, \ldots, X_{N_0-2} = \varepsilon_{N_0-2}, X_{N_0-1} = 1 - \varepsilon_{N_0-1}, X_{N_0} = \varepsilon_2, \ldots,$$

$$X_{2N_0-2} = \varepsilon_{N_0}]$$

$$= P[X_1 = \varepsilon_1, X_2 = \varepsilon_2, \ldots, X_{N_0-2} = \varepsilon_{N_0-2}, X_{N_0-1} = \varepsilon_{N_0-1}, X_{N_0} = \varepsilon_{N_0}, \ldots, X_{2N_0-2} = \varepsilon_{N_0}]$$

$$+P[X_1 = \varepsilon_1, X_2 = \varepsilon_2, \ldots, X_{N_0-2} = \varepsilon_{N_0-2}, X_{N_0-1} = \varepsilon_1, X_{N_0} = \varepsilon_2, \ldots, X_{2N_0-2} = \varepsilon_{N_0}]$$

$$= 0 + 0 = 0$$

which implies that there is a zero-cylinder with length $N_0 - 1$ or $N_0 - 2$. Contradiction. This proves (2).
Assume $N_0 \geq 6$. We have $\varepsilon_1 = \varepsilon_{N_0}$ and $\varepsilon_2 = \varepsilon_{N_0-1}$. If $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 \neq \varepsilon_{N_0-2}$ then we have

$$
\begin{aligned}
&P[X_1 = \varepsilon_1, \ldots, X_{N_0-3} = \varepsilon_{N_0-3}]\, P[X_{N_0-1} = \varepsilon_2, \ldots, X_{2N_0-3} = \varepsilon_{N_0}] = \\
&= P[X_1 = \varepsilon_1, \ldots, X_{N_0-3} = \varepsilon_{N_0-3}, X_{N_0-2} = \varepsilon_{N_0-2}, X_{N_0-1} = \varepsilon_2, \ldots, X_{2N_0-3} = \varepsilon_{N_0}] \\
&+ P[X_1 = \varepsilon_1, \ldots, X_{N_0-3} = \varepsilon_{N_0-3}, X_{N_0-2} = \varepsilon_1, X_{N_0-1} = \varepsilon_2, \ldots, X_{2N_0-3} = \varepsilon_{N_0}] \\
&= 0 + 0 = 0,
\end{aligned}
$$

which implies that there is a zero-cylinder with length $N_0 - 3$ or $N_0 - 1$. The other cases of (3) go analogously and are left as an exercise to the reader. (4) follows immediately from (1), (2) and (3). □ □

We conjecture that all minimal zero-cylinders $[\varepsilon_1 \ldots \varepsilon_{N_0}]$ are symmetric; i.e. $\varepsilon_i = \varepsilon_{N_0+1-i}$ for all i. However not all symmetric cylinders appear as minimal zero-cylinder.
First we show, without using any HSR-techniques, that [1001] can not be the minimal zero-cylinder of a $0-1$ valued two-block-factor.
Then we will show that a $0-1$ valued one-dependent process with [1001] as minimal zero-cylinder has a 2-dimensional Hilbert Space Representation. But then by Theorem 4.3 this process should be a two-block-factor, and we conclude that [1001] can not be a minimal zero-cylinder of a $0-1$ valued one-dependent process.

Theorem 5.3 *Let $(X_N)_{N\in\mathbf{Z}}$ be a $0-1$ valued two-block-factor of an i.i.d. sequence $(Y_N)_{N\in\mathbf{Z}}$. Then* [1001] *can not be a minimal zero-cylinder of $(X_N)_{N\in\mathbf{Z}}$.*

Proof. Let $(Y_N)_{N\in\mathbf{Z}}$ be an i.i.d. sequence of random variables that are uniformly distributed over the unit interval.
Let $A \subset [0,1] \times [0,1]$ be a Lebesgue-measurable set such that

$$X_N = I_A(Y_N, Y_{N+1}) \qquad (N \in \mathbf{Z}).$$

We define the functions V and H on the unit interval

$$V(x) := P[(x, Y_1) \in A] \quad , x \in [0,1]$$

$$H(y) := P[(Y_1, y) \in A] \quad , y \in [0,1]$$

and the sets B_1 and B_2:

$$B_1 := \{x \in [0,1] : V(x) = 0\}$$

$$B_2 := \{y \in [0,1] : H(y) = 0\}.$$

Finally we define the sets $L(y)$.

$$L(y) := \{z \in [0,1] : (y,z) \in A\}\ , y \in [0,1].$$

Let λ and μ be Lebesgue-measure on the unit interval and the unit square respectively.
We will identify sets V and W if $\lambda(V \triangle W) = 0$ and we will identify functions that are equal almost everywhere.
Assume that $\lambda(B_1 \cap B_2) > 0$, then

$$0 = [1001] = P\left[(Y_0, Y_1) \in A, (Y_1, Y_2) \notin A, (Y_2, Y_3) \notin A, (Y_3, Y_4) \in A\right] \geq$$

$$\geq \int_{B_1^c} \int_{L(Y_0)} \int_{B_1 \cap B_2} \int_{B_1^c} \int_{L(Y_3)} I_A(Y_0, Y_1) \cdot I_{A^c}(Y_1, Y_2) \cdot I_{A^c}(Y_2, Y_3) \cdot$$

$$I_A(Y_3, Y_4) dY_4 dY_3 dY_2 dY_1 dY_0.$$

Consider the fact that $\lambda(B_1^c) > 0$ (because if $\lambda(B_1^c) = 0$, then $\lambda(B_1) = 1 \Rightarrow B_1 = [0,1] = \{y : P[(y, Y_4) \in A] = 0\} \Rightarrow \mu(A) = 0$, contradiction).
Further; if $Y_0 \in B_1^c = \{y : P[(y, Y_1) \in A] > 0\}$ then for $L(Y_0) = \{z : (Y_0, z) \in A\}$ we have $\lambda(L(Y_0)) > 0$.
Further $\lambda(B_1 \cap B_2) > 0$ (as we assumed) and $\lambda(B_1^c) > 0$, and $\lambda(L(Y_3)) > 0$ (because $Y_3 \in B_1^c$).
So, we integrate over a set of strictly positive measure.
Further $I_A(Y_0, Y_1) = 1$ because $Y_1 \in L(Y_0)$;
$I_{A^c}(Y_1, Y_2) = 1$ because $Y_2 \in B_1 \cap B_2 \subset B_2$;
$I_{A^c}(Y_2, Y_3) = 1$ because $Y_2 \in B_1 \cap B_2 \subset B_1$;
$I_A(Y_3, Y_4) = 1$ because $Y_4 \in L(Y_3)$.
This implies that the integrand $\equiv 1 \Rightarrow$

$$0 = [1001] \geq \int_{B_1^c} \int_{L(Y_0)} \int_{B_1 \cap B_2} \int_{B_1^c} \int_{L(Y_3)} dY_4 dY_3 dY_2 dY_1 dY_0 > 0$$

Contradiction, so we conclude that

$$\lambda(B_1 \cap B_2) = 0.$$

Now we have

$0 = [1001] \geq$

$$\geq \int_{B_1^c} \int_{L(Y_0)} \int_{B_2} \int_{B_2} \int_{L(Y_3)} I_A(Y_0, Y_1) I_{A^c}(Y_1, Y_2) I_{A^c}(Y_2, Y_3) I_A(Y_3, Y_4) dY_4 dY_3 dY_2 dY_1 dY_0$$

Once more we have

$$\lambda(B_1^c) > 0,$$

$$\lambda(L(Y_0)) > 0 \text{ for } Y_0 \in B_1^c,$$

$$\lambda(B_2) > 0\ ,\ \lambda(B_2) > 0,$$

(If $\lambda(B_2) = 0$ then $H(Y_1) > 0$ for all Y_1. Because $[001] > 0$ this would imply that

$$[1001] = \int\int\int\int H(Y_1)I_{A^c}(Y_1,Y_2)I_{A^c}(Y_2,Y_3)I_A(Y_3,Y_4)dY_4dY_3dY_2dY_1 > 0.$$

Contradiction.)

$$\lambda(L(Y_3)) > 0 \text{ for } Y_3 \in B_2 \subset B_1^c,$$

so, we integrate over a set of strictly positive measure, and

$$I_A(Y_0,Y_1) = 1 \text{ because } Y_1 \in L(Y_0),$$

$$I_{A^c}(Y_1,Y_2) = 1 \text{ because } Y_2 \in B_2,$$

$$I_{A^c}(Y_2,Y_3) = 1 \text{ because } Y_3 \in B_2,$$

$$I_A(Y_3,Y_4) = 1 \text{ because } Y_4 \in L(Y_3),$$

$$\Rightarrow 0 = [1001] \geq \int_{B_1^c}\int_{L(Y_0)}\int_{B_2}\int_{B_2}\int_{L(Y_3)} dY_4dY_3dY_2dY_1dY_0 > 0$$

Contradiction. This proves Theorem 5.3. □ □

Theorem 5.4 *Let $(X_N)_{N\in\mathbf{Z}}$ be a one-dependent process over $\{0,1\}^{\mathbf{Z}}$. Then $[1001]$ can not be a minimal zero-cylinder of $(X_N)_{N\in\mathbf{Z}}$.*

Proof. Let (H,x,x,A_0,A_1) be a HSR of $(X_N)_{N\in\mathbf{Z}}$.
The proof of Theorem 5.1 (last remark before the Lemma) shows that we can assume that $A_1A_0A_0A_1 = 0$. The Remark after Theorem 5.1 shows that we can assume that

$$\begin{aligned} H &= \overline{sp}\,\{x, A_0x, A_0^2x, A_0^3x\} = \\ &= \overline{sp}\,\{x, A_1x, A_1^2x, A_1^3x\}. \end{aligned}$$

We have $A_0A_1v = A_0 < v;x > x - A_0^2v$;

$$\begin{aligned} A_0A_0A_1v &= < v;x > A_0^2x - A_0^3v; \\ A_1A_0A_0A_1v &= < v;x >< A_0^2x;x > x- < v;x > A_0^3x- < A_0^3v;x > x + A_0^4v = 0. \end{aligned}$$

This implies that (for all $v \in H$)

$$A_0^4 v = < A_0^3 v; x > x - < v; x >< A_0^2 x; x > x + < v; x > A_0^3 x.$$

So

$$A_0^4 H \subset \overline{sp}\ \{x, A_0^3 x\}.$$

We have

$$A_0^4 x = (< A_0^3 x; x > - < A_0^2 x; x >)x + A_0^3 x =$$

$$= (P\,[X_1 = X_2 = X_3 = 0] - P\,[X_1 = X_2 = 0])x + A_0^3 x =$$

$$= -P\,[X_1 = X_2 = 0, X_3 = 1]\,x + A_0^3 x$$

Because [1001] is minimal zero-cylinder $P\,[X_1 = X_2 = 0, X_3 = 1] \neq 0$. This implies that

$$x = \frac{A_0^3 x - A_0^4 x}{P\,[X_1 = X_2 = 0, X_3 = 1]} \in \overline{sp}\{A_0^3 x, A_0^4 x\}.$$

So we have

$$\begin{aligned} H &= \overline{sp}\ \{x, A_0 x, A_0^2 x, A_0^3 x\} \subset \\ &\subset \overline{sp}\ \{A_0 x, A_0^2 x, A_0^3 x, A_0^4 x\} = \\ &= A_0 \overline{sp}\ \{x, A_0 x, A_0^2 x, A_0^3 x\} = \\ &= A_0(H). \end{aligned}$$

So $A_0(H) = H$ and $A_0^4(H) = H$.
But $A_0^4 H \subset \overline{sp}\ \{x, A_0^3 x\}$ and we conclude that $\dim(H) \leq 2$. □

In Theorem 4.3 we proved that $(X_N)_{N \in \mathbf{Z}}$ is a two-block-factor of an i.i.d. sequence if $(X_N)_{N \in \mathbf{Z}}$ has a HSR with dimension ≤ 2.
Now Theorem 5.3 finishes the proof. □

6 Generalization to m-dependent processes.

In this section we generalize the concept of Hilbert space representation to m-dependent processes. We will prove that all m-dependent processes admit a representation.

Let H be a real Hilbert space, let $K \geq 2$, $m \geq 2$ be integers, let $A_1, \ldots, A_K : H \to H$ be linear, continuous operators, let $x, y \in H$ be two fixed vectors with $< x; y > = 1$.
We assume that

$$(A_1 + \ldots + A_K)^m h = < h; x > \; y \tag{5}$$

for all $h \in H$, so $(A_1 + \ldots + A_K)^m$ has rank one.
Further we assume that

$$< A_{i_1} \ldots A_{i_N} y; x > \geq 0 \tag{6}$$

for all $N \geq 0$, $i_1, \ldots, i_N \in \{1, \ldots, K\}$.
Also we assume that

$$(A_1 + \ldots + A_K) y = y \tag{7}$$

and

$$(A_1^* + \ldots + A_K^*) x = x. \tag{8}$$

Under this conditions we claim that

$$P[X_1 = i_1, \ldots, X_N = i_N] := < A_{i_1} \ldots A_{i_N} y; x >$$

defines an m-dependent probability measure on $\{1, \ldots, K\}^{\mathbf{Z}}$. We call $(H, x, y, A_1, \ldots, A_K)$ the Hilbert space representation of the m-dependent process $(X_N)_{N \in \mathbf{Z}}$.
We have

$$\sum_{i_N=1}^{K} P[X_1 = i_1, \ldots, X_N = i_N] =$$

$$= \sum_{i_N=1}^{K} < A_{i_1} \ldots A_{i_N} y; x > =$$

$$= < A_{i_1} \ldots A_{i_{N-1}} (A_1 + \ldots + A_K) y; x > =$$

$$= < A_{i_1} \ldots A_{i_{N-1}} y; x > = P[X_1 = i_1, \ldots, X_{N-1} = i_{N-1}]$$

and

$$\sum_{i_1=1}^{K} P[X_1 = i_1, \dots, X_N = i_N] =$$

$$= \sum_{i_1=1}^{K} < A_{i_1} \dots A_{i_N} y; x > = < (A_1 + \dots + A_K) A_{i_2} \dots A_{i_N} y; x > =$$

$$= < A_{i_2} \dots A_{i_N} y; (A_1^* + \dots + A_K^*) x > =$$

$$= < A_{i_2} \dots A_{i_N} y; x > = P[X_2 = i_2, \dots, X_N = i_N].$$

and

$$\sum_{i=1}^{K} P[X_1 = i] = < (A_1 + \dots + A_K) y; x > = < y; x >= 1$$

and

$$\sum_{j_1,\dots,j_m=1}^{K} P[X_1 = i_1, \dots, X_N = i_N, X_{N+1} = j_1, \dots, X_{N+m} = j_m, X_{N+m+1} = i_{N+m+1}, \dots,$$

$$X_{N+m+t} = i_{N+m+t}] =$$

$$= < A_{i_1} \dots A_{i_N} (A_1 + \dots + A_K)^m A_{i_{N+m+1}} \dots A_{i_{N+m+t}} y; x > =$$

$$= < A_{i_1} \dots A_{i_N} < A_{i_{N+m+1}} \dots A_{i_{N+m+t}} y; x > y; x > =$$

$$= < A_{i_1} \dots A_{i_N} y; x > < A_{i_{N+m+1}} \dots A_{i_{N+m+t}} y; x > =$$

$$= P[X_1 = i_1, \dots, X_N = i_N] \cdot P[X_{N+m+1} = i_{N+m+1}, \dots, X_{N+m+t} = i_{N+m+t}]$$

for all $N, t \in \mathbf{N}$ and all $i_1, \dots, i_N, i_{N+m+1}, \dots, i_{N+m+t} \in \{1, \dots, K\}$.
We conclude that $(H, x, y, A_1, \dots, A_K)$ represents an m-dependent process over $\{1, \dots, K)^{\mathbf{Z}}$.

We generalize Theorem 3.1 to m-dependent processes.

Theorem 6.1 *Let $(X_N)_{N \in \mathbf{Z}}$ be an m-dependent process over $\{1, \dots, K\}^{\mathbf{Z}}$ for some $m \geq 2 (m \in \mathbf{N})$.*
Let H_0 be a Hilbert space, let $x \in H_0$ be a vector with $< x; x >= 1$, let $A_1, \dots, A_K : H_0 \to H_0$ be linear, continuous operators such that $(A_1 + \dots + A_K) x = x$.
Assume that

$$< A_{i_1} \dots A_{i_N} x; x > = P[X_1 = i_1, \dots, X_N = i_N]$$

for all $N \in \mathbf{N}$ and all $i_1, \dots, i_N \in \{1, \dots, K\}$.
Then there exists a closed subspace $H \subset H_0$ with $x \in H$, such that $(H, x, x, PA_1, \dots, PA_K)$ is a HSR of $(X_N)_{N \in \mathbf{Z}}$, where $P : H_0 \to H$ is the orthogonal projection from $H_0 \to H$.

Proof. We define the collection $\mathcal{H}$ of those closed subspaces H of H_0 with the properties that $x \in H$ and that for the orthogonal projection $P : H_0 \to H$ holds

$$< PA_{i_1} \dots PA_{i_N}x; x > = P[X_1 = i_1, \dots, X_N = i_N]$$

for all $N \in \mathbf{N}$ and all $i_1, \dots, i_N \in \{1, \dots, K\}$.
We define a partial ordering on $\mathcal{H}$ by

$$H_1 \leq H_2 \text{ if } H_1 \supset H_2.$$

Note that $\mathcal{H} \neq \emptyset$ because $H_0 \in \mathcal{H}$. □

Just as in the proof of Theorem 3.1 we claim that every totally ordered subset of $\mathcal{H}$ has an upper bound.
The proof of this claim is exactly the same as in the proof of Theorem 3.1. Now Zorn's Lemma implies the existence of a maximal element $H \in \mathcal{H}$. Let $P : H_0 \to H$ be the orthogonal projection from H_0 on H. Consider the restricted operators $PA_1 |_H, \dots, PA_K |_H : H \to H$. Let $B_i : H \to H$ be the adjoints of these operators $(i = 1, \dots, K)$. Just as in the proof of Theorem 3.1 we have that

$$H = \overline{sp}\{B_{j_1} \dots B_{j_r}x : r \geq 0, j_1, \dots, j_r \in \{1, \dots, K\}\}$$

and

$$H = \overline{sp}\{PA_{i_1} \dots PA_{i_N}x : N \geq 0, i_1, \dots, i_N \in \{1, \dots, K\}\}.$$

We claim that $(H, x, x, PA_1, \dots, PA_K)$ is a HSR of the m-dependent process $(X_N)_{N \in \mathbf{Z}}$. We have to prove that
(1) $(PA_1 + \dots + PA_K)x = x$,
(2) $(B_1 + \dots + B_K)x = x$ and
(3) $(PA_1 + \dots + PA_K)^m h = < h; x > x$ for all $h \in H$.
(1) holds trivially and (2) and (3) are equivalent to
(2') $< (B_1 + \dots + B_K)x; h > = < x; h >$ for all $h \in H$,
(3') $< (PA_1 + \dots + PA_K)^m h; g > = < h; x > < x; g >$ for all $g, h \in H$.

Just as in the proof of Theorem 3.1 it is sufficient to check (2') and (3') for $h = PA_{i_1} \dots PA_{i_N}x$ and $g = B_{j_1} \dots B_{j_r}x$. For this h and g we have:

$$< (B_1 + \dots + B_K)x; h > = < (B_1 + \dots + B_K)x; PA_{i_1} \dots PA_{i_N}x > =$$

$$= < x; (PA_1 + \dots + PA_K)PA_{i_1} \dots PA_{i_N}x > =$$

$$= \sum_{i=1}^{K} P[X_1 = i, X_2 = i_1, \dots, X_{N+1} = i_N] =$$

$$= P[X_2 = i_1, \dots, X_{N+1} = i_N] = < x; PA_{i_1} \dots PA_{i_N}x > = < x; h >$$

and

$$< (PA_1 + \ldots + PA_K)^m PA_{i_1} \ldots PA_{i_N} x; B_{j_1} \ldots B_{j_r} x > =$$

$$= < PA_{j_r} \ldots PA_{j_1} (PA_1 + \ldots + PA_K)^m PA_{i_1} \ldots PA_{i_N} x; x > =$$

$$\sum_{s_0,\ldots,s_{m-1} \in \{1,\ldots,K\}} P[X_{-r} = j_r, \ldots, X_{-1} = j_1, X_0 = s_0, \ldots, X_{m-1} = s_{m-1},$$

$$X_m = i_1, \ldots, X_{m+N-1} = i_N] =$$

$$= P[X_{-r} = j_r, \ldots, X_{-1} = j_1] \cdot P[X_m = i_1, \ldots, X_{m+N-1} = i_N] =$$

$$= < PA_{j_r} \ldots PA_{j_1} x; x > < PA_{i_1} \ldots PA_{i_N} x; x > =$$

$$= < x; B_{j_1} \ldots B_{j_r} x > < h; x > = < x; g > < h; x > .$$

So (2') and (3') hold, thus the claim is proved and the proof of Theorem 6.1 is finished. □

Now we can generalize Theorem 3.2.

Theorem 6.2 *Let $(X_N)_{N \in \mathbf{Z}}$ be a K-valued m-dependent process (for some $K, m \in \mathbf{N}$). Then there exists a HSR of $(X_N)_{N \in \mathbf{Z}}$.*

Proof. Let $(X_N)_{N \in \mathbf{Z}}$ be an m-dependent process over $\{1, \ldots, K\}^{\mathbf{Z}}$. $(X_N)_{N \in \mathbf{Z}}$ induces a probability measure P on $\{1, \ldots, K\}^{\mathbf{N}}$. We define the Hilbert space $H_0 := L^2(\{1, \ldots, K\}^{\mathbf{N}})$. Let $I \in H_0$ be the function that is identically one. We define the operators $A_1, \ldots, A_K : H_0 \to H_0$ in the same way as in the proof of Theorem 3.2. Analogously to that proof we observe that the conditions of Theorem 6.1 are fulfilled. So Theorem 6.1 implies the existence of a HSR of $(X_N)_{N \in \mathbf{Z}}$. □ □

Many theorems on one-dependent processes can be generalized to m-dependent processes. We generalize the Theorems 3.5, 3.6, 4.1, 4.3 and 4.5.

Theorem 6.3 *Let $(X_N)_{N \in \mathbf{Z}}$ be a K valued m-block-factor of an i.i.d. sequence (for some $K, m \in \mathbf{N}$). Then there exists a HSR of $(X_N)_{N \in \mathbf{Z}}$ with as Hilbert space a subspace of $L^2([0,1]^{m-1})$.*

Proof. The same as the proof of Theorem 3.5. Observe that the conditions of Theorem 6.1 are fulfilled. □ □

Theorem 6.4 *Let $(H, x, y, A_1, \ldots, A_K)$ be a HSR of an m-dependent process $(X_N)_{N \in \mathbf{Z}}$. Let $(Y_N)_{N \in \mathbf{Z}}$ be the reversed process; i.e. $Y_N := X_{-N} (N \in \mathbf{Z})$.*
Then $(H, y, x, A_1^*, \ldots, A_K^*)$ is a HSR of $(Y_N)_{N \in \mathbf{Z}}$.

Proof. This is an easy exercise for the reader. Generalize the proof of Theorem 3.6 to m-dependent processes. □ □

Theorem 6.5 *Let $(H, x, y, A_1, \ldots, A_K)$ be a HSR of an m-dependent process $(X_N)_{N\in\mathbf{Z}}$. Then there exists a vector $x_0 \in H$ and there exist operators $B_1, \ldots, B_K : H \to H$ such that $(H, x_0, x_0, B_1, \ldots, B_K)$ is a HSR of $(X_N)_{N\in\mathbf{Z}}$.*

Proof. Generalize the proof of Theorem 4.1 to m-dependent processes. We leave this as an exercise for the reader. □ □

Theorem 6.6 *Let $(H, x, y, A_1, \ldots, A_K)$ be a HSR of an m-dependent process $(X_N)_{N\in\mathbf{Z}}$.*
(a) If $\dim(H) = 1$, then $(X_N)_{N\in\mathbf{Z}}$ is an i.i.d. sequence.
(b) If $\dim(H) = 2$, then $(X_N)_{N\in\mathbf{Z}}$ is a two-block-factor of an i.i.d. sequence.

Proof. $(A_1 + \ldots + A_K)^m$ has rank one. Because $\dim(H) \leq 2$, this implies that $(A_1 + \ldots + A_K)$ has rank one. But this means that $(H, x, y, A_1, \ldots, A_K)$ is a HSR of a one-dependent process. Now we can apply Theorem 4.3. to prove Theorem 6.6. □ □

Theorem 6.7 *Let $(H, x, y, A_1, \ldots, A_K)$ be a HSR of an m-dependent process $(X_N)_{N\in\mathbf{Z}}$. If the operators $A_1, \ldots, A_K$ commute (i.e. $A_iA_j = A_jA_i$ for all i, j), then $(X_N)_{N\in\mathbf{Z}}$ is an i.i.d. sequence.*

Proof. The proof is left as an easy exercise for the reader (generalize the proof of Theorem 4.5). □

Remark. We conclude from Theorem 6.7 that exchangeable m-dependent processes are i.i.d. sequences. There are more dependence structures (such as Markov, ergodicity, mixing and renewal) that can be translated to properties of operators in Hilbert space representations, see [V.4].

7 Conjectures and open problems.

1. The essential difference between two-block-factors and one-dependent processes that are not two-block-factors is determined by the geometry of the invariant cone.
   More research is necessary to investigate this.

2. A $0 - 1$ valued one-dependent process can have no other minimal zero-cylinders than $[101], [010], [1^N]$ and $[0^N]$ $(N \in \mathbf{N})$. The minimal dimensions are $2, 2, N$ and N respectively.

3. For any $N \in \mathbf{N}$ $(N \geq 3)$ there exists a one-dependent process, that is not a two-block-factor, with minimal dimension equal to N, and without zero-cylinders.

4. For any $N \in \mathbf{N}$ $(N \geq 3)$ there exist a one-dependent process, that is not a two-block-factor, with minimal dimension equal to N, and with a minimal zero-cylinder with length N.

5. For any $N \in \mathbf{N}$ $(N \geq 1)$ there exists a two-block-factor with minimal dimension equal to N, and without zero-cylinders.

6. For any $N \in \mathbf{N}$ $(N \geq 1)$ there exists a two-block-factor with minimal dimension equal to N, and with a minimal zero-cylinder with length N.

7. Under which conditions is a one-dependent Markov process necessarily a two-block-factor?

8. Are one-dependent processes always functions of Markov processes, or even functions of one-dependent Markov processes?

9. Do there exist one-dependent m-block-factors $(m \geq 3)$ that can not be written as a two-block-factor?

10. Is a one-dependent process with an m-dimensional HSR $(m \geq 3)$ always an m-block-factor?

11. Do there exist two-dependent processes that are not two-block-factors of one-dependent processes?

Acknowledgements.
The author thanks Mike Keane of the Delft University of Technology, David Gilat and Jon Aaronson of the Tel Aviv University, František Matúš of the Czechoslovak Academy of Sciences and Svante Janson of the Uppsala University for many

inspiring discussions.
Further the author thanks the Netherlands Organization for Scientific Research (NWO) for the stipendium to work a year at Tel Aviv University (Israel) and Oregon State University (U.S.A.), and he thanks these universities for their hospitality.
Finally the author thanks the University of Groningen for being a good environment for mathematical research.

References

[A.G.K.V.] J. Aaronson, D. Gilat, M. Keane and V. de Valk, *An algebraic construction of a class of one-dependent processes*, Ann. Probab. 17 (1989), 128-143.

[Be.] J. van den Berg, *On some results by S. Janson concerning runs in m-dependent processes*, preprint 1986.

[Di.1] P.H. Diananda, *Some probability limit theorems with statistical applications*, Proc. Cambridge Phil. Soc. 49 (1953), 239-246.

[Di.2] P.H. Diananda, *The central limit theorem for m-dependent variables asymptotically stationary to the second order*, Proc. Cambridge Phil. Soc. 50 (1954), 287-292.

[Di.3] P.H. Diananda, *The central limit theorem for m-dependent variables*, Proc. Cambridge Phil. Soc. 51 (1955), 92-95.

[G.K.V.] A. Gandolfi, M. Keane and V. de Valk, *Extremal two-correlations of two-valued stationary one-dependent processes*, Probab. Theory Related Fields 80 (1989), 475-480.

[G.H.1] F. Götze and C. Hipp, *Asymptotic expansions for potential functions of i.i.d. random fields*, Probab. Theory Related Fields 82 (1989), 349-370.

[G.H.2] F. Götze and C. Hipp, *Local limit theorems for sums of finite range potentials of a Gibbsian random field*, Ann. Probab. 18 (1990), 810-828.

[Ha.] M.G. Haiman, *Valeurs extrémales de suites stationnaires de variables aléatoires m-dépendantes*, Ann. Inst. H. Poincaré Sec. B 17 (1981), 309-330.

[He.1] L. Heinrich, *Asymptotic expansions in the central limit theorem for a special class of m-dependent random fields I*, Math. Nachr. 134 (1987), 83-106.

[He.2] L. Heinrich, *Asymptotic expansions in the central limit theorem for a special class of m-dependent random fields II-Lattice case*, Math. Nachr. 145 (1990), 309-327.

[Ho.Ro.] W. Hoeffding and H. Robbins, *The central limit theorem for dependent random variables*, Duke Math. J. 15 (1948), 773-780.

[Ibr.Li.] I.A. Ibragimov and Y.V. Linnik, *Independent and stationary sequences of random variables*, Wolters Noordhoff, Groningen, 1971.

[Ja.1] S. Janson, *Renewal theory for m-dependent variables*, Ann. Probab. 11 (1983), 558-568.

[Ja.2] S. Janson, *Runs in m-dependent sequences*, Ann. Probab. 12 (1984), 805-818.

[Ma.1] F. Matúš, *Sur les chemins de longeur deux et les produits d'un espace probabilisé*, research report nr. 1586, ÚTIA ČSAV, Prague, 1989.

[Ma.2] F. Matúš, *Inequalities concerning the demi-degrees and numbers of paths*, submitted to Israel J. Math.

[Ma.3] F. Matúš, *Independence structures modelling stochastic and algebraic independences and connectedness*, submitted to Probab. Theory Related Fields.

[O'Br.] G.L. O'Brien, *Scaling transformations for* $\{0,1\}$*-valued sequences*, Z. Wahrsch. Verw. Gebiete 53 (1980), 35-49.

[O'Ci.] C.A. O'Cinneide, *Some properties of one-dependent sequences*, preprint 1987.

[Pe.] V.V. Petrov, *On the central limit theorem for m-dependent variables*, Selected Transl. Math. Statist. and Prob. 9 (1970), 83-88.

[P.R.] B.L.S. Prakasa Rao, *A non-uniform estimate of the rate of convergence in the central limit theorem for m-dependent random fields*, Z. Wahrsch. Verw. Gebiete 58 (1981), 247-256.

[V.1] V. de Valk, *The maximal and minimal* 2*-correlation of a class of* 1*-dependent* $0-1$ *valued processes*, Israel J. Math., 62 (1988), 181-205.

[V.2] V. de Valk, *One-dependent processes*, Dissertation, Delft University of Technology, Delft University Press, 1988.

[V.3] V. de Valk, *One-dependent processes; two-block-factors and non-two-block-factors*, M.C.-tract, C.W.I. Amsterdam, in preparation.

[V.4] V. de Valk, *Two-fold-mixing implies N-fold-mixing for all N for processes with compact Hilbert space representations*, report 637, State University of Utrecht, 1991.

Stochastic Processes and their Applications 46 (1993) 183-198
North-Holland

On regression representations of stochastic processes

Ludger Rüschendorf
Universität Münster, Germany

Vincent de Valk
University of Groningen, Netherlands

Received 6 November 1991
Revised 28 April 1992

We construct a.s. nonlinear regression representations of general stochastic processes $(X_n)_{n\in\mathbb{N}}$. As a consequence we obtain in particular special regression representations of Markov chains and of certain m-dependent sequences. For m-dependent sequences we obtain a constructive method to check, whether these sequences have a monotone $(m+1)$-block factor representation.

AMS Subject Classifications: 60J10, 28D05, 60G10.

representation as function of i.i.d. sequences * generalized two-block factor * m-dependence * Markov regression * Markov chain

1. Markov regression and standard representation

Let $X=(X_n)_{n\in\mathbb{N}}$ be a stochastic, real-valued process. The aim of this section is to construct two types of a.s. regression representations of X by an i.i.d. sequence (U_n). One representation is of the form $X_n=f_n(X_1,\ldots,X_{n-1},U_n)$ a.s.; we call this representation '*Markov regression*' (on X). A second representation is of the form $X_n=f_n(U_1,\ldots,U_n)$ a.s.; we call this regression representation '*standard representation*' (on U). These constructions are the counterpart for autoregressive representations in time series analysis. Here we obtain a nonlinear representation of X_n of the past and of innovations U_n (which are independent and not only orthogonal).

We need a technical proposition about quantile transformations to construct standard representations. We write λ for the Lebesgue measure and $F_-(t):=\lim_{s\uparrow t}F(s)$.

Correspondence to: Vincent de Valk, Department of Mathematics, University of Groningen, P.O. Box 800, 9700 AV Groningen, Netherlands.

Proposition 1 (Ferguson, [3, Lemma 1, p. 216]). *Let X be a real random variable with distribution function F and let U be independent of X, $\mathrm{R}(0,1)$-distributed* ($\mathrm{R}(0,1)$ *is the uniform distribution over the interval* $(0,1)$). *Define for* $\alpha \in (0,1)$,

$$\tilde{F}(x,\alpha) := P[X<x] + \alpha P[X=x]. \tag{1}$$

Then

$$\tilde{F}(X,U) \stackrel{\mathrm{d}}{=} \mathrm{R}(0,1) \quad (\stackrel{\mathrm{d}}{=} \textit{ is equality in distribution}), \tag{2}$$

$$F^{-1}(U) \stackrel{\mathrm{d}}{=} X \quad (F^{-1}(t) := \inf\{s\colon F(s) \geqslant t\}) \tag{3}$$

and

$$X = F^{-1}(\tilde{F}(X,U)) \quad \textit{a.s.} \tag{4}$$

Since a proof of this result seems to be not easy accessible in the literature, we provide a proof of this well-known result.

Proof. Let $D \subset \mathbb{R}$ denote the set of discontinuities of F, then

$$\begin{aligned} &P[\tilde{F}(X,U)\in A] \\ &\quad = P[\tilde{F}(X,U)\in A, X\in D] \\ &\qquad + P[\tilde{F}(X,U)\in A, X\in D^{\mathrm{c}}]P[\tilde{F}(X,U)\in A, X\in D] \\ &\quad = \sum_{x\in D} P[\tilde{F}(x,U)\in A]P[X=x] \\ &\quad = \sum_{x\in D} P[F_{-}(x)+U(F(x)-F_{-}(x))\in A](F(x)-F_{-}(x)) \\ &\quad = \sum_{x\in D} \frac{\lambda(A\cap(F_{-}(x),F(x)])}{F(x)-F_{-}(x)}(F(x)-F_{-}(x)) \\ &\quad = \sum_{x\in D} \lambda(A\cap(F_{-}(x),F(x)]) = \lambda(A\cap\bar{D}), \end{aligned}$$

where $\bar{D} := \bigcup_{x\in D}(F_{-}(x), F(x)]$. Further ($D^{\mathrm{c}}$ is the complement of D)

$$P[\tilde{F}(X,U)\in A, X\in D^{\mathrm{c}}] = P[F(X)\in A, X\in D^{\mathrm{c}}] = \lambda(A\cap\bar{D}^{\mathrm{c}}).$$

In the proof we used that U and $\{X=x\}$ are independent for all $x\in D$. We conclude that

$$P[\tilde{F}(X,U)\in A] = \lambda(A\cap\bar{D}) + \lambda(A\cap\bar{D}^{\mathrm{c}}) = \lambda(A)$$

and this proves (2).

From the definition of the pseudo-inverse follows

$$P(F^{-1}(U)\leqslant t) = P(U\leqslant F(t)) = F(t) = P(X\leqslant t)$$

which proves (3) and

$$\{F^{-1}(\tilde{F}(X,U))\leqslant t\} = \{\tilde{F}(X,U)\leqslant F(t)\} = \{X\leqslant t\} \quad \text{a.s.}$$

which proves (4). $\square$

The a.s. representation in (4) has some useful applications in stochastic ordering [11]. If F is continuous, then $\tilde{F}(X, U) = F(X)$.

We next consider the multivariate generalization of Proposition 1. Let $X = (X_1, \ldots, X_n)$ be a random vector in $\mathbb{R}^n$ and let $F_1, F_{2|1}, \ldots, F_{n|1,\ldots,n-1}$ denote the first marginal distribution function respectively the conditional distribution function of X_k given $X_1, \ldots, X_{k-1}$. Let $V_1, \ldots, V_n$ be i.i.d. R(0, 1)-distributed random variables and define the multivariate quantile transform

$$\begin{aligned} Y_1 &:= F_1^{-1}(V_1), \\ Y_k &:= F_{k|1,\ldots,k-1}^{-1}(V_k \mid Y_1, \ldots, Y_{k-1}), \quad 2 \leq k \leq n. \end{aligned} \tag{5}$$

For this transformation see [8, 9, 10] and [11]. Note that $Y = (Y_1, \ldots, Y_n)$ is of the form $f(V)$ with $V = (V_1, \ldots, V_n)$, where the ith component $f_i(V) = f_i(V_1, \ldots, V_i)$.

Proposition 2. (a) $X \overset{\mathrm{d}}{=} Y$.

(b) *There exists an i.i.d.* R(0, 1)*-sequence* $U = (U_i)_{1 \leq i \leq n}$ *such that*

$$X = f(U) \quad \textit{a.s., as defined above.} \tag{6}$$

Proof. (a) The proof of (a) in the case $n = 2$ is as follows:

$$\begin{aligned} P(Y_1 \leq a, Y_2 \leq b) &= P(Y_1 \leq a, V_2 \leq F_{2|1}(b \mid Y_1)) \\ &= \int_{-\infty}^{a} P(V_2 \leq F_{2|1}(b \mid t))\, \mathrm{d}F_1(t) \\ &= \int_{-\infty}^{a} F_{2|1}(b \mid t)\, \mathrm{d}F_1(t) \\ &= P(X_2 \leq b, X_1 \leq a). \end{aligned}$$

The general case follows by induction.

(b) Since $f(V) \overset{\mathrm{d}}{=} X$ we obtain from Proposition 1 in [7] the existence of a measure preserving transformation $\varphi : (\Omega, \Sigma) \to (\Omega, \Sigma)$ such that

$$X = f(U) \quad \text{a.s.} \tag{7}$$

where $U_i = V_i \circ \varphi$, $1 \leq i \leq n$, are again i.i.d. R(0, 1)-distributed random variables. □

Skorohod (1976) proved for random variables X, Y with values in Borel spaces and a given R(0, 1)-distributed random variable V independent of X, Y the existence of a random variable U and measurable functions f, g such that

$$\begin{aligned} &X = f(Y, U) \quad \text{a.s.,} \\ &U = g(X, Y, V) \quad \text{is independent of } Y. \end{aligned} \tag{8}$$

The following theorem extends this result to stochastic processes. Furthermore, in the case of real stochastic processes we obtain an explicit representation.

Let $X=(X_1, X_2, \ldots)$ be a real valued stochastic process and let $V=(V_1, V_2, \ldots)$ be an i.i.d. sequence of R(0, 1)-distributed random variables, V independent of X. Define

$$
\begin{aligned}
&U_1 := \tilde{F}_1(X_1, V_1) \quad (\tilde{F}_1 \text{ as in (1), } F_1 \text{ distribution function of } X_1),\\
&Z_1 := F_1^{-1}(U_1),
\end{aligned}
\tag{9}
$$

and let for $k \geqslant 2$,

$$
\begin{aligned}
\tilde{F}_{k|1,\ldots,k-1}(x, v \mid z_1, \ldots, z_{k-1}) &:= P(X_k < x \mid Z_1 = z_1, \ldots, Z_{k-1} = z_{k-1})\\
&\qquad + vP(X_k = x \mid Z_1 = z_1, \ldots, Z_{k-1} = z_{k-1}),\\
U_k &:= \tilde{F}_{k|1,\ldots,k-1}(X_k, V_k \mid Z_1, \ldots, Z_{k-1}),\\
Z_k &:= F^{-1}_{k|1,\ldots,k-1}(U_k \mid Z_1, \ldots, Z_{k-1}),
\end{aligned}
\tag{10}
$$

where $F_{k|1,\ldots,k-1}$ is the conditional distribution function of X_k given $X_1, \ldots, X_{k-1}$.

Theorem 3. *Let* $Z=(Z_1, Z_2, \ldots)$ *then*:

(a) $Z = X$ *a.s.*

(b) $U=(U_1, U_2, \ldots)$ *is an i.i.d.* R(0, 1)*-distributed random sequence.*

(c) U_k *and* $(X_1, \ldots, X_{k-1})$ *are independent.*

We call the representation $X_1 = f_1(U_1)$, $X_k = f_k(X_1, \ldots, X_{k-1}, U_k)$ in (9), (10), *Markov-regression representation* of X.

Proof. The equality $Z_1 = X_1$ follows from (4). We continue by induction on k. Assume that $(Z_1, \ldots, Z_k) = (X_1, \ldots, X_k)$ a.s. Since $P^{(U_{k+1}|Z_1=z_1,\ldots,Z_k=z_k)}$ is R(0, 1)-distributed for all $z_1, \ldots, z_k$ we have that U_{k+1} and $(Z_1, \ldots, Z_k) = (X_1, \ldots, X_k)$ a.s. are independent.

From

$$
\begin{aligned}
\{Z_{k+1} \leqslant t\} &= \{F^{-1}_{k+1|1,\ldots,k}(U_{k+1} \mid Z_1, \ldots, Z_k) \leqslant t\}\\
&= \{U_{k+1} \leqslant F_{k+1|1,\ldots,k}(t \mid Z_1, \ldots, Z_k)\}\\
&= \{\tilde{F}_{k+1|1,\ldots,k}(X_{k+1}, V_{k+1} \mid Z_1, \ldots, Z_k) \leqslant F_{k+1|1,\ldots,k}(t \mid Z_1, \ldots, Z_k)\}\\
&= \{X_{k+1} \leqslant t\} \quad \text{a.s.}
\end{aligned}
$$

we conclude that $X_{k+1} = Z_{k+1}$ a.s. Because U_{k+1} and $(X_1, \ldots, X_k)$ are independent, we have that U_{k+1} and $U_1, \ldots, U_k$ (functions of $X_1, \ldots, X_k, V_1, \ldots, V_k$) are independent. $\square$

The existence of a Markov regression representation for processes with values in Borel spaces is immediate from Theorem 3 (but is nonconstructive).

In the case that $(X_n)_{n\in\mathbb{N}}$ is an *m-Markov chain* (for some $m\in\mathbb{N}$), i.e. the conditional distribution of X_{n+m+1} given the past $\{X_1,\dots,X_{m+n}\}$ only depends on $\{X_{n+1},\dots,X_{n+m}\}$ the Markov regression representation in Theorem 3 specializes to:

Corollary 4. *Let $X=(X_n)$ be an m-Markov chain. Then there exists an i.i.d. sequence $U=(U_1,U_2,\dots)$ of* R(0, 1)*-distributed random variables and a sequence of measurable functions (f_n) such that*

$$\begin{aligned}&X_n=f_n(X_{n-m},\dots,X_{n-1},U_n)\quad a.s.\quad(n\geqslant m+1),\\&U_n\ \textit{independent of}\ (X_1,\dots,X_{n-1}).\qquad\square\end{aligned}\tag{11}$$

For the case of a Markov chain $(m=1)$, see [6, p. 155]. By Theorem 3 the method of pathwise constructions of stochastic models is equivalent to constructions in distribution. One can characterize further distributional properties as in Corollary 4. E.g. if (X_n) is a Markov chain and a martingale, then X_n has a representation $X_n=f_n(X_{n-1},U_n)$ with $\int_0^1 f_n(x,u)\,\mathrm{d}u=x$ for all x.

The following alternative construction of a standardization sequence $U=(U_1,U_2,\dots)$ of $X=(X_1,X_2,\dots)$ will be of interest in connection with m-dependent sequences. This i.i.d. sequence U is a.s. equal to the sequence U in Theorem 3. We will explain this in Remark 16.

Let $V=(V_1,V_2,\dots)$ be an i.i.d. R(0, 1)-distributed sequence independent of $X=(X_1,X_2,\dots)$. Let G_1 be the distribution function of X_1 and define

$$\begin{aligned}&U_1:=\tilde G_1(X_1,V_1),\\&U_k:=\tilde G_{k|1,\dots,k-1}(X_k,V_k\,|\,U_1,\dots,U_{k-1})\quad(k\geqslant 2),\end{aligned}\tag{12}$$

where $G_{k|1,\dots,k-1}$ is the conditional distribution function of X_k given $(U_1,\dots,U_{k-1})$. The functions $\tilde G$ are associated to G as in the proof of Theorem 3. Similarly to the proof of Theorem 3 we obtain:

Theorem 5. (a) *(U_k) is an i.i.d.* R(0, 1)*-sequence.*

$$\begin{aligned}\text{(b)}\quad&X_1=G_1^{-1}(U_1),\\&X_k=G_{k|1,\dots,k-1}^{-1}(U_k\,|\,U_1,\dots,U_{k-1}).\qquad\square\end{aligned}\tag{13}$$

We call the representation in (13) the *standard representation* of X.

If for some $m\in\mathbb{N}$,

$$G_{k+m+1|1,\dots,k+m}(t_{k+m+1}\,|\,t_1,\dots,t_{k+m})=g_{k+m+1}(t_{k+1},\dots,t_{k+m+1}),\tag{14}$$

i.e. the conditional distribution of X_{k+m+1} given $U_1,\dots,U_{k+m}$ depends only on $U_{k+1},\dots,U_{k+m}$, we say that X has *m-Markov regression* on U.

Corollary 6. *If X has m-Markov regression on U, then X is a* generalized $(m+1)$-block factor, *i.e. (X_n) has the representation*

$$X_n=f_n(U_{n-m},\dots,U_{n-1},U_n)\quad a.s.,\quad n\geqslant m+1.\qquad\square\tag{15}$$

An interesting problem in probability theory is to find simple sufficient conditions for the existence of an $(m+1)$-block factor representation as in (15) (cf. [13]).

2. Markov chains and *m*-dependence

A process (X_n) is called *m-dependent* $(m \in \mathbb{N})$ if $(X_n)_{n<t}$ and $(X_n)_{n \geq t+m}$ are independent for all $t \in \mathbb{N}$. It is trivial that a *generalized* $(m+1)$*-block factor* $(X_n) = (f_n(U_n, U_{n+1}, \ldots, U_{n+m}))$ a.s. of an i.i.d. sequence (U_n) is m-dependent.

For quite a time it was conjectured that every stationary m-dependent process has a representation as $(m+1)$*-block factor* $(f(U_n, \ldots, U_{n+m}))$ (here f_n is independent of n!). In [2] a two-parameter family of counterexamples is given of stationary one-dependent processes, assuming only two values, which do not have a two-block factor representation $(f(U_n, U_{n+1}))$ of an i.i.d. sequence (U_n). It was shown in [4] that certain extremal 0–1 valued one-dependent stationary processes have a two-block factor representation while in [1] it was shown that a stationary one-dependent Markov chain with not more than four states has a two-block factor representation. There is a counterexample for five states.

In addition to the results on Markov chains in [1] it is proved that one-dependent renewal processes are two-block factors. It will be shown next that a symmetry condition implies that one-dependent Markov chains are already independent.

Proposition 7. *Let* (X_n) *be a stationary, one-dependent* 0–1 *valued Markov chain. Then* (X_n) *is an i.i.d. sequence.*

Proof. We use the short notation

$$[a_1 \cdots a_n] := P[X_1 = a_1, \ldots, X_n = a_n].$$

From $[0] = [00] + [01] = [00] + [10]$ follows that $[01] = [10]$. In our formulas we use the convention $0/0 = 0$. By the stationarity, the one-dependence and the Markov property we have

$$[a_k]^2 = \sum_i [a_k a_i a_k] = \sum_i \frac{[a_k a_i a_k]}{[a_k a_i]}[a_k a_i] = \sum_i \frac{[a_i a_k]}{[a_i]}[a_k a_i] = \sum_i \frac{[a_k a_i]^2}{[a_i]}.$$

Thus we obtain

$$\begin{aligned} 0 &\leq \sum_i \left\{ \frac{[a_k a_i]}{\sqrt{[a_i]}} - [a_k]\sqrt{[a_i]} \right\}^2 \\ &= \sum_i \left\{ \frac{[a_k a_i]^2}{[a_i]} - 2[a_k][a_k a_i] + [a_k]^2 [a_i] \right\} \\ &= [a_k]^2 - 2[a_k]^2 + [a_k]^2 = 0. \end{aligned}$$

This implies $[a_k a_i]/\sqrt{[a_i]} = [a_k]\sqrt{[a_i]}$ for all $a_i a_k$ which is equivalent to $[a_k a_i] = [a_k][a_i]$. Combined with the Markov property this implies independence. $\square$

Remark 8. From the proof it follows that the statement of the proposition also holds for one-dependent Markov chains with countable state space under the condition

$$[a_1 a_2] = [a_2 a_1] \quad \text{for all } a_1, a_2.$$

For any two-valued stationary one-dependent process we have a much stronger reversibility property:

Proposition 9. *Let $(X_n)_n$ be a stationary one-dependent 0–1 valued process. Then*

$$[a_1 \cdots a_n] = [a_n \cdots a_1] \quad \textit{for all } n \textit{ and all } a_1, \ldots, a_n \in \{0, 1\}.$$

Proof. For $n = 2$ the statement follows from $[0] = [00] + [01] = [00] + [10]$, hence $[01] = [10]$ as in the proof of Proposition 7. We use induction on n. We write

$$[1^m] := [\underbrace{1 \cdots 1}_{m \text{ times}}].$$

Assume that the statement holds for n, then for $n+1$ we denote the number of zeroes in $w = a_1 \cdots a_n a_{n+1}$ by $n_0(w)$. We continue by induction on $n_0(w)$.

If $n_0(w) = 0$ then the statement is trivial. Assume that the statement holds for $n_0 \leq k$. If $n_0(w) = k + 1 > 0$ then $w = 1^m 0 v$ for some $m \geq 0$. Then

$$\begin{aligned}[a_1 \cdots a_n a_{n+1}] &= [1^m 0 v] = [1^m][v] - [1^{m+1} v] \\ &= [1^m][a_{m+2} \cdots a_{n+1}] - [1^{m+1} a_{m+2} \cdots a_{n+1}] \\ &= [a_{n+1} \cdots a_{m+2}][1^m] - [a_{n+1} \cdots a_{m+2} 1^{m+1}] \\ &= [a_{n+1} \cdots a_{m+2} 0 1^m] = [a_{n+1} a_n \cdots a_1]\end{aligned}$$

which proves the proposition. □

The statement of Proposition 9 does not hold for one-dependent processes that assume three or more values. If the condition $[a_1 a_2] = [a_2 a_1]$ for all a_1, a_2 does not hold, then the statement of Proposition 7 is no longer valid as the following example shows.

Example 1. Let $(U_n)_n$ be a Bernoulli sequence with $P[U_n = 1] = p = 1 - P[U_n = 0]$ for some $p \in (0, 1)$. Define the two-block factor $(X_n)_n$ by

$$X_n = 2U_n + U_{n+1}.$$

It is easily checked that $(X_n)_n$ is a one-dependent Markov chain with state space $\{0, 1, 2, 3\}$ and transition matrix

$$\begin{pmatrix} 1-p & p & 0 & 0 \\ 0 & 0 & 1-p & p \\ 1-p & p & 0 & 0 \\ 0 & 0 & 1-p & p \end{pmatrix}$$

and apparently $(X_n)_n$ is not an i.i.d. sequence.

Under a symmetry condition we prove a general version of Proposition 7.

Proposition 10. *Let $(X_n)_n$ be a stationary, one-dependent real Markov chain and assume that*

$$P^{(X_1,X_2)} = P^{(X_2,X_1)}. \tag{16}$$

Then $(X_n)_n$ is an i.i.d. sequence.

Proof. Let $f:\mathbb{R}\to(0,1)$ be one to one measurable, then $Y_n := f(X_n)$ also is a one-dependent Markov chain and $(\mathbb{E}\,Y_1)^2 = \mathbb{E}(Y_1 Y_3) = \mathbb{E}[\mathbb{E}(Y_1 Y_3 \mid Y_2)] = \mathbb{E}[\mathbb{E}(Y_1 \mid Y_2)\mathbb{E}(Y_3 \mid Y_2)]$. Since $E[Y_1 \mid Y_2] = g(Y_2)$ for a measurable g we can continue by using the stationarity and (16), $E(g(Y_2)E[Y_3 \mid Y_2]) = Eg(Y_2)Y_3 = Eg(Y_1)Y_2 = Eg(Y_2)Y_1 = E(g(Y_2)E[Y_1 \mid Y_2]) = EZ^2$, where $Z = E[Y_1 \mid Y_2]$. Therefore $\mathbb{E}Z = \mathbb{E}\,Y_1$ and $(\mathbb{E}Z)^2 = \mathbb{E}(Z^2)$ imply that $Z = \mathbb{E}Z$ a.s., i.e.

$$\mathbb{E}(f(X_1) \mid f(X_2) = t) = \mathbb{E}f(X_1) \quad [P^{f(X_2)} \text{ a.s.}]$$

equivalently

$$\mathbb{E}(f(X_1) \mid X_2 = f^{-1}(t)) = \mathbb{E}f(X_1) \quad [P^{X_2} \text{ a.s.}].$$

Since this holds for all f we obtain independence. □

We leave it as an exercise to the reader to prove that the assumption $P^{(X_1,X_2)} = P^{(X_2,X_1)}$ is equivalent to reversibility of the Markov chain, i.e. $P^{(X_1,\dots,X_n)} = P^{(X_n,\dots,X_1)}$ for all n. Of course X_n could take also values in a Borel space. By a modification of the constructions in section one we next show that one-dependent Markov chains have a three-block factor representation.

Theorem 11. *Let $(X_n)_n$ be a real Markov chain. Then there exists an* R(0, 1)*-sequence $(U_n)_n$ and a sequence of functions g_n such that U_n is independent of $X_1, \dots, X_{n-1}, X_{n+1}, \dots$ and*

$$X_n = g_n(U_n, X_{n-1}, X_{n+1}).$$

If $(X_n)_n$ is additionally one-dependent, then there exists an independent sequence $(Y_n)_n$ and a sequence of functions $(f_n)_n$ such that X_n is a three-block factor of $(Y_n)_n$,

$$X_n = f_n(Y_{n-2}, Y_{n-1}, Y_n).$$

Proof. Let F_1 be the distribution function of X_1 and let $F_{n|n-1,n+1}$ $(n \geq 2)$ be the conditional distribution function of X_n given X_{n-1}, X_{n+1}. Define $U_1 := \tilde{F}_1(X_1, V_1)$ and $(n \geq 2)$ $U_n := \tilde{F}_{n|n-1,n+1}((X_n, V_n) \mid X_{n-1}, X_{n+1})$, where (V_n) is an i.i.d. R(0, 1)-sequence independent of $(X_n)_n$.

Because $(U_n \mid X_{n-1}=x_{n-1}, X_{n+1}=x_{n+1})$ is R(0, 1)-distributed for every x_{n-1}, x_{n+1}, the Markov property implies that U_n is independent of $(X_1, \ldots, X_{n-1}, X_{n+1}, \ldots)$. Analogously to Theorem 3 we have

$$X_n = F^{-1}_{n|n-1,n+1}(U_n \mid X_{n-1}, X_{n+1}) := g_n(U_n, X_{n-1}, X_{n+1}).$$

Define $Y_1 := X_1$, $Y_n := (X_{2n-1}, U_{2n-2})$ $(n \geqslant 2)$ and we obtain

$$X_1 = f_1(Y_1), \quad X_2 = f_2(Y_1, Y_2), \quad X_3 = f_3(Y_2),$$

$$X_{2n} = f_{2n}(Y_n, Y_{n+1}), \quad X_{2n+1} = f_{2n+1}(Y_{n+1}), \qquad n \geqslant 1.$$

If (X_n) is one-dependent, then $(Y_n)_n$ is an independent sequence. We can make a decent three-block factor out of this sequence by taking some i.i.d. R(0, 1)-sequence $(T_N)_N$ that is independent of X, Y and U. Define the process $(Z_N)_N$ by

$$Z_{2N+1} := T_{N+1}, \quad N \geqslant 0,$$

$$Z_{2N} := Y_N, \quad N \geqslant 1.$$

It is trivial that

$$X_N = h_N(Z_N, Z_{N+1}, Z_{N+2})$$

for measurable functions h_N. $\square$

Remark 12. From the last proof follows that every one-dependent Markov sequence of length 3 is a two-block factor of an i.i.d. sequence.

3. Standard representation and m-dependence

In this section we want to prove a partial converse of Corollary 6, namely if (under some assumptions) (X_n) has an $(m+1)$-block factor representation, then (X_n) has m-Markov regression on the standard representation U in (12). In this way we obtain a constructive method to check the possibility of an $(m+1)$-block factor representation for some subclasses of m-dependent sequences. This also justifies the notion of standard representation for (12), (13) and implies that the standardization U in (12) is the right standardization for the $(m+1)$-block factor representation problem. We shall deal explicitly with the case $m=1$. We begin with the following example.

Example 2. Let $V=(V_n)_{n\in\mathbb{N}}$ be an i.i.d. R(0, 1)-distributed sequence and define $X_1 = V_1$, $X_n = V_{n-1} + V_n$ $(n \geqslant 2)$. Then $(X_n)_{n\in\mathbb{N}}$ has a two-block factor representation on the standardization $(V_n)_{n\in\mathbb{N}}$. We consider the standardization $(U_n)_n$ of (12). Obviously $U_1 = X_1 = V_1$. Furthermore, $\tilde{G}_{2|1}(x, v \mid v_1) = P[X_2 \leqslant x \mid V_1 = v_1] = P[V_2 \leqslant x - v_1] = x - v_1$, $v_1 \leqslant x \leqslant v_1 + 1$. So $U_2 := \tilde{G}_{2|1}(X_2, V_2 \mid U_1) = X_2 - V_1 = V_2$. By induction we obtain in a similar way $U_n = V_n \ \forall n$, i.e. our standardization (12) produces the right standardization leading to the two-block-factor representation $X_1 = V_1, X_n = V_{n-1} + V_n \ (n \geq 2)$.

Generalizing this example, we say that $f_1(V_1), f_2(V_1, V_2), f_3(V_2, V_3), \ldots$ is a *monotone two-block factor*, if $f_1, f_i(v, \cdot)$ are monotonically nondecreasing for all i, v.

Obviously the standard representation (13) has a monotonicity property as defined here; so this assumption is necessary if the two-block factor representation is identical to the standard representation.

Theorem 13. *Assume that $X_1 = f_1(V_1)$ a.s., $X_k = f_k(V_{k-1}, V_k)$ a.s. has a monotone two-block factor representation and assume that all (conditional) distribution functions $G_1, G_{k|1,\ldots,k-1}$ in (13) are continuous, then the standardization U in (12) is identical to V and the standard representation (13) gives the two-block factor representation.*

Proof. Since $G_1 = \tilde{G}_1$ and $G_{k|1,\ldots,k-1} = \tilde{G}_{k|1,\ldots,k-1}$ we obtain from (12), (13),

$$U_1 = G_1(X_1),$$

where

$$G_1(x) = P(X_1 \leq x) = P(f_1(V_1) \leq x) = P(V_1 \leq g_1(x)) = (g_1 = f_1^{-1}) = g_1(x)$$

and, therefore,

$$U_1 = g_1 \circ f_1(V_1) = V_1 \quad \text{a.s.}$$

$$U_2 = G_{2|1}(X_2 \mid V_1),$$

where

$$\begin{aligned} G_{2|1}(x \mid v_1) &= P(f_2(V_1, V_2) \leq x \mid V_1 = v_1) \\ &= P(f_2(v_1, V_2) \leq x) = P(V_2 \leq g_2(v_1, x)) \\ &= g_2(v_1, x) \quad (g_2(v_1, \cdot) = f_2^{-1}(v_1, \cdot)). \end{aligned}$$

Therefore,

$$\begin{aligned} U_2 &= g_2(V_1, f_2(V_1, V_2)) = V_2, \\ G_{3|12}(x \mid v_1, v_2) &= P(f_3(V_2, V_3) \leq x \mid V_1 = v_1, V_2 = v_2) \\ &= P(f_3(v_2, V_3) \leq x) = g_3(v_2, x) \end{aligned}$$

implying that

$$U_3 = g_3(V_2, f_3(V_2, V_3)) = V_3 \quad \text{a.s.}$$

The general case now follows from induction. So we obtain that our standardization yields the right standardization for the two-block factor representation, which is obtained by (13), since obviously using $U = V$ a.s.

$$\begin{aligned} G_{k|1,\ldots,k-1}(\cdot \mid U_1, \ldots, U_{k-1}) &= G_{k|1,\ldots,k-1}(\cdot \mid V_1, \ldots, V_{k-1}) \\ &= G_{k|k-1}(\cdot \mid V_{k-1}). \qquad \square \end{aligned}$$

If the conditional distribution functions $G_{k|1,\dots,k-1}$ are not continuous, it is not possible to reconstruct (V_i) from $X=(X_i)$. We next show that the standardization (12), (13) can be applied to a version $\bar{X}$ of X.

Theorem 14. *If X has a monotone two-block factor representation $X=f(V)$ a.s., then there exists an i.i.d.* R(0, 1)*-sequence $(\bar{U})=(\bar{U}_i)$ such that the standard representation of $\bar{X}:=f(\bar{U})$ reproduces $\bar{U}$ and $\bar{X}=f(\bar{U})$.*

Proof. Let $f_1, f_k(v_{k-1}, \cdot)$ be monotonically nondecreasing for all k, v_{k-1} with $X_1=f_1(V_1)$, $X_k=f_k(V_{k-1}, V_k)$, $k\geqslant 2$.

Let $(\bar{V}_i)$ be an i.i.d. R(0, 1)-distributed sequence independent of (V_i) and consider the standard representation (12), with $\bar{U}_1:=\tilde{G}_1(X_1, \bar{V}_1)$, where

$$\begin{aligned}\tilde{G}_1(x,\alpha)&=P(X_1<x)+\alpha P(X_1=x)\\&=P(f_1(V_1)<x)+\alpha P(f_1(V_1)=x)\\&=P(V_1<f_1^{-1}(x))+\alpha P(V_1\in f_1^{-1}\{x\})\\&=f_1^{-1}(x)+\alpha\lambda(f_1^{-1}\{x\});\end{aligned}$$

$$f_1^{-1}(x)=\inf\{y: f_1(y)\geqslant x\},\qquad f_1^{-1}\{x\}=\{y: f_1(y)=x\}.$$

Therefore

$$\bar{U}_1=f_1^{-1}\circ f_1(V_1)+\bar{V}_1\lambda(f_1^{-1}\{f_1(V_1)\})=f_1^{-1}(X_1)+\bar{V}_1\lambda(f_1^{-1}\{X_1\}). \tag{17}$$

Define $X_1':=f_1(\bar{U}_1)=X_1$, $X_2':=f_2(\bar{U}_1, V_2)$ then $(\bar{U}_1, V_2, V_3, \dots)$ are i.i.d., R(0, 1)-distributed and

$$X^{(1)}:=(X_1', X_2', X_3, X_4, \dots)\overset{\mathrm{d}}{=}(X_1, X_2, X_3, X_4, \dots)=X.$$

In the next step consider

$$\begin{aligned}\tilde{G}_{2|1}(x,\alpha\,|\,u_1)&=P(X_2'<x\,|\,\bar{U}_1=u_1)+\alpha P(X_2'=x\,|\,\bar{U}_1=u_1)\\&=P(f_2(u_1, V_2)<x)+\alpha P(f_2(u_1, V_2)=x)\\&=P(V_2<f_2^{-1}(u_1, x))+\alpha P(V_2\in\{f_2^{-1}(u_1, x)\})\\&=f_2^{-1}(u_1, x)+\alpha\lambda(\{f_2^{-1}(u_1, x)\})\end{aligned}$$

the generalized inverse is taken w.r.t. the second component. Then our standard construction gives

$$\begin{aligned}\bar{U}_2&:=\tilde{G}_{2|1}((X_2', \bar{V}_2)\,|\,\bar{U}_1)\\&=f_2^{-1}(\bar{U}_1, f_2(\bar{U}_1, V_2))+\bar{V}_2\lambda(\{f_2^{-1}(\bar{U}_1, f_2(\bar{U}_1, V_2)\})\\&=f_2^{-1}(\bar{U}_1, X_2')+V_2\lambda(\{f_2^{-1}(\bar{U}_1, X_2')\}).\end{aligned} \tag{18}$$

Since $(\bar{U}_1, \bar{U}_2)$ are functions of $(V_1, \bar{V}_1, V_2, \bar{V}_2)$ the sequence $(\bar{U}_1, \bar{U}_2, V_3, V_4, \dots)$ is i.i.d., R(0, 1)-distributed. Define

$$X^{(2)}:=(f_1(\bar{U}_1), f_2(\bar{U}_1, \bar{U}_2), f_3(\bar{U}_2, V_3), f_4(V_3, V_4), \dots) \tag{19}$$

then $X^{(2)}\overset{\mathrm{d}}{=}X$.

We apply our standard construction to the third component $X_3' := f_3(\bar{U}_2, V_3)$ of $X^{(2)}$ to obtain $\bar{U}_3 := \tilde{G}_{3|1,2}((X_3', \bar{V}_3) \mid \bar{U}_1, \bar{U}_2)$, where

$$\begin{aligned}\tilde{G}_{3|12}(x, \alpha \mid u_1, u_2) &= P(f_3(\bar{U}_2, V_3) < x \mid \bar{U}_1 = u_1, \bar{U}_2 = u_2) \\ &\quad + \alpha P(f_3(\bar{U}_2, V_3) = x \mid (\bar{U}_1 = u_1, \bar{U}_2 = u_2) \\ &= P(V_3 < f_3^{-1}(u_2, x)) + \alpha P(V_3 \in \{f_3^{-1}(u_2, x)\}).\end{aligned}$$

Therefore,

$$\bar{U}_3 = f_3^{-1}(\bar{U}_2, f_3(\bar{U}_2, V_3)) + \bar{V}_3 \lambda(\{f_3^{-1}(\bar{U}_2, f_3(\bar{U}_2, V_3))\}).$$

Again $(\bar{U}_1, \bar{U}_2, \bar{U}_3, V_4, V_5, V_6 \ldots) \stackrel{\mathrm{d}}{=} (V_1, V_2, V_3, \ldots)$ and

$$X^{(3)} = (f_1(\bar{U}_1), f_2(\bar{U}_1, \bar{U}_2), f_3(\bar{U}_2, \bar{U}_3), f_4(\bar{U}_3, V_4), f_5(V_4, V_5), \ldots) \stackrel{\mathrm{d}}{=} X$$

and we can continue this process by induction. Thus we obtain that for a version $\bar{X}$ of X we have the two-block-factor representation

$$\bar{X}_1 = f_1(\bar{U}_1), \quad \bar{X}_2 = f_2(\bar{U}_1, \bar{U}_2), \quad \bar{X}_3 = f_3(\bar{U}_2, \bar{U}_3), \ldots, \tag{20}$$

where the $(\bar{U}_i)$ are obtained from our modified standardization process.

Next we apply the standardization (12) to $\bar{X}$ to obtain

$$\begin{aligned}U_1 := \tilde{G}_1(\bar{X}_1, \bar{V}_1) &= f_1^{-1}(\bar{X}_1) + \bar{V}_1 \lambda(f_1^{-1}(\{\bar{X}_1\})) \\ &= f_1^{-1}(f_1(\bar{U}_1)) + \bar{V}_1 \lambda(f_1^{-1}\{f_1(\bar{U}_1)\}) \\ &= f_1^{-1}(X_1) + \bar{V}_1 \lambda(f_1^{-1}(X_1)) = \bar{U}_1,\end{aligned}$$

i.e. the standardization reproduces $\bar{U}_1$. In the next step

$$U_2 = \tilde{G}_{2|1}((\bar{X}_1, \bar{V}_2) \mid U_1) = \bar{U}_2,$$

$$U_3 = \tilde{G}_{3|1,2}((\bar{X}_3, \bar{V}_3) \mid U_1, U_2) = \tilde{G}_{3|1,2}((\bar{X}_3, \bar{V}_3) \mid \bar{U}_1, \bar{U}_2) = \bar{U}_3,$$

and so on. □

So in general from the two-block factor representation $X = f(V)$ we construct by a modification of the standardization procedure a version $\bar{X}$ of X with a two-block factor representation $\bar{X} = f(\bar{U})$. The standardization (12), applied to this representation reproduces $\bar{U}$ i.e. $U = \bar{U}$ and (13), our standard regression representation, reproduces this two-block factor representation of $\bar{X}$.

Remark 15. Obviously a result similar to Theorem 13, 14 also holds for $(m+1)$-block factor representations. While Theorem 13 is constructive, Theorem 14 indicates the applicability of the standard construction to a (not known) version of X.

Remark 16. The i.i.d. sequence U in Theorem 3 is a.s. equal to the i.i.d. sequence U in Theorem 5. The proof is essentially the same as the proof of Theorem 13. We leave it as an exercise to the reader. The consequence of this observation is that the Standard Representation $X_n = f_n(X_1, \ldots, X_{n-1}, U_n)$ can also be obtained by iterating the Markov Regression $X_n = g_n(U_1, \ldots, U_n)$; i.e. $X_2 = f_2(X_1, U_2) = f_2(f_1(U_1), U_2) = g_2(U_1, U_2)$ and $X_3 = f_3(X_1, X_2, U_3) = f_3(f_1(U_1), f_2(f_1(U_1), U_2), U_3) = g_3(U_1, U_2, U_3)$ etc.

The question now is: how restrictive is the assumption of a monotone two-block factor representation?

Example 3. (a) Let (V_i) be an i.i.d. R(0, 1)-sequence and consider the two-block factor $X_1 = V_1$, $X_2 = V_1 - V_2$, $X_3 = V_2 - V_3, \ldots$. We obtain a monotone two-block factor representation by defining $U_1 = V_1$, $U_i := 1 - V_i$, $i \geq 2$. Then

$$X_1 = U_1, \quad X_2 = U_1 + U_2 - 1, \quad X_3 = U_3 - U_2, \quad X_4 = U_4 - U_3, \ldots, \tag{21}$$

is a monotone two-block factor representation.

(b) If $X_1 = V_1$, $X_2 = (V_1 - \frac{1}{2})V_2$, $X_3 = (V_2 - \frac{1}{2})V_3 \ldots$, then define

$$U_1 = V_1, \qquad U_i = \begin{cases} V_i & \text{if } V_{i-1} \geq \frac{1}{2}, \\ 1 - V_i & \text{if } V_{i-1} < \frac{1}{2}, \end{cases} \quad i \geq 2.$$

It is easy to check that (U_i) is an i.i.d. R(0, 1)-sequence and we obtain the monotone representation (in distribution) $\bar{X}$ of X,

$$\bar{X}_1 = U_1, \qquad \bar{X}_i = \begin{cases} (U_{i-1} - \frac{1}{2})U_i & \text{if } U_{i-1} \geq \frac{1}{2}, \\ (U_{i-1} - \frac{1}{2})(1 - U_i) & \text{if } U_{i-1} < \frac{1}{2}, \end{cases} \quad i \geq 2. \tag{22}$$

(c) If more generally than in (b) $X_1 = f_1(V_1)$, $X_i = f_i(V_{i-1}, V_i)$, $f_1 \uparrow$ and for all $v_{i-1}, i, f_i(v_{i-1}, \cdot)$ is either monotonically nondecreasing or nonincreasing (i.e. $f_i(v_{i-1}, \cdot)\uparrow$ for $v_{i-1} \in V_i^+$ and $f_i(v_{i-1}, \cdot)\downarrow$ for $v_{i-1} \in V_i^-$) then define a sequence

$$U_1 := V_1, \qquad U_i := \begin{cases} V_i & \text{if } U_{i-1} \in V_i^+, \\ 1 - V_i & \text{if } U_{i-1} \in V_i^-, \end{cases} \quad i \geq 2.$$

Then (U_i) is an i.i.d. R(0, 1)-sequence and with $g_1 = f_1$,

$$g_i(v_{i-1}, v_i) = \begin{cases} f_i(v_{i-1}, v_i) & \text{if } v_{i-1} \in V_i^+, \\ f_i(v_{i-1}, 1 - v_i) & \text{if } v_{i-1} \in V_i^-, \end{cases}$$

the sequence $g_1(U_1)$, $g_2(U_1, U_2), \ldots$ has the same distribution as X. Therefore, X has a monotone two-block factor representation.

For the general question we use the following proposition.

Proposition 17. *Let (V_n) be an i.i.d.* R(0, 1)*-sequence and $X_1 = f_1(V_1)$, $X_n = f_n(V_{n-1}, V_n)$, $n \geq 2$, a generalized two-block factor. Furthermore, let $(\bar{V}_n)$ be an i.i.d.* R(0, 1)*-sequence independent of (V_n). Then there exist an i.i.d.,* R(0, 1)*-sequence (U_n), $U_n = h_n(V_{n-1}, V_n, \bar{V}_n)$ independent of $(V_1, \ldots, V_{n-1})$ and functions (g_n) such that*

$$\begin{aligned} &X_1 = g_1(U_1), \qquad X_n = g_n(V_{n-1}, U_n), \quad n \geq 2, \qquad \textit{and} \\ &g_1, g_n(v_{n-1}, \cdot) \ \textit{monotonically nondecreasing } \forall n, v_{n-1}. \end{aligned} \tag{23}$$

Proof. Let G_1 be the distribution function of X_1 and let $G_{n|n-k,\ldots,n-1}$ be the conditional distribution function of X_n given $V_{n-k}, \ldots, V_{n-1}$. Define

$$\begin{aligned} &U_1 := \tilde{G}_1(X_1, \bar{V}_1), \\ &U_n := \tilde{G}_{n|1,\ldots,n-1}(X_n, \bar{V}_n \mid V_1, \ldots, V_{n-1}), \quad n \geq 2. \end{aligned} \tag{24}$$

Since the conditional distribution of U_n given $V_1 = v_1, \dots, V_{n-1} = v_{n-1}$ is R(0, 1) for all $v_1, \dots, v_{n-1}$ we have that U_n is independent of $(V_1, \dots, V_{n-1})$. Since $U_k = h_k(V_1, \dots, V_k, \bar{V}_1, \dots, \bar{V}_k)$, this implies that U_n is independent of $U_1, \dots, U_{n-1}$. From (4) we conclude that

$$X_1 = G_1^{-1}(U_1), \qquad X_n = G_{n|1,\dots,n-1}^{-1}(U_k \mid V_1, \dots, V_{n-1}), \quad n \geq 2.$$

Actually, $\tilde{G}_{n|1,\dots,n-1} = \tilde{G}_{n|n-1}$ since

$$\begin{aligned} G_{n|1,\dots,n-1}(x \mid v_1, \dots, v_{n-1}) &= P(X_n \leq x \mid V_1 = v_1, \dots, V_{n-1} = v_{n-1}) \\ &= P(f_n(V_{n-1}, V_n) \leq x \mid V_1 = v_1, \dots, V_{n-1} = v_{n-1}) \\ &= P(f_n(v_{n-1}, V_n) \leq x) = G_{n|n-1}(x \mid v_{n-1}) \end{aligned}$$

(and similarly for $\tilde{G}_{n|1,\dots,n-1}$). So we have $X_n = G_{n|n-1}^{-1}(U_n \mid V_{n-1}) = g_n(V_{n-1}, U_n)$, where $g_n(v_{n-1}, \cdot)$ is nondecreasing. □

Obviously from (23) we obtain a monotone two-block factor representation if $V_{n-1} = h(U_{n-1})$ for some function h. In general we obtain the following weakened monotone representation property.

Corollary 18. *Let (W_n) be an i.i.d.* R(0, 1)*-sequence independent of $(V_n), (\bar{V}_n)$ and let $X_1 = f_1(V_1)$, $X_i = f_i(V_{i-1}, V_i)$, $i \geq 2$, be a generalized two-block factor. Then there exists an* R(0, 1)*-sequence $\bar{U}_i = \tilde{h}_i(U_i, V_i, W_i)$ such that $\bar{U}_i$ is independent of U_i and*

$$X_1 = g_1(U_1), \quad X_2 = g_2(U_1, \bar{U}_1, U_2), \quad X_3 = g_3(U_2, \bar{U}_2, U_3), \dots, \tag{25}$$

where $g_1, g_i(u_i, \bar{u}_i, \cdot)$ are monotonically nondecreasing.

Proof. From Proposition 17 we have a monotone representation, $X_1 \doteq h_1(U_1)$, $X_n = h_n(V_{n-1}, U_n)$, $n \geq 2$. We apply (8) to obtain $V_i = \tilde{g}_i(U_i, \bar{U}_i)$ where $\bar{U}_i = \tilde{h}_i(U_i, V_i, W_i)$ is independent of U_i. Together we obtain (25). □

Generally, we can not assert that (U_n, V_n) is independent of $(V_1, \dots, V_{n-1})$ (we only have separately the independence of U_n respectively V_n of $(V_1, \dots, V_{n-1})$). In the case that (U_n, V_n) is independent of $(V_1, \dots, V_{n-1})$ we obtain that in the representation (25) the sequence

$$U_1, \bar{U}_1, U_2, \bar{U}_2, \dots \text{ is an i.i.d. R}(0, 1)\text{-sequence.} \tag{26}$$

Example 4. Let (V_i) be an i.i.d. R(0, 1)-sequence and let $X_1 = (V_1 - \frac{1}{2})^2$, $X_i = V_{i-1} \cdot (V_i - \frac{1}{2})^2$, $i \geq 2$ be a generalized two-block-factor. Then the construction of (25) is the following: $F_{X_1}(x) = 2\sqrt{x}$, $g_1(y) = (\frac{1}{2}y)^2$ and $U_1 = 2|V_1 - \frac{1}{2}|$. Let ε_i be random signs defined by $\varepsilon_i = +1$ if $V_i \geq \frac{1}{2}$ and $\varepsilon_i = -1$, else, and define $U_i = 2|V_i - \frac{1}{2}|$. Then $V_i = \frac{1}{2} + \frac{1}{2}\varepsilon_i U_i$ (and we can formally write ε_i as function of an R(0, 1)-random variable $\bar{U}_i$). Obviously, (ε_i, U_i) is independent of $V_1, \dots, V_{i-1}$ and we obtain from (25) the weakened monotone representation

$$X_1 = g_1(U_1), \quad X_2 = (\tfrac{1}{2} + \tfrac{1}{2}\varepsilon_1 U_1) g_1(U_2), \dots. \tag{27}$$

Proposition 19. *There exists a generalized two-block factor which does not have a monotone two-block factor representation.*

Proof. Let (V_i) be an i.i.d. R(0, 1)-sequence and let $X_1 = |V_1 - \frac{1}{2}|$, $X_i = V_{i-1}V_i$, $i \geqslant 2$. In order to show that (X_i) does not admit a monotone two-block factor representation we apply Theorem 13. So we calculate the standardization (U_i) from (12) and we show that the standard representation is not a two-block factor. Since $G_1(x) = P(X_1 \leqslant x) = 2x$, we obtain $U_1 = 2|V_1 - \frac{1}{2}|$. Furthermore,

$$G_{2|1}(x \mid u) = P(X_2 \leqslant x \mid U_1 = u) = \frac{1}{2}\left[\left(\frac{2x}{1+u} \wedge 1\right) + \left(\frac{2x}{1-u}\right) \wedge 1\right],$$

i.e.

$$U_2 = \frac{1}{2} \cdot \left[\frac{2V_1V_2}{1+2|V_1 - \frac{1}{2}|} \wedge 1 + \frac{2V_1V_2}{1-2|V_1-\frac{1}{2}|} \wedge 1\right].$$

With some calculations we obtain

$$G_{3|1,2}(x \mid u_1, u_2) = P(X_3 \leqslant x \mid U_1 = u_1, U_2 = u_2)$$

$$= \begin{cases} \frac{1}{2} \cdot \left[\frac{x}{(1-u_1)u_2} \wedge 1 + \frac{x}{(1+u_1)u_2} \wedge 1\right] & \text{if } u_2 \leqslant \frac{1}{1+u_1}, \\ \frac{x}{2u_2 - 1} \wedge 1 & \text{if } u_2 > \frac{1}{1+u_1}. \end{cases}$$

From this we conclude that (X_i) does not have a monotone two-block factor representation. □

References

[1] J. Aaronson, D. Gilat and M. Keane, On the structure of one-dependent Markov shifts, preprint (1990).

[2] J. Aaronson, D. Gilat, M.S. Keane and V. de Valk, An algebraic construction of a class of one-dependent processes, Ann. Probab. 17 (1989) 128–143.

[3] T.S. Ferguson, Mathematical Statistics. A Decision Theoretic Approach (Academic Press, New York, 1967).

[4] A. Gandolfi, M. Keane and V. de Valk, Extremal two-correlations of two-valued stationary one-dependent processes, Probab. Theory Rel. Fields 80 (1989) 475–480.

[5] D.L. Hanson, On the representation problem for stationary stochastic processes with trivial tail field, J. Math. Mech. 12 (1963) 293–301.

[6] D. Pfeiffer, Einführung in die Extremwertstatistik (Teubner, Stuttgart, 1989).

[7] S.T. Rachev and L. Rüschendorf, A transformation property of minimal metrics, Theory Probab. Appl. 35 (1990) 131–137.

[8] M. Rosenblatt, Remarks on a multivariate transformation, Ann. Math. Statist. 23 (1953) 470–472.

[9] M. Rosenblatt, Stationary Markov chains and independent random variables, J. Math. Mech. 9 (1960) 945–949.

[10] M. Rosenblatt, Addendum to: "Stationary Markov chains and independent random variables", J. Math. Mech. 11 (1962) 317.
[11] L. Rüschendorf, Stochastically ordered distributions and monotonicity of the OC-function of sequential probability ratio tests, Math. Operationsforsch. Statist. Ser. Statist. 12 (1981) 327-338.
[12] A.V. Skorohod, On a representation of random variables, Theory Probab. Appl. 21 (1976) 628-631.
[13] V. de Valk, One-dependent processes, dissertation, Delft Univ. of Technology (Delft, 1988).

AUTHOR INDEX

SUBJECT INDEX

CWI TRACTS

1 D.H.J. Epema. *Surfaces with canonical hyperplane sections.* 1984.

2 J.J. Dijkstra. *Fake topological Hilbert spaces and characterizations of dimension in terms of negligibility.* 1984.

3 A.J. van der Schaft. *System theoretic descriptions of physical systems.* 1984.

4 J. Koene. *Minimal cost flow in processing networks, a primal approach.* 1984.

5 B. Hoogenboom. *Intertwining functions on compact Lie groups.* 1984.

6 A.P.W. Böhm. *Dataflow computation.* 1984.

7 A. Blokhuis. *Few-distance sets.* 1984.

8 M.H. van Hoorn. *Algorithms and approximations for queueing systems.* 1984.

9 C.P.J. Koymans. *Models of the lambda calculus.* 1984.

10 C.G. van der Laan, N.M. Temme. *Calculation of special functions: the gamma function, the exponential integrals and error-like functions.* 1984.

11 N.M. van Dijk. *Controlled Markov processes; time-discretization.* 1984.

12 W.H. Hundsdorfer. *The numerical solution of nonlinear stiff initial value problems: an analysis of one step methods.* 1985.

13 D. Grune. *On the design of ALEPH.* 1985.

14 J.G.F. Thiemann. *Analytic spaces and dynamic programming: a measure theoretic approach.* 1985.

15 F.J. van der Linden. *Euclidean rings with two infinite primes.* 1985.

16 R.J.P. Groothuizen. *Mixed elliptic-hyperbolic partial differential operators: a case-study in Fourier integral operators.* 1985.

17 H.M.M. ten Eikelder. *Symmetries for dynamical and Hamiltonian systems.* 1985.

18 A.D.M. Kester. *Some large deviation results in statistics.* 1985.

19 T.M.V. Janssen. *Foundations and applications of Montague grammar, part 1: Philosophy, framework, computer science.* 1986.

20 B.F. Schriever. *Order dependence.* 1986.

21 D.P. van der Vecht. *Inequalities for stopped Brownian motion.* 1986.

22 J.C.S.P. van der Woude. *Topological dynamix.* 1986.

23 A.F. Monna. *Methods, concepts and ideas in mathematics: aspects of an evolution.* 1986.

24 J.C.M. Baeten. *Filters and ultrafilters over definable subsets of admissible ordinals.* 1986.

25 A.W.J. Kolen. *Tree network and planar rectilinear location theory.* 1986.

26 A.H. Veen. *The misconstrued semicolon: Reconciling imperative languages and dataflow machines.* 1986.

27 A.J.M. van Engelen. *Homogeneous zero-dimensional absolute Borel sets.* 1986.

28 T.M.V. Janssen. *Foundations and applications of Montague grammar, part 2: Applications to natural language.* 1986.

29 H.L. Trentelman. *Almost invariant subspaces and high gain feedback.* 1986.

30 A.G. de Kok. *Production-inventory control models: approximations and algorithms.* 1987.

31 E.E.M. van Berkum. *Optimal paired comparison designs for factorial experiments.* 1987.

32 J.H.J. Einmahl. *Multivariate empirical processes.* 1987.

33 O.J. Vrieze. *Stochastic games with finite state and action spaces.* 1987.

34 P.H.M. Kersten. *Infinitesimal symmetries: a computational approach.* 1987.

35 M.L. Eaton. *Lectures on topics in probability inequalities.* 1987.

36 A.H.P. van der Burgh, R.M.M. Mattheij (eds.). *Proceedings of the first international conference on industrial and applied mathematics (ICIAM 87).* 1987.

37 L. Stougie. *Design and analysis of algorithms for stochastic integer programming.* 1987.

38 J.B.G. Frenk. *On Banach algebras, renewal measures and regenerative processes.* 1987.

39 H.J.M. Peters, O.J. Vrieze (eds.). *Surveys in game theory and related topics.* 1987.

40 J.L. Geluk, L. de Haan. *Regular variation, extensions and Tauberian theorems.* 1987.

41 Sape J. Mullender (ed.). *The Amoeba distributed operating system: Selected papers 1984-1987.* 1987.

42 P.R.J. Asveld, A. Nijholt (eds.). *Essays on concepts, formalisms, and tools.* 1987.

43 H.L. Bodlaender. *Distributed computing: structure and complexity.* 1987.

44 A.W. van der Vaart. *Statistical estimation in large parameter spaces.* 1988.

45 S.A. van de Geer. *Regression analysis and empirical processes.* 1988.

46 S.P. Spekreijse. *Multigrid solution of the steady Euler equations.* 1988.

47 J.B. Dijkstra. *Analysis of means in some non-standard situations.* 1988.

48 F.C. Drost. *Asymptotics for generalized chi-square goodness-of-fit tests.* 1988.

49 F.W. Wubs. *Numerical solution of the shallow-water equations.* 1988.

50 F. de Kerf. *Asymptotic analysis of a class of perturbed Korteweg-de Vries initial value problems.* 1988.

51 P.J.M. van Laarhoven. *Theoretical and computational aspects of simulated annealing.* 1988.

52 P.M. van Loon. *Continuous decoupling transformations for linear boundary value problems.* 1988.

53 K.C.P. Machielsen. *Numerical solution of optimal control problems with state constraints by sequential quadratic programming in function space.* 1988.

54 L.C.R.J. Willenborg. *Computational aspects of survey data processing.* 1988.

55 G.J. van der Steen. *A program generator for recognition, parsing and transduction with syntactic patterns.* 1988.

56 J.C. Ebergen. *Translating programs into delay-insensitive circuits.* 1989.

57 S.M. Verduyn Lunel. *Exponential type calculus for linear delay equations.* 1989.

58 M.C.M. de Gunst. *A random model for plant cell population growth.* 1989.

59 D. van Dulst. *Characterizations of Banach spaces not containing l^1.* 1989.

60 H.E. de Swart. *Vacillation and predictability properties of low-order atmospheric spectral models.* 1989.

61 P. de Jong. *Central limit theorems for generalized multilinear forms.* 1989.

62 V.J. de Jong. *A specification system for statistical software.* 1989.

63 B. Hanzon. *Identifiability, recursive identification and spaces of linear dynamical systems, part I.* 1989.

64 B. Hanzon. *Identifiability, recursive identification and spaces of linear dynamical systems, part II.* 1989.

65 B.M.M. de Weger. *Algorithms for diophantine equations.* 1989.

66 A. Jung. *Cartesian closed categories of domains.* 1989.

67 J.W. Polderman. *Adaptive control & identification: Conflict or conflux?.* 1989.

68 H.J. Woerdeman. *Matrix and operator extensions.* 1989.

69 B.G. Hansen. *Monotonicity properties of infinitely divisible distributions.* 1989.

70 J.K. Lenstra, H.C. Tijms, A. Volgenant (eds.). *Twenty-five years of operations research in the Netherlands: Papers dedicated to Gijs de Leve.* 1990.

71 P.J.C. Spreij. *Counting process systems. Identification and stochastic realization.* 1990.

72 J.F. Kaashoek. *Modeling one dimensional pattern formation by anti-diffusion.* 1990.

73 A.M.H. Gerards. *Graphs and polyhedra. Binary spaces and cutting planes.* 1990.

74 B. Koren. *Multigrid and defect correction for the steady Navier-Stokes equations. Application to aerodynamics.* 1991.

75 M.W.P. Savelsbergh. *Computer aided routing.* 1992.

76 O.E. Flippo. *Stability, duality and decomposition in general mathematical programming.* 1991.

77 A.J. van Es. *Aspects of nonparametric density estimation.* 1991.

78 G.A.P. Kindervater. *Exercises in parallel combinatorial computing.* 1992.

79 J.J. Lodder. *Towards a symmetrical theory of generalized functions.* 1991.

80 S.A. Smulders. *Control of freeway traffic flow.* 1993.

81 P.H.M. America, J.J.M.M. Rutten. *A parallel object-oriented language: design and semantic foundations.* 1992.

82 F. Thuijsman. *Optimality and equilibria in stochastic games.* 1992.

83 R.J. Kooman. *Convergence properties of recurrence sequences.* 1992.

84 A.M. Cohen (ed.). *Computational aspects of Lie group representations and related topics. Proceedings of the 1990 Computational Algebra Seminar at CWI, Amsterdam.* 1991.

85 V. de Valk. *One-dependent processes.* 1994.

86 J.A. Baars, J.A.M. de Groot. *On topological and linear equivalence of certain function spaces.* 1992.

87 A.F. Monna. *The way of mathematics and mathematicians.* 1992.

88 E.D. de Goede. *Numerical methods for the three-dimensional shallow water equations.* 1993.

89 M. Zwaan. *Moment problems in Hilbert space with applications to magnetic resonance imaging.* 1993.

90 C. Vuik. *The solution of a one-dimensional Stefan problem.* 1993.

91 E.R. Verheul. *Multimedians in metric and normed spaces.* 1993.

92 J.L.M. Maubach. *Iterative methods for non-linear partial differential equations.* 1993.

93 A.W. Ambergen. *Statistical uncertainties in posterior probabilities.* 1993.

94 P.A. Zegeling. *Moving-grid methods for time-dependent partial differential equations.* 1993.

95 M.J.C. van Pul. *Statistical analysis of software reliability models.* 1993.

96 J.K. Scholma. *A Lie algebraic study of some integrable systems associated with root systems.* 1993.

97 J.L. van den Berg. *Sojourn times in feedback and processor sharing queues.* 1993.

98 A.J. Koning. *Stochastic integrals and goodness-of-fit tests.* 1993.

99 B.P. Sommeijer. *Parallelism in the numerical integration of initial value problems.* 1993.

100 J. Molenaar. *Multigrid methods for semiconductor device simulation.* 1993.

101 H.J.C. Huijberts. *Dynamic feedback in nonlinear synthesis problems.* 1994.

102 J.A.M. van der Weide. *Stochastic processes and point processes of excursions.* 1994.

MATHEMATICAL CENTRE TRACTS

1 T. van der Walt. *Fixed and almost fixed points.* 1963.

2 A.R. Bloemena. *Sampling from a graph.* 1964.

3 G. de Leve. *Generalized Markovian decision processes, part I: model and method.* 1964.

4 G. de Leve. *Generalized Markovian decision processes, part II: probabilistic background.* 1964.

5 G. de Leve, H.C. Tijms, P.J. Weeda. *Generalized Markovian decision processes, applications.* 1970.

6 M.A. Maurice. *Compact ordered spaces.* 1964.

7 W.R. van Zwet. *Convex transformations of random variables.* 1964.

8 J.A. Zonneveld. *Automatic numerical integration.* 1964.

9 P.C. Baayen. *Universal morphisms.* 1964.

10 E.M. de Jager. *Applications of distributions in mathematical physics.* 1964.

11 A.B. Paalman-de Miranda. *Topological semigroups.* 1964.

12 J.A.Th.M. van Berckel, H. Brandt Corstius, R.J. Mokken, A. van Wijngaarden. *Formal properties of newspaper Dutch.* 1965.

13 H.A. Lauwerier. *Asymptotic expansions.* 1966, out of print; replaced by MCT 54.

14 H.A. Lauwerier. *Calculus of variations in mathematical physics.* 1966.

15 R. Doornbos. *Slippage tests.* 1966.

16 J.W. de Bakker. *Formal definition of programming languages with an application to the definition of ALGOL 60.* 1967.

17 R.P. van de Riet. *Formula manipulation in ALGOL 60, part 1.* 1968.

18 R.P. van de Riet. *Formula manipulation in ALGOL 60, part 2.* 1968.

19 J. van der Slot. *Some properties related to compactness.* 1968.

20 P.J. van der Houwen. *Finite difference methods for solving partial differential equations.* 1968.

21 E. Wattel. *The compactness operator in set theory and topology.* 1968.

22 T.J. Dekker. *ALGOL 60 procedures in numerical algebra, part 1.* 1968.

23 T.J. Dekker, W. Hoffmann. *ALGOL 60 procedures in numerical algebra, part 2.* 1968.

24 J.W. de Bakker. *Recursive procedures.* 1971.

25 E.R. Paërl. *Representations of the Lorentz group and projective geometry.* 1969.

26 European Meeting 1968. *Selected statistical papers, part I.* 1968.

27 European Meeting 1968. *Selected statistical papers, part II.* 1968.

28 J. Oosterhoff. *Combination of one-sided statistical tests.* 1969.

29 J. Verhoeff. *Error detecting decimal codes.* 1969.

30 H. Brandt Corstius. *Exercises in computational linguistics.* 1970.

31 W. Molenaar. *Approximations to the Poisson, binomial and hypergeometric distribution functions.* 1970.

32 L. de Haan. *On regular variation and its application to the weak convergence of sample extremes.* 1970.

33 F.W. Steutel. *Preservation of infinite divisibility under mixing and related topics.* 1970.

34 I. Juhász, A. Verbeek, N.S. Kroonenberg. *Cardinal functions in topology.* 1971.

35 M.H. van Emden. *An analysis of complexity.* 1971.

36 J. Grasman. *On the birth of boundary layers.* 1971.

37 J.W. de Bakker, G.A. Blaauw, A.J.W. Duijvestijn, E.W. Dijkstra, P.J. van der Houwen, G.A.M. Kamsteeg-Kemper, F.E.J. Kruseman Aretz, W.L. van der Poel, J.P. Schaap-Kruseman, M.V. Wilkes, G. Zoutendijk. *MC-25 Informatica Symposium.* 1971.

38 W.A. Verloren van Themaat. *Automatic analysis of Dutch compound words.* 1972.

39 H. Bavinck. *Jacobi series and approximation.* 1972.

40 H.C. Tijms. *Analysis of (s,S) inventory models.* 1972.

41 A. Verbeek. *Superextensions of topological spaces.* 1972.

42 W. Vervaat. *Success epochs in Bernoulli trials (with applications in number theory).* 1972.

43 F.H. Ruymgaart. *Asymptotic theory of rank tests for independence.* 1973.

44 H. Bart. *Meromorphic operator valued functions.* 1973.

45 A.A. Balkema. *Monotone transformations and limit laws.* 1973.

46 R.P. van de Riet. *ABC ALGOL, a portable language for formula manipulation systems, part 1: the language.* 1973.

47 R.P. van de Riet. *ABC ALGOL, a portable language for formula manipulation systems, part 2: the compiler.* 1973.

48 F.E.J. Kruseman Aretz, P.J.W. ten Hagen, H.L. Oudshoorn. *An ALGOL 60 compiler in ALGOL 60, text of the MC-compiler for the EL-X8.* 1973.

49 H. Kok. *Connected orderable spaces.* 1974.

50 A. van Wijngaarden, B.J. Mailloux, J.E.L. Peck, C.H.A. Koster, M. Sintzoff, C.H. Lindsey, L.G.L.T. Meertens, R.G. Fisker (eds.). *Revised report on the algorithmic language ALGOL 68.* 1976.

51 A. Hordijk. *Dynamic programming and Markov potential theory.* 1974.

52 P.C. Baayen (ed.). *Topological structures.* 1974.

53 M.J. Faber. *Metrizability in generalized ordered spaces.* 1974.

54 H.A. Lauwerier. *Asymptotic analysis, part 1.* 1974.

55 M. Hall, Jr., J.H. van Lint (eds.). *Combinatorics, part 1: theory of designs, finite geometry and coding theory.* 1974.

56 M. Hall, Jr., J.H. van Lint (eds.). *Combinatorics, part 2: graph theory, foundations, partitions and combinatorial geometry.* 1974.

57 M. Hall, Jr., J.H. van Lint (eds.). *Combinatorics, part 3: combinatorial group theory.* 1974.

58 W. Albers. *Asymptotic expansions and the deficiency concept in statistics.* 1975.

59 J.L. Mijnheer. *Sample path properties of stable processes.* 1975.

60 F. Göbel. *Queueing models involving buffers.* 1975.

63 J.W. de Bakker (ed.). *Foundations of computer science.* 1975.

64 W.J. de Schipper. *Symmetric closed categories.* 1975.

65 J. de Vries. *Topological transformation groups, 1: a categorical approach.* 1975.

66 H.G.J. Pijls. *Logically convex algebras in spectral theory and eigenfunction expansions.* 1976.

68 P.P.N. de Groen. *Singularly perturbed differential operators of second order.* 1976.

69 J.K. Lenstra. *Sequencing by enumerative methods.* 1977.

70 W.P. de Roever, Jr. *Recursive program schemes: semantics and proof theory.* 1976.

71 J.A.E.E. van Nunen. *Contracting Markov decision processes.* 1976.

72 J.K.M. Jansen. *Simple periodic and non-periodic Lamé functions and their applications in the theory of conical waveguides.* 1977.

73 D.M.R. Leivant. *Absoluteness of intuitionistic logic.* 1979.

74 H.J.J. te Riele. *A theoretical and computational study of generalized aliquot sequences.* 1976.

75 A.E. Brouwer. *Treelike spaces and related connected topological spaces.* 1977.

76 M. Rem. *Associons and the closure statement.* 1976.

77 W.C.M. Kallenberg. *Asymptotic optimality of likelihood ratio tests in exponential families.* 1978.

78 E. de Jonge, A.C.M. van Rooij. *Introduction to Riesz spaces.* 1977.

79 M.C.A. van Zuijlen. *Emperical distributions and rank statistics.* 1977.

80 P.W. Hemker. *A numerical study of stiff two-point boundary problems.* 1977.

81 K.R. Apt, J.W. de Bakker (eds.). *Foundations of computer science II, part 1.* 1976.

82 K.R. Apt, J.W. de Bakker (eds.). *Foundations of computer science II, part 2.* 1976.

83 L.S. van Benthem Jutting. *Checking Landau's "Grundlagen" in the AUTOMATH system.* 1979.

84 H.L.L. Busard. *The translation of the elements of Euclid from the Arabic into Latin by Hermann of Carinthia (?), books vii-xii.* 1977.

85 J. van Mill. *Supercompactness and Wallman spaces.* 1977.

86 S.G. van der Meulen, M. Veldhorst. *Torrix I, a programming system for operations on vectors and matrices over arbitrary fields and of variable size.* 1978.

88 A. Schrijver. *Matroids and linking systems.* 1977.

89 J.W. de Roever. *Complex Fourier transformation and analytic functionals with unbounded carriers.* 1978.

90 L.P.J. Groenewegen. *Characterization of optimal strategies in dynamic games.* 1981.

91 J.M. Geysel. *Transcendence in fields of positive characteristic.* 1979.

92 P.J. Weeda. *Finite generalized Markov programming.* 1979.

93 H.C. Tijms, J. Wessels (eds.). *Markov decision theory.* 1977.

94 A. Bijlsma. *Simultaneous approximations in transcendental number theory.* 1978.

95 K.M. van Hee. *Bayesian control of Markov chains.* 1978.

96 P.M.B. Vitányi. *Lindenmayer systems: structure, languages, and growth functions.* 1980.

97 A. Federgruen. *Markovian control problems; functional equations and algorithms.* 1984.

98 R. Geel. *Singular perturbations of hyperbolic type.* 1978.

99 J.K. Lenstra, A.H.G. Rinnooy Kan, P. van Emde Boas (eds.). *Interfaces between computer science and operations research.* 1978.

100 P.C. Baayen, D. van Dulst, J. Oosterhoff (eds.). *Proceedings bicentennial congress of the Wiskundig Genootschap, part 1.* 1979.

101 P.C. Baayen, D. van Dulst, J. Oosterhoff (eds.). *Proceedings bicentennial congress of the Wiskundig Genootschap, part 2.* 1979.

102 D. van Dulst. *Reflexive and superreflexive Banach spaces.* 1978.

103 K. van Harn. *Classifying infinitely divisible distributions by functional equations.* 1978.

104 J.M. van Wouwe. *Go-spaces and generalizations of metrizability.* 1979.

105 R. Helmers. *Edgeworth expansions for linear combinations of order statistics.* 1982.

106 A. Schrijver (ed.). *Packing and covering in combinatorics.* 1979.

107 C. den Heijer. *The numerical solution of nonlinear operator equations by imbedding methods.* 1979.

108 J.W. de Bakker, J. van Leeuwen (eds.). *Foundations of computer science III, part 1.* 1979.

109 J.W. de Bakker, J. van Leeuwen (eds.). *Foundations of computer science III, part 2.* 1979.

110 J.C. van Vliet. *ALGOL 68 transput, part I: historical review and discussion of the implementation model.* 1979.

111 J.C. van Vliet. *ALGOL 68 transput, part II: an implementation model.* 1979.

112 H.C.P. Berbee. *Random walks with stationary increments and renewal theory.* 1979.

113 T.A.B. Snijders. *Asymptotic optimality theory for testing problems with restricted alternatives.* 1979.

114 A.J.E.M. Janssen. *Application of the Wigner distribution to harmonic analysis of generalized stochastic processes.* 1979.

115 P.C. Baayen, J. van Mill (eds.). *Topological structures II, part 1.* 1979.

116 P.C. Baayen, J. van Mill (eds.). *Topological structures II, part 2.* 1979.

117 P.J.M. Kallenberg. *Branching processes with continuous state space.* 1979.

118 P. Groeneboom. *Large deviations and asymptotic efficiencies.* 1980.

119 F.J. Peters. *Sparse matrices and substructures, with a novel implementation of finite element algorithms.* 1980.

120 W.P.M. de Ruyter. *On the asymptotic analysis of large-scale ocean circulation.* 1980.

121 W.H. Haemers. *Eigenvalue techniques in design and graph theory.* 1980.

122 J.C.P. Bus. *Numerical solution of systems of nonlinear equations.* 1980.

123 I. Yuhász. *Cardinal functions in topology - ten years later.* 1980.

124 R.D. Gill. *Censoring and stochastic integrals.* 1980.

125 R. Eising. *2-D systems, an algebraic approach.* 1980.

126 G. van der Hoek. *Reduction methods in nonlinear programming.* 1980.

127 J.W. Klop. *Combinatory reduction systems.* 1980.

128 A.J.J. Talman. *Variable dimension fixed point algorithms and triangulations.* 1980.

129 G. van der Laan. *Simplicial fixed point algorithms.* 1980.

130 P.J.W. ten Hagen, T. Hagen, P. Klint, H. Noot, H.J. Sint, A.H. Veen. *ILP: intermediate language for pictures.* 1980.

131 R.J.R. Back. *Correctness preserving program refinements: proof theory and applications.* 1980.

132 H.M. Mulder. *The interval function of a graph.* 1980.

133 C.A.J. Klaassen. *Statistical performance of location estimators.* 1981.

134 J.C. van Vliet, H. Wupper (eds.). *Proceedings international conference on ALGOL 68.* 1981.

135 J.A.G. Groenendijk, T.M.V. Janssen, M.J.B. Stokhof (eds.). *Formal methods in the study of language, part I.* 1981.

136 J.A.G. Groenendijk, T.M.V. Janssen, M.J.B. Stokhof (eds.). *Formal methods in the study of language, part II.* 1981.

137 J. Telgen. *Redundancy and linear programs.* 1981.

138 H.A. Lauwerier. *Mathematical models of epidemics.* 1981.

139 J. van der Wal. *Stochastic dynamic programming, successive approximations and nearly optimal strategies for Markov decision processes and Markov games.* 1981.

140 J.H. van Geldrop. *A mathematical theory of pure exchange economies without the no-critical-point hypothesis.* 1981.

141 G.E. Welters. *Abel-Jacobi isogenies for certain types of Fano threefolds.* 1981.

142 H.R. Bennett, D.J. Lutzer (eds.). *Topology and order structures, part 1.* 1981.

143 J.M. Schumacher. *Dynamic feedback in finite- and infinite-dimensional linear systems.* 1981.

144 P. Eijgenraam. *The solution of initial value problems using interval arithmetic; formulation and analysis of an algorithm.* 1981.

145 A.J. Brentjes. *Multi-dimensional continued fraction algorithms.* 1981.

146 C.V.M. van der Mee. *Semigroup and factorization methods in transport theory.* 1981.

147 H.H. Tigelaar. *Identification and informative sample size.* 1982.

148 L.C.M. Kallenberg. *Linear programming and finite Markovian control problems.* 1983.

149 C.B. Huijsmans, M.A. Kaashoek, W.A.J. Luxemburg, W.K. Vietsch (eds.). *From A to Z, proceedings of a symposium in honour of A.C. Zaanen.* 1982.

150 M. Veldhorst. *An analysis of sparse matrix storage schemes.* 1982.

151 R.J.M.M. Does. *Higher order asymptotics for simple linear rank statistics.* 1982.

152 G.F. van der Hoeven. *Projections of lawless sequences.* 1982.

153 J.P.C. Blanc. *Application of the theory of boundary value problems in the analysis of a queueing model with paired services.* 1982.

154 H.W. Lenstra, Jr., R. Tijdeman (eds.). *Computational methods in number theory, part I.* 1982.

155 H.W. Lenstra, Jr., R. Tijdeman (eds.). *Computational methods in number theory, part II.* 1982.

156 P.M.G. Apers. *Query processing and data allocation in distributed database systems.* 1983.

157 H.A.W.M. Kneppers. *The covariant classification of two-dimensional smooth commutative formal groups over an algebraically closed field of positive characteristic.* 1983.

158 J.W. de Bakker, J. van Leeuwen (eds.). *Foundations of computer science IV, distributed systems, part 1.* 1983.

159 J.W. de Bakker, J. van Leeuwen (eds.). *Foundations of computer science IV, distributed systems, part 2.* 1983.

160 A. Rezus. *Abstract AUTOMATH.* 1983.

161 G.F. Helminck. *Eisenstein series on the metaplectic group, an algebraic approach.* 1983.

162 J.J. Dik. *Tests for preference.* 1983.

163 H. Schippers. *Multiple grid methods for equations of the second kind with applications in fluid mechanics.* 1983.

164 F.A. van der Duyn Schouten. *Markov decision processes with continuous time parameter.* 1983.

165 P.C.T. van der Hoeven. *On point processes.* 1983.

166 H.B.M. Jonkers. *Abstraction, specification and implementation techniques, with an application to garbage collection.* 1983.

167 W.H.M. Zijm. *Nonnegative matrices in dynamic programming.* 1983.

168 J.H. Evertse. *Upper bounds for the numbers of solutions of diophantine equations.* 1983.

169 H.R. Bennett, D.J. Lutzer (eds.). *Topology and order structures, part 2.* 1983.