A note on general sliding window processes

Noga Alon [*] Ohad N. Feldheim [†]

September 19, 2014

Abstract

Let $f : \mathbb{R}^k \to [r] = \{1, 2, \ldots, r\}$ be a measurable function, and let $\{U_i\}_{i\in\mathbb{N}}$ be a sequence of i.i.d. random variables. Consider the random process $Z_i = f(U_i, ..., U_{i+k-1})$. We show that for all q, there is a positive probability, uniform in f, that $Z_1 = Z_2 = ... = Z_q$. A continuous counterpart is that if $f : \mathbb{R}^k \to \mathbb{R}$, and U_i and Z_i are as before, then there is a positive probability, uniform in f, for $Z_1, ..., Z_q$ to be monotone. We prove these theorems, give upper and lower bounds for this probability, and generalize to variables indexed on other lattices.

The proof is based on an application of combinatorial results from Ramsey theory to the realm of continuous probability.

1 Introduction

The objective of this note is to bring to the attention of probabilists an application of tools from Ramsey theory to probabilistic questions on sliding window processes. These results can be further extended with relative ease to other noise type variables (see Section 4). Let $f : \mathbb{R}^n \to \mathbb{R}$ be a measurable function, and let $\{U_i\}_{i\in\mathbb{N}}$ be a sequence of i.i.d. random variables. Consider the random process $Z_i^f = f(U_i, ..., U_{i+k-1})$. Such processes are called *k-block factors*.

Our main observation is the following result dealing with functions f with finite image.

Theorem 1. *For every $k, q, r \in \mathbb{N}$ there exists $p = p_{k,q,r} > 0$ such that for every measurable $f : \mathbb{R}^k \to \{1, \ldots, r\}$ the following holds:*

$$\mathbb{P}(Z_1^f = Z_2^f = \cdots = Z_q^f) > p$$

[*]Sackler School of Mathematics and Blavatnik School of Computer Science, Tel Aviv University, Tel Aviv 69978, Israel. Email: nogaa@tau.ac.il. Research supported in part by a USA-Israeli BSF grant, by an ISF grant, by the Hermann Minkowski Minerva Center for Geometry at Tel Aviv University and by the Israeli I-Core program.

[†]Department of Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel. Email: ohad_f@netvision.net.il. Partially supported by a grant from the Israel Science Foundation.

In other words, the probability of every k-factor to be constant on a discrete interval of length q is bounded away from zero.

The particular case $r = 2$ which motivated our interest in the problem, is presented in the following corollary.

Corollary 2. *For every $k, q \in \mathbb{N}$ there exists $p = p_{k,q} > 0$ such that for every measurable $f : \mathbb{R}^k \to \{0, 1\}$ the following holds:*

$$\mathbb{P}(Z_1^f = Z_2^f = \cdots = Z_q^f) > p$$

A more general corollary is the following continuous counterpart:

Theorem 3. *For every $k, q \in \mathbb{N}$ there exists $p = p_{k,q} > 0$ such that for every measurable $f : \mathbb{R}^k \to \mathbb{R}$, one of the following holds:*

$$\begin{array}{ll} \textit{either} & \mathbb{P}(Z_1^f < Z_2^f < \cdots < Z_q^f) > p, \\ \textit{or} & \mathbb{P}(Z_1^f = Z_2^f = \cdots = Z_q^f) > p, \\ \textit{or} & \mathbb{P}(Z_1^f > Z_2^f > \cdots > Z_q^f) > p. \end{array}$$

Proof of Theorem 3 using Theorem 1. Let $k, q \in \mathbb{N}$, and let $f : \mathbb{R}^k \to \mathbb{R}$. We define a new function $g : \mathbb{R}^{k+1} \to \{-1, 0, 1\}$ by

$$g(x_1, ..., x_{k+1}) = \begin{cases} -1 & f(x_1, ..., x_k) > f(x_2, ..., x_{k+1}) \\ 0 & f(x_1, ..., x_k) = f(x_2, ..., x_{k+1}) \\ 1 & f(x_1, ..., x_k) < f(x_2, ..., x_{k+1}) \end{cases}$$

By Theorem 1 there exists a positive $p = p_{k+1,q-1,3}$ such that one of the following holds:

$$\begin{array}{ll} \text{either} & \mathbb{P}(Z_1^g = Z_2^g = \cdots = Z_{q-1}^g = -1) > p, \\ \text{or} & \mathbb{P}(Z_1^g = Z_2^g = \cdots = Z_{q-1}^g = 0) > p, \\ \text{or} & \mathbb{P}(Z_1^g = Z_2^g = \cdots = Z_{q-1}^g = 1) > p. \end{array}$$

The theorem follows. $\square$

The decay of p as a function of k in the above theorems is very fast. In particular the p which Theorem 1 yields is $\frac{1}{M^{k+q-1}}$ for M satisfying

$$M = 1 + \underbrace{2^{2^{\cdot^{\cdot^{\cdot^{2^{q^r}}}}}}}_{k-2 \text{ times}}.$$

Such a tower dependency is in fact essential as the following proposition shows:

Theorem 4. *For all large enough k, there exists $f : [0,1]^k \to \{0,1\}$ such that for the process $\{Z_i^f\}$ defined as above with respect to uniform variables U_i on $[0,1]$ the following holds:*

$$\mathbb{P}(Z_1^f = Z_2^f = \cdots = Z_{2k}^f) < \frac{9k^2}{M},$$

where

$$M = \underbrace{2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}}_{\substack{k-2 \\ times}}{}^{\lceil k/\sqrt{8}\rceil}.$$

2 Background and motivation

The research of k-block factors originated as a part of a wider attempt to understand m-*dependent* processes. These generalize independent processes in discrete time, by requiring that every two events which are separated by a time-interval with length more than m will be independent. Such processes arise naturally as scaling limits in renormalization theory (see for example [2]). Clearly, every k-block factor is $(k-1)$-dependent. For a while the converse was also conjectured to hold, to the extent that in certain papers, results on k-block factors are presented as results on $(k-1)$-dependent processes, conditioned on the validity of the conjecture (see for example [6]).

While for Gaussian processes, every m-dependent process is indeed an $m+1$-block factor, we now know that for general m-dependent processes this is not true. Ibragimov and Linnik have already stated in 1971 that there should exist a 1-dependent process which is not a 2-block factor, but provided no example. The first example was published by Aaronson and Gilat in [1] in 1987. Later, in [3], Burton, Goulet and Meester showed that there exists a 1-dependent process which is not a k-block factor for any k.

One property of binary block factors, i.e., block factors with range $\{0,1\}$, which have been extensively studied, is the probability of observing r consecutive occurrences of the value b in the process. This event is called an r-*run* of b-s. Janson, in [6], studied the convergence of the statistics of runs of zeros in a k-factor in which every two ones are guaranteed to be separated by $k-1$ zeros. De Valk, in [8], computed the minimal and maximal possible probability of a 2-run of ones given the marginal probability of seeing the value one. Such studies give rise to the following natural question: is it possible to create a binary k-block factor for some k which almost surely has neither an r-run of zeros nor an r-run of ones? Here we show that this is impossible. The result is twofold. On one hand, the probability of seeing an arbitrarily long run is bounded away from zero. On the other hand, it can be extremely small.

3 Proof of the results

This section is dedicated to the proofs of Theorems 1 and 4. For this purpose we shall use a classical result on de-Bruijn graphs (Theorem A), whose proof we present for completeness sake. Throughout this section we shall set the distribution of $\{U_i\}$ to be uniform on $(0,1)$. Observe that this restriction does not limit the generality of our proofs, since every random variable is the image of a uniform variable on $(0,1)$ through some function.

For a directed graph G let $\chi(G)$, the chromatic number of G, denote the minimal number of colors required to color the vertices of G so that no two adjacent vertices get the same color.

Define $D(k,m)$, the increasing k-dimensional de-Bruijn graph of m symbols, to be the directed graph whose vertices are all the strictly increasing sequences of length k with elements in $\{1,\dots,m\}$, such that there is a directed edge from the sequence $\{a_1,...,a_k\}$ to the sequence $\{b_1,...,b_k\}$ if and only if $b_i = a_{i+1}$ for all $i \in \{1,\dots,k-1\}$.

We shall make use of the fact that $D(k+1,m)$ is the directed line-graph of $D(k,m)$. That is - that the map $\phi : (a_1,...,a_{k+1}) \to ((a_1,...,a_k),(a_2,...,a_{k+1}))$ is a bijection, mapping every vertex of $D(k+1,m)$ to an edge of $D(k,m)$.

Theorem A. $\log_2 \chi(D(k,m)) \le \chi(D(k+1,m))$.

Proof. Using the fact that $D(k+1,m)$ is the directed line-graph of $D(k,m)$, we get that a proper vertex coloring of $D(k+1,m)$ is equivalent to an edge coloring of $D(k,m)$ in which there is no monochromatic directed path of length 2. Thus, it is enough to show that for every such coloring of $E(D(k,m))$ using q colors, there exists a proper coloring of $V(D(k,m))$ using 2^q colors.

Let $C : E(D(k,m)) \to \{1,\dots,q\}$ be an edge-coloring of $D(k,m)$ as above. Construct $C' : V(D(k,m)) \to \mathcal{P}\{1,\dots,q\}$ using the subsets of $\{1,\dots,q\}$ as colors in the following way. Define $C'(u) = \{C(u,v) \ : \ (u,v) \in E(D(k,m))\}$. To see that C' is a proper vertex coloring, observe that if $C'(u) = C'(v)$ and $(u,v) \in E(D(k,m))$ then $C(u,v) \in C'(v)$ which implies the existence of (v,w) such that $C(v,w) = C(u,v)$, in contradiction to our premises. $\square$

Since clearly $\chi(D(1,m)) = m$, we get that for $k \ge 2$,

$$\chi(D(k,m)) \ge \log_2^{(k-1)}(m), \tag{1}$$

where $\log_2^{(k)}$ represents k iterations of the function $\log_2$.

We now use the following theorem by Chvátal [4].

Theorem B (Chvátal)**.** *Let D be a directed graph and let $q,r \in \mathbb{N}$ satisfy $\chi(D) > q^r$; then any edge-coloring of D with r colors contains a monochromatic directed path of q edges.*

Combining this with the fact that $D(k,m)$ is the directed line-graph of $D(k-1,m)$ and with (1), we draw the following corollary.

Corollary 5. *Given $k, q, r \in \mathbb{N}$, let M be an integer such that $\log_2^{(k-2)}(M) > q^r$. Then any r-coloring of the vertices of $D(k, M)$ contains a monochromatic directed path of q vertices.*

Using this we are ready to prove Theorem 1.

Proof of Theorem 1. Let $r \in \mathbb{N}$, let $f : [0,1]^k \to \{1, \ldots, r\}$ be a measurable function, U_i uniform on $[0,1]$ and $\{Z_i^f = f(U_i, ..., U_{i+k-1})\}_{i \in \mathbb{Z}}$. Choose $M = M(k, q, r)$ as in Corollary 5 to get:

$$\mathbb{P}\left(X_1 = \cdots = X_q\right) = \int_0^1 dx_1 \cdots \int_0^1 dx_{q+k-1} \ \mathbb{1}\left\{f(x_1, \ldots, x_k) = \cdots = f(x_q, \ldots, x_{q+k-1})\right\}$$
$$= \int_0^1 dy_1 \cdots \int_0^1 dy_M \ \frac{(M-k-q+1)!}{M!} \sum_{1 \le j_1 < \cdots < j_{q+k-1} \le M} \mathbb{1}\left\{f(y_{j_1}, \ldots, y_{j_k}) = \cdots = f(y_{j_q}, \ldots, y_{j_{q+k-1}})\right\}.$$

where the equality in the second line is obtained by first picking M random i.i.d. values in $[0,1]$ and then by assigning a random set of $q + k - 1$ of them to the variables $x_1, \ldots, x_{q+k-1}$ uniformly at random.

Now, for a given $\bar{y} = (y_1, \ldots, y_M) \in [0,1]^M$, the inner sum counts the number of monochromatic directed paths in $D(k, M)$ for the coloring

$$c(a_1, \ldots, a_k) = f(y_{a_1}, \ldots, y_{a_k}).$$

This is an r-coloring, therefore by the above corollary, this inner sum is at least 1. We conclude that

$$\mathbb{P}\left(X_1 = \cdots = X_q\right) \ge \frac{(M-k-q+1)!}{M!} > \frac{1}{M^{k+q-1}},$$

as required. $\square$

3.1 Tower dependency is essential

This subsection contains the proof of Theorem 4.

For $i > 1$, the 2-tower function, t_i, denotes the function satisfying $t_i(k) = 2^{t_{i-1}(k)}$, and $t_1(k) = k$. Also, recall the notation $D(k, m)$ of the increasing k-dimensional de-Bruijn graph of m symbols which is defined in the beginning of this section.

In our proof we use the following lemma of Moshkovitz and Shapira (see [7, Corallary 3]).

Lemma C. *There exists $n_0 \in \mathbb{N}$ such that for any $k \ge 3$, $q \ge 2$ and $n > n_0$, there exists an edge coloring of $D(k, t_{k-1}(\lceil n^{q-1}/\sqrt{8}\rceil))$ with q colors which contains no monochromatic path of n edges.*

Recalling that edge colorings of $D(k-1, m)$ are the same as vertex colorings of $D(k, m)$, and plugging $q = 2$, $n = k$ in Lemma C, we get the following useful proposition.

Proposition 6. *For every large enough k, there exists a vertex 2-coloring of $D(k, t_{k-2}(\lceil k/\sqrt{8}\rceil))$ such that no path of length k is monochromatic.*

We are now ready to prove Theorem 4.

Proof of Theorem 4. Let $M = t_{k-2}(\lceil k/\sqrt{8}\rceil)$, and let g be a vertex 2-coloring of $D(k, M)$ by the colors $\{0, 1\}$ such that no path of length k is monochromatic, which exists by Proposition 6. Define $h : \{1, \dots, M\}^k \to \{0, 1\}$ as follows:

$$h(z_1, \dots, z_k) = \begin{cases} g(z_1, \dots, z_k) & z_1 < \cdots < z_k \\ g(z_k, \dots, z_1) & z_1 > \cdots > z_k \\ 0 & \exists i \neq j \text{ s.t. } z_i = z_j \\ \alpha(z_2, z_3) & \text{otherwise,} \end{cases} \tag{2}$$

where $\alpha(x, y)$ takes the value 0 if $x < y$, and 1 otherwise.

Let $z_1, \dots, z_{3k}$ be distinct integers in $\{1, \dots, M\}$. We claim that the following is impossible:

$$h(z_1, \dots, z_k) = h(z_2, \dots, z_{k+1}) = \cdots = h(z_{2k+1}, \dots, z_{3k}). \tag{3}$$

Assuming the contrary, we study two cases.

The first case is when $z_2, \dots, z_{2k}$ is monotone. In this case h is equal to g along the path

$$(z_2, \dots, z_{k+1}), (z_3, \dots, z_{k+2}), \dots, (z_{k+1}, \dots, z_{2k})$$

(by the first or second case of the definition (2)). This is a contradiction, since g cannot be constant along a path of length k in $D(k, M)$.

In the complimentary case there exists a local extremum among $z_3, \dots, z_{2k-1}$, i.e., there exists $i \in \{3, \dots, 2k-1\}$ such that either $z_i > \max\{z_{i-1}, z_{i+1}\}$ or $z_i < \min\{z_{i-1}, z_{i+1}\}$. Thus, the values

$$\begin{aligned} &h(z_{i-2}, z_{i-1}, \dots, z_{i+k-3}) = \alpha(z_{i-1}, z_i), \text{ and} \\ &h(z_{i-1}, z_i, \dots, z_{i+k-2}) = \alpha(z_i, z_{i+1}) \end{aligned}$$

are not equal, which also leads to a contradiction.

Now, observe that taking uniform distribution over $1, \dots, M$ the probability that $(z_i)_{i \in \{1,\dots,3k\}}$ are distinct is greater than

$$\prod_{j=0}^{3k-1} \left(1 - \frac{j}{M}\right) > 1 - \frac{9k^2}{M}.$$

We may therefore define $f : [0, 1]^k \to \{0, 1\}$ to be $f(x_1, \dots, x_k) = h(\lceil Mx_1 \rceil, \dots, \lceil Mx_k \rceil)$ and get

$$\mathbb{P}(Z_1^f = Z_2^f = \cdots = Z_{2k}^f) < \frac{9k^2}{M}$$

as required. $\square$

4 Generalizations

Theorem 1 can be generalized with relative ease to factors of other spaces such as the discrete lattice $\mathbb{Z}^d$ or the infinite binary tree. In this section we prove a general theorem which can be used to handle these cases and demonstrate its use. However, unlike Theorems 1 and 3, the proof method described here often yields bounds which are far from being tight.

Let $k, q, r \in \mathbb{N}$ and let $Z = \{Z_v\}_{v\in[q]}$ be a collection of i.i.d. random variables. We write $\mathcal{S}_{k,q} := \{S \subset [q] \ : \ |S| = k\}$. Given a function $f : \mathbb{R}^k \to [r]$ and $\{s_1, \ldots, s_k\} \in \mathcal{S}_q$ such that $s_1 < \cdots < s_k$, we write $f_Z(S) := f(Z_{s_1}, \ldots, Z_{s_k})$. The following is a generalization of Theorem 1.

Theorem 7. *There exists $p_{k,q,r} > 0$ such that for any function $f : \mathbb{R}^k \to [r]$ the following holds.*

$$\mathbb{P}\big(\forall S, S' \in \mathcal{S}_{k,q} \text{ we have } f_Z(S) = f_Z(S')\big) > p.$$

Proof. The proof of Theorem 7 is similar to the proof of Theorem 1. For any $M > q$ we write

$$\begin{aligned}
&\mathbb{P}\big(\forall S, S' \in \mathcal{S}_{k,q} \text{ we have } f(S) = f(S')\big)\\
&= \int_0^1 dx_1 \cdots \int_0^1 dx_q \ \mathbb{1}\left\{\forall\, {}^{1\le i_1<i_2<\cdots<i_k\le q}_{1\le j_1<j_2<\cdots<j_k\le q}, f(x_{i_1}, \ldots, x_{i_k}) = f(x_{j_1}, \ldots, x_{j_k}))\right\}\\
&= \int_0^1 dy_1 \cdots \int_0^1 dy_M \sum_{1\le P_1<\cdots<P_q\le M} \binom{M}{q}^{-1} \mathbb{1}\left\{\forall\, {}^{1\le i_1<i_2<\cdots<i_k\le q}_{1\le j_1<j_2<\cdots<j_k\le q}, f(y_{P_{i_1}}, \ldots, y_{P_{i_k}}) = f(y_{P_{j_1}}, \ldots, y_{P_{j_k}})\right\}.
\end{aligned}$$

We then use Ramsey theorem for hypergraphs (which plays here the role of Corollary 5.)

Theorem D (Ramsey, c.f., e.g., [5])**.** *For all $k, q, r \in \mathbb{N}$ there exists large $M \in \mathbb{N}$ such that for every r-coloring of the k-subsets of $[M]$ there exists a $Y \subset [M]$ of size q, which satisfies that all k-subsets of Y are colored by the same color.*

Viewing f as a coloring of the k-subsets of $y_1, \ldots y_M$, this allows us to deduce that for large enough M there exists a subset $Y = \{y_{P_1}, \ldots, y_{P_q}\}$ such that f is monochromatic on all the k-subsets of Y. We deduce that $p_{k,q,r} > \binom{M}{q}^{-1}$. □

Applications.

We can use Theorem 7 to generalize Theorems 1 and 3 to other spaces, such as the discrete lattice $\mathbb{Z}^d$ or the infinite binary tree. Let us demonstrate this on the following example. Let $\{U_{x,y}\}_{x,y\in\mathbb{Z}}$ be a sequence of i.i.d. random variables. Given $k, r \in \mathbb{N}$ and $f : \mathbb{R}^{k^2} \to [r]$ we write

$$Z^f_{i,j} = f\begin{pmatrix} U_{i,j} & U_{i+1,j} & \cdots & U_{i+k-1,j} \\ U_{i,j+1} & U_{i+1,j+1} & \cdots & U_{i+k-1,j+1} \\ \cdots & \cdots & \cdots & \cdots \\ U_{i,j+k-1} & U_{i+1,j+k-1} & \cdots & U_{i+k-1,j+k-1} \end{pmatrix}.$$

We get the following.

Corollary 8. *Let $k, q, r \in \mathbb{N}$. There exists $p = p_{k,q,r}$ such that for every function $f : \mathbb{R}^{k^2} \to [r]$, we have*

$$\mathbb{P}(\forall i, j, i', j' \in [q] \ : Z^f_{i,j} = Z^f_{i',j'}) > p$$

To see this we order the elements of $[q + k - 1]^2$ lexicographically, i.e, $U(i, j)$ precedes $U(i', j')$ if and only if either $i > i'$ or $i = i'$ and $j > j'$. Under this order $[q + k - 1]^2$ is isomorphic to $[(q+k-1)^2]$. We then observe that f is always applied to elements whose indices are ordered by this order. Theorem 7 implies that the probability that f assigns the same value to all ordered subsets of size k^2 of $[q+k-1]^2$ is bounded away from zero uniformly for all functions f, which is a stronger claim than that of Corollary 8.

5 Acknowledgments

We wish to thank Péter Mester and Boris Tsirelson for useful discussions, which ignited and fanned our interest in the problem.

References

[1] J. Aaronson, D. Gilat, "On the structure of stationary one dependent processes", School of Mathematical Sciences, Tel Aviv University, Israel, 1987.

[2] G.L. O'Brien, "Scaling transformations for 0,1-valued sequences", *Zeit. Wa. hr.* 53, 35–19, 1980.

[3] R. M. Burton, M. Goulet, R. Meester, "On 1-Dependent Processes and k-Block Factors", *Ann. Prob.* 21:4, 2157–2168, 1993.

[4] V. Chvátal, "Monochromatic paths and circuits in edge-colored graphs", *J. Combinatorial Theory Ser. B* 10, 299–300, 1971.

[5] R. L. Graham, B. L. Rothschild and J. H. Spencer, *Ramsey Theory* (Second Edition), Wiley, New York, 1990.

[6] S. Janson. "Runs in m-dependent sequences", *Ann. Prob.* 12:5, 805–818, 1984.

[7] G. Moshkovitz, A. Shapira, "Integer Partitions and a New Proof of the Erdos-Szekeres Theorem", *Eprint arXiv*:1206.4001 [math.CO].

[8] V. De Valk, "The maximal and minimal 2-correlation of a class of 1-dependent $0-1$ valued processes", *Israel J. of Math.* 62:2, 181–205, 1988.