Local Dependence and Persistence in Discrete Sliding Window Processes

Ohad N. Feldheim
Joint work with Noga Alon

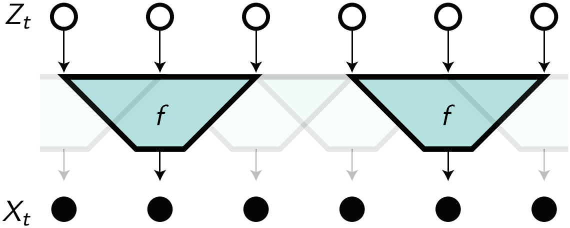
Weizmann Institute of Science

July 2014
Technishe Universitat Darmstadt

Sliding Window Processes

$\{Z_t\}_{t\in\mathbb{Z}} :=$ i.i.d. uniform on $[0,1]$.
$f : [0,1]^k \to \{0,\ldots,r-1\}$ measurable.

$$\{X_t\}_{t\in\mathbb{Z}} := f(Z_t, Z_{t+1}, \ldots, Z_{t+k-1}).$$



Such a process is called k-**block factor**. If $r=2$ we call it a **binary** k-block factor.

Applications

Sliding window processes have many real-life applications, e.g.,

Linguistics, Vocoding:	• Model for voiceless phonemes
Cryptography:	• Encryption schemes with parallel decryption
Computer science:	• Data processes by stateless machines • Distributive ring computation

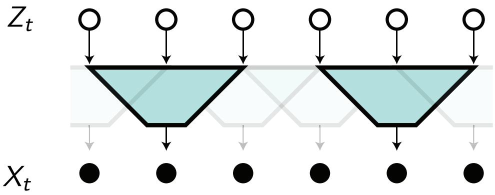
Local dependence

k-dependence for stationary processes

If every E_- which is $\{X_t\}_{t<0}$ measurable,
and every E_+ which is $\{X_t\}_{t\geq k}$ measurable are independent,
then $\{X_t\}$ is said to be k-**dependent**.

Observation

$k+1$-block factors are stationary k-dependent.



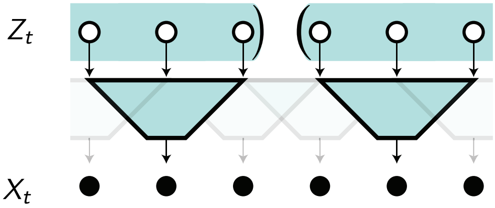
Local dependence

k-dependence for stationary processes

If every E_- which is $\{X_t\}_{t<0}$ measurable,
and every E_+ which is $\{X_t\}_{t\geq k}$ measurable are independent,
then $\{X_t\}$ is said to be k-**dependent**.

Observation

$k+1$-block factors are stationary k-dependent.



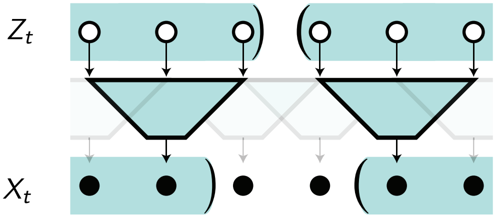
Local dependence

k-dependence for stationary processes

If every E_- which is $\{X_t\}_{t<0}$ measurable,
and every E_+ which is $\{X_t\}_{t\geq k}$ measurable are independent,
then $\{X_t\}$ is said to be k-**dependent**.

Observation

$k+1$-block factors are stationary k-dependent.



Some previous results on block factors

2-block factors

Katz, 1971 Computed $\max \mathbb{P}(X_1 = X_2 = 1)$ given $\mathbb{P}(X_1 = 1)$.
De Valk, 1988 Computed $\min \mathbb{P}(X_1 = X_2 = 1)$ given $\mathbb{P}(X_1 = 1)$ and showed uniqueness of the minimal and maximal processes. He did this also for general 1-dependent processes.

Some previous results on block factors

2-block factors

Katz, 1971 Computed $\max \mathbb{P}(X_1 = X_2 = 1)$ given $\mathbb{P}(X_1 = 1)$.
De Valk, 1988 Computed $\min \mathbb{P}(X_1 = X_2 = 1)$ given $\mathbb{P}(X_1 = 1)$ and showed uniqueness of the minimal and maximal processes. He did this also for general 1-dependent processes.

k-block factors

Janson, 1984: Explored several examples of binary k-block factors with at least $k-1$ zeroes between consecutive ones, and showed convergence of the gaps between consecutive ones for such processes.

Persistence

A natural definition of **persistence** in a frame of size q, for processes with discrete image:

$$P_q^X = \mathbb{P}(X_1 = X_2 = \cdots = X_q)$$

Coincides with the usual definition of persistence, if

$$f(Z_1, \ldots, Z_k) = \mathbb{1}\{g(Z_1, \ldots, Z_k) > 0\},$$

for some function g.

Persistence

A natural definition of **persistence** in a frame of size q, for processes with discrete image:

$$P_q^X = \mathbb{P}(X_1 = X_2 = \cdots = X_q)$$

Coincides with the usual definition of persistence, if

$$f(Z_1, \ldots, Z_k) = \mathbb{1}\{g(Z_1, \ldots, Z_k) > 0\},$$

for some function g.

Observation

X is non-constant k-dependent $\to \exists c > 0$ s. t. $P_q^X < e^{-cq}$

Persistence

A natural definition of **persistence** in a frame of size q, for processes with discrete image:

$$P_q^X = \mathbb{P}(X_1 = X_2 = \cdots = X_q)$$

Coincides with the usual definition of persistence, if

$$f(Z_1, \ldots, Z_k) = \mathbb{1}\{g(Z_1, \ldots, Z_k) > 0\},$$

for some function g.

Observation

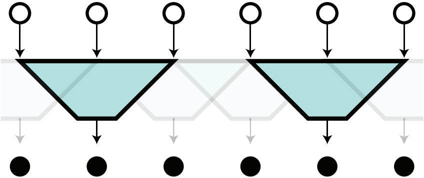
X is non-constant k-dependent $\to \exists c > 0$ s. t. $P_q^X < e^{-cq}$

But what about a lower bound?

Lower bound if $Z_t \in \{0, \ldots, \ell - 1\}$

Observation

If we had $Z_t \in \{0, \ldots, \ell - 1\}$ it would imply $\ell^{-(q+k-1)} < P_q^X$.



Somewhat unusual question

Usually: low correlation $\not\to$ lower bound on persistence.

Somewhat unusual question

Usually: low correlation $\not\to$ lower bound on persistence.

Lower bound on block-factor persistence $\longleftrightarrow$

There is a universal constant $p_{k,q}$ such that every symmetric real sliding window process $\{X_t\}_{t\in\mathbb{Z}}$ with a given window size k must have:

$$\mathbb{P}\big(X_1, \ldots, X_q \in [0,\infty)\big) > p_{k,q}$$

Somewhat unusual question

Usually: low correlation $\not\to$ lower bound on persistence.

Lower bound on block-factor persistence $\longleftrightarrow$

There is a universal constant $p_{k,q}$ such that every symmetric real sliding window process $\{X_t\}_{t\in\mathbb{Z}}$ with a given window size k must have:

$$\mathbb{P}\big(X_1, \ldots, X_q \in [0, \infty)\big) > p_{k,q}$$

There is a block factor with $P_q = 0$ for some q $\longleftrightarrow$

Somewhat unusual question

Usually: low correlation $\not\to$ lower bound on persistence.

Lower bound on block-factor persistence $\longleftrightarrow$

There is a universal constant $p_{k,q}$ such that every symmetric real sliding window process $\{X_t\}_{t\in\mathbb{Z}}$ with a given window size k must have:

$$\mathbb{P}\left(X_1,\dots,X_q \in [0,\infty)\right) > p_{k,q}$$

There is a block factor with $P_q = 0$ for some q $\longleftrightarrow$

Each of N players, standing in a row is assigned a random number uniform in $[0,1]$. By looking only on the numbers in their q neighborhood, using a symmetric algorithm, the players can divide themselves to consecutive pairs and triplets.

Our results

Let $k, q \in \mathbb{N}$. For $f : \mathbb{R}^k \to \{0,1\}$ write $X_t^f = f(Z_t, \dots, Z_{t+k-1})$ where Z_t are i.i.d, and write

$$p_q^{\min} = \inf_f \{\mathbb{P}(X_1^f = X_2^f = \cdots = X_q^f)\}$$

Theorem (Alon, F.)

$$\frac{1}{\left(T_{k-2}(q^2)\right)^{k+q-1}} < p_q^{\min} < \frac{1}{T_{k-2}\left(\frac{q}{100}\right)},$$

where

$$T_\ell(x) := \underbrace{2^{2^{2^{\cdot^{\cdot^{\cdot^{2^x}}}}}}}_{\ell \text{ times}}$$

Our results

Let $k, q \in \mathbb{N}$. For $f : \mathbb{R}^k \to \{0,1\}$
write $X_t^f = f(Z_t, \dots, Z_{t+k-1})$ where Z_t are i.i.d, and write

$$p_q^{\min} = \inf_f \{\mathbb{P}(X_1^f = X_2^f = \dots = X_q^f)\}$$

Theorem (Alon, F.)

$$\frac{1}{\left(T_{k-2}(q^2)\right)^{k+q-1}} < p_q^{\min} < \frac{1}{T_{k-2}(\frac{q}{100})},$$

where

$$T_\ell(x) := \underbrace{2^{2^{2^{\cdot^{\cdot^{\cdot^{2^x}}}}}}}_{\ell \text{ times}}$$

- Heavily involves Ramsey theory.

Is it possible to extend to k-dependent processes?

For upper bound on p_q^X we used only k-dependence. Can we do the same for the lower bound?

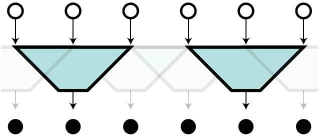
Is it possible to extend to k-dependent processes?

For upper bound on p_q^X we used only k-dependence. Can we do the same for the lower bound?

- **Does k-dependence imply being a $k+1$-block factor?**

Are the two properties equivalent

Does k-dependence imply being a $k+1$-block factor?



$k+1$-block factor

For $Z_t \sim U[0,1]$ i.i.d.
$\exists f : \mathbb{R} \to L$ such that

$$\{X_t\} \stackrel{\text{law}}{=} \{f(Z_t, Z_t, \ldots, Z_{t+k})\}$$

k-dependent

If E_- is $\{X_t\}_{t<0}$ measurable and E_+ is $\{X_t\}_{t\geq k}$ measurable, then

$$\mathbb{P}(E_-)\mathbb{P}(E_+) = \mathbb{P}(E_- \cap E_+)$$

History of this question

Does k-dependence imply being a $k+1$-block factor? (Ibragimov and Linnik '71)

History of this question

Does k-dependence imply being a $k+1$-block factor? (Ibragimov and Linnik '71)

- True for Gaussian processes.

History of this question

Does k-dependence imply being a $k+1$-block factor? (Ibragimov and Linnik ’71)

- True for Gaussian processes.
- In ’84 it was still conjectured to be true.

History of this question

The Annals of Probability
1984, Vol. 12, No. 3, 805–818

RUNS IN *m*-DEPENDENT SEQUENCES

BY SVANTE JANSON

Uppsala University

• • •

To obtain complete results we will impose one further condition.

(*) There exists a sequence $\{\xi_i\}$ of i.i.d. random variables and a measurable function α such that $I_i = \alpha(\xi_{i-m}, \cdots, \xi_i)$.

Obviously, any sequence $\{I_i\}$ satisfying (*) is m-dependent. It seems to be unknown whether the converse holds, i.e. whether every m-dependent stationary sequence may be thus represented. Hence it is conceivable that this condition is redundant.

• • •

History of this question

Does k-dependence imply being a $k+1$-block factor? (Ibragimov and Linnik '71)

- True for Gaussian processes.
- In '84 it was still conjectured to be true.
- Although Ibragimov and Linnik stated in '71 that a counter example should exist.

History of this question

Does k-dependence imply being a $k+1$-block factor? (Ibragimov and Linnik '71)

- True for Gaussian processes.
- In '84 it was still conjectured to be true.
- Although Ibragimov and Linnik stated in '71 that a counter example should exist.
- In '87 Aaronson and Gilat came up with a counter example, showing a 1-dependent process which is not a 2-block factor.

History of this question

Does k-dependence imply being a $k+1$-block factor? (Ibragimov and Linnik '71)

- True for Gaussian processes.
- In '84 it was still conjectured to be true.
- Although Ibragimov and Linnik stated in '71 that a counter example should exist.
- In '87 Aaronson and Gilat came up with a counter example, showing a 1-dependent process which is not a 2-block factor.
- In '93 Burton, Goulet and Meester found a 1-dependent process which is not a k-factor for any k.

History of this question

Does k-dependence imply being a $k+1$-block factor? (Ibragimov and Linnik '71)

- True for Gaussian processes.
- In '84 it was still conjectured to be true.
- Although Ibragimov and Linnik stated in '71 that a counter example should exist.
- In '87 Aaronson and Gilat came up with a counter example, showing a 1-dependent process which is not a 2-block factor.
- In '93 Burton, Goulet and Meester found a 1-dependent process which is not a k-factor for any k.
- In that year Tsirelson showed a quantum mechanical example of 1-dependent non-2-block factor process.

Is it possible to extend to k-dependent processes?

For upper bound on p_q^X we used only k-dependence. Can we do the same for the lower bound?

- **Does k-dependence imply being a $k+1$-block factor? - No.**

Is it possible to extend to k-dependent processes?

For upper bound on p_q^X we used only k-dependence. Can we do the same for the lower bound?

- **Does k-dependence imply being a $k+1$-block factor? - No.**
- **Can we extend our results to k-dependent processes?**

Is it possible to extend to k-dependent processes?

For upper bound on p_q^X we used only k-dependence. Can we do the same for the lower bound?

- **Does k-dependence imply being a $k+1$-block factor? - No.**
- **Can we extend our results to k-dependent processes? - No.**

Finitely dependent coloring

Theorem (Holroyd and Liggett 2014)

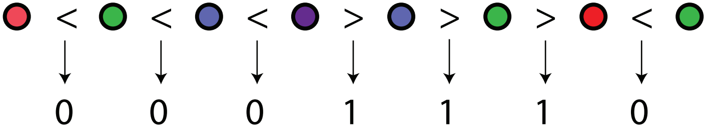
There exists a 1-dependent stationary random proper coloring of $\mathbb{Z}$ with 4 colors.

Finitely dependent coloring

Theorem (Holroyd and Liggett 2014)

There exists a 1-dependent stationary random proper coloring of $\mathbb{Z}$ with 4 colors.

Writing 0 whenever a color comes before a color of lower value and 1 otherwise, we get a 2-dependent process, with $p_4^X = 0$.

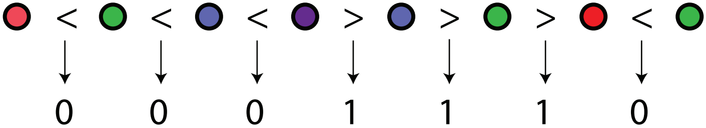


Finitely dependent coloring

Theorem (Holroyd and Liggett 2014)

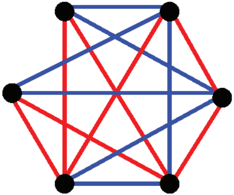
There exists a 1-dependent stationary random proper coloring of $\mathbb{Z}$ with 4 colors.

Writing 0 whenever a color comes before a color of lower value and 1 otherwise, we get a 2-dependent process, with $p_4^X = 0$.



$\rightarrow$ There is no lower bound on persistence for 2-dependent processes.

Proof Idea



Formula for persistence

We would like to calculate: $\mathbb{P}(X_1 = \cdots = X_q)$
Writing $w := q + k - 1$ we have,

$$= \int_0^1 dx_1 \cdots \int_0^1 dx_w \; \mathbb{1}\{f(x_1, \ldots, x_k) = \cdots = f(x_q, \ldots, x_w)\}$$

Probabilistic reformulation

Let $\{Z_t\}_t \in \mathbb{Z}$ be i.i.d. uniform random variables.

Observation

$$(Z_1, \ldots, Z_w) \overset{\text{law}}{=} (Z_{\sigma(1)}, \ldots, Z_{\sigma(w)})$$

where $\sigma \in S_M$ for some $M > w$.

Thus,

$$\int_{\bar{x}\in[0,1]^w} \mathbb{1}\{f(x_1,\ldots,x_k)=\cdots=f(x_q,\ldots,x_w)\}$$

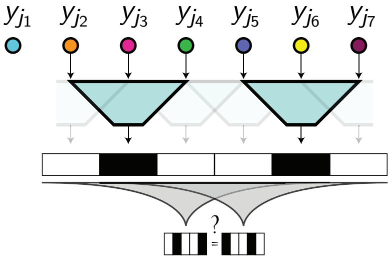
$$= \int_{\bar{y}\in[0,1]^M} \frac{(M-w)!}{M!} \sum_{1\le j_1<\cdots<j_w\le M} \mathbb{1}\{f(y_{j_1},\ldots,y_{j_k})=\cdots=f(y_{j_q},\ldots,y_{j_w})\}.$$

Probabilistic reformulation

We must therefore bound this sum combinatorially from below.

$$\sum_{\leq j_1<\cdots<j_w\leq M} \mathbb{1}\left\{f(y_{j_1},\ldots,y_{j_k})=\cdots=f(y_{j_q},\ldots,y_{j_w})\right\}.$$

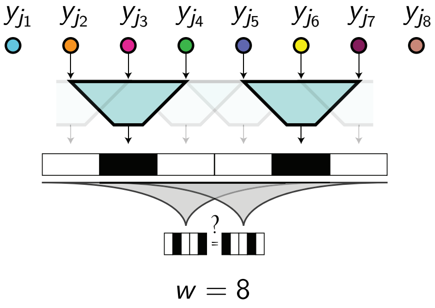
Probabilistic reformulation

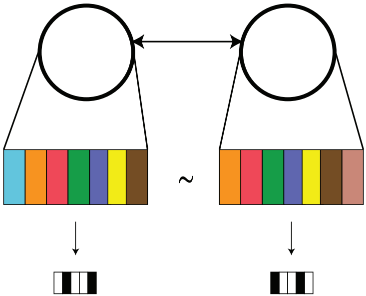


$w = 8$

$$\sum_{\leq j_1 < \cdots < j_w \leq M} \mathbb{1}\left\{ f(y_{j_1}, \ldots, y_{j_k}) = \cdots = f(y_{j_q}, \ldots, y_{j_w}) \right\}.$$

Probabilistic reformulation





$$\sum_{\leq j_1 < \cdots < j_w \leq M} \mathbb{1}\left\{ f(y_{j_1},\ldots,y_{j_k}) = \cdots = f(y_{j_q},\ldots,y_{j_w}) \right\}.$$

Combinatorial reformulation

Let $k, q \in \mathbb{N}$. We define a graph D_M^w whose vertices are increasing sequences of elements in $\{1 \dots M\}$ of length w, and

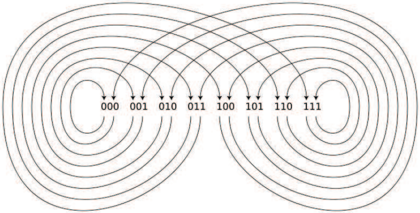
$$\bar{x} \sim \bar{y} \iff \forall_{i \in \{2,\dots,w\}} \left(x_i = y_{i-1} \right).$$

Combinatorial reformulation

Let $k, q \in \mathbb{N}$. We define a graph D_M^w whose vertices are increasing sequences of elements in $\{1 \ldots M\}$ of length w, and

$$\bar{x} \sim \bar{y} \ \longleftrightarrow \ \forall_{i \in \{2,\ldots,w\}} \big(x_i = y_{i-1} \big).$$

This is called a De-Bruijn graph. We ask if it can be properly colored.



Combinatorial reformulation

Let $k, q \in \mathbb{N}$. We define a graph D_M^w whose vertices are increasing sequences of elements in $\{1 \ldots M\}$ of length w, and

$$\bar{x} \sim \bar{y} \;\longleftrightarrow\; \forall_{i \in \{2,\ldots,w\}} \big(x_i = y_{i-1}\big).$$

This is called a De-Bruijn graph. We ask if it can be properly colored.

Reduced problem

Must show: There exists $M = M_{k,q}$ s.t. there is no proper coloring of D_M^{w-1} with 2^q colors.

Ramsey Theory

Theorem (implied by Chvátal)

For every k, d, if M is big enough, then there is no proper coloring of D_M^k with d colors.

Ramsey Theory

Theorem (implied by Chvátal)

For every k, d, if M is big enough, then there is no proper coloring of D_M^k with d colors.

Time does not permit giving exact details...

Ramsey Theory

Theorem (implied by Chvátal)

For every k, d, if M is big enough, then there is no proper coloring of D_M^k with d colors.

Time does not permit giving exact details...
Similar to the classical Ramsey results

Theorem (Ramsey)

For every d, there exists M such that K_M cannot be properly colored by d colors.

Thank you.