Localisation-Resistant Random Words with Small Alphabets

Cyril Gavoille[1(✉)], Ghazal Kachigar[1,2], and Gilles Zémor[2]

[1] LaBRI, University of Bordeaux, Bordeaux, France
{gavoille,ghazal.kachigar}@labri.fr
[2] IMB, University of Bordeaux, Bordeaux, France
zemor@math.u-bordeaux.fr

Abstract. We consider q-coloured words, that is words on $\{1,\ldots,q\}$ where no two consecutive letters are equal. Motivated by multipartite colouring games with nonsignalling resources, we are interested in random q-coloured words satisfying a *k-localisability* property. More precisely, the probability of containing any given pair of words as subwords spaced at least k letters apart can depend only on their lengths. We focus on the issue of the smallest alphabet size q for which a probability distribution for such random words can exist. For $k = 1$, we prove a lower bound of $q \geqslant 4$. The bound is optimal because there exists a suitable distribution for random 4-colourings that was constructed by Holroyd and Liggett in 2015. Our lower bound can be generalized to k-localisable random words where the letters of each subword of $k+1$ letters must be pairwise different. We show that the alphabet size in this case must be at least $(k+1)\cdot(1+1/k)^k$.

AQ1 AQ2

Keywords: Random words · Stochastic colouring process · Hard-core process · Colouring game

1 Introduction

Multipartite Colouring Game. Let us consider the following general multipartite graph colouring game. We are given a graph G with nodes $v_1,\ldots,v_n$, a colour bound q, and m players $P_1,\ldots,P_m$. The referee virtually places each player P_i at a node v_j and gives them a personalised input. This information depends on the variant of the game. For example, it can consist of the index j that the player P_i is placed on. Each player then has to output a colour for its node, i.e., an integer taken from $\{1,\ldots,q\}$. The players win if the resulting node colouring is a q-colouring of the coloured subgraph of G, i.e., the subgraph induced by the nodes hosting at least one player. More precisely, colours must differ for adjacent players and coincide for players that have been placed on the same node, if any.

Supported by the French ANR projects ANR-16-CE40-0023 (DESCARTES) and ANR-18-CE47-0010 (QUDATA).

R. Mercaş and D. Reidenbach (Eds.): WORDS 2019, LNCS 11682, pp. 1–14, 2019.
https://doi.org/10.1007/978-3-030-28796-2_15

Players are allowed to agree on a joint strategy beforehand, which may depend on G, q and m. Once placed on their node with the referee's input, players are not allowed to communicate in any way. So, the output colour can only depend on the joint strategy and on the referee's input. Unless specified by the referee's input, players are not aware whether other players stand at the same node, or which players are adjacent to them.

The main question is to understand how small the colour bound q can be so that the m players can still win the game for G, under given assumptions on their joint strategy and on the referee's inputs.

This question is related to fundamental problems in graph theory, distributed computing, and quantum information. To illustrate, assume that the referee's input consists of the node index where each player is placed. For $m = n$, the smallest q is precisely the chromatic number of G since the referee can force the players to output a q-colouring of the whole graph, and a strategy for the players consists of agreeing beforehand on any given q-colouring. For $m = 2$, and if the two players share quantum resources (materialised say by entangled particles), this leads to the notion of *quantum chromatic number*. Interestingly, this variant of chromatic number can be smaller than the classical one. For instance, there is a graph with 18 nodes and chromatic number 5 on which the two players Alice and Bob can win this colouring game with only 4 colours [8].

Two-partite games with quantum resources (sometimes called pseudo-telepathy games) are well studied in Computer Science and in Physics. However, multipartite quantum games with a large number m of players are much less understood. There are multipartite games where quantum superiority can be proved, and also outperformed (in terms of winning probability) by general *nonsignalling* resources [1,2,7,12]. Such exotic resources, which are not predicted to exist according to current physical theories, allow the players to use *any* non-local correlations in their outputs without any communication.

Links to Distributed Computing. Colouring a network with a minimal amount of communication is a fundamental symmetry-breaking problem studied in distributed computing (see [4,10,13,15,16] for recent breakthroughs, and [3] for a book dedicated to this field). In this setting and in brief, each player acts at a single node as a processor that can exchange messages with its neighbours in some underlying graph. They must output a colouring of the graph after a limited number of synchronous rounds of communication. In this model, a.k.a. the LOCAL model, there are distributed algorithms for q-colouring n-node paths[1] that require $O(\log^* n - \log^* q)$ rounds[2] of communication for $q > 2$, and this is tight. More precisely, [26,27] showed that after collecting the IDs of k neighbours around each node, i.e., $k = 2t$ numbers after t two-sided rounds of communication in the path, any possibly randomised q-colouring algorithm must satisfy $q = \Omega(\log^{(k)} n)$.

[1] This holds also for cycles, and more generally for graphs of maximum degree Δ with $q > \Delta$.

[2] We write $\log^* n = \min\{i : \log_2^{(i)} n \leqslant 1\}$, the inverse function of a power-2 tower.

Since after t rounds every node has been able to communicate with nodes at distance at most t, these t-round algorithms imply that information about nodes at distance t suffices to provide a q-colouring. Yet, the colouring problem in the LOCAL model can be viewed as a particular setting of the general multipartite colouring game where the referee's placement is a permutation ($m = n$) and where the input for each player is the t-neighbourhood of its node[3]. Whether the number of rounds t can be significantly reduced if quantum resources are available is a widely open question [14,24,25] even for path graphs.

Lastly, we notice that the colouring game as described above can be further extended to *locally checkable labelling* games where the goal for the players is to output a label taken from a predefined set and satisfying some local constraints in the graph. This captures not only colouring problems, but also maximal independent set, dominating set, weak 2-colouring[4], and many others [6,28].

Our Contribution. In this paper we consider the multipartite q-colouring game for the path $v_1 - v_2 - \cdots - v_n$ where the referee's placement is a permutation σ on $\{1, \ldots, n\}$ not revealed to the players. So player P_i is placed at position $\sigma(i)$ on the path, i.e., at the node $v_{\sigma(i)}$. The referee's input for P_i consists only of the index $\sigma^{-1}(\sigma(i)+1)$ of the player placed on its right neighbouring node[5], $v_{\sigma(i)+1}$. As a result, each player outputs a colour given its own index and the one of its right neighbour. The players win the game if they produce a q-colouring of the path. Thus in this game a player can coordinate only with its right neighbour and its colour cannot depend on its position $\sigma(i)$.

As explained above, this game can be seen as the q-colouring problem in the distributed LOCAL model where each processor has received information only from its right neighbour, after what may be called a half-round of communication. From the above lower bound with $k = 1$, we must have $q = \Omega(\log n)$ for every joint strategy based on classical resources including shared randomness.

The question we want to address is what is the minimum number of colours q that can be achieved if players are allowed to use quantum resources, and more generally any nonsignalling resources. It should be stressed that quantum resources allow each player to use non-local correlations that may in fact beat the previous $\Omega(\log n)$ lower bound on q. We prove in this paper that $q \geqslant 4$ for n large enough, and this is optimal.

To formally state our main theorem, we interpret the colouring resulting from a run of the game as a random word $X_1 X_2 \cdots X_n$, where each letter X_i is a random variable ranging in $\{1, \ldots, q\}$ and corresponding to the colour output by the player at node v_i. Here randomness may come from the kind of resource used by the players in their joint strategy (e.g., a quantum state) that is revealed at the time they output the colour (e.g., measurement).

[3] Say, the list of player's IDs and the edges list between them.

[4] In this problem, each player must produce one out two possible colours such that at least one of its adjacent node receives a different colour.

[5] Player at v_n receives index 0. Alternatively, we may assume that an extra player P_0 is always placed at a virtual node v_{n+1} and does not take part in the colouring game.

Given an interval $I = [a, b]$, we use the notation X_I for the subword $X_a X_{a+1} \cdots X_{b-1} X_b$. Define the *distance* between any two intervals I, J as $\inf\{|i - j| : i \in I, j \in J\}$. (The distance between I and $J = \varnothing$ is $+\infty$ by convention). Note that the two subwords X_I and X_J are separated by k letters in $X_1 \cdots X_n$ if and only if I and J are at distance $k + 1$. We say that a word is *coloured* if any two consecutive letters are distinct.

In order to lower bound q for any probability distribution for random q-coloured words coming from such games, we introduce the notion of k-localisability defined as follows:

Definition 1. *A probability distribution for a random word $X_1 \cdots X_n$ is* k-localisable *if, for all intervals $I, J \subseteq \{1, \ldots, n\}$ at distance more than k, the distribution of (X_I, X_J) can only depend on $\{|I|, |J|\}$.*

Informally, this means that the probability of having two given words S and T in a random word depends neither on their absolute positions, nor on their order, nor on their distance in the word, as long as the number of letters between them is at least k.

Coming back to our colouring game on the path where players are only aware of their immediate right neighbour, the word distribution resulting of any winning strategy based on nonsignalling resources must be 1-localisable. This is because otherwise two players at nodes v_i and v_j sufficiently far apart could collectively retrieve information about i, j or $|i - j|$ from their colour distribution. From the rules of the game this is not possible without *signals* (i.e., communication). This holds for any nonsignalling theory including quantum mechanics. Note however that a k-localisable colour distribution does not forbid non-local correlation.

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

As we will see in the next paragraph, the lower bound of Theorem 1 is tight. This is actually a consequence of the random 4-colouring given in [20].

Our approach to prove Theorem 1 is to study random binary words $Y_1 \cdots Y_n$ obtained from a random q-coloured word $X_1 \cdots X_n$ by fixing any colour $c \in \{1, \ldots, q\}$ and by setting $Y_i = 1$ if $X_i = c$, and $Y_i = 0$ otherwise. Observe that $Y_1 \cdots Y_n$ codes an independent set of the n-node path, and let us call an *independent-set* word any binary word that does not contain any two consecutive ones. Such random words can also be seen as *hard-core* processes where the variable Y_i indicates the presence of a radius-1 hard-core particle at position i on the discrete line.

The lower bound of Theorem 1 is actually a corollary of our following main technical contribution. It gives a fine analysis of the marginal probabilities of having a given number of ones in fixed positions for 1-localisable random independent-set words, a result interesting in its own right. We let $\mathsf{c}_n = \frac{1}{n+1}\binom{2n}{n}$ denote the n-th Catalan number.

Theorem 2. *Let p_i denote the probability of having i ones in the positions indexed by the odd integers $1, 3, \ldots, 2i-1$, for a random independent-set word of length $n \geqslant 2i$. Let $\ell = \lfloor n/2 \rfloor$. Then, for every even n:*

i. *Every 1-localisable probability distribution for random independent-set words of length n satisfies, for each $i \in \{0, \ldots, \ell\}$, $p_i \leqslant \mathsf{c}_{\ell-i+1}/\mathsf{c}_{\ell+1}$.*
ii. *There exists a 1-localisable probability distribution for random independent-set words of length n such that, for each $i \in \{0, \ldots, \ell\}$, $p_i = \mathsf{c}_{\ell-i+1}/\mathsf{c}_{\ell+1}$.*

By marginalising, it is easy to derive from Theorem 2(i) that $p_i \leqslant \mathsf{c}_{\lceil n/2 \rceil - i+1}/\mathsf{c}_{\lceil n/2 \rceil+1}$ for every length n, and not only for even n.

Let us explain why Theorem 1 follows from Theorem 2. The first observation is that any letter transformation $Y_i = f(X_i)$ preserves the k-localisability of the distribution as long as f does not depend on i. Now, given any 1-localisable distribution for a random q-coloured word $X_1 \cdots X_n$, consider the most frequent colour c, so appearing with probability at least $1/q$ in the random word. The random independent-set word $Y_1 \cdots Y_n$ as defined above has a 1-localisable distribution. And the probability of having a one at any fixed position in $Y_1 \cdots Y_n$ is $p_1 \geqslant 1/q$. However, from Theorem 2(i) applied to $Y_1 \cdots Y_n$, we get that $p_1 \leqslant \mathsf{c}_5/\mathsf{c}_6 = 7/22$ whenever $n \geqslant 10$ noting that $\mathsf{c}_\ell/\mathsf{c}_{\ell+1} = (\ell+2)/(4\ell+2)$. Thus, we obtain $1/q \leqslant p_1 \leqslant 7/22$, implying that $q > 3$ as claimed in Theorem 1.

Related Works. The notion of k-localisability introduced in this paper is a natural notion for the study of multipartite colouring games on paths with quantum resources (and beyond). A related notion in probability theory is the well-known k-dependence of random variables [17,22] studied for more than seven decades. A probability distribution for random variables $X_1 \cdots X_n$ is *k-dependent* if, for all intervals $I, J \subseteq \{1, \ldots, n\}$ at distance more than k, the variables X_I and X_J are independent. Clearly, 0-dependence is the same as independence.

Recall that a probability distribution for a random word $X_1 \cdots X_n$ is *stationary* if, for every interval $I \subseteq \{1, \ldots, n\}$, the distribution of X_I can depend only on $|I|$. It is not difficult to see that any stationary k-dependent distribution is also k-localisable: for k-dependent distributions, the distribution of (X_I, X_J) is the product of the marginals which, by stationarity, can depend only on $|I|$ and on $|J|$. However, the reverse is false. Although every k-localisable distribution is stationary (setting $J = \varnothing$ in the definition), there exist k-localisable distributions that are not k-dependent. For instance $X_i = \sigma(i)$ for a uniform random permutation σ of $\{1, \ldots, n\}$ defines a 0-localisable distribution that is not k-dependent for every k. Indeed, $\mathbb{P}(X_I = S) = (n - |I|)!/n!$ and $\mathbb{P}(X_I = S, X_J = T) = (n - (|I| + |J|))!/n!$ for any two disjoint intervals I, J (so at distance more than k for some $k \geqslant 0$). However, $\mathbb{P}(X_I = S, X_J = T) \neq \mathbb{P}(X_I = S) \cdot \mathbb{P}(X_J = T)$ for every k. Furthermore, the random binary word defined by $Y_i = X_i \bmod 2$ is still 0-localisable and once again not k-dependent for every k.

Interestingly, the notion of 0-localisability corresponds to the notion of *exchangeability* [5,9], in connection with the celebrated de Finetti Theorem that explains the relationship between exchangeability and independence. Random

variables are (finitely) exchangeable if they are invariant under permutations of their indices, i.e., if $\mathbb{P}(X_{\sigma(1)}, \cdots, X_{\sigma(n)}) = \mathbb{P}(X_1, \cdots, X_n)$ for any permutation σ on $\{1, \ldots, n\}$.

Until very recently, no stationary k-dependent distribution for q-coloured words of growing length n was known, even for large q. It is easy to see that $k \geqslant 1$ and $q \geqslant 3$ are required. Indeed, $k \geqslant 1$ since X_i and X_{i+1} cannot be independent. And $q \geqslant 3$, since for 2-colouring $\mathbb{P}(X_i = X_j)$ depends on the parity of $|i - j|$ that can be much larger than k. In fact, for large enough n, it has been proved in [21] that no stationary 1-dependent 3-colouring exists. This result is actually implied by our Theorem 2(i).

The relationship between k and q has been investigated in [18–20]. In particular, in [20] a stationary 1-dependent 4-colouring is constructed, as well as a 2-dependent 3-colouring of words of infinite length. The construction is based on recursion formulae extending a suitable colouring of a word of length n to a word of length $n + 1$. This stationary 1-dependent 4-colouring implies that a 1-localisable 4-colouring exists. Thus our lower bound in Theorem 1 is tight.

Overview. Let $\mathcal{I}_n \subset \{0, 1\}^n$ be the set of all independent-set words of length n, i.e., the binary words of length n with no two consecutive ones. As explained in the previous paragraph, Theorem 1 is a corollary of Theorem 2(i). So we focus on 1-localisable distributions for binary words of $\mathcal{I}_n$.

In a first step, we show that, for every 1-localisable probability distribution $\mathbb{P}$, the probability $\mathbb{P}(s)$ of every binary word s of length n can always be written as a linear combination with integral coefficients of the p_i's, i.e., the probabilities of a random word having i ones in positions $1, 3, \ldots, 2i - 1$. This leads to a system of linear inequalities with $O(n)$ variables p_i and with $O(|\mathcal{I}_n|)$ constraints. We can in principle find the maximum value of p_1 by solving such a linear programming problem. Unfortunately, $|\mathcal{I}_n|$ grows exponentially in n since it satisfies a Fibonacci recurrence. This approach may at first seem intractable.

However, we show that the $O(|\mathcal{I}_n|)$ constraints are highly redundant and that there is a subset of only $O(n)$ constraints strictly equivalent to the original ones. Hence, we end up with a much smaller linear programming problem with $n/2$ variables and $n/2$ constraints which moreover turns out to be sufficiently structured so as to admit a closed-form solution.

Section 2 is dedicated to deriving this structured linear programming problem with p_1 as the linear objective function that we are maximising. Section 3 addresses the problem of solving this linear program. We are first able to derive a feasible solution for the linear program involving a binomial formula for the Catalan numbers (namely Corollary 1). We then show that the feasible solution we found at the previous step is indeed the optimal one by using the duality theorem for linear programming. We also show that this particular solution maximises simultaneously all the p_i's, which will prove Theorem 2(i & ii).

Due to space limitations, proofs and intermediate lemmas will appear in the full version.

2 Localisable Distribution on Independent-Set Words

A small worked-out example will go a long way towards explaining what the present and following section are about. Let $\mathbb{P}$ be a 1-localisable probability distribution on $\mathcal{I}_4$, and let $X_1X_2X_3X_4$ be a random word with this distribution. We define $p_1 = \mathbb{P}(X_1 = 1)$ and $p_2 = \mathbb{P}(X_1 = X_3 = 1)$. Let us now consider the probabilities of the 8 individual words of $\mathcal{I}_4$:

$$0000, 1000, 0100, 0010, 0001, 1010, 0101, 1001$$

We have $\mathbb{P}(1010) = p_2$ by definition, and 1-localisability tells us that $\mathbb{P}(X_1 = 1, X_3 = 1) = \mathbb{P}(X_1 = 1, X_4 = 1) = \mathbb{P}(X_2 = 1, X_4 = 1)$. Hence, $\mathbb{P}(1010) = \mathbb{P}(1001) = \mathbb{P}(0101) = p_2$. Now we also have: $p_1 = \mathbb{P}(1000) + \mathbb{P}(1010) + \mathbb{P}(1001)$. Hence the value of $\mathbb{P}(1000)$, which is readily seen to be the same as $\mathbb{P}(0001)$:

$$\mathbb{P}(1000) = \mathbb{P}(0001) = p_1 - 2p_2.$$

From $\mathbb{P}(X_2 = 1) = \mathbb{P}(0100) + \mathbb{P}(0101)$ and $\mathbb{P}(X_2 = 1) = \mathbb{P}(X_1 = 1)$ we get the value of $\mathbb{P}(0100)$ and similarly of $\mathbb{P}(0010)$:

$$\mathbb{P}(0100) = \mathbb{P}(0010) = p_1 - p_2.$$

The only probability of an individual word that is unaccounted for is $\mathbb{P}(0000)$. Writing that all probabilities of all individual words sum to 1, we get:

$$\mathbb{P}(0000) = 1 - 4p_1 + 3p_2.$$

We may now notice two things. Any 1-localisable distribution on $\mathcal{I}_4$ is entirely determined by the two values p_1 and p_2. Conversely, any probability distribution defined as above by the two values p_1 and p_2 is 1-localisable. Finally, given any two positive numbers p_1 and p_2, such a probability distribution is well-defined if and only if all the linear expressions in p_1, p_2 that we have just computed take positive values. In other words, the values of p_1, p_2 for which there exists a 1-localisable probability distribution on $\mathcal{I}_4$ such that $\mathbb{P}(X_1 = 1) = p_1$ and $\mathbb{P}(X_1 = X_3 = 1) = p_2$, are exactly the solutions of the system of linear inequalities:

$$p_1, p_2 \geqslant 0, \quad p_1 - 2p_2 \geqslant 0, \quad p_1 - p_2 \geqslant 0, \quad \text{and} \quad 1 - 4p_1 + 3p_2 \geqslant 0.$$

Determining the largest allowable value of p_1 consists therefore in solving the associated linear program for the objective function p_1. In the present example we find that the maximum value is $p_1 = 2/5$. Our goal is to prove that the phenomena that we observe on this small example carry over to the general case of 1-localisable distributions on $\mathcal{I}_n$. We will then solve the general linear program associated with the maximisation of p_1.

We will find it convenient to write expressions such as $\mathbb{P}(10\star\star)$ for the value $\mathbb{P}(X_1 = 1) = p_1$. More generally, for a distribution $\mathbb{P}$ for binary words of length n, two words s, t, and an integer $i \geqslant 0$ such that $|s| + i + |t| = n$, we will write:

$$\mathbb{P}(s \star^i t) \;=\; \sum_{u \in \{0,1\}^i} \mathbb{P}(s\,u\,t).$$

We now focus on the case of even n, and set $n = 2\ell$. It will be useful to introduce an algebraic formalism that will enable us to manipulate the general linear program and identify redundant linear inequalities.

Consider ℓ variables $p_1, \ldots, p_\ell$. Consider a function $\Lambda_n : \{0,1\}^n \to \mathbb{Z}[p_1, \ldots, p_\ell]$, and define $p_0 = \sum_{s\in\{0,1\}^n} \Lambda_n(s)$. We define the following rule for extending the domain of Λ_n to $\{0,1,\star\}^n$:

(R0) $\Lambda_n(s\star t) = \Lambda_n(s0t) + \Lambda_n(s1t)$ for every s, t such that $|s| + |t| = n - 1$.

Repeated application of rule (R0) until only the symbol $\star$ remains on the left-hand side gives that $\Lambda_n(\star^n) = \sum_{s\in\{0,1\}^n} \Lambda_n(s) = p_0$.

We also define the following properties:

(R1) $\Lambda_n(s) = 0$ if $s \in \{0,1\}^n \setminus \mathcal{I}_n$.
(R2) $\Lambda_n(s\star t\star) = \Lambda_n(s\star\star t) = \Lambda_n(\star s\star t)$ for s, t such that $|s| + |t| = n - 2$.
(R3) $\Lambda_n(s\star\star t) = \Lambda_n(t\star\star s)$ for s, t such that $|s| + |t| = n - 2$.
(R4) $\Lambda_n((1\star)^i\star^{n-2i}) = p_i$ for $i \in \{1, \ldots, \ell\}$.

Lemma 1. *For every p_0, there is a unique function Λ_n on $\{0,1\}^n$ satisfying (R1), (R2) and (R4) and $p_0 = \sum_{s\in\{0,1\}^n} \Lambda_n(s)$. For every $s \in \mathcal{I}_{2\ell}$, $\Lambda_n(s)$ is a linear function of $p_1, \ldots, p_\ell$. Furthermore, Λ_n satisfies Property (R3).*

From now on, we consider only functions Λ_n that satisfy (R1) through (R4).

We now introduce the system of linear inequalities:

System 1. $p_i \geqslant 0$ and $\Lambda_{2\ell}(s) \geqslant 0$, for all $i \in \{1, \ldots, \ell\}$ and $s \in \mathcal{I}_{2\ell}$.

We then have the relatively straightforward result:

Theorem 3. *Let $p_1, \ldots, p_\ell \in [0,1]$. There exists a 1-localisable probability distribution $\mathbb{P}$ on $\mathcal{I}_{2\ell}$ such that $\mathbb{P}((1\star)^i\star^{2\ell-2i}) = p_i$ for all $i \in \{1, \ldots, \ell\}$ iff System 1 is satisfied with $p_0 = 1$. We then have $\Lambda_{2\ell}(s) = \mathbb{P}(s)$.*

Let $\mathcal{S}_n = \{(10)^k 0^{n-2k} : k \in \{0, \ldots, \ell\}\} \subset \mathcal{I}_n$. We define the following subsystem of System 1:

System 2. $p_i \geqslant 0$ and $\Lambda_{2\ell}(s) \geqslant 0$, for all $i \in \{1, \ldots, \ell\}$ and $s \in \mathcal{S}_{2\ell}$.

We have the following:

Lemma 2. *For every $s \in \mathcal{I}_n$, there is $(a_t)_{t\in\mathcal{S}_n}$, $a_t \in \mathbb{N}$, such that $\Lambda_n(s) = \sum_{t\in\mathcal{S}_n} a_t \Lambda_n(t)$.*

Using this, we prove that

Proposition 1. *System 1 is equivalent to System 2, i.e., any solution of one is also a solution of the other.*

Thus, one can focus on the much more manageable System 2. We have the following expressions for the Λ-values of the elements of $\mathcal{S}_n$:

Lemma 3. $\Lambda_n((10)^k 0^{n-2k}) = \sum_{i=0}^{\ell-k} (-1)^i \binom{2\ell-2k+1-i}{i} p_{k+i}$, *for $k \in \{0, \ldots, \ell\}$.*

3 Solving the LP System

To summarise, we have shown so far that the existence of a 1-localisable probability distribution on $\mathcal{I}_{2\ell} \subset \{0,1\}^{2\ell}$ is equivalent to the solvability of a system of $O(|\mathcal{I}_n|) \sim \exp(\Omega(n))$ inequalities $\Lambda_{2\ell}(s) \geqslant 0$ for $s \in \mathcal{I}_{2\ell}$. Moreover, every $\Lambda_{2\ell}(s)$, for $s \in \mathcal{I}_{2\ell}$, is a linear function of $p_i = \Lambda_{2\ell}((1\star)^i \star^{n-2i})$ for $1 \leqslant i \leqslant \ell$. We obtain therefore a system of *linear* inequalities. We furthermore showed that there is a size-ℓ subset $\mathcal{S}_{2\ell}$ of $\mathcal{I}_{2\ell}$ such that the inequalities corresponding to its members imply all inequalities for all the members of $\mathcal{I}_{2\ell}$.

Since we are interested in the values that can be taken by $p_1, \ldots, p_\ell$, in particular p_1 and its maximum value, Lemma 3 tells us that we are now faced with the explicit linear programming problem defined by $p_0 = 1$ and:

$$\text{maximise} \quad p_1 \quad \text{subject to:}$$

$$\begin{cases} p_i \geqslant 0, & i \in \{1, \ldots, \ell\} \\ \sum_{i=0}^{\ell-k} (-1)^i \binom{2\ell-2k+1-i}{i} p_{k+i} \geqslant 0, & k \in \{0, \ldots, \ell-1\}\,. \end{cases} \tag{1}$$

Once we know this maximum value of p_1, we set the value of p_1 to be something less than or equal to this maximum value. It turns out that this gives rise to another linear programming problem which is very similar in form to the first one, and where the goal is now to maximise p_2. We repeat this procedure until we get the maximum value of every p_i when the values of p_j for $j < i$ are set to something less than or equal to their maximum possible value. Indeed, we will show that we have the following, which implies directly Theorem 2:

Theorem 4. *Any solution $(p_1, \ldots, p_\ell) \in \mathbb{R}^\ell$ to the system of inequalities (1) satisfies $p_i \leqslant (\mathsf{c}_\ell / \mathsf{c}_{\ell+1}) \cdot p_{i-1} \leqslant (\mathsf{c}_{\ell-i+1} / \mathsf{c}_{\ell+1}) \cdot p_0$, possibly with equality.*

We now need linear programming notation:

Definition 2. *Let $m, n \in \mathbb{N}$, $c_i, b_j, a_{i,j} \in \mathbb{R}$ for $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$. Let $\boldsymbol{c} = (c_1, \ldots, c_n)^\mathsf{T}$, $\boldsymbol{b} = (b_1, \ldots, b_m)^\mathsf{T}$, $\mathbf{x} = (x_1, \ldots, x_n)^\mathsf{T}$ and $\boldsymbol{A} = (a_{i,j})$ be an $m \times n$ matrix. A problem of the form:*

$$\textit{Maximise } \mathbf{c}^\mathsf{T}\mathbf{x}\textit{, subject to } \mathbf{A}\mathbf{x} \leqslant \mathbf{b} \textit{ and } \mathbf{x} \geqslant \mathbf{0},$$

is called an LP problem in standard form. *The linear expression $\mathbf{c}^\mathsf{T}\mathbf{x}$ is called the* objective function, *$\mathbf{A}\mathbf{x} \leqslant \mathbf{b}$ and $\mathbf{x} \geqslant \mathbf{0}$ are called the* constraints, *the latter being more specifically non-negativity constraints.*

The corresponding dual problem *is defined as the following problem on m variables $(y_1, \ldots, y_m)^\mathsf{T} = \mathbf{y}$:*

$$\textit{Minimise } \mathbf{b}^\mathsf{T}\mathbf{y}\textit{, subject to } \mathbf{A}^\mathsf{T}\mathbf{y} \geqslant \mathbf{c} \textit{ and } \mathbf{y} \geqslant \mathbf{0}.$$

We will need the duality theorem, see for instance [11, Chap. 5].

Theorem 5 (Duality Theorem). *If the primal problem has an optimal solution* $\mathbf{x}^* = (x_1^*, \ldots, x_n^*)^\mathsf{T}$, *then the dual problem has an optimal solution* $\mathbf{y}^* = (y_1^*, \ldots, y_m^*)^\mathsf{T}$ *such that* $\mathbf{c}^\mathsf{T}\mathbf{x}^* = \mathbf{b}^\mathsf{T}\mathbf{y}^*$. *Furthermore, if* $\mathbf{x}$ *and* $\mathbf{y}$ *are feasible solutions to the primal and the dual problem respectively, such that* $\mathbf{c}^\mathsf{T}\mathbf{x} = \mathbf{b}^\mathsf{T}\mathbf{y}$, *then this common value optimises both objective functions.*

The solution to the linear program (1) will be a consequence of the following:

Theorem 6. *Consider an* $n \times n$ *matrix* $\mathbf{A_n}$ *which is of the form*

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & \cdots & a_{1,n-1} & a_{1,n} \\ -1 & a_{2,2} & \cdots & \cdots & a_{2,n-1} & a_{2,n} \\ 0 & -1 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & -1 & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & \ldots & 0 & -1 & a_{n,n} \end{bmatrix}.$$

Consider the LP maximisation problem $\mathbf{P_n}$ *associated with* $(\mathbf{A_n}, \mathbf{c_n}, \mathbf{b_n}, \mathbf{x_n})$, *where* $\mathbf{b_n} = (b, 0, \ldots, 0)^\mathsf{T}$, $\mathbf{c_n} = (c, 0, \ldots, 0)^\mathsf{T}$ *and* $\mathbf{x_n} = (x_1, \ldots, x_n)^\mathsf{T}$ *are vectors of length* n. *Then, the optimal value of the objective function of* $\mathbf{P_n}$ *is obtained by solving the special case* $\mathbf{A_n x_n} = \mathbf{b_n}$. *And this optimal value is* $\frac{u_n}{u_{n+1}} bc$, *where the sequence* $(u_k)_{k \geqslant 1}$ *is defined by* $u_1 = 1$ *and* $u_{k+1} = \sum_{i=1}^{k} a_{n-k+1,n-k+i}\, u_{k+1-i}$.

Some comments are in order: in matrix form, the linear program (1) is exactly of the form envisaged by Theorem 6 with $b = c = 1$. We will therefore obtain the maximum of p_1 predicted by Theorem 4 by applying Theorem 6 and by proving that the associated sequence $(u_n)_{n \geqslant 1}$ is the sequence of Catalan numbers.

Our goal is therefore to prove Theorem 6. In other words the aim is to solve the following LP maximisation problem $\mathbf{P_n}$ associated with $(\mathbf{A_n}, \mathbf{c_n}, \mathbf{b_n}, \mathbf{x_n})$.

Problem 1. *Maximise* $\mathbf{c_n}^\mathsf{T}\mathbf{x_n}$, *subject to* $\mathbf{A_n x_n} \leqslant \mathbf{b_n}$ *and* $\mathbf{x_n} \geqslant \mathbf{0}$.

We will first compute the special solution given by $\mathbf{A_n x_n} = \mathbf{b_n}$. We obtain:

Proposition 2. *The value of the objective function of Problem 1 in the case where* $\mathbf{A_n x_n} = \mathbf{b_n}$ *is* $\frac{u_n}{u_{n+1}} bc$.

Proposition 2 follows from the following intermediate results.

Lemma 4. *In the case where* $\mathbf{A_n x_n} = \mathbf{b_n}$, *there is* $(\mu_j)_{1 \leqslant j \leqslant n}$ *such that* $x_{n-j} = \mu_j x_{n-j+1}$ *for* $j \neq n$, $b = \mu_n x_1$.

Corollary 1. *Let* $u_1 = 1$ *and* $u_i = \prod_{j=1}^{i-1} \mu_i$ *for* $i \in \{2, \ldots, n+1\}$. *Then, we have the recurrence relation* $u_{k+1} = \sum_{i=1}^{k} a_{n-k+1,n-k+i}\, u_{k+1-i}$.

Corollary 2. *The sequence* $(u_j)_{1 \leqslant j \leqslant n+1}$ *as defined in Corollary 1 satisfies* $x_{n-j} = \frac{u_{j+1}}{u_j} x_{n-j+1}$ *for* $j \neq n$ *and* $b = \frac{u_{n+1}}{u_n} x_1$.

Proposition 2 now follows from Corollary 2 by remarking that the objective function is cx_1.

We now wish to prove that the value of p_1 given by Proposition 2 actually maximises p_1. To this end we consider the dual of Problem 1, namely:

Problem 2. *Minimise* $\mathbf{b_n}^\mathsf{T}\mathbf{y_n}$, *subject to* $\mathbf{A_n}^\mathsf{T}\mathbf{y_n} \geqslant \mathbf{c_n}$ *and* $\mathbf{y_n} \geqslant \mathbf{0}$.

Once again we solve a particular instance of this problem, namely $\mathbf{A_n}^\mathsf{T}\mathbf{y_n} = \mathbf{c_n}$. We will show that

Proposition 3. *The value of the objective function of Problem 2 in the case where* $\mathbf{A_n}^\mathsf{T}\mathbf{y_n} = \mathbf{c_n}$ *is* $\frac{u_n}{u_{n+1}}bc$.

It will be useful to now define a sequence of LP maximisation problems $(\mathbf{P_k})_{k\leqslant n}$ associated with $(\mathbf{A_k}, \mathbf{c_k}, \mathbf{b_k}, \mathbf{x_k})$, where $\mathbf{A_{k-1}}$ is the $(k-1)\times(k-1)$ submatrix at the bottom right of $\mathbf{A_k}$. In other words, $\mathbf{A_k} = (a^k_{i,j})$ where:

$$a^n_{i,j} = a_{i,j} \quad \text{and} \quad a^{k-1}_{i,j} = a^k_{i+1,j+1}$$

and where $\mathbf{b_{k-1}} = (x_{n-k}, 0, \ldots, 0)^\mathsf{T}$ and $\mathbf{x_{k-1}} = (x_{n-k}, \ldots, x_n)^\mathsf{T}$. We will now write $a^n_{i,j}$ instead of $a_{i,j}$ because we shall need to modify the superscript n later.

We next prove the following results.

Proposition 4. *There are* $(U_{k,n})_{1\leqslant k\leqslant n+1}$ *and* $(V_{k,n})_{1\leqslant k\leqslant n+1}$ *such that*

$$\begin{aligned} y_k &= U_{k,n}\, y_1 + V_{k,n} \quad \text{for } 1 \leqslant k \leqslant n \\ 0 &= U_{n+1,n}\, y_1 + V_{n+1,n} \end{aligned}$$

Furthermore,

$$(1)\ V_{k,n} = -cU_{k-1,n-1},\ \text{and}\ (2)\ U_{k,n} = \sum_{\substack{0\leqslant j\leqslant k-1 \\ 0=i_0<\cdots<i_j=k-1}} \prod_{0\leqslant u\leqslant j-1} a^n_{i_u+1,i_{u+1}},\ \text{for } k \geqslant 2.$$

Corollary 3. *We have* $y_1 = \frac{U_{n,n-1}}{U_{n+1,n}}c$ *in the case of equality in* $\mathbf{A_n}^\mathsf{T}\mathbf{y_n} = \mathbf{c_n}$.

Lemma 5. $U_{n+1,n} = u_{n+1}$, *where the sequence* $(u_n)_{n\geqslant 1}$ *is as in Corollary 1.*

Corollary 3 and Lemma 5 together prove Proposition 3, remarking that the objective function is by_1. Since we have found a solution to Problem 1 that gives the value $(u_n/u_{n+1})bc$ for its objective function and we have also found a solution to its dual problem, namely Problem 2, that gives the very same value for the dual objective function, the Duality Theorem implies that this common value maximises both objective functions. This proves therefore Theorem 6.

It remains to compute the specific value of the sequence (u_n) in the case of the LP problem (1). We have:

Proposition 5. *The sequence* (u_n) *is exactly the sequence of Catalan numbers.*

Thus, the maximum value that can be taken by p_1 is $\mathsf{c}_\ell/\mathsf{c}_{\ell+1}$.

The proof of Theorem 4 is completed by determining the optimal values of the remaining variables $p_2, \ldots, p_\ell$. This amounts to solving linear systems $\mathbf{P}_k$ for decreasing values of k, so that all the preceding techniques and results apply.

4 Generalisation and Conclusion

In this paper we have introduced k-localisable distributions for random words. They generalise the notion of exchangeability for random variables in much the same way as k-dependence generalises independence. Furthermore, we believe this notion is of great interest for the study of multipartite games with nonsignalling resources (capturing quantum resources). This raises fundamental questions in graph theory (through chromatic numbers) and distributed computing (through symmetry-breaking problems). We have given a fine-grained analysis of 1-localisable distributions for independent-set words, implying an optimal lower bound of $q \geqslant 4$ for 1-localisable random q-coloured words.

Using the same approach, we can extend Theorem 1 to d-distance q-coloured words in which $d+1$ consecutive letters must receive pairwise distinct colours (Theorem 1 is for $d = 1$). This also corresponds to d-distance q-colouring of a path, a well-known notion in graph theory [23,29]. As d-distance chromatic number of the path is $d+1$, we must have $q \geqslant d+1$. Observe that there is no k-localisable d-distance q-colouring for $k < d$ since $\mathbb{P}(X_i = X_j)$ depends on whether $|i-j| = d$ (it must be 0) or $|i-j| > d$ (it must be >0 for $q < n$). So, we investigate the case $k = d$, the d-localisable distributions for random d-distance q-coloured words of length n. We can show that the minimum number of colours must be $q \geqslant (d+1)\cdot(1+1/d)^d$ for n large enough, generalising Theorem 1. Using the same approach as for $d = 1$, we consider distance-d independent-set words (i.e., binary words with no two ones at distance $\leqslant d$). We show, using the same technique, that the probability of having a one in any fixed position is upper bounded by the ratio of two consecutive Fuss-Catalan numbers of parameters $n/(d+1)$, a generalisation of Catalan numbers, whose limit is $d^d/(d+1)^{d+1}$. This will appear in the full version of the paper.

A step further would be to extend the results to combinatorial structures other than words. The notion of k-localisability extends naturally to graphs as follows. Here each node v of a graph G gets a random variable X_v. Let X_S, for every subset S of nodes, denote the collection of random variables X_s with $s \in S$, and let G_S denote any graph isomorphic to $G[S]$, the subgraph of G induced by S. Then, a probability distribution for random variables (X_v) with support the nodes of G is k-localisable if, for every two subsets I, J at distance more than k in G such that G_I and G_J are connected, the distribution of (X_I, X_J) can depend only on $\{G_I, G_J\}$. The notion of independent-set word transfers also to binary variables encoding independent sets in G. The study of k-localisable q-colourings (or independent-sets) on graphs would have potential applications to understanding the possibilities of distributed quantum computing.

References

1. Almeida, M.L., Bancal, J.-D., Brunner, N., Acín, A., Gisin, N., Pironio, S.: Guess your neighbor's input: a multipartite nonlocal game with no quantum advantage. Phys. Rev. Lett. **104** (2010). https://doi.org/10.1103/PhysRevLett.104.230404

2. Arfaoui, H., Fraigniaud, P.: What can be computed without communications? In: Even, G., Halldórsson, M.M. (eds.) SIROCCO 2012. LNCS, vol. 7355, pp. 135–146. Springer, Berlin (2012). https://doi.org/10.1007/978-3-642-31104-8_12
3. Barenboim, L., Elkin, M.: Distributed graph coloring: fundamentals and recent developments. Synth. Lect. Distrib. Comput. Theory **4**(1) (2013). https://doi.org/10.2200/S00520ED1V01Y201307DCT011
4. Barenboim, L., Elkin, M., Pettie, S., Schneider, J.: The locality of distributed symmetry breaking. In: 53rd Annual IEEE Symposium on Foundations of Computer Science (FOCS), pp. 321–330. IEEE Computer Society Press, October 2012. https://doi.org/10.1109/FOCS.2012.60
5. Brandão, F.G.S.L., Harrow, A.W.: Quantum de Finetti theorems under local measurements with applications. In: 45th Annual ACM Symposium on Theory of Computing (STOC), pp. 861–870. ACM Press, June (2013). https://doi.org/10.1145/2488608.2488718
6. Brandt, S., et al.: LCL problems on grids. In: 35th Annual ACM Symposium on Principles of Distributed Computing (PODC), pp. 101–110. ACM Press, July 2016. https://doi.org/10.1145/3087801.3087833
7. Brassard, G., Broadbent, A., Hänggi, E., Méthot, A.A., Wolf, S.: Classical, quantum and nonsignalling resources in bipartite games. Theor. Comput. Sci. **486**, 61–72 (2013). https://doi.org/10.1016/j.tcs.2012.12.017
8. Cameron, P.J., Montanaro, A., Newman, M.W., Severin, S., Winter, A.: On the quantum chromatic number of a graph. Electron. J. Comb. **14**, R81 (2007)
9. Caves, C.M., Fuchs, C.A., Schack, R.: Unknown quantum states: the quantum de Finetti representation. J. Math. Phys. **43**, 4537 (2001). https://doi.org/10.1063/1.1494475
10. Chang, Y.-J., Li, W., Pettie, S.: An optimal distributed $(\delta+1)$-coloring algorithm? In: 50th Annual ACM Symposium on Theory of Computing (STOC), pp. 445–456. ACM Press, June 2018. https://doi.org/10.1145/3188745.3188964
11. Chvátal, V.: Linear Programming. W. H. Freeman, New York (1983)
12. Czekaj, Ł., Pawłowski, M., Vértesi, T., Grudka, A., Horodecki, M., Horodecki, R.: Quantum advantage for distributed computing without communication. Phys. Rev. A **92**, 032122 (2015). https://doi.org/10.1103/PhysRevA.92.032122
13. Fraigniaud, P., Heinrich, M., Kosowski, A.: Local conflict coloring. In: 57th Annual IEEE Symposium on Foundations of Computer Science (FOCS), pp. 625–634. IEEE Computer Society Press, October 2016. https://doi.org/10.1109/FOCS.2016.73
14. Gavoille, C., Kosowski, A., Markiewicz, M.: What can be observed locally? In: Keidar, I. (ed.) DISC 2009. LNCS, vol. 5805, pp. 243–257. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-04355-0_26
15. Ghaffari, M., Kuhn, F., Maus, Y.: On the complexity of local distributed graph problems. In: 49th Annual ACM Symposium on Theory of Computing (STOC), pp. 784–797. ACM Press, June 2017. https://doi.org/10.1145/3055399.3055471
16. Harris, D.G., Schneider, J., Su, H.-H.: Distributed $(\Delta+1)$-coloring in sublogarithmic rounds. In: 48th Annual ACM Symposium on Theory of Computing (STOC), pp. 465–478. ACM Press, June 2016. https://doi.org/10.1145/2897518.2897533
17. Hoeffding, W., Robbins, H.: The central limit theorem for dependent variables. Duke Math. J. **15**, 773–780 (1948). https://doi.org/10.1215/S0012-7094-48-01568-3
18. Holroyd, A.E., Hutchcroft, T., Levy, A.: Finitely dependent cycle coloring. Electron. Commun. Probab. **23**, 1–8 (2018). https://doi.org/10.1214/18-ECP118

19. Holroyd, A.E., Liggett, T.M.: Symmetric 1-dependent colorings of the integers. Electron. Commun. Probab. **20**, 1–8 (2015). https://doi.org/10.1214/ECP.v20-4070
20. Holroyd, A.E., Liggett, T.M.: Finitely dependent coloring. Forum Math., Pi **4**, 1–43 (2016). https://doi.org/10.1017/fmp.2016.7
21. Holroyd, A.E., Schramm, O., Wilson, D.B.: Finitary coloring. Ann. Probab. **45**, 2867–2898 (2017). https://doi.org/10.1214/16-AOP1127
22. Işlak, U.: Asymptotic normality of random sums of m-dependent random variables. Stat. Probab. Lett. **109**, 22–29 (2016). https://doi.org/10.1016/j.spl.2015.10.015
23. Kramer, F., Kramer, H.: A survey on the distance-colouring of graphs. Discrete Math. **308**, 422–426 (2008). https://doi.org/10.1016/j.disc.2006.11.059
24. Le Gall, F., Magniez, F.: Sublinear-time quantum computation of the diameter in CONGEST networks. In: 37th Annual ACM Symposium on Principles of Distributed Computing (PODC), pp. 337–347. ACM Press, July 2018. https://doi.org/10.1145/3212734.3212744
25. Le Gall, F., Nishimura, H., Rosmanis, A.: Quantum advantage for the LOCAL model in distributed computing. Technical report, October 2018. arXiv:1810.10838v1 [quant-ph]
26. Linial, N.: Locality in distributed graphs algorithms. SIAM J. Comput. **21**, 193–201 (1992). https://doi.org/10.1137/0221015
27. Naor, M.: A lower bound on probabilistic algorithms for distributive ring coloring. SIAM J. Discrete Math. **4**, 409–412 (1991). https://doi.org/10.1137/0404036
28. Naor, M., Stockmeyer, L.: What can be computed locally. SIAM J. Comput. **24**, 1259–1277 (1995). https://doi.org/10.1137/S0097539793254571
29. Niranjan, P.K., Kola, S.R.: The k-distance chromatic number of trees and cycles. AKCE Int. J. Graphs Comb. (2017, in press). https://doi.org/10.1016/j.akcej.2017.11.007