# The Action of a Few Permutations on r-Tuples Is Quickly Transitive

**Joel Friedman,[1] Antoine Joux,[2] Yuval Roichman,[3,*] Jacques Stern,[4] Jean-Pierre Tillich[5,†]**

[1] *Department of Mathematics, University of British Columbia, Vancouver V6T 1Z2, Canada; e-mail: jf@math.ubc.ca*

[2] *CELAR, France; e-mail: joux@ssig.celar.fr*

[3] *Department of Applied Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02138; e-mail: yuval@math.mit.edu*

[4] *Ecole Normale Supérieure, 45 rue d'ULM, 75005 Paris, France; e-mail: stern@dmi.ens.fr*

[5] *GREYC, Université de Caen, France; e-mail: tillich@info.unicaen.fr*

*Received 4 January 1996; revised 23 January 1996; accepted 30 January 1997*

**ABSTRACT:** We prove that for every $r$ and $d \geq 2$ there is a $C$ such that for most choices of $d$ permutations $\pi_1, \pi_2, \ldots, \pi_d$ of $S_n$, the following holds: for any two $r$-tuples of distinct elements in $\{1, \ldots, n\}$, there is a product of less than $C \log n$ of the $\pi_i$s which map the first $r$-tuple to the second. Although we came across this problem while studying a rather unrelated cryptographic problem, it belongs to a general context of which random Cayley graph quotients of $S_n$ are good expanders. © 1998 John Wiley & Sons, Inc. Random Struct. Alg., 12, 335–350, 1998

## 1. INTRODUCTION

Choose $d \geq 2$ permutations $\pi_1, \ldots, \pi_d$ at random in the symmetric group $S_n$. Consider now two $r$-tuples $(u_1, u_2, \ldots, u_r)$ and $(v_1, v_2, \ldots, v_r)$ of distinct elements of $\{1, 2, \ldots, n\}$. Is there always a short product of these $\pi_i$s which maps the first

* Partially supported by University of British Columbia.

† Part of this work was done while the author was visiting the University of British Columbia, BC, Canada.

Correspondence to: Joel Friedman

© 1998 John Wiley & Sons, Inc. CCC 1042-9832/98/040335-16

$r$-tuple to the second one? We prove in what follows that for almost all choices of these permutations we only need products of length at most $C \log n$, $C$ being a constant depending on $d$ and $r$. When such a condition is met by $d$ permutations $\pi_1, \pi_2, \ldots, \pi_d$, we say that the action of these permutations is $C$ quickly $r$-transitive.

This issue has been raised by the study of the security of some low cost cryptographic devices constructed from permutation automata (see Section 3). We will exhibit a probabilistic algorithm which reconstructs the secrete device, and which can be shown to run in polynomial time by using the aforementioned result. This show that such schemes are highly insecure.

Actually the results obtained here are more general than that, and should be put in the broader context of whether or not Cayley graphs over $S_n$ are good expanders for a fixed number of random generators, and/or have a small diameter and mixing time. This is quite an important open problem, for a survey see [5, 16, 17]. A solution of this problem is of great theoretical importance, while a positive solution would be useful for instance for generating random permutations quickly.

Here we take a first step toward a solution of the above open problem. We show that for bounded $r$ the random Schreier graphs $S_n/S_{n-r}$ (i.e., the quotients $S_n/S_{n-r}$ of Cayley graphs over $S_n$) are good expanders, have a small diameter and mixing time.

More precisely we study the following random model: we choose independently $d$ permutations of the numbers from 1 to $n, \pi_1, \pi_2, \ldots, \pi_d$, each permutation equally likely. We construct a directed graph, $G = (V, E)$ with vertex set the set of $r$-tuples of distinct elements of $\{1, \ldots, n\}$ and there is a directed edge from $(u_1, u_2, \ldots, u_r)$ to $(v_1, v_2, \ldots, v_r)$ iff $(v_1, v_2, \ldots, v_r) = (\pi_i(u_1), \pi_i(u_2), \ldots, \pi_i(u_r))$ for one of those $\pi_i$s. We denote this probability space of random directed graphs $\mathscr{G}_{n,d,r}$. Such graphs are $d$-regular (and may have multiple edges or self-loops). We will consider the associated space of undirected graphs $\mathscr{G}^*_{n,d,r}$ too. This space is simply obtained from the first one by changing each directed graph into an undirected one, just by replacing each directed edge of the former graph by an undirected edge. The latter space is therefore formed by undirected $2d$-regular graphs.

It should be noted that $r = 1$ corresponds to the common probabilistic model of $2d$-regular graphs, $\mathscr{G}_{n,2d}$ (as studied in [9, 12], for example), and that $r = n$ is just the common probabilistic model of random $2d$-regular Cayley graphs over $S_n$.

We will show that for every fixed $r$ and for all $d \geq 2$ almost all graphs in $\mathscr{G}^*_{n,d,r}$ have a small second eigenvalue when the number of vertices becomes large, and that this implies that almost all graphs of $\mathscr{G}^*_{n,d,r}$ and $\mathscr{G}_{n,d,r}$ are good expanders, and also have a small diameter and mixing time.

The counting methods used by Bollobas, for example, to prove this kind of results for random regular graphs (see [6, 7, 8]) seem to fail in our case. In fact we prove our results by generalizing Broder–Shamir [9] and Friedman [12] spectral approach.

Moreover we are interested here in obtaining results about the expansion properties of directed graphs too.

Besides being a step toward the study of random Cayley graphs over $S_n$:

1. Those results cover the case of graphs of very small degree, which was not really settled by previous works of A. Broder, J. Friedman, and E. Shamir

(especially the case $d = 2$ which is worth studying!). For instance our results show that as soon as $d \geq 2$, random $2d$-regular graphs have almost always a small second largest eigenvalue, and are therefore good certifiable expanders.

2. We address here the issue of the expansion properties of directed graphs too. We provide here tools to achieve such results, provided that the directed graphs we are interested in have the same indegree and outdegree for each vertex. Although most of the theory on expanders has been developed for undirected graphs—we wish to lay emphasis on the fact that for some applications, expansion properties of directed graphs have to be estimated, this is the case, for example, in [23, 25]—and in this article (see Section 3).

Finally we mention that $\mathscr{G}^*_{n,d,r}$ arises naturally from $\mathscr{G}^*_{n,d} = \mathscr{G}^*_{n,d,1}$, as does $\mathscr{G}_{n,d,r}$ from $\mathscr{G}_{n,d} = \mathscr{G}_{n,d,1}$, in the following way. Given maps of graphs, $G \to B$ and $H \to B$, there is a fiber product (see [13]) $G \times_B H$ with a natural map to $B$. We can similarly form $G^r = G \times_B \cdots \times_B G$. If $G \to B$ is a covering map, then $G^r$ naturally breaks into certain components, such as the diagonal component and the component $G^{\wedge r}$ whose vertices are those $r$-tuples of $G$ vertices which are all distinct. In our case a $G \in \mathscr{G}^*_{n,d}$ comes with a natural map to $B = W_d$, the bouquet of $d$ self-loops (the map given via the $d$ permutations); $G^{\wedge r}$ would just be a random element in $\mathscr{G}^*_{n,d,r}$. Similarly a random $G \in \mathscr{G}_{n,d}$ gives rise to $G^{\wedge r}$, which is the same as a random element of $\mathscr{G}_{n,d,r}$.

## 2. THE MAIN THEOREM

### 2.1. The Main Theorem and Its Consequences

Before describing our main theorem, let us introduce some notations. Let us recall that the adjacency matrix $A = (a_{ij})$ of a graph with $n$ vertices is the $n \times n$ matrix indexed by the vertices of the graph, such that entry $a_{ij}$ is the number of edges between $i$ and $j$. In our case the adjacency matrix $A = (a_{IJ})$ of the graph $G^*$ of $\mathscr{G}^*_{n,d,r}$ obtained by choosing the permutations $\pi_1, \pi_2, \ldots, \pi_d$, is defined by,

$$a_{IJ} = \#\{l \mid \pi_l(I) = J\} + \#\{l \mid \pi_l(J) = I\};$$

here we understand that $I$ and $J$ are $r$-tuples of distinct numbers, and each $\pi_i$ operates on $r$-tuples in the obvious way. Let us note that each self-loop counts twice for the corresponding (diagonal) entry of the adjacency matrix. $G^*$ is a $2d$-regular graph, therefore its adjacency matrix has real eigenvalues $2d = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$ with $N = n(n-1)\cdots(n-r+1)$; let $\rho = \max_{i \geq 2} |\lambda_i| = \max(\lambda_2, -\lambda_N)$. Our main result asserts that $\rho$ is almost always well separated from $2d$.

**Theorem 2.1.** *For fixed $d, r$ we have*

$$\mathrm{E}\{\rho^k\} \leq \left[ 2d \left( \frac{\sqrt{2d-1}}{d} \right)^{1/(r+1)} \left( 1 + O\frac{\log\log n}{\log n} \right) \right]^k \quad \text{if } k \leq 2\lfloor (r+1)\log_{d^2/(2d-1)} n \rfloor$$

$$\mathrm{Prob}\left\{ \rho \leq (1+\epsilon) 2d \left( \frac{\sqrt{2d-1}}{d} \right)^{1/(r+1)} \right\} = 1 - o_\epsilon(1) \quad \text{for every } \epsilon > 0.$$

However, as corollaries we obtain that almost all graphs of $\mathscr{G}^*_{n,d,r}$ are good expanders, and the same property holds for almost all graphs of $\mathscr{G}_{n,d,r}$. The result for directed graphs follows from an argument relating the (edge-)expansion properties of a directed graphs to the (edge-)expansion properties of its associated undirected graph, obtained by replacing each directed edge by an undirected edge, and then relating the (edge-)expansion properties of the undirected graph to its second largest eigenvalue in absolute value.

More precisely:

**Theorem 2.2.** *For every fixed $d \geq 2$, $r \geq 1$, and real $\epsilon$, the probability that the graph $\mathscr{G}_{n,d,r}$ is a $c$-expander tends to 1, as $n$ tends to infinity, for $c = ((1-\epsilon)/2)(1 - (\sqrt{2d-1}/d)^{1/(1+r)})$.*

A $c$-expander is defined as follows:

**Definition 2.3.** *A directed graph $G(V, E)$ with $n$ vertices is a $c$-expander if, for any subset $X$ of vertices with size $\leq n/2$ the following inequalities hold*

$$|N^+(X)| \geq c|X| \quad \text{and} \quad |N^-(X)| \geq c|X|,$$

*where $N^+(X)$ (respectively, $N^-(X)$) denotes the set of vertices not in $X$ which are endpoints (respectively, initial points) of an edge with initial point (respectively, endpoint) in $X$.*

Note that Theorem 2.2 is far from being optimal, especially for $r = 1$, where the standard counting argument as used, for example, in [7, 12] for undirected graphs gives us sharper estimates on the expansion constant. It must be noted here that this argument applies to the case $d = 2$ to the directed graph model with $r = 1$, and shows that $c > 0.16$, whereas the bound of the Theorem 2.2 gives only $c > 0.034$ (see [15]). Unfortunately, this counting argument seems to fail for $r > 1$. Nevertheless, for most practical applications, this theorem is actually sufficient (see, for example, Section 3).

*Remark 2.4.* It should be pointed out that by using Theorem 2.1 and classical results on Markov chains (see Section 2 of [24]) imply that for ever fixed $d \geq 2$ and $r \geq 1$, random walks on graphs of $\mathscr{G}^*_{n,d,r}$ are rapidly mixing for almost all choices of such graphs, and that the mixing time of such random walks is not more than $O(\log n)$ almost surely. The same result holds for the directed graph model. This is a consequence of a result obtained by Mihail in [19] (this result is recalled in Theorem 1 in [24] too) and of Theorem 2.2.

### 2.2. Proof of the Main Theorem

In this section we prove Theorem 2.1. Throughout this section we view $r$ as fixed. We also use the following notation.

*Notation.* Let $N = n(n-1)\cdots(n-r+1)$ be the number of vertices of the graphs $\mathcal{G}^*_{n,d,r}$ or $\mathcal{G}_{n,d,r}$ we consider here. We note $\Pi$ the alphabet $\{\pi_1, \pi_1^{-1}, \pi_2, \pi_2^{-1}, \ldots, \pi_d, \pi_d^{-1}\}$, where the $\pi_i$s are permutations in $S_n$.

We begin by describing the general approach, which follows essentially the approach initiated by Broder and Shamir in [9].

The idea of the proof is to get a rather tight upper bound on $\mathrm{E}\{\rho^{2k}\}$ for rather large values of $k$. This is obtained by upper bounding this quantity by the expectation of the number of closed walks of length $2k$ minus $(2d)^{2k}$, when we choose a random graph from $\mathcal{G}^*_{n,d,r}$.

This is justified by what follows: if we call $A = (a_{ij})$ the adjacency matrix of a random graph of $\mathcal{G}^*_{n,d,r}$, then an entry $b_{ij}$ of $A_k^2$ represents the number of walks on the graph of length $2k$ from $i$ to $j$. Therefore,

$$\text{Number of closed walks of length } 2k = \mathrm{Tr}(A^{2k}) = \sum_{i=1}^{N} \lambda_i^{2k} \geq (2d)^{2k} + \rho^{2k}.$$

Let us note that the expectation $\mathrm{E}\{i \to^{2k} i\}$ of the number of walks starting from a vertex $i$ and ending at the same vertex, can be seen as the probability of the following event. We first choose a random word $w = w_1 w_2 \cdots w_{2k}$ in $\Pi^{2k}$ (all the $(2d)^{2k}$ possible words are chosen with the same probability $1/(2d)^{2k}$), and then we assign the letters $\pi_i$ a permutation of $S_n$ chosen uniformly at random. We have $\mathrm{E}\{i \to^{2k} i\} = (2d)^{2k}\,\mathrm{Prob}\{w_1 w_2 \cdots w_{2k}(i) = i\}$.

So for each word $w$ of length $2k$ over the alphabet $\Pi$, and for each vertex $v$ of the graph $\mathcal{G}^*_{n,d,r}$ let $P(w,v)$ denote the probability that when $\pi_1, \ldots, \pi_d$ are assigned permutations at random, the walk determined by $w$ starting in $v$ ends in $v$. Clearly $P(w,v) = P(w)$ is independent of $v$. Hence we have

$$\begin{aligned}
\mathrm{E}\{\rho^{2k}\} &\leq \left( \sum_{i}^{N} \mathrm{E}\left\{ i \overset{2k}{\to} i \right\} \right) - (2d)^{2k} \\
&\leq \left( \sum_{w \in \Pi^{2k},\, 1 \leq i \leq N} P(w,i) \right) - (2d)^{2k} \\
&\leq N \left( \sum_{w \in \Pi^{2k}} P(w) \right) - (2d)^{2k} \\
&\leq N \sum_{w \in \Pi^{2k}} \left( P(w) - \frac{1}{N} \right). \qquad (1)
\end{aligned}$$

Our task is reduced now to estimate the quantity $P(w) - 1/N$. First we will classify the words $w$ which give us the same probability to close the walk.

We say that a word, $w \in \Pi^*$ is *irreducible* if $w$ contains no consecutive occurrence of $\pi$ and $\pi^{-1}$ for an $\pi \in \Pi$. For each word, $w \in \Pi^{2k}$, we obtain a unique irreducible word, $w'$, by canceling all consecutive occurrences of $\pi$ and $\pi^{-1}$ in $w$; we say that $w$ *reduces* to $w'$. In this case we clearly have $P(w) = P(w')$. So to estimate $P(w)$ it suffices to do so when $w$ is irreducible.

Similarly, if $w = w_1 \cdots w_{2k}$ with $w_i \in \Pi$, then if $w_1 = w_{2k}^{-1}$ we have that $P(w) = P(w')$ with $w' = w_2 \cdots w_{2k-1}$. So we say that $w = w_1 \cdots w_{2k}$ is *strongly irreducible* if

$w$ is irreducible and $w_1 \neq w_{2k}^{-1}$. Again, repeatedly canceling first and last letters of $w$ if they are inverses (and of consecutive $\pi$ and $\pi^{-1}$ occurrences) gives us a unique strongly irreducible word, $w'$, for which $P(w) = P(w')$. We say that $w$ *strongly reduces to* $w'$, and it suffices to estimate $P(w)$ when $w$ is strongly irreducible.

We say that $w \in \Pi^*$ is *periodic* if it is of the form $u^m$ for some $u \in \Pi^*$ and $m \geq 2$. Our main lemma used to estimate $P(w)$ is the following:

**Lemma 2.5.** *Let $w$ be a strongly irreducible word of length $2s$ with $s > 1$ such that $w$ is not periodic. Then*

$$P(w) = \frac{1}{n^r} + O\left(\frac{s^{2r+2}}{n^{r+1}}\right).$$

This lemma will be proved in Appendix A. We finish the proof of the theorem assuming this lemma.

Let $k$ be fixed. Let $R$ be the set of words in $\Pi^{2k}$ which strongly reduce to a word which is empty or periodic, and let

$$p_{2k} = \frac{|R|}{|\Pi^{2k}|} = \frac{|R|}{(2d)^{2k}}$$

be the probability that a random word in $\Pi^{2k}$ belongs to $R$. Clearly we have

$$\begin{aligned}\sum_{w \in \Pi^{2k}} \frac{P(w) - \frac{1}{N}}{(2d)^{2k}} &\leq p_{2k} \max_{w \in R}\left(P(w) - \frac{1}{N}\right) + (1 - p_{2k}) \max_{w \notin R}\left(P(w) - \frac{1}{N}\right) \\ &\leq p_{2k} + \max_{w \notin R}\left(P(w) - \frac{1}{N}\right). \qquad (2)\end{aligned}$$

In Lemma 2.5 we estimated the right-hand side of the above expression; the left-hand side can be estimated by the following lemma.

**Lemma 2.6.** *We have*

$$p_{2k} \leq (k+1)\left(\frac{2d-1}{d^2}\right)^k.$$

The proof of this statement is in Appendix B.

Applying Lemmas 2.5 and 2.6 to Eq. (2) we get

$$\sum_{w \in \Pi^{2k}} \frac{P(w) - \frac{1}{N}}{(2d)^{2k}} \leq (k+1)\left(\frac{2d-1}{d^2}\right)^k + O\frac{k^{2r+2}}{n^{r+1}}.$$

Combining this with Eq. (1), and using $N \le n^r$, yields

$$\mathrm{E}\{\rho^{2k}\} \le \left( O\frac{k^{2r+2}}{n} + n^r(k+1)\left(\frac{2d-1}{d^2}\right)^k \right)(2d)^{2k}.$$

Taking $k$ to be the greatest integer $K$ less than $(r+1)\log_{d^2/(2d-1)} n$ yields

$$\left(\mathrm{E}\{\rho^{2K}\}\right)^{1/(2K)} \le 2d\, 2^{1/(2K)} \max\left( \left[ O\frac{K^{2r+2}}{n} \right]^{1/(2K)}, \left[ n^r(K+1)\left(\frac{2d-1}{d^2}\right)^K \right]^{1/(2K)} \right)$$

$$\le 2d\left(\frac{\sqrt{2d-1}}{d}\right)^{1/(r+1)} \left(1 + O\frac{\log\log n}{\log n}\right).$$

Finally, Hölder's inequality implies that for any $k \le 2K$,

$$\mathrm{E}\{\rho\} \le \left(\mathrm{E}\{\rho^k\}\right)^{1/k} \le \left(\mathrm{E}\{\rho^{2K}\}\right)^{1/(2K)} \le 2d\left(\frac{\sqrt{2d-1}}{d}\right)^{1/(r+1)} \left(1 + O\frac{\log\log n}{\log n}\right), \tag{3}$$

which completes the proof of the first statement of the main theorem. The second claim is just a consequence of Markov's inequality,

$$\mathrm{Prob}\{\rho > \alpha\} \le \frac{\mathrm{E}\{\rho^{2l}\}}{\alpha^{2l}},$$

by putting $l = K$ and $\alpha = (1+\epsilon)2d(\sqrt{2d-1}/d)^{1/(r+1)}$, and using inequality (3).

### 2.3. The Link Between Expansion Properties of $\mathscr{G}_{n,d,r}$ and $\mathscr{G}^*_{n,d,r}$

To obtain this link, we will not compare directly the expansion constant of a directed graph $G$ and its associated undirected graph $G^*$ (which is the graph obtained from the directed one by replacing each directed edge by an undirected edge), but we will compare their isoperimetric number first. This number is defined as follows.

**Definition 2.7.** *The isoperimetric number $i$ of a directed graph $G(V,E)$ with $n$ vertices is the largest number, $i$, such that for any subset $X$ of vertices with size $\le n/2$ the following inequalities hold*

$$|\partial^+(X)| \ge i|X|,$$
$$|\partial^-| \ge i|X|,$$

*where $\partial^+(X)$ (respectively, $\partial^-(X)$) denotes the set of edges with initial point (respectively, endpoint) in $X$ and endpoint (respectively, initial point) in $V \setminus X$. The isoperimetric number $i^*$ of a undirected graph $G^*(V,E)$ with $n$ vertices is the largest number,*

*$i^*$, such that for any subset $X$ of vertices with size $\leq n/2$ one has*

$$|\partial(X)| \geq i^*|X|,$$

*where $\partial(X)$ denotes the set of edges between $X$ and $V \setminus X$.*

For a *regular* directed graph $G$ we have:

**Lemma 2.8.** *Let $G$ be a regular directed graph, and $G^*$ its associated undirected graph. Let $i$ and $i^*$ be the isoperimetric numbers of $G$ and $G^*$, respectively. Then*

$$i = \frac{i^*}{2}.$$

*Proof.* For a directed regular graph, it is readily seen that for any subset, $X$, of vertices of the graph $|\partial^+(X)| = |\partial^-(X)|$. With the associated undirected graph, we have $|\partial(X)| = |\partial^+(X)| + |\partial^-(X)|$. Therefore $|\partial(X)| = 2|\partial^+(X)| = 2|\partial^-(X)|$ and these equalities imply the lemma. ■

Now we can bring in the connection between the isoperimetric number $i^*$ of an undirected regular graph $G^*(V, E)$ and the second smallest eigenvalue of the Laplacian of the graph $\lambda(G^*)$ (recall here that the Laplacian is the matrix $\text{diag}_{v \in V}(\deg(v)) - A$, where $A$ is the adjacency matrix of the graph $G^*(V, E)$),

$$i^* \geq \frac{\lambda(G^*)}{2}.$$

This inequality can be found in [20] (see Theorem 4.1), and is essentially due to N. Alon and V. D. Milman (see [3]).

Now, it just remains to give a connection between this parameter and the isoperimetric number $i$. Since for every subset of vertices $X$ of a directed regular graph $G(V, E)$ of degree $d$, one has

$$|N^+(X)| \geq \frac{1}{d}|\partial^+(X)| \quad \text{and} \quad |N^-(X)| \geq \frac{1}{d}|\partial^-(X)|,$$

one deduces that such a graph is an $i/d$-expander.

From these consideration we can deduce the lemma:

**Lemma 2.9.** *Let $G$ be a regular directed graph of degree $d$, $G^*$ its associated undirected graph. Let $\lambda$ be the second largest eigenvalue in absolute value of the adjacency matrix of $G^*$. Then $G$ is a $c$-expander, where $c = \frac{1}{2} - \lambda/4d$.*

*Proof.* It is readily seen that $G^*$ is a $2d$-regular graph. Therefore the second smallest eigenvalue of the Laplacian of this graph is greater (or equal) than $2d - \lambda$. The preceding discussion gives that the isoperimetric number $i$ of $G$ satisfies $i = i^*/2 \geq (2d - \lambda)/4$ (where $i^*$ is the isoperimetric number of $G^*$). Consequently $G$ is at lest a $(\frac{1}{2} - \lambda/4d)$-expander. ■

Corollary 2.2 appears therefore as a consequence of Lemma 2.9 and Theorem 2.1.

## 3. AN APPLICATION TO CRYPTOGRAPHY

With the development of memory card technology, and in particular the development of prepaid cards, that give access to some service, the protection of the service issuer against fraud is becoming a critical issue. However, for low-cost applications, the service provider might not afford to replace his memory cards by smart cards containing classical cryptographic protocols for identification of genuine cards. Still, it might be possible to devise (classical) identification protocols to improve the security offered by memory cards, while keeping their cost within reasonable bounds. In particular, permutation automata have been considered as offering a general design methodology for such purposes. We recall here some definitions and describe an identification protocol which, albeit never published, has been circulating in the smart-card community.

**Definition 3.1.** *An automaton is a tuple* $(Q, B, \delta, q_0)$ *where*:

1. $Q$ *is a finitely nonempty set of states*,
2. $B$ *is a finite nonempty set of input symbols or basic actions*,
3. $\delta$ *is the next-state function which maps* $Q \times B$ *into* $Q$,
4. $q_0 \in Q$ *is the initial state*.

**Definition 3.2.** $A$ permutation *automaton is such that*, *for every action* $b$, *the function* $\delta(\cdot, b)$ *is a permutation of* $Q$.

We let $A = B^*$ be the set of finite sequences of basic actions, and we extend the domain of the function $\delta$ to $A$ in the usual way.

We now consider the special case $B = \{0, 1\}$, $Q = [1 \cdots n]$ and $q_0 = 1$. Moreover, we fix a parameter $L$ and we let $B^L$ be the subset of all words of length $L$ in $A$.

The basic idea of the identification protocol is to install a secret permutation automaton in each memory card. All these automata are generated from a master key, and the card reader can reconstruct it before starting the identification. During the identification itself, the card reader sends a random query $w$ from $B^L$ to the memory card, which computes the image of $q_0$ by $w$, and outputs this result. The card reader checks this result, accepts the card and issues the service if it is correct, and rejects otherwise.

What is important in the above, is the fact that the length of the queries is fixed. In the designer's mind, this was presumably enough to prevent a statistical analysis of the outputs. This was even expected to remain true if the user was allowed to make repeated experiments with the automaton. Thus our cryptographic problem can be interpreted as a problem of learning theory for automata (see, for example, [1, 2, 4, 11, 14, 21, 22]). However, for all the attacks based on learning theory that we are aware of, the number of experiments needed to construct automata which stimulates the identification grows with the length of the query. Thus, these attacks can be defeated by limiting the number of identifications that a single card can

perform. Yet, we show that even in this context, the above identification algorithm is insecure by describing and algorithm that allows to reconstruct the given automaton after a few queries (the number of queries does not depend on their size). The algorithm is straightforward by the main achievement of the paper is an actual proof that, with high probability, the algorithm succeeds with only a polynomial number of queries, when the permutation automaton is chosen at random. This is by no means obvious, and relies on Theorem 2.2. We first describe the algorithm. $k$ and $l$ are parameters to be specified later.

*Initialization step.* Fix a set $U = \{u_1, \ldots, u_t\}$ of elements of $A$, all of the same length $k < L$. And given any state $q$ in $Q$, let $E_q$ be the equivalence relation on $U$ defined by

$$\delta(q, u_i) = \delta(q, u_j).$$

*Sampling step.* Pick at random (without repetitions) elements $a_i$ of $A$ of (small) length $\leq l$ (s.t. $k + l < L$); set $q_i := \delta(q_0, a_i)$. Repeat until the set of equivalence relations $E_{q_i}$ has $n$ distinct members. Renumber the chosen elements so that $E_{q_1}, \ldots, E_{q_n}$ are distinct and discard the other values. *Comment*: *computing $E_{q_i}$ can be done by comparing the answers given by the automaton to the concatenation $a_i^\frown u_j^\frown w_i$ where $w_i$ is a word independent of $j$ such that $a_i^\frown u_j^\frown w_i$ has length $L$ and $j$ ranges over* $\{1, \ldots, t\}$.

*Computing step.* For $i := 1$ to $n$ and for each $b$ in $B$ compute $E_{\delta(q_i, b)}$. *Comment*: *this can be done by comparing the answers given by the automaton to the concatenation $a_i^\frown b^\frown u_j^\frown w$ where $w$ is any fixed word of appropriate length and $j$ ranges over* $\{1, \ldots, t\}$.

*Identification step.* Choose random words $W_i$ of length $L$. Compute $E_{\delta(q_0, W_i)}$ using the table built in the computing step and identify this equivalence relation with the output of the automaton under $W_i$. Repeat until all output states have been identified.

The sampling step of this algorithm is the most crucial one, if it can construct $n$ elements $a_i$ such that all $E_{q_i}$ are distinct then the whole algorithm will succeed, otherwise it will not. In order to prove this fact, let us remark that $E_{q_i} \neq E_{q_j}$ implies $q_i \neq q_j$, thus if we construct $n$ different values $E_{q_i}$, then the $q_i$s range over all possible values in $Q$. Moreover, the equivalence classes $E_q$ induced by $U$ uniquely identify each state $q$, and the rest of the algorithm will succeed.

We can easily characterize the properties that are needed for the sampling step to succeed; they are the following:

1. Any state can be reached from $q_0$ by applying a sequence of actions of length $\leq l$.
2. There exists $U$ that *discriminates* the given automaton, i.e., such that two different states $q \neq q'$ have different equivalence relations $E_q \neq E_{q'}$.

We now claim that the two above properties are the consequence of expansion properties of directed random graphs $G_1$ and $G_3$ of $\mathcal{G}_{n,2,1}$ and $\mathcal{G}_{n,2,3}$, respectively,

both defined by the two (random) permutations $\pi_1 = \delta(\cdot, 0)$ and $\pi_2 = \delta(\cdot, 1)$. From Corollary 2.2 we know that $G_1$ is almost always an $\alpha_1$-expander and $G_3$ an $\alpha_3$-expander (the corresponding values of $\alpha_i$ are given by the corollary). We need the following theorem, that states that expanders have small diameters:

**Theorem 3.3.** *If $G$ is an $\alpha$-expander with $v$ vertices, then the diameter of $G$ is smaller than*

$$2\big(1 + \log_{1+\alpha}(v)\big).$$

A proof of this statement can be found, for example, in [25]. The idea of the proof is to look at two distinct vertices $x$ and $y$, and to consider the ball of radius $r$ centered at $x$ (the vertices attained by a directed path of length $\leq r$ starting from $x$) and the "inverse" ball of radius $r$ centered at $y$ (the vertices from which we can attain $y$ by a directed path of length $\leq r$). The crux is that the size is exponential in $r$ (the exponent is at least $1 + \alpha$) therefore the two balls must intersect for a radius which is logarithmic in $v$- and this gives a directed path traversing the point of intersection and going from $x$ to $y$, of length at most twice the radius of the balls.

In particular, $G_1$ has diameter smaller than $2(1 + \log_{1+\alpha_1}(n))$, thus the first property needed for the algorithm to succeed holds as soon as $l > 2(1 + \log_{1+\alpha_1}(n))$. Moreover, the number of elements of length smaller than $l$ is polynomial in $n$ if we choose $l = 2(2 + \log_{1+\alpha_1}(n))$. Thus the sampling step will take polynomial time, once $U$ is correctly chosen.

We now want to prove that the small diameter of $G_3$ implies the second property. Let us remark that in order to prove this property it suffices for any pair of states $(x, y)$ to produce a pair of words of the same length (smaller than $k$) $w_1$ and $w_2$ such that $\delta(x, w_1) = \delta(x, w_2)$ and $\delta(y, w_1) \neq \delta(y, w_2)$. $w_1$ and $w_2$ can be completed to length $k$ by appending any fixed word of appropriate length at their ends. We construct $U$ as the union of all words $w_1(x, y)$ and $w_2(x, y)$.

Let $d_3$ denote the diameter of $G_3$, and given a pair $(x, y)$, choose $(z, r, s)$ such that $x, y, z, r, s$ are all distinct (we suppose that $n \geq 5$). Since $G_3$ has diameter $d_3$, there exists a word $m_1$ of length $\leq d_3$ that goes from edge $(x, y, z)$ to edge $(y, z, r)$ and likewise a word $m_2$ of length $\leq d_3$ that goes from edge $(x, y, z)$ to edge $(y, z, s)$. Let $w_1 = m_1 \hat{} m_2$ and $w_1 = m_2 \hat{} m_1$, then clearly $w_1$ and $w_2$ have the same length and,

$$\delta(x, w_1) = z = \delta(x, w_2) \quad \text{and} \quad \delta(y, w_1) = s \neq r = \delta(y, w_2).$$

Thus the second property holds if, $k \geq 2d_3 = 4(1 + \log_{1+\alpha_3}(n))$. Moreover, it suffices to choose $U = B^k$, the set of all words of length $k$, whose size is polynomial in $n$ if we choose $k = 4(1 + \log_{1+\alpha_3}(n))$.

Thus, we have constructed a polynomial time algorithm that reconstructs the secret automaton of a given card as soon as $n$ is large enough for the expansion properties of $G_1$ and $G_3$ to hold. Moreover, we have implemented this algorithm for a small value $n = 8$ and even in this case it succeeds with a good probability, and with less than 30 queries.

## APPENDIX A. PROOF OF THE MAIN LEMMA

In this section we complete the proof of Theorem 2.1 by proving the main lemma, Lemma 2.5.

Here we review the Broder–Shamir–Friedman (of [9] and [12]) approach to understanding $P(w)$. So fix an irreducible word, $w$, of length $2s$ over $\Pi$, and consider the trajectory of $(1,2,\ldots,r)$ under $w$. To estimate $P(w)$ we will consider only the part of $\pi_1,\ldots,\pi_d$ determined by this trajectory.

Specifically, let $w = w_1 \cdots w_{2s}$ with $w_i \in \Pi$, and consider the random variables,

$$\begin{array}{llll} t_1 = w_1(1), & t_2 = w_1(2), & \cdots & t_r = w_1(r), \\ t_{r+1} = w_1 w_2(1), & t_{r+2} = w_1 w_2(2), & \cdots & t_{2r} = w_1 w_2(r), \\ \cdots & \cdots & \cdots & \cdots \\ t_{(2r-1)s+1} = w_1 w_2 \cdots w_{2s}(1), & & \cdots & t_{2rs} = w_1 w_2 \cdots w_{2s}(r), \end{array}$$

in this order. For example, $t_1$ takes on each of values $\{1,\ldots,n\}$ with probability $1/n$; $t_2$ takes on each of the values $\{1,\ldots,n\} - \{t_1\}$ with probability $1/(n-1)$. However, as another example consider $t_{r+1}$ with $t_1,\ldots,t_r$ being determined; then there are a number of possibilities: if $w_2 = w_1$, then $t_{r+1}$ may already be determined; if $w_2 \neq w_1$ then since $w_2 \neq w_1^{-1}$ we have that $t_{r+1}$ takes on each of values $\{1,\ldots,n\}$ with probability $1/n$. When $t_j$s value is determined by the previous values of $t_1,\ldots,t_{j-1}$, we say that $t_j$ is a *forced choice*; otherwise we say that $t_j$ is a *free choice*. If $t_j$ is a free choice, then clearly $t_j$ takes on one of $n-m$ values each with probability $1/(n-m)$ for some $m \leq 2rs-1$. If a free choice happens to be a previously occurring vertex, i.e., $1,\ldots,r$ or $t_1,\ldots,t_{j-1}$, we say that $t_j$ is a *coincidence*; this will occur with probability $\leq (r+j-1)/(n-j+1)$. Notice that at least $r$ coincides must occur if we have that $(1,2,\ldots,r)$ returns to $(1,2,\ldots,r)$ under $w$, i.e., that $t_{(2s-1)r+j} = j$ for $j = 1,\ldots,r$.

The first estimate we need is that the probability that $\geq r+1$ coincidences occurring is small. Namely, the probability is less than

$$\binom{2rs}{r+1}\left[\frac{r+2rs-1}{n-2rs+1}\right]^{r+1} = O\frac{s^{2r+2}}{n^{r+1}}.$$

The second estimate we need involves the more complicated analysis of how $r$ coincidences may occur; the analysis is greatly simplified by requiring $w$ to be a strongly irreducible word not of the form $u^m$ for any $m \geq 2$. So consider the graph, $G$, determined by the choice of the $t_j$s, i.e., consisting of all vertices $1,\ldots,r$ and $t_1,\ldots,t_{2rs}$, and all edges $(j,t_j)$ for $1 \leq j \leq r$ and edges $(t_j, t_{j+r})$ for which $t_{j+r}$ is a free choice, for $1 \leq j \leq (2s-1)r$. This graph may have self-loops and multiple edges. Our main claim is:

**Lemma A.1.** *Let $r$ coincidences occur and assume $w$ is not of the form $u^m$ for any $m \geq 2$. If $(1,\ldots,r)$ returns to $(1,\ldots,r)$ under $w$, then the graph determined by the $t_j$s must consist of $r$ distinct loops of length $2s$ originating and terminating at each of $1,\ldots,r$. In particular, in this case we have that $t_1,\ldots,t_{2rs-r}$ are free choices without coincidence, and the last $r$ $t_j$s are free choices with coincidence.*

*Proof.* Consider the degrees of the vertices of the graph, $G = (V, E)$, determined by the $t_j$s. It is easy to see that the number of coincidences is precisely,

$$|E| - |V| + r = r + \tfrac{1}{2} \sum_{v \in V} (\deg(v) - 2).$$

Furthermore, given that $(1, \ldots, r)$ returns to $(1, \ldots, r)$ under $w$, and that the first and last letter of $w$ are not inverses of each other, then the degree of each of $1, \ldots, r$ is at least two. Furthermore, since $w$ is reduced it follows that the degree of each other vertex is at least two (for the time it is first entered and first left). Finally, each coincidence raises the degree of a vertex by one (not counting the $r$ coincidences corresponding to the last letter of $w$ at the vertices $1, \ldots, r$. It follows from the above formula that there are no other coincidences, and that the degree of each vertex in $G$ is 2. So if $G$ does not consist of $r$ loops, one for each vertex $1, \ldots, r$, then there must exist a path from one of these vertices to another, say from 1 to 2. If this path is of length $q$, then since all paths are nonbacktracking walks (since $w$ is irreducible) in an everywhere degree two graph, $w$ must equal $w'$ where $w'$ is cyclic shift of $w$ by $q$. However, then $w = u^m$ with $m = |w|/\mathrm{GCD}(|w|, q)$, implying that $m \geq 2$ since $q < |w|$, and therefore contradicting the assumption of the theorem. ∎

At this point we can bound $P(w)$ for $w$ not periodic by the probability of the last $r$ free choices being coincidences. After $t_{2rs-r}$ have been determined, at most $2rs - r$ values of each permutation $\pi_i$ have been determined. So the probability that the last $r$ $t_j$s are precisely $1, \ldots, r$ (in that order) is no more than

$$\left(\frac{1}{n - 2rs + r}\right)\left(\frac{1}{n - 2rs + r - 1}\right) \cdots \left(\frac{1}{n - 2rs + 1}\right) \leq \frac{1}{(n - 2rs)^r}.$$

Adding this to the probability of having $\geq r + 1$ coincidences yields the main lemma.

## APPENDIX B. PROOF OF LEMMA 2.6

First of all let us bound the probability that a random word in $\Pi^{2k}$ reduces to the empty word. Let us note $q_{k,s}$ the probability that a random word in $\Pi^k$ reduces to an irreducible word of length $s$.

**Sublemma B.1.** *We have*

$$q_{2k,0} \leq \left(\frac{2d - 1}{d^2}\right)^k.$$

This lemma follows from the inequality given in Section 2 of [10]. The same result can be obtained (asymptotically) by using Lemma 3.1 of [18].

Let us now estimate the probability that a random word in $\Pi^{2k}$ strongly reduces to a periodic word. However, as in [9] we will first consider the case of an irreducible word. More precisely:

**Sublemma B.2.** *The probability that an irreducible word of length $2s$ chosen uniformly at random strongly reduces to a periodic word is less than*

$$\frac{4s}{3(2d-1)^s}.$$

*Proof.* We proceed as in [9]. The number of irreducible words which have the form $w_a w_b^i w_a^{-1}$ (where the word $w_a$ is of fixed length $t$, and $w_b$ of fixed length $l$) is less than: $2d(2d-1)^{t+l-1}$ (this is because the word $w_a w_b$ must be irreducible). Therefore the probability we want to estimate is less than (for $d > 1$),

$$\frac{1}{2d(2d-1)^{2s-1}} \sum_{0 \le t \le s-1} \sum_{1 \le l \le s-t} 2d(2d-1)^{t-l+1} \tag{B1}$$

$$\le \frac{1}{2d(2d-1)^{2s-1}} \sum_{0 \le t \le s-1} (2d)^2 (2d-1)^{s-2} \tag{B2}$$

$$= \frac{1}{2d(2d-1)^{2s-1}} s(2d)^2 (2d-1)^{s-2} \tag{B3}$$

$$\le \frac{4s}{3(2d-1)^s}. \quad \blacksquare \tag{B4}$$

Let us consider the set of words $w$ of length in $\Pi^{2k}$ which reduce to a word $w'$ of fixed length $2s$. If we choose such a word uniformly at random, the word $w'$ is uniformly distributed over the irreducible words of length $2s$. From this remark and Sublemmas B.1 and B.2 we deduce that

$$p_{2k} \le \left(\frac{2d-1}{d^2}\right)^k + \sum_{s=1}^{k} q_{2k,2s} \frac{4s}{3(2d-1)^s}. \tag{B5}$$

This leads us to consider the following generation function: $f_l(z) = \sum_{s=0}^{l} q_{l,s} z^s$.

**Sublemma B.3.** *For $0 < z < 1$,*

$$f_l(z) \le z\left(\frac{(2d-1)z^2+1}{2dz}\right)^{l-1}.$$

*Proof.* It is readily seen that

$$q_{2l,2s} = \frac{1}{2d} q_{2l-1,2s+1} + \frac{2d-1}{2d} q_{2l-1,2s-1}, \tag{B6}$$

$$q_{2l+1,2s+1} = \frac{1}{2d} q_{2l,2s+2} + \frac{2d-1}{2d} 1_{2l,2s} \quad \text{for } s \geq 1, \tag{B7}$$

$$q_{2l+1,1} = q_{2l,0} + \frac{1}{2d} q_{2l,2}, \tag{B8}$$

$$q_{2l+1,2s} = q_{2l,2s+1} = 0. \tag{B9}$$

From (B9) it follows $f_{2l}(z) = \sum_{s=0}^{l} q_{2l,2s} z^{2s}$ and $f_{2l+1}(z) = \sum_{s=0}^{l} q_{2l+1,2s+1} z^{2s+1}$. From (B6) we obtain

$$f_{2l}(z) = \frac{(2d-1)z^2 + 1}{2\,dz} f_{2l-1}(z). \tag{B10}$$

From (B7) and (B8) we obtain

$$f_{2l+1}(z) = \frac{(2d-1)z^2 + 1}{2\,dz} f_{2l}(z) - \frac{1-z^2}{2\,dz} q_{2l,0}. \tag{B11}$$

From $f_1(z) = z$, relations (B10) and (B11) we deduce our sublemma. ■

We are now ready to prove Lemma 2.6. By using inequality (B5) and Sublemma B.3, we obtain for $d > 1$,

$$\begin{aligned} p_{2k} &\leq \left(\frac{2d-1}{d^2}\right)^k + \sum_{s=1}^{k} q_{2k,2s} \frac{4s}{3(2d-1)^s} \\ &\leq \left(\frac{2d-1}{d^2}\right)^k + \frac{4\,dk}{3(2d-1)} \left(\frac{2d-1}{d^2}\right)^k \\ &\leq (k+1)\left(\frac{2d-1}{d^2}\right)^k. \end{aligned}$$

## REFERENCES


[1] D. Angluin, On the complexity of minimum inference of regular sets, *Inform. and Control*, **39**, 302–320 (1978).

[2] D. Angluin and C. H. Smith, Inductive inference, theory and methods, *Comput. Surveys*, **15**(3), 237–269 (1983).

[3] N. Alon and V. D. Milman, $\lambda_1$, isoperimetric inequalities for graphs and superconcentrators, *J. Combin. Theory Ser. B*, **38**, 73–88 (1985).

[4] L. Babai, Transparent proofs and limits to approximation, preprint, 1994.

[5] L. Babai, G. Hetyei, W. M. Kantor, A. Lubotzky, and A. Seres, On the diameter of finite groups, *Thirty-First Annual Symposium on Foundations of Computer Science*, pp. 857–865, 1990.

[6] B. Bollobas, *Random Graphs*, Academic Press, London, 1985.

[7] B. Bollobas, The isoperimetric number of random regular graphs, *European J. Combin.*, **9**, 241–244 (1988).

[8] B. Bollobas and W. F. de la Vega, The diameter of random-regular graphs, *Combinatorica*, **2**, 125–134 (1982).

[9] A. Broder and E. Shamir, On the second eigenvalue of random regular graphs, *Twenty-Eighth Annual Symposium on Foundations of Computer Science*, pp. 286–294, 1987.

[10] C. Delorme, Counting closed paths in trees, Technical Report No. 516, Laboratoire de recherche en informatique Orsay, University of Paris-Sud, Paris, Sept. 1989 (in French).

[11] Y. Freund, M. Kearns, D. Ron, R. Rubinfeld, R. E. Schapire, and L. Sellie, Efficient learning of typical finite automata from random walks, *Twenty-Fifth Symposium on the Theory of Computing*, 1993, pp. 315–324.

[12] J. Friedman, On the second eigenvalue and random walks in random $d$-regular graphs, *Combinatorica*, **11**(4), 331–362 (1991).

[13] J. Friedman, Some geometric aspects of graphs and their eigenfunctions, *Duke Math. J.*, **69**, 487–525 (1993).

[14] E. M. Gold, Complexity of automaton identification from given data, *Inform. and Control*, **37**, 302–320 (1978).

[15] A. Joux, J. Stern, and J. P. Tillich, Inferring finite automata by queries of fixed length, Technical report of the Ecole Normale Supérieure.

[16] A. Lubotzky, *Discrete Groups*, *Expanding Graphs and Invariant Measures*, Progress in Mathematics, Birkhäuser, Basel, 1994, Vol. 125.

[17] A. Lubotzky, Cayley graphs: eigenvalues, expanders and random walks, *London Math. Soc. Lecture Notes Ser.*, **218** 155–189 (1995).

[18] B. McKay, The expected eigenvalue distribution of a large regular graph, *Linear Algebra Appl.*, **40**, 203–216 (1981).

[19] M. Mihail, Conductance and convergence of Markov chains—a combinatorial treatment of expanders, *Proceedings of the Thirtieth Annual Symposium on Foundations of Computer Science*, 1989.

[20] B. Mohar, Isoperimetric number of graphs, *J. Combin. Theory Ser. B*, 274–291 (1989).

[21] R. L. Rivest and R. E. Schapire, Diversity based inference of finite automata, *Proceedings of the Twenty-Eighth Annual Symposium on the Foundations of Computer Science*, 1987, pp. 78–87.

[22] R. L. Rivest and R. E. Schapire, Inference of finite automata using homing sequences, *Proceedings of the Twenty-First Symposium on the Theory of Computing*, 1989, pp. 411–420.

[23] J. P. Tillich and G. Zémor, Group-theoretic hash functions, in *Proceedings of the First French–Israeli Workship in Algebraic Coding 1993*, Lecture Notes in Computer Science, Springer-Verlag, Berlin/New York, Vol. 781, pp. 90–110.

[24] U. Vazirani, Rapidly mixing markov chains, *Proc. Sympos. Appl. Math.*, **44**, 99–121 (1991).

[25] G. Zémor, Hash functions and Cayley graphs, *Design*, *Codes and Cryptography*, to appear.