White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of York.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/79774/

Version: Accepted Version

Article:
Bate, Michael E. and Connor, Stephen B. (2018) Mixing Time and Cutoff for a Random Walk on the Ring of Integers mod n. Bernoulli. pp. 993-1009. ISSN: 1350-7265

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

arXiv:1407.3580v2 [math.PR] 25 Feb 2016

Mixing Time and Cutoff for a Random Walk on the Ring of Integers mod n

Michael E. Bate and Stephen B. Connor

26th February 2016

Abstract

We analyse a random walk on the ring of integers mod n, which at each time point can make an additive 'step' or a multiplicative 'jump'. When the probability of making a jump tends to zero as an appropriate power of n we prove the existence of a total variation pre-cutoff for this walk. In addition, we show that the process obtained by subsampling our walk at jump times exhibits a true cutoff, with mixing time dependent on whether the step distribution has zero mean.

Keywords: random walk; mixing time; cutoff phenomenon; pre-cutoff; group representation theory; integers mod n; random number generation

2010 Mathematics Subject Classification:
Primary 60J10

1 Introduction

In this note we consider a random walk $X = X^{(n)}$ on $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$ (where n is *odd*) defined as follows: $X_0 = 0$, and for $t \geq 1$,

$$X_t = \begin{cases} X_{t-1} + \xi'_t \mod n & \text{with probability } 1 - p_n \\ 2X_{t-1} \mod n & \text{with probability } p_n, \end{cases} \tag{1}$$

where $\{\xi'_t\}$ are a set of i.i.d. random variables with finite support $B \subset \mathbb{Z}$, whose distribution does not vary with n. We denote the mean and variance of ξ' by μ and σ^2 respectively. We will refer to an 'addition' move as a 'step', and to a 'multiplication' move as a 'jump'. To ensure that X is irreducible we assume that the group $\langle B_n, + \rangle$ is not a proper subgroup of $\mathbb{Z}_n$ for any odd n, where $B_n = \{z \mod n : z \in B\}$. Furthermore, since n is odd, multiplication by 2 is an invertible operation, and thus X is ergodic with uniform equilibrium distribution π_n on $\mathbb{Z}_n$.

Define the total variation distance from π_n of a probability distribution P on $\mathbb{Z}_n$ by

$$\|P - \pi_n\| = \max_{A \subset \mathbb{Z}_n} |P(A) - \pi_n(A)| = \frac{1}{2} \sum_{s \in \mathbb{Z}_n} |P(s) - 1/n| \,.$$

The *ε-mixing time* of X is defined for any $\varepsilon \in [0, 1]$ as

$$t_{\text{mix}}(\varepsilon) = \min\{t : \|\mathbb{P}(X_t \in \cdot) - \pi_n(\cdot)\| \leq \varepsilon\} \,,$$

and the value of $t_{\text{mix}}(1/4)$ is commonly referred to as the 'mixing time' of X.

A number of authors have previously considered random processes of the form

$$X_t = a_t X_{t-1} + b_t \mod n;$$

these processes are similar to schemes used for random number generation, a link which has naturally motivated interest in bounding their mixing times. A nice introduction to the area can be found in Terras (1999, Chapter 6). The earliest such work appears to be that of Chung et al. (1987), in which $a_t = a = 2$ and b_t is chosen uniformly from $\{-1, 0, 1\}$: they show that $O(\log n \log\log n)$ steps suffice for this walk to mix, and that $O(\log n \log\log n)$ steps are also necessary for n of the form $2^m - 1$; on the other hand, for almost all odd n, $1.02 \log_2 n$ steps suffice. This (deterministic) act of doubling each time causes the process to mix significantly faster than when $a_t = 1$ for all t where, if b_t is uniform on a finite set (and assuming that the resulting process is irreducible), the mixing time is of order n^2 (Diaconis, 1988; Saloff-Coste, 2004).

Rather more general results have been established in a series of works by Hildebrand. It is shown in his thesis (Hildebrand, 1990, Chapter 3) that if $a_t = a$ for all t, and for fairly general choices of b_t (which don't depend on n), $O(\log n \log\log n)$ steps suffice, and in fact for almost all n, $O(\log n)$ steps suffice. When a_t is allowed to vary with t, a general upper bound for the mixing time is proved in Hildebrand (1993): using a recursive relation involving discrete Fourier transforms (of which more below), he shows that (unless $a_t = 1$ always, $b_t = 0$ always, or a_t and b_t can each take on only one value) $O((\log n)^2)$ time steps are always sufficient. Other related results can be found in Hildebrand (1994a,b).

A particularly interesting feature of these processes is the quantitatively different behaviour that can be obtained by making small changes to the distribution of a_t and b_t. For example, Chung et al. (1987) remark upon the following curiosity to be found when $a_t = 2$ and b_t is supported on $\{-1, 0, 1\}$ with $\mathbb{P}(b_t = 1) = \mathbb{P}(b_t = -1) = q$: if $q = 1/4$ or $q = 1/2$ then $O(\log n)$ steps suffice to make the total variation distance small; however, if $q = 1/3$ then $O(\log n \log\log n)$ steps may be required. Similarly, Hildebrand (1990, Chapter 5) considers the situation where b_t is uniform on ± 1 and a_t is supported on $\{2, (n+1)/2\}$, with $\mathbb{P}(a_t = 2) = p \in (0, 1)$: the mixing time is shown to be at most $O((\log n)^m)$, where m is 2 if $p = 1/2$, and 1 otherwise. If the distribution of b_t is altered to uniform on $\{-1, 0, 1\}$ then $O((\log n \log\log n)^m)$ steps suffice.

The principal difference between these earlier works and the process defined in (1) is that here we allow the probability of a 'jump', p_n, to depend on n. In particular, we are able to show that if p_n tends to zero as a power of n, then our process exhibits a total variation *pre-cutoff*.

Definition 1. A sequence of chains $\{X^{(n)}\}_{n\in\mathbb{N}}$, with ε-mixing times $\{t^{(n)}_{\text{mix}}(\varepsilon)\}$, is said to exhibit a *pre-cutoff* if it satisfies

$$\sup_{0<\varepsilon<1/2} \limsup_{n\to\infty} \frac{t^{(n)}_{\text{mix}}(\varepsilon)}{t^{(n)}_{\text{mix}}(1-\varepsilon)} < \infty\,.$$

A stronger condition than pre-cutoff is that the sequence of chains exhibits a total variation *cutoff*.

Definition 2. A sequence of Markov chains $\{X^{(n)}\}_{n\in\mathbb{N}}$ is said to exhibit a *total variation cutoff* at time T_n with *window size* w_n if $w_n = o(T_n)$ and

$$\lim_{c\to\infty} \liminf_{n\to\infty} \|\mathbb{P}(X^{(n)}_{T_n - cw_n} \in \cdot) - \pi_n(\cdot)\| = 1$$
$$\lim_{c\to\infty} \limsup_{n\to\infty} \|\mathbb{P}(X^{(n)}_{T_n + cw_n} \in \cdot) - \pi_n(\cdot)\| = 0\,.$$

Intuitively this says that as n gets large the convergence to equilibrium, measured using total variation distance, happens in a negligible window of order w_n around the cutoff time T_n. We remark that it is possible for the 'right' and 'left' window sizes in the above definition to be of different orders – see Connor (2010) for an example. There has been much interest in studying the mixing times of Markov chains and proving the existence of cutoff phenomena: see Levin et al. (2009) and Diaconis (2011) for recent introductions to the area, or Saloff-Coste (2004) for a more analytical overview. In addition a number of natural sequences of Markov chains are known to exhibit pre-cutoff, with the question of whether they in fact exhibit a cutoff still open; these include card shuffles using cyclic-to-random transpositions (Mossel et al., 2004) or random-to-random insertions (Subag, 2013), and a Gibbs sampler on the n-simplex (Smith, 2014).

Throughout the remainder of this paper we shall simply write X for $X^{(n)}$, with the understanding that we are in reality considering a sequence of processes on state spaces $\mathbb{Z}_n$ of increasing size. The main obstruction to analysing X using standard techniques for random walks on groups is that the distribution of X_k is not given by convolution of k independent increment distributions. This problem can be overcome by (initially) restricting attention to the process Y which is produced by subsampling X at jump times. Denote the jump times of X by $\tau_1, \tau_2, \ldots$, and let $\tau_0 = 0$; then $Y_k := X_{\tau_k}$, with $Y_0 = X_0 = 0$. This process clearly satisfies $Y_k = Y'_k \mod n$, where

$$Y'_k = \sum_{i=1}^{k} 2^{k+1-i} S'_i \quad \text{and} \quad S'_i = \sum_{t=\tau_{i-1}+1}^{\tau_i - 1} \xi'_t\,. \tag{2}$$

Here (and throughout) we use the convention that random variables with a prime take values in $\mathbb{Z}$, while those without take values in $\mathbb{Z}_n$. Thus $S_i = S'_i \mod n$ is the change in X due to steps taken between jump times τ_{i-1} and τ_i. Like X, Y is ergodic with uniform equilibrium distribution. From (2) it is clear that the distribution of Y_k is given by convolution of the distributions corresponding to the independent increments $\{2^{k+1-i}S_i\}$, and this will prove essential to our method for establishing an upper bound on the mixing time of both X and Y in Section 3.

In order to state our main results, we first need to establish a little more notation: we shall write $\sigma^2_{S'} = \mathrm{Var}\,(S'_i)$ and $T_n = \log_2(n/\sigma_{S'})$. Note that the length of time between jumps of X has a Geometric(p_n) distribution, and a straightforward application of the conditional variance formula shows that

$$\sigma^2_{S'} = \frac{(1-p_n)(\mu^2 + p_n\sigma^2)}{p_n^2}\,. \tag{3}$$

Thus if $p_n \to 0$,

$$T_n \sim \begin{cases} \log_2\left(\frac{n\sqrt{p_n}}{\sigma}\right) & \text{if } \mu = 0 \\ \log_2\left(\frac{np_n}{|\mu|}\right) & \text{otherwise.} \end{cases} \tag{4}$$

Our main results are as follows.

Theorem 3. *Suppose that $p_n = 1/(2n^\alpha)$ for some $\alpha \in (0,2)$ such that $T_n \to \infty$ as $n \to \infty$. Then Y exhibits a total variation cutoff at time T_n, with cutoff window of size $O(1)$. Indeed, for sufficiently large $c > 0$,*

1. $\liminf_{n\to\infty} \| \mathbb{P}(Y_{T_n - c} \in \cdot) - \pi_n(\cdot)\| \geq 1 - 4^{1-c/3}$;
2. $\limsup_{n\to\infty} \| \mathbb{P}(Y_{T_n + c} \in \cdot) - \pi_n(\cdot)\| \leq O(2^{-c})$.

Note that the mixing time of Y is relatively insensitive to the distribution of the step lengths: as can be seen from (4), the cutoff time T_n essentially depends on ξ' only through its mean, μ; in the case of zero drift the mixing time is asymptotically $(1-\alpha/2)\log_2 n$ $(0 < \alpha < 2)$, while if $\mu \neq 0$ the mixing is slightly faster, with cutoff at $(1-\alpha)\log_2 n$ $(0 < \alpha < 1)$.

Theorem 4. *Suppose that p_n satisfies the same condition as in Theorem 3. Let $T_n^L = (2\ln 2)n^\alpha T_n$, $T_n^R = 2n^\alpha T_n$, $w_n^L = 2n^\alpha$ and $w_n^R = 2n^\alpha\sqrt{T_n}$. Then for sufficiently large $c > 0$,*

1. $\liminf_{n\to\infty} \|\mathbb{P}(X_{T_n^L - cw_n^L} \in \cdot) - \pi_n(\cdot)\| \geq 1 - ae^{-c/2}$, *for some finite constant a;*
2. $\limsup_{n\to\infty} \|\mathbb{P}(X_{T_n^R + cw_n^R} \in \cdot) - \pi_n(\cdot)\| \leq 2/c^2$.

In particular, since $T_n^L/T_n^R = O(1)$, $w_n^L = o(T_n^L)$ and $w_n^R = o(T_n^R)$, Theorem 4 shows that X exhibits a pre-cutoff, with mixing time $t_{\text{mix}}^{(n)}(\varepsilon) = \Omega(n^\alpha \ln n)$.

Throughout the rest of the paper we shall work under the assumptions of Theorem 3. In Section 2 we establish lower bounds on the mixing time of both X and Y (proving part 1 of Theorems 3 and 4). In Section 3 we prove the corresponding upper bounds. Section 4 contains some concluding remarks and open questions.

2 Lower bounds

As is typical for many problems of this sort, finding lower bounds for the mixing times of our two processes is significantly easier than establishing upper bounds. The general approach in each case is to find a suitably large subset of the state space which our chain has negligible chance of hitting before the time of interest.

2.1 Lower bound for Y

An elementary calculation using (2) shows that $\mathbb{E}\left[Y'_{T_n - c}\right] \sim \text{sgn}(\mu)2^{1-c}n$, where we define $\text{sgn}(0) = 0$. Furthermore, $\text{Var}(Y'_k) = 4(4^k - 1)\sigma_{S'}^2/3$, and so $\text{Var}\left(Y'_{T_n - c}\right) \leq 4^{1-c}n^2/3$ for sufficiently large n. Now consider the interval

$$A_n(c) = \left\{ z \in \mathbb{Z}_n \,:\, |z - \mathbb{E}\left[Y'_{T_n - c}\right]| > d_c n \right\},$$

for some value $d_c \in (0, 1/2)$ which we shall choose later, and where $|\cdot|$ represents the usual distance between two numbers mod n. Note that $\pi_n(A_n(c)) \geq 1 - 2d_c - 1/n$, and that (subject to this condition) this set has been chosen to be as far away as possible from $\mathbb{E}\left[Y'_{T_n - c}\right]$.

We now use the fact that Y is equal to $Y' \mod n$, along with Chebychev's inequality, to bound the probability that Y_{T_n-c} belongs to our chosen set $A_n(c)$:

$$\mathbb{P}\left(Y_{T_n-c} \in A_n(c)\right) \leq \mathbb{P}\left(|Y'_{T_n-c} - \mathbb{E}\left[Y'_{T_n-c}\right]| > d_c n\right) \leq \frac{\mathrm{Var}\left(Y'_{T_n-c}\right)}{(d_c n)^2} \leq \frac{4^{1-c}}{3d_c^2}\,.$$

Thus the set $A_n(c)$ satisfies

$$\pi_n(A_n(c)) - \mathbb{P}\left(Y_{T_n-c} \in A_n(c)\right) \geq 1 - 2d_c - 1/n - \frac{4^{1-c}}{3d_c^2}\,.$$

Finally, taking $d_c = \left(4^{1-c}/3\right)^{1/3}$ yields the claimed left hand window of the cutoff in part 1 of Theorem 3:

$$\liminf_{n\to\infty} \|\,\mathbb{P}\left(Y_{T_n-c} \in \cdot\right) - \pi_n(\cdot)\| \geq 1 - \left(\frac{9}{4^{c-1}}\right)^{1/3} \geq 1 - 4^{1-c/3}\,.$$

2.2 Lower bound for X

Let X' be the random walk on $\mathbb{Z}$ defined as follows:

$$X'_0 = 0; \qquad X'_t = \begin{cases} X'_{t-1} + \xi'_t & \text{with probability } 1-p_n \\ 2X'_{t-1} & \text{with probability } p_n. \end{cases} \tag{5}$$

That is, $X = X' \mod n$.

A natural approach to lower bound the mixing time for X would be to replicate the above argument for Y, using Chebychev's inequality applied to X'. However, the random number of jumps by time $T_n^L - cw_n^L$ causes the variance of X' at this time to be too large for this to work. Instead, we proceed by bounding the expectation of $|X'_{T_n^L - cw_n^L}|$ and then using Markov's inequality to show that $X_{T_n^L - cw_n^L}$ has negligible chance of belonging to a region of the state space situated 'opposite' $X_0 = 0$. We begin by proving the following lemma.

Lemma 5. *There exists a constant $a < \infty$ such that at time $T_n^L - cw_n^L$,*

$$\mathbb{E}\left[|X'_{T_n^L - cw_n^L}|\right] \leq ane^{-c}\,. \tag{6}$$

Proof. Let $J_n(-c)$ be the number of jumps in X' by time $T_n^L - cw_n^L$. For large n, the distribution of $J_n(-c)$ is well approximated by a Poisson$(m_n(-c))$ distribution, with

$$m_n(-c) = p_n(T_n^L - cw_n^L) = \ln(n/\sigma_{S'}) - c\,, \tag{7}$$

(using the definition of T_n^L in Theorem 4 and the equality $p_n = 1/(2n^\alpha)$). Note that under the assumptions of Theorem 4, $m_n(-c) \to \infty$ as $n \to \infty$. Conditional on the event $\{J_n(-c) = k\}$ we can express $X'_{T_n^L - cw_n^L}$ as follows:

$$X'_{T_n^L - cw_n^L} \,|\, \{J_n(-c) = k\} = \sum_{i=1}^{k+1} 2^{k+1-i} S_i'^{(k)}\,, \tag{8}$$

where $S_i'^{(k)} \stackrel{d}{=} S'_i \,|\, \{J_n(-c) = k\}$. That is, $S_i'^{(k)}$ is the additive increment in X' between jump times τ_{i-1} and τ_i, with $\tau_0 := 0 < \tau_1 < \cdots < \tau_k \leq \tau_{k+1} := T_n^L - cw_n^L$. It is clear that for

$i = 1, \ldots, k$ the random variables $S_i'^{(k)}$ have a common distribution, and that $\mathbb{E}[|S_{k+1}'^{(k)}|] \leq \mathbb{E}[|S_1'^{(k)}|]$ (since it is possible to have $\tau_k = \tau_{k+1}$). It follows from (8) that for $k \geq 1$,

$$\mathbb{E}\left[|X'_{T_n^L - cw_n^L}| \,\middle|\, J_n(-c) = k \right] \leq 2^{k+1}\, \mathbb{E}\left[|S_1'^{(k)}| \right] . \tag{9}$$

We now deal with the cases of zero and non-zero μ separately.

- Case 1: $\mu \neq 0$.

Suppose that $k \geq 1$. Let b be an integer satisfying $B \subseteq [-2^b, 2^b]$, where (recall that) B is the support of ξ'. Then

$$\begin{aligned} \mathbb{E}\left[|S_1'^{(k)}| \right] &\leq 2^b\, \mathbb{E}\left[\tau_1 - 1 \,|\, J_n(-c) = k \right] \\ &\leq \frac{2^b (T_n^L - cw_n^L)}{k} \leq \frac{2^{b+1}(T_n^L - cw_n^L)}{k+1} , \end{aligned} \tag{10}$$

where the second inequality follows from the symmetry observation made just before (9), and the last one uses the assumption that $k \geq 1$. Combining (9) and (10) we see that for $k \geq 1$,

$$\mathbb{E}\left[|X'_{T_n^L - cw_n^L}| \,\middle|\, J_n(-c) = k \right] \leq \frac{2^{b+k+2}(T_n^L - cw_n^L)}{k+1} .$$

Furthermore, note that this also (trivially) holds when $k = 0$.

Now average over the distribution of $J_n(-c)$:

$$\begin{aligned} \mathbb{E}\left[|X'_{T_n^L - cw_n^L}| \right] &\leq 2^{b+1}(T_n^L - cw_n^L) \sum_{k=0}^{\infty} \frac{e^{-m_n(-c)} m_n(-c)^k 2^{k+1}}{k!(k+1)} \\ &\leq \frac{2^{b+1}(T_n^L - cw_n^L)}{m_n(-c)} e^{m_n(-c)} = \frac{2^{b+1}}{p_n} \frac{ne^{-c}}{\sigma_{S'}} \sim 2^{b+1} n e^{-c} |\mu| \end{aligned}$$

for large n, thanks to the expression for $m_n(-c)$ in (7) and the relationship between p_n and $\sigma_{S'}$ in (3). Taking $a = 2^{b+1}|\mu|$ gives the required result.

- Case 2: $\mu = 0$.

In this case we know that $\mathbb{E}\left[S_1'^{(k)} \right] = 0$. Furthermore,

$$\mathrm{Var}\left(S_1'^{(k)} \right) \leq \mathbb{E}\left[\sigma^2(\tau_1 - 1) \,|\, J_n(-c) = k \right] \leq \frac{2\sigma^2 T_n^L}{k+1} ,$$

using the same reasoning that led to (10). Chebychev's inequality then yields, for any positive x:

$$\mathbb{E}\left[|S_1'^{(k)}| \right] \leq x + \int_x^{\infty} \mathbb{P}\left(|S_1'^{(k)}| > s \right) \mathrm{d}s \leq x + \frac{2\sigma^2 T_n^L}{(k+1)x} .$$

Substituting $x = 2\sigma\sqrt{T_n^L/(k+1)}$ we obtain

$$\mathbb{E}\left[|S_1'^{(k)}| \right] \leq 3\sigma \sqrt{\frac{T_n^L}{k+1}} .$$

Combining this with (8) we see that

$$\mathbb{E}\left[|X'_{T_n^L - cw_n^L}|\right] \leq 6\sigma\sqrt{T_n^L}\,\mathbb{E}\left[\frac{2^{J_n(-c)}}{\sqrt{J_n(-c)+1}}\right]. \tag{11}$$

Now consider the final expectation in (11). Recalling that, for large n, $J_n(-c)$ is well-approximated by a Poisson distribution with mean $m_n(-c)$, we can bound this as follows:

$$\begin{aligned}\mathbb{E}\left[\frac{2^{J_n(-c)}}{\sqrt{J_n(-c)+1}}\right] &\leq e^{m_n(-c)} \sum_{k=0}^{m_n(-c)} \frac{(2m_n(-c))^k e^{-2m_n(-c)}}{k!} \\ &\quad + \frac{1}{\sqrt{m_n(-c)}} \sum_{k>m_n(-c)} \frac{(2m_n(-c))^k e^{-m_n(-c)}}{k!} \\ &\leq e^{m_n(-c)} \left[\mathbb{P}\left(\Lambda < m_n(-c)\right) + \frac{1}{\sqrt{m_n(-c)}}\right], \end{aligned} \tag{12}$$

where $\Lambda \sim \text{Poisson}(2m_n(-c))$. A final application of Chebychev's inequality tells us that $\mathbb{P}\left(\Lambda < m_n(-c)\right) \leq 2/m_n(-c)$, and since $m_n(-c) \to \infty$ as $n \to \infty$ this term in (12) is negligible. Combining (11) and (12) we therefore arrive at our desired result: for large n we can write

$$\mathbb{E}\left[|X'_{T_n^L - cw_n^L}|\right] \leq 12\sigma\sqrt{T_n^L}\frac{e^{m_n(-c)}}{\sqrt{m_n(-c)}} = \frac{12\sigma}{\sqrt{p_n}}\frac{ne^{-c}}{\sigma_{S'}} \sim 12ne^{-c}$$

thanks once again to the expressions for $m_n(-c)$ and $\sigma_{S'}$ in (7) and (3). □

Recall that our aim in this section is to lower bound

$$\|\mathbb{P}(X_{T_n^L - cw_n^L} \in \cdot) - \pi_n(\cdot)\| = \sup_A \left(\pi_n(A) - \mathbb{P}(X_{T_n^L - cw_n^L} \in A)\right). \tag{13}$$

Define the set $D_n(c)$ to be those points in $\mathbb{Z}_n$ whose distance from 0 (measured in the usual way between two numbers in $\mathbb{Z}_n$) is greater than $e^{-c/2}n$. Note that

$$\pi_n(D_n(c)) \geq 1 - 2e^{-c/2} - 1/n\,.$$

Using Markov's inequality and Lemma 5 we obtain:

$$\mathbb{P}\left(X_{T_n^L - cw_n^L} \in D_n(c)\right) \leq \mathbb{P}\left(|X'_{T_n^L - cw_n^L}| > e^{-c/2}n\right) \leq ae^{-c/2}\,.$$

Thus

$$\pi_n(D_n(c)) - \mathbb{P}(X_{T_n^L - cw_n^L} \in D_n(c)) \geq 1 - (2+a)e^{-c/2} - 1/n$$

and so

$$\liminf_{n\to\infty} \|\mathbb{P}(X_{T_n^L - cw_n^L} \in \cdot) - \pi_n(\cdot)\| \geq 1 - (2+a)e^{-c/2}\,.$$

3 Upper bounds

In this section we prove the second parts of Theorems 3 and 4. Most of the work here is required to prove the result for the subsampled chain Y. Indeed, the result for X follows quickly from this, as we now demonstrate.

Let $J_n(c)$ denote the number of jumps in X by time $T_n^R + cw_n^R$: for large n this is once again well approximated by a Poisson random variable, this time with mean $m_n(c) = T_n + c\sqrt{T_n}$. Assuming Theorem 3 to be true, we know that the subsampled chain Y is well mixed after $T_n + c$ jumps; we proceed by considering whether or not this number of jumps has occurred by our time of interest, $T_n^R + cw_n^R$. For ease of display, in the next few lines we shall write $\tau = \tau_{T_n+c}$ for the (random) time at which X jumps for the $(T_n + c)$-th time.

$$\begin{aligned}\|\mathbb{P}(X_{T_n^R+cw_n^R} \in \cdot) - \pi_n(\cdot)\| \leq \mathbb{E}\left[\|\mathbb{P}(X_{T_n^R+cw_n^R} \in \cdot) - \pi_n(\cdot)\|\,;\, \tau \leq T_n^R + cw_n^R\right] \\ + \mathbb{P}\left(\tau > T_n^R + cw_n^R\right) \qquad (14)\end{aligned}$$

due to total variation being bounded above by 1. We now use the fact that total variation is non-increasing over time to bound the expectation term:

$$\begin{aligned}\mathbb{E}\left[\|\mathbb{P}(X_{T_n^R+cw_n^R} \in \cdot) - \pi_n(\cdot)\|\,;\, \tau \leq T_n^R + cw_n^R\right] &= \mathbb{E}\left[\sum_{k=1}^{T_n^R+cw_n^R} \|\mathbb{P}(X_{T_n^R+cw_n^R} \in \cdot) - \pi_n(\cdot)\|\mathbf{1}_{[\tau=k]}\right] \\ &\leq \mathbb{E}\left[\sum_{k=1}^{T_n^R+cw_n^R} \|\mathbb{P}(X_k \in \cdot) - \pi_n(\cdot)\|\mathbf{1}_{[\tau=k]}\right] \\ &= \mathbb{E}\left[\|\mathbb{P}(X_\tau \in \cdot) - \pi_n(\cdot)\|\,;\, \tau \leq T_n^R + cw_n^R\right] \\ &\leq \|\,\mathbb{P}\left(X_\tau \in \cdot\right) - \pi_n(\cdot)\| \\ &= \|\,\mathbb{P}\left(Y_{T_n+c} \in \cdot\right) - \pi_n(\cdot)\|\,. \qquad (15)\end{aligned}$$

The first term in (14) is thus controlled by part 2 of Theorem 3. Furthermore, using Chebychev's inequality once again:

$$\begin{aligned}\mathbb{P}\left(\tau > T_n^R + cw_n^R\right) = \mathbb{P}\left(J_n(c) < T_n + c\right) &\leq \mathbb{P}\left(|J_n(c) - m_n(c)| \geq c\left(\sqrt{T_n} - 1\right)\right) \\ &\leq \frac{m_n(c)}{c^2 T_n + O(\sqrt{T_n})} \to \frac{1}{c^2} \quad \text{as } n \to \infty.\end{aligned}$$

Putting all of the above together we complete the proof of Theorem 4: for sufficiently large $c > 0$,

$$\limsup_{n\to\infty} \|\mathbb{P}(X_{T_n^R+cw_n^R} \in \cdot) - \pi_n(\cdot)\| \leq \frac{2}{c^2}\,.$$

3.1 Upper bounds and representation theory

Our basic method for obtaining upper bounds on the mixing times of our processes is to employ the techniques developed by Diaconis and Shahshahani (1981) for analysing random walks on groups. Given a probability Q on a finite group G, and a representation ρ of G, we can form the Fourier transform $\hat{Q}(\rho)$ of Q at ρ by setting

$$\hat{Q}(\rho) := \sum_{g\in G} Q(g)\rho(g)\,.$$

The following Upper Bound Lemma (Diaconis, 1988) then allows one to compute an explicit upper bound for the total variation distance between a probability Q on G and the uniform distribution π.

Lemma 6. *Given a probability Q on a finite group G, we have*

$$\|Q-\pi\|^2 \leq \frac{1}{4}\sum \deg(\rho)\mathrm{tr}(\hat{Q}(\rho)\hat{Q}(\rho)^*),$$

where $A^* = (\overline{a_{ji}})$ denotes the complex conjugate transpose of the matrix $A = (a_{ij})$, tr *denotes the trace function on square matrices, and the sum is taken over all non-trivial irreducible representations ρ of G.*

Since the Fourier transform behaves well with respect to convolution, this lemma provides a practical tool for bounding the mixing time of a random walk on a group. Although Y is not strictly a random walk on the additive group $(\mathbb{Z}_n, +)$, the measure giving the distribution of Y_k *can* be expressed as the convolution of measures. Here the representation theory is particularly straightforward: the Upper Bound Lemma becomes

$$\|Q-\pi\|^2 \leq \frac{1}{4}\sum_{s=1}^{n-1} |\hat{Q}(\rho_s)|^2\,, \tag{16}$$

where the representations $\rho_0, \rho_1, \ldots, \rho_{n-1}$ all have degree 1, and are completely determined by the following equations:

$$\rho_s(1) := e^{i\frac{2\pi}{n}s} \text{ for } 0 \leq s \leq n-1\,.$$

Recall from (2) that (with $Y_0 = 0$), $Y_k = \sum_{j=1}^{k} 2^j S'_{k+1-j} \mod n$. The measure P_k giving the distribution of Y_k is the convolution of the measures λ_j given by $\lambda_j(2^j a \mod n) = \mathbb{P}(S_1 = a)$ for every j, a, so we begin by calculating the Fourier transforms of the λ_j. To ease notation, for each $1 \leq j \leq k$ and $0 \leq s \leq n-1$, set

$$\omega_{s,j} = \rho_s(2^j) = e^{i\frac{2\pi}{n}2^j s}$$

and note that for any j, s we have $\omega_{s,j}^n = 1$. Then for each $0 \leq s \leq n-1$,

$$\begin{aligned}\hat{\lambda}_j(\rho_s) &= \sum_{a=0}^{n-1} \omega_{s,j}^a\, \mathbb{P}(S_1 = a) = \sum_{a=0}^{n-1} \omega_{s,j}^a \sum_{d\in\mathbb{Z}} \mathbb{P}\left(S'_1 = a+dn\right)\\ &= \sum_{d\in\mathbb{Z}}\sum_{a=0}^{n-1} \omega_{s,j}^{a+dn}\, \mathbb{P}\left(S'_1 = a+dn\right) = \sum_{a\in\mathbb{Z}} \omega_{s,j}^a\, \mathbb{P}\left(S'_1 = a\right) = G_{S'}(\omega_{s,j})\,,\end{aligned}$$

where $G_{S'}$ is the probability generating function (PGF) of S'. It follows from its definition in (2) as a random sum of random step lengths that this satisfies

$$G_{S'}(\omega_{s,j}) = \frac{p_n}{1-(1-p_n)G_{\xi'}(\omega_{s,j})}\,, \tag{17}$$

where $G_{\xi'}$ is the PGF of ξ'.

When we substitute into the Upper Bound Lemma 6, we are interested in the modulus squared of such expressions, by Equation (16). The modulus of the top line squared is p_n^2, and the modulus of the bottom line squared is

$$\begin{aligned}(1-(1-p_n)G_{\xi'}(\omega_{s,j}))&(1-(1-p_n)\overline{G_{\xi'}(\omega_{s,j})})\\ &= 1-(1-p_n)\left(G_{\xi'}(\omega_{s,j})+\overline{G_{\xi'}(\omega_{s,j})}\right)+(1-p_n)^2 G_{\xi'}(\omega_{s,j})\overline{G_{\xi'}(\omega_{s,j})}\\ &= 1-2(1-p_n)\mathrm{Re}\left(G_{\xi'}(\omega_{s,j})\right)+(1-p_n)^2|G_{\xi'}(\omega_{s,j})|^2\,.\end{aligned}$$

Combining all of the above leads to the following upper bound for the total variation distance at time k:

$$\|\,\mathbb{P}\,(Y_k\in\cdot)-\pi_n(\cdot)\|^2 \le \frac{1}{4}\sum_{s=1}^{n-1}\prod_{j=1}^{k}\frac{p_n^2}{1-2(1-p_n)\mathrm{Re}\left(G_{\xi'}(\omega_{s,j})\right)+(1-p_n)^2|G_{\xi'}(\omega_{s,j})|^2}\,. \quad (18)$$

3.2 Strategy for analysing the upper bound

In order to establish a cutoff for Y, we need to control the right hand side of (18) around time $T_n=\log_2(n/\sigma_{S'})$. To that end, we define for $c\in\mathbb{N}$ a function $U_n(c)$ by

$$U_n(c)=\sum_{s=1}^{n-1}\prod_{j=1}^{T_n+c}\phi_n(s,j) \quad (19)$$

where

$$\phi_n(s,j):=\frac{p_n^2}{1-2(1-p_n)\mathrm{Re}\left(G_{\xi'}(\omega_{s,j})\right)+(1-p_n)^2|G_{\xi'}(\omega_{s,j})|^2}\in(0,1]\,, \quad (20)$$

and note that, thanks to (18), Theorem 3 will be proved if we can show that (for odd n) $\limsup_{n\to\infty}U_n(c)\le O(4^{-c})$.

Our strategy for bounding $U_n(c)$ involves identifying for each $1\le s\le n-1$ enough values j for which $\phi_n(s,j)$ is sufficiently small to provide a useful upper bound. In order to do this, it is convenient to first reparametrise, so we let Z_n be a random variable uniformly distributed on the set $\{s/n : s=1,\dots,n-1\}\subset[0,1]$. Then we may write

$$U_n(c)=\mathbb{E}\left[f_n(Z_n,T_n+c)\right],\quad \text{where } f_n(x,t):=(n-1)\prod_{j=1}^{t}\phi_n(nx,j)\,. \quad (21)$$

The second step is to split the analysis of the function f_n into two stages by splitting the range of x into two pieces. In order to do this, let L be an integer satisfying $2\alpha L>1$, and once again let b be an integer satisfying $B\subseteq[-2^b,2^b]$. We define a finite lattice $\mathcal{L}$ of points in $[0,1]$ by

$$\mathcal{L}=\left\{\frac{k}{2^{L+b}}\ :\ k=0,\dots,2^{L+b}\right\}.$$

Now choose some $\varepsilon\in(0,1/(2^{L+b}))$, and define the set $\mathcal{L}_\varepsilon$ to be the intersection of $[0,1]$ with

$$\bigcup_{x\in\mathcal{L}}\left[x-\frac{\varepsilon}{2},x+\frac{\varepsilon}{2}\right].$$

Importantly, $\mathcal{L}_\varepsilon$ depends only on α, B and ε, but not on n. We now proceed to bound $f_n(x,T_n+c)$ by considering in turn the cases where x does and does not belong to the set $\mathcal{L}_\varepsilon$.

3.3 Controlling f_n for $x \notin \mathcal{L}_\varepsilon$

For $x \notin \mathcal{L}_\varepsilon$ we see that $2\pi 2^j a x \neq 0 \mod 2\pi$ for any $j = 1, 2, \ldots, L$ and $a \in B$. Thus $\cos(2\pi 2^j a x)$ is bounded away from 1 for all such x and j, and we can write

$$\mathrm{Re}\left(G_{\xi'}(e^{i2\pi 2^j x})\right) = \sum_{a=0}^{2^b} \mathbb{P}\left(|\xi'| = a\right) \cos\left(2\pi 2^j a x\right) \leq 1 - \kappa(x)\,,$$

for all $j = 1, \ldots, L$, where $\kappa(x)$ is strictly positive.

Substituting this into the expression for ϕ_n in (20), and lower-bounding the modulus squared of a complex number by the square of its real part, we obtain:

$$\begin{aligned}
\phi_n(nx, j) &\leq \frac{p_n^2}{1 - 2(1-p_n)\mathrm{Re}\left(G_{\xi'}(e^{i2\pi 2^j x})\right) + (1-p_n)^2 |G_{\xi'}(e^{i2\pi 2^j x})|^2} \\
&\leq \frac{p_n^2}{1 - 2(1-p_n)\mathrm{Re}\left(G_{\xi'}(e^{i2\pi 2^j x})\right) + (1-p_n)^2 \mathrm{Re}\left(G_{\xi'}(e^{i2\pi 2^j x})\right)^2} \\
&= \left(\frac{p_n}{1 - (1-p_n)\mathrm{Re}\left(G_{\xi'}(e^{i2\pi 2^j x})\right)}\right)^2 \leq \left(\frac{p_n}{1 - (1-p_n)(1-\kappa(x))}\right)^2 = O(p_n^2)\,.
\end{aligned}$$

Since $\phi_n(nx, j) \in (0, 1]$, it follows that for $x \notin \mathcal{L}_\varepsilon$ and for n sufficiently large that $T_n + c \geq L$,

$$f_n(x, T_n + c) = (n-1) \prod_{j=1}^{T_n + c} \phi_n(nx, j) \leq (n-1) \prod_{j=1}^{L} \phi_n(nx, j) \leq O(n^{1-2\alpha L})\,.$$

Thanks to our choice of $L > 1/2\alpha$ we can now use Fatou's Lemma to deduce that

$$\limsup_{n\to\infty} \mathbb{E}\left[f_n(Z_n, T_n + c); Z_n \notin \mathcal{L}_\varepsilon\right] = 0\,. \tag{22}$$

3.4 Controlling f_n for $x \in \mathcal{L}_\varepsilon$

It remains to deal with $\mathbb{E}\left[f_n(Z_n, T_n + c); Z_n \in \mathcal{L}_\varepsilon\right]$. We begin by writing (for any $t \in \mathbb{N}$)

$$\begin{aligned}
\mathbb{E}\left[f_n(Z_n, t); Z_n \in \mathcal{L}_\varepsilon\right] = \frac{1}{n-1} &\sum_{k=1}^{2^{L+b}-1} \sum_{r \geq 1} f_n\left(\frac{r}{n}, t\right) \mathbf{1}_{\left[|\frac{k}{2^{L+b}} - \frac{r}{n}| \leq \frac{\varepsilon}{2}\right]} \\
&+ \frac{1}{n-1} \sum_{r \geq 1} \left(f_n\left(\frac{r}{n}, t\right) + f_n\left(1 - \frac{r}{n}, t\right)\right) \mathbf{1}_{\left[\frac{r}{n} \leq \frac{\varepsilon}{2}\right]}\,,
\end{aligned} \tag{23}$$

where the last sum deals with the two end intervals in $\mathcal{L}_\varepsilon$. Consider first of all the double sum here. Since n is odd, the shortest possible distance between any point r/n and the lattice point $k/2^{L+b}$ is $1/(n2^{L+b})$; as f_n is non-negative we can therefore upper bound the inner sum by summing over a lattice of size $1/(n2^{L+b})$ centred around $k/2^{L+b}$ as follows:

$$\sum_{r \geq 1} f_n\left(\frac{r}{n}, t\right) \mathbf{1}_{\left[|\frac{k}{2^{L+b}} - \frac{r}{n}| \leq \frac{\varepsilon}{2}\right]} \leq \sum_{\substack{r=-\infty \\ r \neq 0}}^{\infty} f_n\left(\frac{k}{2^{L+b}} - \frac{r}{n 2^{L+b}}, t\right)\,. \tag{24}$$

Thus (23) can be bounded as follows:

$$\mathbb{E}\left[f_n(Z_n,t); Z_n \in \mathcal{L}_\varepsilon\right] \le \frac{1}{n-1}\sum_{k=1}^{2^{L+b}-1}\sum_{\substack{r=-\infty\\ r\neq 0}}^{\infty} f_n\left(\frac{k}{2^{L+b}} - \frac{r}{n2^{L+b}}, t\right) + \frac{2}{n-1}\sum_{r=1}^{\infty} f_n\left(\frac{r}{n}, t\right), \tag{25}$$

where we have used the symmetry of the functions f_n at either end of the interval $[0,1]$ to rewrite the expression for the end intervals. Now replace t by T_n+c, and consider the function f_n in the double sum above:

$$f_n\left(\frac{k}{2^{L+b}} - \frac{r}{n2^{L+b}}, T_n + c\right) = (n-1)\prod_{j=1}^{T_n+c}\phi_n\left(\frac{nk-r}{2^{L+b}}, j\right) \le (n-1)\phi_n\left(\frac{nk-r}{2^{L+b}}, T_n+c\right). \tag{26}$$

Here we have bounded the product by a single term, once again making use of the fact that ϕ_n takes values in $(0,1]$. Since $\phi_n(s,j)$ involves s and j only through the function $G_{\xi'}(\omega_{s,j})$, where $\omega_{s,j} = \exp(2\pi i 2^j s/n)$, we have (for sufficiently large n) that the bound in (26) is a function of

$$\exp\left(2\pi i\frac{2^{T_n+c}}{n}\left(\frac{nk-r}{2^{L+b}}\right)\right) = \exp\left(-2\pi i\frac{2^{T_n+c}r}{n2^{L+b}}\right) = \exp\left(-\frac{2^{1+c-(L+b)}\pi i r}{\sigma_{S'}}\right).$$

The second equality simply uses the definition of T_n, while the first results from shifting the argument of the exponential function by $2\pi i k 2^{T_n+c-(L+b)}$. (For large enough n this is an integer multiple of $2\pi i$, thanks to the finiteness of L and b and the assumption that $T_n \to \infty$.)

Writing

$$\theta_{nrc} = \frac{2^{1+c-(L+b)}\pi r}{\sigma_{S'}},$$

(where recall that $\sigma_{S'}$ depends on n) we therefore need to upper bound the function

$$\phi_n\left(\frac{nk-r}{2^{L+b}}, T_n+c\right) = \frac{p_n^2}{1-2(1-p_n)\mathrm{Re}\left(G_{\xi'}(e^{-i\theta_{nrc}})\right) + (1-p_n)^2|G_{\xi'}(e^{-i\theta_{nrc}})|^2}.$$

Now note that

$$G_{\xi'}(e^{-i\theta_{nrc}}) = \mathbb{E}\left[e^{-i\xi'\theta_{nrc}}\right] \quad \text{and thus} \quad \mathrm{Re}\left(G_{\xi'}(e^{-i\theta_{nrc}})\right) = \mathbb{E}\left[\cos\left(\xi'\theta_{nrc}\right)\right].$$

Similarly,

$$|G_{\xi'}(e^{-i\theta_{nrc}})|^2 = \mathbb{E}\left[\cos\left(\xi'\theta_{nrc}\right)\right]^2 + \mathbb{E}\left[\sin\left(\xi'\theta_{nrc}\right)\right]^2.$$

Since $p_n \to 0$ as $n\to\infty$, we see from (3) that $\sigma_{S'}\to\infty$ and so $\theta_{nrc}\to 0$. Using the Taylor expansions of cosine and sine the above can be approximated by

$$\mathbb{E}\left[\cos\left(\xi'\theta_{nrc}\right)\right] = 1 - \frac{(\mu^2+\sigma^2)\theta_{nrc}^2}{2} + O(\theta_{nrc}^4); \quad \mathbb{E}\left[\sin\left(\xi'\theta_{nrc}\right)\right] = \mu\theta_{nrc} + O(\theta_{nrc}^3).$$

Neglecting terms of $O(\theta_{nrc}^3)$ we arrive at

$$\phi_n\left(\frac{nk-r}{2^{L+b}}, T_n + c\right) \sim \frac{p_n^2}{1-(1-p_n)\left[2-(\mu^2+\sigma^2)\theta_{nrc}^2\right]+(1-p_n)^2\left[1-(\mu^2+\sigma^2)\theta_{nrc}^2+\mu^2\theta_{nrc}^2\right]}$$
$$= \frac{p_n^2}{p_n^2+(1-p_n)(\mu^2+\sigma^2 p_n)\theta_{nrc}^2} = \frac{1}{1+4^{1+c-(L+b)}\pi^2 r^2}\,.$$

We now combine this bound with that in (26) and insert into (25) (using an identical argument for the second sum there):

$$\limsup_{n\to\infty}\mathbb{E}\left[f_n(Z_n,t); Z_n\in\mathcal{L}_\varepsilon\right] \le \sum_{k=1}^{2^{L+b}-1}\sum_{\substack{r=-\infty\\ r\neq 0}}^{\infty}\frac{1}{1+4^{1+c-(L+b)}\pi^2 r^2} + 2\sum_{r=1}^{\infty}\frac{1}{1+4^{1+c-(L+b)}\pi^2 r^2}$$
$$= 2^{L+b}\left(2^{L+b-(1+c)}\coth(2^{L+b-(1+c)})-1\right)$$
$$\sim 2^{L+b}\,\frac{4^{L+b-(1+c)}}{3} \quad \text{as } c\to\infty, \tag{27}$$

where we have made use of the well-known identity (Apostol, 1974, p. 334)

$$\coth(x) = \frac{1}{x} + 2x\sum_{r=1}^{\infty}\frac{1}{x^2+\pi^2 r^2} \qquad (x>0)\,.$$

Combining (22) and (27) yields the required result

$$\limsup_{n\to\infty} U_n(c) = \limsup_{n\to\infty}\mathbb{E}\left[f_n(Z_n,T_n+c); Z_n\in\mathcal{L}_\varepsilon\right] \le O(4^{-c}) \quad \text{as } c\to\infty,$$

and thanks to the comment after (20), this completes the proof of part 2 of Theorem 3.

4 Concluding remarks

We have shown that the subsampled process Y exhibits a cutoff when the probability p_n of jumping takes the form $p_n = 1/(2n^\alpha)$, for a range of α which depends upon the mean of our step distribution ($\alpha\in(0,2)$ when $\mu=0$, and $\alpha\in(0,1)$ otherwise). Furthermore, our original chain of interest X exhibits a pre-cutoff, with mixing time $t_{\text{mix}}^{(n)}(\varepsilon) = \Omega(n^\alpha \ln n)$.

We have not yet said much about the mixing time of either process when α takes values on the boundary of these intervals, however. If $\alpha = 0$ then part 1 of Theorems 3 and 4 (which do not rely on p_n tending to zero) still hold; however, our argument for upper bounding the mixing time of Y (and hence of X) breaks down, since a sufficiently fine lattice $\mathcal{L}$ does not exist. (An upper bound of $O(\ln n \ln\ln n)$ can be obtained for the mixing time of X by employing the method of Chung et al. (1987).) On the other hand, if α takes the value at the upper boundary of the relevant interval then $n/\sigma_{S'} = O(1)$, and thus T_n is asymptotically independent of n: in this case it is relatively easy to show that Y mixes in constant time (and so no longer exhibits a cutoff), and that X has mixing time of $\Omega(n^\alpha)$.

It is of course possible to generalise the process considered in this paper in a number of ways. For example, changing the form of p_n to $1/(\beta n^\alpha)$ for some constant $\beta > 1$ has no effect on the cutoff result for Y. Similarly, changing the transitions of X so that jumps involve

multiplying by some (fixed) $k \geq 2$ (and considering only those n for which the resulting process still has a uniform equilibrium distribution) presumably has the effect of changing the base of the logarithm in the cutoff time T_n for Y from 2 to k; Theorem 4 should also still hold, with the factor of $\ln 2$ in T_n^L being replaced by $(\ln k)/(k-1)$. More interesting would be an analysis of a process X for which the multiplication factor is not deterministic, and for which the resulting subsampled chain Y does not have a distribution given by simple convolution; for example where jumps in X correspond to multiplication by a_t (again with probability $p_n \to 0$), with a_t being uniformly chosen from the set $\{2, (n+1)/2\}$.

Acknowledgements

Some of the ideas in this work arose during investigations into random walks on $\mathbb{Z}_n$ by Sam Wright, who was supported by Nuffield Science Undergraduate Research Bursary URB/40605. The authors would also like to express their gratitude to John Payne, whose numerical calculations for a particular instance of our process provided some useful early insights into the behaviour of the upper bound for Y in Section 3.

References

Apostol, T. M. (1974). *Mathematical analysis* (2nd ed.). Addison Wesley Publishing Company.

Chung, F., P. Diaconis, and R. Graham (1987). Random walks arising in random number generation. *The Annals of Probability 15*(3), 1148–1165.

Connor, S. B. (2010). Separation and coupling cutoffs for tuples of independent Markov processes. *Latin American Journal of Probability and Mathematical Statistics 7*, 65–77.

Diaconis, P. (1988). *Group representations in probability and statistics*, Volume 11 of *Lecture Notes - Monograph Series*. Institute of Mathematical Statistics.

Diaconis, P. (2011). The Mathematics of Mixing Things Up. *Journal of Statistical Physics 144*(3), 445–458.

Diaconis, P. and M. Shahshahani (1981). Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete 57*, 159–179.

Hildebrand, M. (1990). *Rates of convergence of some random processes on finite groups*. Ph. D. thesis, Dept. Mathematics, Harvard University.

Hildebrand, M. (1993). Random processes of the form $X_{n+1} = a_n X_n + b_n \ (mod\ p)$. *Annals of Probability 21*(2), 710–720.

Hildebrand, M. (1994a). Random walks supported on random points of $\mathbb{Z}/n\mathbb{Z}$. *Probab. Theory Related Fields 100*, 191–203.

Hildebrand, M. (1994b). *Some random processes related to affine random walks*. IMA Preprint Series, 1210.

Levin, D. A., Y. Peres, and E. L. Wilmer (2009). *Markov chains and mixing times*. American Mathematical Soc.

Mossel, E., Y. Peres, and A. Sinclair (2004). Shuffling by semi-random transpositions. *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science (FOCS'04)*, 572–581.

Saloff-Coste, L. (2004). Random walks on finite groups. In *Probability on discrete structures*, Volume 110 of *Encyclopaedia Math. Sci.*, pp. 263–346. Springer.

Smith, A. (2014). A Gibbs sampler on the n-simplex. *Ann. Appl. Probab. 24*(1), 114–130.

Subag, E. (2013). A lower bound for the mixing time of the random-to-random insertions shuffle. *Electronic Journal of Probability 18*, 1–20.

Terras, A. (1999). *Fourier Analysis on Finite Groups and Applications*. London Mathematical Society Student Texts. Cambridge University Press.

Department of Mathematics
University of York
York, YO10 5DD
UK