Using stopping times to bound mixing times *

Igor Pak†

October 31, 2004

1 Introduction

In the past decades Monte Carlo methods proved to be of a great importance in both theory and applications. There are several main techniques for bounding the convergence of Markov chains used in the analysis of algorithms. While they are powerful in some examples, they are weak in others. Here we will show how the use of a largely forgotten strong stationary time technique due to Aldous and Diaconis (see [1]) can improve some old bounds and give new bounds in cases when other techniques fail.

We will restrict ourselves to random walks (r.w.) on groups, which can defined as nearest neighbor random walks on Cayley graphs. Among others, we obtain bound $O(n^{2.5})$ on a mixing time for a random walk on a group of upper triangular matrices. A bound $O(n^{3.5} \log\log n)$ is given for a group of real orthogonal matrices generated by random rotations. While in the first case only a weaker $O(n^3 \log n)$ bound was known (see [8]), no polynomial bound was known in the second case.

We also show that at least theoretically one can always *speed up* the random walk on a group G by changing the probability of generators as well as by considering liftings to nonreversible Markov chains on $G \times G$. The first approach improves mixing from a general bound $O(|S|\Delta^2 \log |G|)$ to $O(\Delta^2 \log |G|)$, while the second improves further to $O(\Delta \log |G|)$ (Δ denotes the diameter). The first result solves an open problem of Aldous and Fill (see [2]), while the second can be generalized to *all* Markov chains by use of multicommodity flows (joint work with L. Lovász).

*Supported by NSF Postdoctoral Research Fellowship.
†Department of Mathematics, Yale University

Let G be a finite or compact group, and let S be a set of generators. A *random walk* $\mathcal{W}(G,S,p)$ on G generated by S, with a holding probability p is defined as $X_0 = id$, $X_{t+1} = X_t \cdot s$, where $s = id$ with probability p, or s is uniform in S with probability $(1-p)$. Let $Q^t(g) = \mathbf{P}(X_t = g) \to 1/|G|$ as $t \to \infty$. Define a *mixing time*: $\mathbf{mix} = \min\{k \mid Q^k(g) > 1/2|G| \ \forall\, g \in G\}$.

Let τ be a stopping time defined by a randomized stopping rule. We say that τ is *strong uniform* if its stopping state X_τ is uniformly distributed on G and is independent on τ : $\mathbf{P}(X_\tau = g \mid \tau = k) = 1/|G|$ for all $k > 0$, $g \in G$. The results of Aldous and Diaconis (see [1, 4, 6]) imply that $\mathbf{mix} \le 2\,E(\tau)$ given τ is strong uniform. Not unlike coupling, the main practical problem now is to find a good construction of a strong uniform times. While this is difficult in most cases, it is known to be theoretically possible (see [2, 4]).

2 Results for specific random walks

Let $G = U(n;\mathbb{F}_q)$ be a group of upper triangular matrices over $\mathbb{F}_q$ (i.e. matrices with 1 on diagonal and 0 below diagonal.) Let $M_{i,j}(a)$ be a matrix with a at (i,j) and 0 elsewhere. Let $S_1 = \{E_{i,j}(a) \mid 1 \le i < j \le n, a \in \mathbb{F}_q^*\}$, $S_2 = \{E_{i,i+1}(a) \mid 1 \le i < n, a \in \mathbb{F}_q^*\}$, where $E_{i,j}(a) = I + M_{i,j}(a)$. Everywhere below $O(\cdot)$ depends only on n and does not depend on q.

Theorem 2.1 For a r.w. $\mathcal{W}(G, S_1, 1/q)$ we have $\mathbf{mix} = O(n^2 \log n)$.

Theorem 2.2 For a r.w. $\mathcal{W}(G, S_2, 1/q)$ we have **mix** $= O(n^{2.5})$, given $q = \Omega(n^2)$

The first result has been recently generalized to all nilpotent groups and certain solvable groups (see [3, 6]). The second result improves a bound by Stong (see [8]). Both results are proved via explicit constructions of the strong uniform times.

Let $G_1 = SO(n;\mathbb{R})$, $G_2 = SO(n;\mathbb{F}_q)$. Let $R_{i,j}(\alpha)$ be a rotation by a uniform angle $\alpha \in [0, 2\pi]$ of a random basis 2-dimensional plane (i.e. chosen uniformly among $\binom{n}{2}$.) Let $S_1 = \{R_{i,j}(a) \,|\, 1 \le i < j \le n, a \in [0, 2\pi]\}$, and S_2 is defined analogously by using elements of $SO(2;\mathbb{F}_q)$ instead of rotations.

Theorem 2.3 For a r.w. $\mathcal{W}(G_1, S_1)$ we have **mix** $= O(n^{3.5} \log\log n)$.

Theorem 2.4 For a r.w. $\mathcal{W}(G_2, S_2, 1/q)$ we have **mix** $= O(n^3 \log\log n)$, given $q = \Omega(n^3)$.

Both results are proved in [7] using strong uniform times. No polynomial upper bound were known in the first case. Theorem 4 is a analog for an orthogonal group of the fast mixing in case $G = SL(n;\mathbb{F}_q)$ which give rise to expander constructions.

3 Speeding up the walk

There is a weak general bound **mix** $= O(|S|\Delta^2 \log |G|)$ which follows from $O(1/\Delta|S|)$ expansion property of Cayley graphs (see [2]). Here $S = S^{-1}$, and Δ is the diameter of G in terms of S. Rather than take a uniform distribution on S, we can use a different distribution P on $S + id$. Denote such r.w. by $\mathcal{W}(G, S, P)$.

Theorem 3.1 For any finite G, and any $S \subset G$ there exist a distribution P on $S + id$ such that for a r.w. $\mathcal{W}(G, S, P)$ we have **mix** $= O(\Delta^2 \log |G|)$.

This results solves an open problem 7.18 in [2]. In fact, we can inexplicitly construct such a P. Fix any set of decompositions of all elements of G as a shortest prodect in terms of S. Now take $P(s)$ to be proportional to the average number of times $s \in S$ is used in these decompositions.

We say that a Markov chain $\mathcal{M}$ on H is a *lifting* of $\mathcal{W}(G, S, p)$ if there is a projections $\psi : H \to G$ such that $\psi(\mathcal{M}) = \mathcal{W}$.

Theorem 3.2 For any finite G, and any $S \subset G$ there exist a lifting of $\mathcal{W}(G, S, p)$ to a Markov chain $\mathcal{M}$ on $G \times G$, such that **mix** $= O(\Delta|S| \log |G|)$.

The proof is based on an explicit, while purely theoretical construction of a stopping time and is a subject of an ongoing project with László Lovász (cf. [5]).

References

[1] D. Aldous, P. Diaconis, *Strong uniform times and finite random walks*, Advances in Applied Math., 8 (1987), 69–97.

[2] D. J. Aldous and J. Fill, *Time-reversible Markov chains and random walks on graphs*, (book in preparation)

[3] A. Astashkevich, I. Pak, *Random walks on nilpotent and solvable groups*, preprint (1998)

[4] P. Diaconis, *Group Representations in Probability and Statistics*, IMS, Hayward, California (1988)

[5] L. Lovász, P. Winkler, *Mixing Times*, Microsurveys in Discrete Probability (ed. D. Aldous and J. Propp), DIMACS Series, AMS, 1998

[6] I. Pak, *Random walks on groups: Strong uniform time approach*, Ph.D. thesis, Harvard University (1997)

[7] I. Pak, *On random rotations*, preprint (1998)

[8] R. Stong, *Random walk on the upper triangular matrices*, Ann. Prob., 23 (1995), 1939–1949