Localisation-Resistant Random Words with Small Alphabets

Cyril GAVOILLE[2] **Ghazal KACHIGAR**[1,2] Gilles ZÉMOR[1]

[1] Institut de Mathématiques de Bordeaux, [2] LaBRI

University of Bordeaux, France

WORDS 2019, Loughborough - 09/09/2019

Introduction

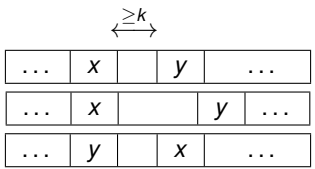
Object of study: Probability distributions on (some subset of) $\{1,\dots,q\}^n$ and $\{0,1\}^n$.

Examples

q-coloured words: $\mathcal{C}_{q,n} = \{x_1 \dots x_n \in \{1,\dots,q\}^n \mid x_i \neq x_{i+1}\}$.

Independent-set words: $\mathcal{I}_n = \{y_1 \dots y_n \in \{0,1\}^n \mid y_i = 1 \Rightarrow y_{i-1} = y_{i+1} = 0\}$.

Localisation resistance: No information is revealed about location of subword(s).



all occur with the same probability.

Introduction

Main results (informal)

- Complete description of localisation-resistant probability distributions on independent-set words.
- We need $q \geq 4$ for n large enough in order for a localisation-resistant probability distribution on q-coloured words to exist.

Motivation

- q-colouring using local resources is a fundamental problem in Distributed Computing **[Pettie et al.- STOC '18]**.
- Open question: are there faster algorithms if we have access to quantum resources ?
- q-colouring using quantum resources must be localisation-resistant **[Gavoille, Kosowki & Markiewicz '09]**.

Some definitions

$x_1 \dots x_n$ a random word on a small alphabet.

Stationarity: "Shift invariance for probability distributions".

For all $I, J \subset \{1, \dots, n\}$ intervals at distance at least k of each other:

k-dependence: $\Pr(x_I, x_J) = \Pr(x_I) \cdot \Pr(x_J)$.
k-localisability: $\Pr(x_I, x_J)$ depends only on $|I|$, $|J|$.

Facts

- 0-dependence = independence.
- 0-localisability = exchangeability.
- k-localisability $\Rightarrow$ stationarity.
- k-dependence + stationarity $\Rightarrow$ k-localisability.

k-dependence VS k-localisability

k-dependence + stationarity $\Rightarrow$ k-localisability.

BUT

k-localisability $\not\Rightarrow$ m-dependence for some m.

Example

$S = \{\text{all permutations of } \{1, \ldots, n\}\}$
Pr: uniform distribution on S.
This distribution is 0-localisable since

$$\Pr(x_I, x_J) = \frac{(n - |I| - |J|)!}{n!}$$

But not k-dependent for any $k \leq n$ because

$$\Pr(x_I) \cdot \Pr(x_J) = \frac{(n - |I|)!}{n!} \cdot \frac{(n - |J|)!}{n!} \neq \Pr(x_I, x_J)$$

The q-colouring problem

Recall that the set of q-coloured words of length n is:

$$\mathcal{C}_{q,n} = \{x_1 \dots x_n \in \{1, \dots, q\}^n \mid x_i \neq x_{i+1}\}$$

1-dependent colouring

[Holroyd & Liggett '15, '16]:

- There is a stationary 1-dependent q-colouring for all $q \geq 4$ for every $n \in \mathbb{N}$.
- There is no stationary 1-dependent 3-colouring for n large enough.

Because "k-dependent + stationary $\Rightarrow$ k-localisable", this implies

1-localisable colouring

There is a 1-localisable q-colouring for all $q \geq 4$ for every $n \in \mathbb{N}$.

The q-colouring problem

1-localisable colouring

There is a 1-localisable q-colouring for all $q \geq 4$ for every $n \in \mathbb{N}$.

Is 1-localisable 3-colouring possible?

Main result

Every 1-localisable probability distribution for random q-coloured words of length n requires $q \geq 4$ for n large enough.

q-coloured words and independent-set words

Recall

q-coloured words: $\mathcal{C}_{q,n} = \{x_1 \dots x_n \in \{1, \dots, q\}^n \mid x_i \neq x_{i+1}\}$.

Independent-set words: $\mathcal{I}_n = \{y_1 \dots y_n \in \{0,1\}^n \mid y_i = 1 \Rightarrow y_{i-1} = y_{i+1} = 0\}$.

Choose any colour c and define a function $f = (f_1, \dots, f_n)$ on $\mathcal{C}_{q,n}$ as follows:

$$y_i = f_i(x_i) = 1 \text{ if } x_i = c \qquad \text{and} \qquad y_i = f_i(x_i) = 0 \text{ if } x_i \neq c$$

Then

- The image of f is exactly $\mathcal{I}_n$.
- f preserves k-dependence, k-localisability, stationarity.
- $\Pr(y_i = 1) \geq 1/q$.

Goal

Find $\max \Pr(y_i = 1)$ for a 1-localisable probability distribution on $\mathcal{I}_n$.

An example

We aim to find $\max \Pr(y_i = 1)$ for a 1-localisable probability distribution on $\mathcal{I}_4$.
Let $p_1 = \Pr(y_i = 1)$, $p_2 = \Pr(y_i = 1, y_{i+2} = 1)$.

$y_1y_2y_3y_4$	$\Pr(y_1y_2y_3y_4)$
0101	
1010	
1001	
0100	
0010	
1000	
0001	
0000	

$\Pr(1010) = \Pr(1010) = \Pr(1001) = p_2$.

An example

We aim to find $\max \Pr(y_i = 1)$ for a 1-localisable probability distribution on $\mathcal{I}_4$.
Let $p_1 = \Pr(y_i = 1)$, $p_2 = \Pr(y_i = 1, y_{i+2} = 1)$.

$y_1y_2y_3y_4$	$\Pr(y_1y_2y_3y_4)$
0101	p_2
1010	p_2
1001	p_2
0100	
0010	
1000	
0001	
0000	

$\Pr(0100) + p_2 = \Pr(0100) + \Pr(0101) = \Pr(010\star) = p_1$.

An example

We aim to find $\max \Pr(y_i = 1)$ for a 1-localisable probability distribution on $\mathcal{I}_4$.
Let $p_1 = \Pr(y_i = 1)$, $p_2 = \Pr(y_i = 1, y_{i+2} = 1)$.

$y_1y_2y_3y_4$	$\Pr(y_1y_2y_3y_4)$
0101	p_2
1010	p_2
1001	p_2
0100	$p_1 - p_2$
0010	$p_1 - p_2$
1000	
0001	
0000	

$\Pr(1000) + 2p_2 = \Pr(1000) + \Pr(1010) + \Pr(1001) = \Pr(10\star\star) = p_1$.

An example

We aim to find $\max \Pr(y_i = 1)$ for a 1-localisable probability distribution on $\mathcal{I}_4$.
Let $p_1 = \Pr(y_i = 1)$, $p_2 = \Pr(y_i = 1, y_{i+2} = 1)$.

$y_1y_2y_3y_4$	$\Pr(y_1y_2y_3y_4)$
0101	p_2
1010	p_2
1001	p_2
0100	$p_1 - p_2$
0010	$p_1 - p_2$
1000	$p_1 - 2p_2$
0001	$p_1 - 2p_2$
0000	

$\Pr(0000) = 1$ minus the rest.

An example

We aim to find $\max \Pr(y_i = 1)$ for a 1-localisable probability distribution on $\mathfrak{I}_4$.
Let $p_1 = \Pr(y_i = 1)$, $p_2 = \Pr(y_i = 1, y_{i+2} = 1)$.

$y_1y_2y_3y_4$	$\Pr(y_1y_2y_3y_4)$
0101	p_2
1010	p_2
1001	p_2
0100	$p_1 - p_2$
0010	$p_1 - p_2$
1000	$p_1 - 2p_2$
0001	$p_1 - 2p_2$
0000	$1 - 4p_1 + 3p_2$

An example

We aim to find $\max \Pr(y_i = 1)$ for a 1-localisable probability distribution on $\mathcal{I}_4$.
Let $p_1 = \Pr(y_i = 1)$, $p_2 = \Pr(y_i = 1, y_{i+2} = 1)$.

$y_1y_2y_3y_4$	$\Pr(y_1y_2y_3y_4)$
0101	p_2
1010	p_2
1001	p_2
0100	$p_1 - p_2$
0010	$p_1 - p_2$
1000	$p_1 - 2p_2$
0001	$p_1 - 2p_2$
0000	$1 - 4p_1 + 3p_2$

Solving $p_i \geq 0$ and $\Pr(y_1y_2y_3y_4) \geq 0$ we find the maximum value to be $p_1 = 2/5$.

General case

Set $p_1 := \Pr(1\star\cdots\star)$, $p_2 := \Pr(1\star 1\star\cdots\star)$, $\ldots$, p_ℓ.
Find maximum value of p_1 such that $p_i \geq 0$ and $\Pr(y_1 \ldots y_n) \geq 0$ (where $n = 2\ell - 1$ or $n = 2\ell$).

- Each $\Pr(y_1 \ldots y_n) \geq 0$ is uniquely determined as a linear function of $p_1, \ldots, p_\ell$.
- We thus need to solve a **linear programming** problem.
- **Problem**: Exponential (in ℓ) number of constraints.

Equivalent subsystem

$\Pr(0^n) \geq 0$
$\Pr((10)0^{n-2} \geq 0$
$\Pr((10)^2 0^{n-4}) \geq 0$
$\ldots$
$\Pr((10)^{\ell-1} 0^{n-2\ell+2)} \geq 0$

$\Rightarrow$ Linear programming problem in ℓ constraints and ℓ variables.

Formulation using matrices and vectors

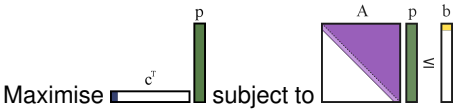
Maximise $\mathbf{c}^{\mathsf{T}}\mathbf{p}(=p_1)$, subject to $\mathbf{Ap} \leq \mathbf{b}$ and $\mathbf{p} \geq \mathbf{0}$.

Here the matrix and the vectors are of the form:

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & \dots & a_{1,\ell-1} & a_{1,\ell} \\ -1 & a_{2,2} & \dots & \dots & a_{2,\ell-1} & a_{2,\ell} \\ 0 & -1 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & -1 & a_{\ell-1,\ell-1} & a_{\ell-1,\ell} \\ 0 & 0 & \dots & 0 & -1 & a_{\ell,\ell} \end{bmatrix}$$

$a_{i,j} = (-1)^{i+j}\binom{2\ell+2-(i+j)}{j-i+1}$, $\mathbf{c} = (1,0,\dots,0)$, $\mathbf{b} = (1,0,\dots,0)$, $\mathbf{p} = (p_1,\dots,p_\ell)$.

Linear programming



Theorem

For linear programming problems that have the general form above, the optimal value for $\mathbf{c}^{\mathsf{T}}\mathbf{p}$ $(= p_1)$ is $\frac{u_\ell}{u_{\ell+1}}$ where the sequence $(u_k)_{k\geq 1}$ is defined by $u_1 = 1$ and

$$u_{k+1} = \sum_{i=1}^{k} a_{\ell-k+1,\ell-k+i}\, u_{k+1-i}$$

This optimal value is obtained by solving the special case $\mathbf{Ap} = \mathbf{b}$.

We prove this theorem using the **duality theorem** for linear programming.

Applying the theorem on the previous slide to our problem ($a_{i,j} = (-1)^{i+j}\binom{2\ell+2-(i+j)}{j-i+1}$) we get

Theorem

The optimal value for p_1 is $\frac{u_\ell}{u_{\ell+1}}$ where $u_1 = 1$ and $u_{k+1} = \sum_{i=1}^{k}(-1)^{i+1}\binom{2k+1-i}{i}u_{k+1-i}$.

This is exactly the sequence of **Catalan numbers** $(c_n)_{n\in\mathbb{N}}$, $c_n = \frac{1}{n+1}\binom{2n}{n}$.

Finally

$$\frac{c_\ell}{c_{\ell+1}} = (\ell+2)/(4\ell+2), \text{ hence the optimal value of } p_1 \to 1/4 \text{ as } \ell \to \infty.$$

Further results and questions

1. **(In paper)** If we plug in a feasible value for p_1, we get a linear programming problem for maximising p_2 that has the same general form and we may apply the same theorem and so on for p_3, etc. This way we get $p_i \leq (c_{\ell-i+1}/c_{\ell+1}) \to 1/4^i$ as $\ell \to \infty$.

2. **(Proven, not in paper)** Let

$$\mathcal{I}_{k,n} = \{y_1 \dots y_n \in \{0,1\}^n \mid y_i = 1 \Rightarrow y_{i-j} = y_{i+j} = 0, 1 \leq j \leq k\}$$

We have a similar result for k-localisable probability distributions on $\mathcal{I}_{k,n}$ with the sequence of **Fuss-Catalan numbers**.

3. **(Open questions)** What about combinatorial structures other than words? E.g. labellings on graphs?