On sampling generating sets of finite groups and product replacement algorithm. (Extended Abstract)

IGOR PAK*, SERGEY BRATUS†

1 Introduction

Let G be a finite group. A sequence of k group elements $(g_1, \ldots, g_k)$ is called a *generating k-tuple* of G if the elements generate G (we write $\langle g_1, \ldots, g_k \rangle = G$). Let $\mathcal{N}_k(G)$ be the set of all generating k-tuples of G, and let $N_k(G) = |\mathcal{N}_k(G)|$.

We consider two related problems on generating k-tuples. Given G and $k > 0$,

1) *Determine* $N_k(G)$

2) *Generate random element of* $\mathcal{N}_k(G)$*, each with probability* $1/N_k(G)$

The problem of determining the structure of $\mathcal{N}_k(G)$ is of interest in several contexts. The counting problem goes back to Philip Hall, who expressed $N_k(G)$ as a Möbius type summation of $N_k(H)$ over all maximal subgroups $H \subset G$ (see [23]). Recently the counting problem has been studied for large simple groups where remarkable progress has been made (see [25, 27]). In this paper we analyze N_k for solvable groups and products of simple groups.

The sampling problem, while often used in theory as a tool for approximate counting, recently began a life of its own. In this paper we will present an algorithm for *exact sampling* in case when G is nilpotent.

When little about structure of G is known, one can only hope for approximate sampling. In [11] Celler et al. proposed a *product replacement* Markov chain on $\mathcal{N}_k(G)$ which is conjectured to be rapidly mixing to a uniform stationary distribution. The subject was further investigated in [6, 12, 17, 16], while the conjecture is fully established only when $G \simeq \mathbb{Z}_p$, p is a prime. We prove rapid mixing for all abelian groups G. Also, we disprove the folklore conjecture that the group elements in generating k-tuples are (nearly) uniformly distributed.

Finally, we would like to remark that the generating k-tuples occur in connection with the so-called random random walks, which are the ordinary random walks on G with random support. The analysis of these "average case" random walks was inspired by Aldous and Diaconis in [1] and was continued in a number of papers (see e.g. [19, 33, 36, 39]). We explain how the sampling problem can be used to test convergence of random random walks.

*Department of Mathematics, Yale University, New Haven, CT 06520, E-mail: **pak@math.yale.edu**

†Department of Mathematics, Northeastern University, Boston, MA 02115, E-mail: **sbratus@ccs.neu.edu**

2 Counting problem

Let G be a finite group. By $|G|$ denote the order of G. As in the introduction, let $N_k(G) = |\mathcal{N}_k(G)|$ be the number of generating k-tuples $\langle g_1, \ldots, g_k \rangle = G$. It is often convenient to consider the probability $\varphi_k(G)$ that k uniform independent group elements generate G :

$$\varphi_k(G) = \frac{N_k(G)}{|G|^k}$$

Theorem 2.1 *For any finite group G, $1 > \epsilon > 0$, we have*

$$\varphi_k(G) > 1 - \epsilon$$

given $k > \log_2 |G| + 1 + \log_2 1/\epsilon$.

This is a slight improvement over a more general classical result by Erdős and Rényi in [20].

Define $\kappa(G)$ to be the minimal possible number of generators of G. In other words,

$$\kappa(G) = \min\{k \,|\, N_k(G) > 0\}$$

The problem of evaluating $\kappa(G)$ has been of intense interest for classes of groups as well as for individual groups (see [14]).

It is known that $\kappa(G) = 2$ for all simple, nonabelian groups, and that $\kappa(G) \leq n/2$ for $G \subset S_n$, with equality achieved when $G \simeq \mathbb{Z}_2^{n/2}$, and n is even. Also, it is easy to see that $\kappa \leq \log_2 |G|$, with equality for $G \simeq \mathbb{Z}_2^n$.

Define $\vartheta(G)$ to be the smallest k such that at least $1/4$ of the random k-tuples $(g_1, \ldots, g_k)$ generate the whole group. In other words,

$$\vartheta(G) = \min\{k \,|\, \varphi_k(G) > \frac{1}{4}\}$$

Note that Theorem 2.1 immediately implies that

$$\vartheta(G) \leq \log_2 |G| + 1$$

By definition $\vartheta(G)/\kappa(G) \geq 1$. It is unclear, however, how big this ratio can be.

Here are a few known results. When G is simple, it is known that $\varphi_2(G) \to 1$ as $|G| \to \infty$. For $G = A_n$, this is a famous result of Dixon (see [18]), while for Chevalley groups the result was conjectured by Kantor, Lubotzky (see [25])

and recently proved by Liebeck and Shalev (see [27]). This immediately implies that $\vartheta(G) < C$ for any simple group G and some universal constant C. It was also noted in [17] that if G is a p-group, then $\vartheta(G) \leq \kappa(G) + 1$. The following result is a significant generalization.

Theorem 2.2
1) *If G is solvable, then $\vartheta(G) \leq \kappa(G) + 1$.*

2) *If G is a direct product of simple groups, then $\vartheta \leq \kappa(G) + C \log\log |G|$ for some universal constant C.*

3 Sampling problem

There are several ways a finite group G can be presented as an input to the algorithm. Regardless of the presentation of G, denote by μ the time necessary for group operations (multiplication, taking an inverse, comparison with id[1]). Denote by ρ the complexity of generating a (nearly) uniform group element (call it *random generation*). It is also convenient to denote by η the time to check whether given k group elements generate a group. We call this the *generation test*.

We start with *permutation groups* which are defined as subgroups of a permutation group S_n. The group is presented by a set of generators. This is the best understood class of groups with efficient management, random elements generation, generation test, etc., based on the fundamental algorithms by C. Sims (see e.g. [38, 13, 28]. In particular one has $\rho = O(\mu n)$, and $\eta = O(\mu n^4)$ by reducing the problem to group membership.

A *matrix group* is a group defined as a subgroup of $GL(n; q)$. This is a harder class of groups to work with (see [24, 6]). Recently there has been some advance work done in this setting (see [7, 10, 31, 29]). Still, polynomial time management for matrix groups is yet to be discovered.

One of the most general and widely accepted is the *black box* setting (see e.g. [6]) in which group elements are encoded by binary strings of fixed length n (possibly in many ways). A black box can multiply elements, take inverses, and compare elements with identity. This presentation of a group generalizes both permutation and matrix groups. In a pioneering work [4], Babai was able to find a polynomial time algorithm for generating (nearly) uniform group elements. The product replacement algorithm was designed to give a *practical* algorithm for random generation. These algorithms were used in a number of subsequent works, particularly on recognition of various classes of finite groups (see [8, 9, 26, 31]). Following Babai (see [4]), there is no subexponential in n algorithm which can perform the generation test.

Now consider sampling problems (see introduction) from the computational point of view. We immediately obtain the following result.

Theorem 3.1 *Let G be a black box group with a generation test oracle, and a random generation oracle. Let $k \geq \vartheta(G)$. Then there exists a randomized algorithm for sampling from $\mathcal{N}_k(G)$ in time $O(\rho k + \eta)$.*

Indeed, given $k \geq \vartheta(G)$, we can always sample from $\mathcal{N}_k(G)$ by simply generating a uniform k-tuple and testing whether it generates the whole group G. At the moment, the problem is open for $\kappa(G) \leq k < \vartheta(G)$. We do not believe that there is an efficient sampling algorithm for all k and for general black box groups. However, such algorithms do exist in cases when the group is already *recognized*, i.e.. when a black box group is provided with an isomorphism $\pi : G \to G^{'}$ to a group in a canonical form (see below).

Theorem 3.2 *Let G be a finite nilpotent group defined as in the preceding paragraph. Then there exists a randomized algorithm for sampling from $\mathcal{N}_k(G)$ with running time $k\rho(1+o(1))$, which requires $k \log_2 |G| \, (1+o(1))$ random bits.*

By random bits we mean, roughly speaking, the number of coin flips required in the algorithm. Clearly, this number cannot be smaller that $\log_2 N_k(G)$. To demonstrate the strength of the algorithm, consider the case $G = \mathbb{Z}_2^n$. Then $\kappa = n$ and $\mathcal{N}_n(G)$ is in one-to-one correspondence with the set of nonsingular matrices $GL(n; \mathbb{Z}_2)$. It is known that $\varphi_n(G) = c > 1/4$ (see e.g. [30, 32]). The standard approach to sampling from $GL(n; \mathbb{Z}_2)$ would be sampling *any* matrix and then checking by Gaussian elimination whether it is nonsingular. The expected number of random bits required for that is $\frac{1}{c}\lceil \log_2(n^2) \rceil$, while our algorithm requires only $\log_2 n^2(1 + o(1))$ random bits. The problem of saving random bits when sampling from $GL(n; \mathbb{F}_q)$ was considered earlier by Randall (see [35]) and the first author (see [32]). Thus Theorem 4 can be thought of as an advance generalization of these results.

We conclude with the remark that for large enough k sampling from $\mathcal{N}_k(G)$ can be done in a generality of black box groups by using the product replacement algorithm (see below).

4 Product replacement algorithm

The *product replacement algorithm* is an important recent advancement in symbolic algebra. In [11] Celler et al. defined a Markov chain X_t on $\mathcal{N}_k(G)$ as follows. Let $X_t = (g_1, \ldots, g_k) \in \mathcal{N}_k(G)$. Define $X_{t+1} = (g_1, \ldots, h_j, \ldots, g_k)$, where $h_j = g_j g_i^{\pm 1}$ or $h_j = g_i^{\pm 1} g_j$, where the pair (i, j), $1 \leq i, j \leq k$, $i \neq j$ is chosen uniformly; the multiplication order and the ± 1 degree are determined by independent flips of a fair coin. By $\widetilde{\kappa}(G)$ denote the maximum size of the minimum generating set (i.e. of the set such that no generator can be omitted). The authors showed (cf. [16]) that when $k \geq \kappa + \widetilde{\kappa}$ this Markov chain is reversible, aperiodic and irreducible, and has a uniform stationary distribution. Thus the chain is ergodic and can be used for approximate sampling from $\mathcal{N}_k(G)$, $k > \kappa(G) + \widetilde{\kappa}$. The empirical tests seem to indicate that the chain mixes rapidly (see [11]).

At the moment it is not known whether the Markov chain converges in time polynomial of k, $\log |G|$, but this is conjectured to be true for suitable values of k. Several weaker versions of this claim has appeared. The only case when the conjecture has been established is due to Diaconis and Saloff–Coste, who proved the claim for $\mathbb{Z}_p$ when p is a prime (see [17]). In [12] Chung and Graham showed that the mixing time is polynomial in k and $|G|$. Babai in [6] showed that the diameter of the underlying graph is $O(\log^2 |G|)$ given $k > 2 \log_2 |G|$. At the moment the best general results are due to Diaconis and Saloff–Coste (see [16]). The authors show that the mixing time is bounded by a polynomial of

[1] For some presentations, such as the presentation by generators and relations, the last task can be non-trivial. The black box model discussed below makes the assumption that the identity test, i.e. comparison with id, can be performed efficiently.

several parameters, including $\Delta_k(G)$, $\binom{k}{\kappa(G)}$, and $1/\varphi_k(G)$, where $\Delta_k(G)$ is the *maximal diameter* of generating k-tuples of G. Unfortunately, some of these parameters can be very large (or are not known to be relatively small), and it is easy to see that this upper bound in [16] is always greater than an upper bound for the natural random walk with *any* k generators.

We prove the rapid mixing for abelian groups. We need several definitions. Fix a starting generating k-tuple $(g_1, \ldots, g_k)$. Define by Q^t the distribution of the product replacement Markov chain after t steps. By U denote the uniform distribution on $\mathcal{N}_k(G)$. Define the *total variation distance* $d(t) = \|Q^t - U\|_{tv}$ of the product replacement Markov chain as follows:

$$d(t) = \max_{A \in \mathcal{N}_k(G)} |Q^t(A) - U(A)|$$

$$= \frac{1}{2} \sum_{(g) \in \mathcal{N}_k(G)} \left| Q^t((g)) - \frac{1}{N_k(G)} \right|$$

Theorem 4.1 *Let G be abelian group, $k > \log|G| + 2$. Then*

$$d(t) < e^{-x} \quad \textit{for} \quad t > C_1 k^2 \log^2|G| (k^2 \log|G| + C_2 x)$$

where C_1, C_2 are universal constants independent of G, k.

While Theorem 4.1 proves the conjecture that the product replacement Markov chain mixes rapidly, in our opinion it is somewhat premature to believe that the conjecture holds in the general case.

Let us point out that if one knows how to sample generating k-tuples, one can also test how close the random replacement Markov chain is to a stationary distribution. Indeed, one can simply compare any given statistics on $\mathcal{N}_k(G)$ on samples obtained by the *exact* sampling and on samples obtained by the random replacement algorithm. The authors in [11] use a chi-square statistics, while this checking method allows more freedom.

To conclude, let us return to the original motivation of [11]. The authors invented the product replacement algorithm with the sole purpose of performing random generation. The authors proposed to generate a (nearly) uniform generating k-tuple and then output a uniformly chosen component. While the authors acknowledge that one can be sure that this would work only when $\varphi_k(G)$ is close to 1, it is widely believed that this would work in practice for all $k \geq \kappa(G) + 1$. We confront this belief by giving an example when no condition $k \geq \kappa(G) + C$, where C is a universal constant, would work.

Observe that there can be two types of error when we generate a (nearly) uniform group element as above. The first type comes from the distribution Q^t being far from the uniform distribution U on $\mathcal{N}_k(G)$. The second one comes from having group elements in generating k-tuples distributed not uniformly. While before we dealt with the first type of error, we will show that the second error can be large in some examples. Note that by symmetry all the elements in generating k-tuples have the same distribution, so it suffices to consider the first element only.

Let $G \simeq A_n^{r(n)}$, where A_n is an alternating group, $n \geq 5$ and $r(n)$ is the maximal degree such that G is generated by two elements. It was shown by Hall (see [23]) that $r(5) = 19$, and by Kantor and Lubotzky (see [25]), using the result of Dixon (see [18]) that $n!/8 \leq r(n) \leq n!/4$. It was observed in [25] that for $k = O(\sqrt{n})$ we have $\varphi_k(G) \to 0$ as $n \to \infty$. We prove a stronger claim.

Theorem 4.2 *Let $G \simeq A_n^{r(n)}$, $k = o(n)$. Denote by P_k the probability distribution on G of the first element in generating k-tuples $(g) \in \mathcal{N}_k(G)$. Then*

$$\|P_k - U\|_{tv} \to 0, \quad \textit{as } n \to \infty$$

We should add that the latter type of error can be avoided by either one of the following tricks. We can add a fixed generating set to our k-tuple and allow the other elements to be multiplied by them. This gives us a uniform limiting distribution on G^k. Similarly, we can add an extra group element we call *sink*, which we allow to be multiplied by the remaining elements, but never use it to multiply the others. In the end, we output this sink element. This procedure was communicated to us by Charles Leedham-Green.

It is interesting to compare these procedures. While the former procedure seems to work slower than the latter, it has an advantage that it outputs more of (nearly) uniform and independent group elements, which is useful in a number of applications (see the previous section).

Finally, consider an oriented graph on $\mathcal{N}_k(G)$ with edges corresponding to product replacement moves. We remark that in general case, when $\kappa(G) \leq k < \kappa(G) + \widetilde{\kappa}$ it is not clear even whether this graph is strongly connected. The question has been completely resolved for abelian groups by Diaconis and Graham (see [15]). It is conjectured that the graph is already strongly connected when $k = 3$ and $G = S_n$, and this has been checked for $n \leq 5$. We hope to return to this problem in subsequent publications.

5 Random random walks

Let G be a finite group, and let $(g_1, \ldots, g_k) \in \mathcal{N}_k(G)$ be a generating k-tuple. A *random walk* X_t on G is defined by $X_0 = id$, $X_{t+1} = X_t \cdot g_i$, where i is chosen uniformly in $[1, \ldots, k]$. One can think of the walk X_t as of a nearest neighbor random walk on a Cayley graph.

It is known that under minor conditions the random walk converges to a uniform distribution on G (see e.g. [2]). An important problem is to estimate how long will it take to converge to stationary. Formally, let $Q^t(g) = \mathbf{P}(X_t = g)$ be a probability distribution of the walk after t steps. Recall the total variation distance $d(t) = \|Q^t - U\|_{tv}$. Usually estimating $d(t)$ is a hard problem, from both theoretical and computational points of view. Good estimates in cases of importance normally require very special knowledge of the behavior of a random walk. In [1] Aldous and Diaconis proposed to study the "average case" random walks, and conjectured that they must have fast convergence. Such random walks with random support are often called *random random walks*.

A breakthrough was made by Dou and Hildebrand, who confirmed the conjecture for superlogarithmic values of k. Roughly, they showed that after $t > C \frac{a}{a-1} \log_k |G|$ steps we have $E(d(t)) \to 0$ as $|G| \to \infty$, given $k > \log^a |G|$. Different proofs and better bounds in special cases, such as abelian groups, were obtained by subsequent investigators (see [3,

33, 34, 36, 39]). For fairly small k, such as $k = o(\log_2 |G|)$, the problem is largely unresolved. Say, for $G = S_n$ it is believed that $t = \Omega(n^3 \log n)$ implies $d(t) \to 0$ as $n \to \infty$ for *any* generating k-tuple, $k = Const \geq 2$ (see above). However, no polynomial bound is known even for random random walks, the best one being that of Babai and Hetyei (see [5]).

Now, given this poor state of the art for $k = o(\log_2 |G|)$ one may wish to collect experimental evidence for behavior of random random walks. That's where one can apply the sampling procedures. Note also that in general, if we can compute $d(t)$ for random walks generated by random k-tuples, there is no need to check whether this is a generating k-tuple. Indeed, if a k-tuple does not generate G, the corresponding Cayley graph is disconnected and $s(t) = 1$ for all $t > 0$. Thus if $k = \Omega(\vartheta(G) \log(1/\epsilon))$, then $Q_k(G) > 1 - \epsilon$ and if $\epsilon \to 0$ we have the expectation over all k-tuples $E(d(t)) \to 0$ if and only if so does the expectation taken over all generating k-tuples.

6 Proofs of Theorems

6.1 Proof of Theorem 2.1 (sketch)

Fix a finite group G. Consider the following random process. Pick a uniform group element $g_1 \in G$. If $H_1 = \langle g_1 \rangle \neq G$, pick a uniform group element $g_2 \in G$. If $H_2 = \langle g_1, g_2 \rangle \neq G$, pick a another groups element, etc. Denote by τ the first time we generate the whole group. We claim that for all k and all G, $|G| \leq 2^r$, the probability $\mathbf{P}(\tau = k)$ is minimal when $G \simeq \mathbb{Z}_2^r$. Indeed, regardless of the group structure, the probability that $H_i \neq H_{i+1}$ for a given i is $1 - |H_i|/|G|$. Notice that then $|H_{i+1}|/|H_i| \geq 2$ with the equality always achieved when $G \simeq \mathbb{Z}_2^r$. Therefore $\mathbf{P}(\tau = k)$ is minimized in this case.

Now observe that $\varphi_k(G) = \mathbf{P}(\tau \leq k)$. Thus $\varphi_k(G)$ is minimized when $G \simeq \mathbb{Z}_2^r$, and it remains to compute $\varphi_k(\mathbb{Z}_2^r)$. Think of the k-tuple of elements of $\mathbb{Z}_2^r$, $k \geq r$, as of a $k \times r$-matrix with elements in $\{0, 1\}$. Such a matrix corresponds to a generating k-tuple if it has rank r. We obtain (cf. [32]):

$$\varphi_k(\mathbb{Z}_2^r) = \frac{2^k - 1}{2^k} \frac{2^k - 2}{2^k} \cdots \frac{2^k - 2^{r-1}}{2^k}$$

$$\geq \frac{1}{2} \left(1 - \frac{1}{2^{k-r+1}}\right)$$

This implies the result.

6.2 Proof of Theorem 2.2 (sketch)

Let G be a solvable group. Following [21], denote by $F_1, \ldots, F_h$ the different (with respect to G-isomorphisms) types of simple G-groups that occur among the factors in a chief series of G. Let α_i be the number of factors of type F_i that have a complement, and β_i be the number of those of type F_i that do not possess a complement. Let E_i be the field of endomorphisms of F_i. Further, let

$$\zeta_i = \begin{cases} 0 & \text{if } F_i \text{ is fixed element-wise by } G \\ 1 & \text{otherwise} \end{cases}$$

Assume $|F_i| = p_i^{\lambda_i}$, p_i prime, and ω_i is the degree of the endomorphism field E_i over its base field of characteristic p_i. Then Theorem 5 in [21] implies

$$N_k(G) = \prod_{i=1}^{h} p_i^{\lambda_i \beta_i k} \cdot \prod_{i=1}^{h} \left(p_i^{\lambda_i k} - p_i^{\lambda_i \zeta_i}\right) \cdot$$

$$\cdot \left(p_i^{\lambda_i k} - p_i^{\lambda_i \zeta_i + \omega_i}\right) \cdot \cdots \cdot \left(p_i^{\lambda_i k} - p_i^{\lambda_i \zeta_i + (\alpha_i - 1)\omega_i}\right)$$

We obtain

$$\kappa(G) = \max_i \left\lceil \frac{(\alpha_i - 1)\omega_i}{\lambda_i} + \zeta_i \right\rceil$$

When $k \geq \kappa(G)$ we have

$$\varphi_k(G) = \prod_{i=1}^{h} \prod_{j=0}^{\alpha_i - 1} \left(1 - p_i^{\lambda_i(\zeta_i - k) + j\omega_i}\right)$$

When $k = \kappa(G) + 1$ this product is bounded from below by the product

$$\varphi_{\kappa+1} \geq \prod_p \prod_{i=2}^{\infty} \left(1 - \frac{1}{p^i}\right)$$

where the first product is taken over all primes p. By Euler's pentagonal theorem and using Lemma 2.3.2 in [32] we obtain

$$\varphi_{\kappa+1}(G) > \prod_p \frac{1 - 1/p - 1/p^2}{1 - 1/p} > \prod_p 1 - \frac{2}{p^2} > \frac{1}{4}$$

which proves the first part of the theorem.

The second part follows from a similar analysis based on the general results in [25]. We return to this problem in the proof of Theorem 3.2.

6.3 Proof of Theorem 3.1 (sketch)

First, generate k independent elements of G. Run a generation test $O(\log 1/\epsilon)$ to check with probability of error ϵ whether these elements indeed generate G. If not, start over. Now, since $k \geq \kappa(G)$ the latter will happen with probability at most $3/4 + \epsilon$. Take $\epsilon = 1/8$. The expected number of trial becomes $1/1 - (3/4 + \epsilon) = 8$. This concludes the proof.

6.4 Proof of Theorem 3.2 (sketch)

Briefly, the case when $G \simeq \mathbb{Z}_p^r$, p prime, is described in [32]. The case when $G \simeq \mathbb{Z}_{p^m}^r$ is no different since to generate $\mathbb{Z}_{p^m}$ it is sufficient and necessary to generate an element $a \in \mathbb{Z}_{p^m}$ which is not a zero divisor. This is equivalent to generating $\mathbb{Z}_p$.

For a general abelian group G, decompose it as a product of p-groups H_p over all primes p. As we observed in the previous paragraph, in each case the problem of generating random generating k-tuple in H_p is equivalent to the problem of generating random generating k-tuple in $\mathbb{Z}_p^{r_p}$ for some integer r_p. It can be shown that given such a set of generating k-tuples for all p one can combine them into a single generating k-tuple of G.

Finally, if G is nilpotent it is a known result (see e.g. [22]) that k-tuples that generate $G/[G, G]$ also generate the whole group. Thus the problem is reduced to a corresponding problem for the group $G/[G, G]$, which is abelian. We skip the details.

6.5 Proof of Theorem 4.1 (sketch)

Let G be a finite abelian group, $n = |G|$. Let $(g_1, \ldots, g_k) \in \mathcal{N}_k(G)$ be a generating k-tuple. Consider a random subproduct h of the following form:

$$h = g_1^{a_1} \cdot \cdots \cdot g_k^{a_k}$$

where integers a_i are independent uniform in $[0, n-1]$. We claim that h is uniform in G. The proof is an easy induction on k.

Now, suppose an integer sequence $(b_1, \ldots, b_k)$ satisfies the condition that $(b_1 \bmod p, \ldots, b_k \bmod p)$ is (nearly) uniform in $[0, p-1]^k$ for all primes $p < 2|G|$. Conclude from this that $(b_1 \bmod n, \ldots, b_k \bmod n)$ is also (nearly) uniform in $[0, n-1]^k$.

Let $k \geq \lceil \log |G| \rceil + 1 + \log(1/\epsilon) \geq \kappa(G)$. Since G is abelian, the relative order of generators in each product is irrelevant. Run a product replacement Markov chain for t steps, starting at $(g_1, \ldots, g_k)$. We get each element of the form

$$h_i = g_1^{b_{i,1}} \cdot \cdots \cdot g_k^{b_{i,k}}$$

Consider a $k \times k$ matrix $B = (b_{i,j})$. By definition of the product replacement Markov chain, this matrix is a product of t random elementary transvections $E_{r,l} \in SL(n, \mathbb{Z})$ which are the matrices with ones on the diagonal, one in position (r, l), and zeros elsewhere. Taking a random walk of these matrices $\bmod p$ we obtain that the matrix $B \bmod p$ is (nearly) uniform in $SL(n, p)$ after $t = O(k^4 \log^3 p)$ steps by the result of [17].

Note that for abelian groups we have $\kappa(G) = \widetilde{\kappa}(G)$, i.e. given $k > \kappa(G)$ at least one of the generators can be omitted. Delete the corresponding column in B which gives us B'. By Theorem 2.1 and above arguments it is easy to see that $B' \bmod p$ is (nearly) uniform in $Mat(k, p)$. Thus its rows are (nearly) uniform and independent in $[0, p-1]^{k-1}$. By the remarks in the first paragraph this implies that taken $\bmod n$ we obtain rows that are (nearly) uniform and independent in $[0, n-1]^{k-1}$, which is exactly what we need.

Using the full power of Theorem 4.1 in [17] and after some straightforward technical computations we obtain the result.

6.6 Proof of Theorem 4.2 (sketch)

Consider the structure of $\mathcal{N}_k(G)$, where $G = A_n^N$, $N = r(n)$. Let $(g_1, \ldots, g_k) \in \mathcal{N}_k(G)$. Consider the number x of permutations σ_i in $g_1 = (\sigma_1, \ldots, \sigma_N)$ such that $\sigma_i(1) = 1$. Think of x as of a random variable on $\mathcal{N}_k(G)$. If g_1 were (nearly) uniform in G, x would be distributed as the number of successes in independent Bernoulli trials with probability of success $1/n$. We claim that x has a somewhat shifted distribution, with the probability that a permutation σ_i satisfies $\sigma_i(1) = 1$ being of the order $1/n - C/n^{k-1}$.

The idea is based on the known results about the generating k-tuples of A_n. It can be easily deduced from the more general results in [37] that $\varphi_k(A_n) \sim 1 - C/n^{k-1}$, as $n \to \infty$, and where C is independent of n. Notice that, conditioned on $\sigma_i(1) = 1$, the probability that the i-th permutation in all g_j, $j = 2, \ldots, k$ satisfies $\sigma_i(1) = 1$, has the same order $1 - C/n^{k-2}$ as the probability $\varphi_k(A_n)$. Deduce from this that the probability of $\sigma_i(1) = 1$ is of the order $1/n - C/n^{k-1}$. Note that these events are no longer independent. Observe, however, that results in [25] imply that a generating k-tuple $(g_1, \ldots, g_N)$ must contain every orbit of the action of A_n on $\mathcal{N}_k(A_n)$ exactly once. Thus the correlation becomes exponentially small and for our purposes these can be viewed as independent events.

Now given two series on length N of Bernoulli trials outcomes, with probabilities $p_1 = 1/n$ and $p_2 = 1/n - C/n^{k-1}$ one can distinguish between them given $n^k = o(N)$. Indeed, use Chernoff bound to split the number of heads below and above $M = N(p_1 + p_2)/2$, i.e. to show that with high probability the first trial will have more than M successes, while the second less than M successes. This implies that given $k = o(\log_n N) = o(n)$ the total variation distance $\to 1$ as $n \to \infty$. This concludes the proof.

7 Acknowledgments

We would like to thank L. Babai, R. Beals, G. Cooperman, P. Diaconis, W. Feit, L. Finkelstein, M. Hildebrand, W. Kantor, C. Leedham-Green, L. Lovasz, A. Lubotzky, G. Margulis, G.–C. Rota, L. Saloff–Coste, A. Shalev, R. Stanley, Van Vu and E. Zelmanov for helpful remarks.

The first author was supported by NSF Postdoctoral Mathematical Research Fellowship.

References

[1] Aldous, D., and Diaconis, P. Shuffling cards and stopping times. *Amer. Math. Monthly 93*, 5 (1986), 333–348.

[2] Aldous, D., and Fill, J. Reversible markov chains and random walks on graphs. monograph in preparation, 1996.

[3] Alon, N., and Roichman, Y. Random Cayley graphs and expanders. *Random Structures Algorithms 5*, 2 (1994), 271–284.

[4] Babai, L. Local expansion of vertex–transitive graphs and random generation in finite groups. In *Proc. 23rd ACM STOC* (1991), pp. 164–174.

[5] Babai, L. Automorphism groups, isomorphism, reconstruction. In *Handbook of combinatorics, Vol. 1, 2.* Elsevier, Amsterdam, 1995, pp. 1447–1540.

[6] Babai, L. Randomization in group algorithms: conceptual questions. In *Groups and computation, II (New Brunswick, NJ, 1995).* Amer. Math. Soc., Providence, RI, 1997, pp. 1–17.

[7] Beals, R., and Babai, L. Las Vegas algorithms for matrix groups. In *34th Annual Symposium on Foundations of Computer Science (Palo Alto, CA, 1993).* IEEE Comput. Soc. Press, Los Alamitos, CA, 1993, pp. 427–436.

[8] Bratus, S., Cooperman, G., Finkelstein, L., and Linton, S. Constructive recognition of a black box group isomorphic to GL(n, q). monograph in preparation, 1998.

[9] Bratus, S., and Pak, I. Fast constructive recognition of a black box group isomorphic to S_n or A_n. preprint, 1997.

[10] Celler, F., and Leedham-Green, C. R. A non-constructive recognition algorithm for the special linear and other classical groups. In *Groups and computation, II (New Brunswick, NJ, 1995)*. Amer. Math. Soc., Providence, RI, 1997, pp. 61–67.

[11] Celler, F., Leedham-Green, C. R., Murray, S. H., Niemeyer, A. C., and O'Brien, E. A. Generating random elements of a finite group. *Comm. Algebra 23*, 13 (1995), 4931–4948.

[12] Chung, F. R. K., and Graham, R. L. Random walks on generating sets for finite groups. *Electron. J. Combin. 4*, 2 (1997), Research Paper 7, approx. 14 pp. (electronic). The Wilf Festschrift (Philadelphia, PA, 1996).

[13] Cooperman, G., and Finkelstein, L. Combinatorial tools for computational group theory. In *Groups and computation (New Brunswick, NJ, 1991)*. Amer. Math. Soc., Providence, RI, 1993, pp. 53–86.

[14] Coxeter, H. S. M., and Moser, W. O. J. *Generators and relations for discrete groups*, third ed. Springer-Verlag, New York, 1972. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 14.

[15] Diaconis, P., and Graham, R. The graph of generating sets of an abelian group. to appear, 1997.

[16] Diaconis, P., and Saloff-Coste, L. Walks on generating sets of abelian groups. *Probab. Theory Related Fields 105*, 3 (1996), 393–421.

[17] Diaconis, P., and Saloff-Coste, L. Walks on generating sets of groups. *Invent. Math. 134*, 2 (1998), 251–299.

[18] Dixon, J. D. The probability of generating the symmetric group. *Math. Z. 110* (1969), 199–205.

[19] Dou, C., and Hildebrand, M. Enumeration and random random walks on finite groups. *Ann. Probab. 24*, 2 (1996), 987–1000.

[20] Erdős, P., and Rényi, A. Probabilistic methods in group theory. *J. Analyse Math. 14* (1965), 127–138.

[21] Gaschütz, W. Die Eulersche Funktion endlicher auflösbarer Gruppen. *Illinois J. Math. 3* (1959), 469–476.

[22] Hall, Marshall, J. *The theory of groups.* Chelsea Publishing Co., New York, 1976. Reprinting of the 1968 edition.

[23] Hall, P. The eulerian functions of a group. *Quart. J. Math. 7* (1936), 134–151.

[24] Kantor, W. M. Simple groups in computational group theory. In *Proceedings of the International Congress of Mathematicians, Vol. II (Berlin, 1998)*, vol. 1998, pp. 77–86 (electronic).

[25] Kantor, W. M., and Lubotzky, A. The probability of generating a finite classical group. *Geom. Dedicata 36*, 1 (1990), 67–87.

[26] Kantor, W. M., and Seress, A. Black box classical groups. preprint, 1997.

[27] Liebeck, M. W., and Shalev, A. The probability of generating a finite simple group. *Geom. Dedicata 56*, 1 (1995), 103–113.

[28] Luks, E. M. Computing the composition factors of a permutation group in polynomial time. *Combinatorica 7*, 1 (1987), 87–99.

[29] Luks, E. M. Computing in solvable matrix groups. In *Proc. 33rd IEEE FOCS* (1992), pp. 111–120.

[30] Neumann, P. M., and Praeger, C. E. Cyclic matrices over finite fields. *J. London Math. Soc. (2) 52*, 2 (1995), 263–284.

[31] Niemeyer, A. C., and Praeger, C. E. A recognition algorithm for classical groups over finite fields. *Proc. London Math. Soc. (3) 77*, 1 (1998), 117–169.

[32] Pak, I. When and how n chose k. In *DIMACS Series*, vol. 43. Amer. Math. Soc., Providence, RI, 1998, pp. 191–238.

[33] Pak, I. Random walks on finite groups with few random generators. *Electron. J. Prob. 4* (1999), 1–11.

[34] Pak, I., and Vu, V. On finite geometric random walks. preprint, 1998.

[35] Randall, D. Efficient generation of random nonsingular matrices. *Random Structures Algorithms 4*, 1 (1993), 111–118.

[36] Roichman, Y. On random random walks. *Ann. Probab. 24*, 2 (1996), 1001–1011.

[37] Shalev, A. Probabilistic groups theory. St. Andrews Lectures, Bath, 1997.

[38] Sims, C. Computation with permutation groups. In *Proc. Second Symp. Symb. Alg. Man.* (1971), ACM, pp. 23–28.

[39] Wilson, D. B. Random random walks on $\mathrm{Z}_2^{\mathrm{d}}$. *Probab. Theory Related Fields 108*, 4 (1997), 441–457.