Electronic Journal of Probability

Electron. J. Probab. **29** (2024), article no. 58, 1–14.
ISSN: 1083-6489 https://doi.org/10.1214/24-EJP1110

Comparing limit profiles of reversible Markov chains*

Evita Nestoridi†

Abstract

We introduce a technique for comparing the limit profile behavior of two reversible, commuting Markov chains on the same space, that share the same stationary distribution. We apply this technique to prove that the limit profile of star transpositions at time $t = n \log n + cn$ is equal to $d_{\text{T.V.}}(\text{Poiss}(1+e^{-c}), \text{Poiss}(1))$ by comparing to the limit profile of random transpositions, as studied in [29]. We also provide examples of important commuting Markov chains, whose limit profile behavior is unknown, which could give new directions for research.

Keywords: Markov chain mixing; cutoff; limit profiles; star transpositions; comparison.
MSC2020 subject classifications: 60C05; 60J10.
Submitted to EJP on July 25, 2023, final version accepted on March 6, 2024.

1 Introduction

Cutoff, a phenomenon according to which a Markov chain converges abruptly to the stationary measure, has been a central question in Markov chain mixing (see [26] for a nice exposition on the history of cutoff and recent developements). Recently, there has been exciting progress towards answering the even sharper question of determining the limit profile of various Markov chains. The limit profile captures the exact shape of the distance of the Markov chain from stationarity. In recent progress, Teyssier [29] derived an exact formula for the limit profile of random transpositions, improving the seminal result of Diaconis and Shahshahani [11] and proving a conjecture of Matthews [21]. He extended the Fourier transform arguments of [11] to use representation theory in order to study limit profiles of conjugacy class walks. In [22], Teyssier's representation theory technique is generalized to study the limit profile of any general reversible Markov chain using its entire spectrum.

This paper focuses on determining the limit profile of a reversible Markov chain by comparing to another Markov chain on the same configuration space whose limit profile is known. The goal of this paper is to prove that the star transposition shuffle exhibits the same limit profile behavior as random transpositions. The assumptions needed for

*Supported by NSF Grant DMS-2052659.
†Stony Brook University, United States. E-mail: evrydiki.nestoridi@stonybrook.edu

this new theory are that the two Markov chains are simultaneously diagonalizable, and that they both exhibit ℓ_2 and total variation distance cutoff. The technique introduced in the present paper differs from the traditional comparison theory developed by Diaconis and Saloff-Coste ([7], [8]); it compares total variation distances via purely algebraic techniques, while the Diaconis and Saloff-Coste technique compares ℓ^2 norms via path techniques.

We now give a general introduction on mixing times, cutoff and limit profiles. Let X be a finite state space with $|X| = n$, and let P be the transition matrix of an aperiodic and irreducible Markov chain. In other words, the entry $P^t(x, y)$ is the probability of the walk starting at x and being at y after t steps, for every $t \in \mathbb{N}$. The measure $P_x^t(\cdot) = P^t(x, \cdot)$ converges to a unique measure $\pi(\cdot)$ on X as t goes to infinity. We study this convergence with respect to total variation distance, which is defined as

$$d_x(t) = \|P_x^t - \pi\|_{\text{T.V.}} := \frac{1}{2} \sum_{y \in X} |P_x^t(y) - \pi(y)|.$$

We set

$$d(t) = \max_{x \in X} \{d_x(t)\}.$$

Definition 1.1. *The mixing time with respect to the total variation distance is defined as*

$$t_{\text{mix}}(\varepsilon) := \min\{t : d(t) \leq \varepsilon\},$$

for every $\varepsilon \in (0, 1)$.

Cutoff describes a phase transition: as we run the family of Markov chains, the total variation distance is almost equal to 1, and then suddenly it drops and approaches zero as n grows. We now give the formal definition of cutoff.

Definition 1.2. *A family of Markov chains is said to have cutoff at time* t_n *with window* $w_n = o(t_n)$ *if and only if*

$$\lim_{c \to \infty} \lim_{n \to \infty} d^{(n)}(t_n - cw_n) = 1 \text{ and } \lim_{c \to \infty} \lim_{n \to \infty} d^{(n)}(t_n + cw_n) = 0,$$

where $d^{(n)}(t)$ *denotes the total variation distance of the* n*–th Markov chain.*

Given a Markov chain exhibiting cutoff, one can ask for more precise control on the exact distance from stationarity. This is known as the limit profile, defined as:

$$\Phi_x(c) := \lim_{n \to \infty} d_x^{(n)} (t_n + cw_n), \text{ for all } c \in \mathbb{R}.$$

If this limit does not exist, similar definitions apply for the $\limsup$ and the $\liminf$.

The limit profile is known for only a few Markov chains, such as the riffle shuffle [1], the asymmetric exclusion process on the segment [3], the simple exclusion process on the cycle [18], and the simple random walk on Ramanujan graphs [20], etc. Teyssier [29] determined the limit profile for random transpositions. Using representation theory of the symmetric group S_n, he used Fourier transform arguments for studying limit profiles that work for random walks on groups using a generating set that is a conjugacy class. In particular, he proved that for random transpositions, if $t = \frac{1}{2} n \log n + cn$, then

$$\Phi_x(c) = \|\text{Poiss}(1 + e^{-2c}) - \text{Poiss}(1)\|_{\text{T.V.}},$$

for every $x \in S_n$ and $c \in \mathbb{R}$. In [22], this limit profile behavior is proven to hold for the k-cycle card shuffle, under the assumption that $k = o(n/\log n)$.

Our main application studies the limit profile of the star transpositions shuffle, which is not generated by a conjugacy class. One step of the star transpositions shuffle consists of picking a card of the deck uniformly at random and transposing it with the top card. Flatto, Odlyzko and Wales [14] found the eigenvalues of the transition matrix using representation theory of the symmetric group. Diaconis analysed the eigenvalue behavior in [4] to prove that star transpositions exhibit cutoff at $n \log n$ with window of order n. The following theorem discusses the limit profile of star transpositions.

Theorem 1.3. *For the star transpositions card shuffle at time $t = n(\log n + c)$, we have that*

$$\Phi_x(c) = d_{\text{T.V.}}(\text{Poiss}(1 + e^{-c}), \text{Poiss}(1)),$$

for every $x \in S_n$ and $c \in \mathbb{R}$.

The proof of Theorem 1.3 requires a new idea, since Teyssier's technique works for conjugacy invariant random walks and the more general technique analyzed in [22] assumes the knowledge of certain eigenfunctions. The main ingredient of the proof of Theorem 1.3 is that star transpositions and random transpositions are simultaneously diagonalizable. This allows us to only use the eigenvalues, found in [14], and not the eigenfunctions of star transpositions. We also need to use the eigenvalues [11] and the limit profile [29] of random transpositions.

To prove Theorem 1.3, we develop a technique that allows us to compare the limit profile of the shuffle in question to the limit profile of the random transpositions shuffle. The following main lemma of this paper allows us to compare limit profiles of reversible Markov chains on the same space X, with the same stationary measure.

Lemma 1.4. *Let P and Q be the transition matrices of two reversible Markov chains on a finite space X that share the same eigenbasis and stationary measure π. In particular, we denote by $f_i : X \to \mathbb{C}$ the orthonormal eigenvectors satisfying*

$$Pf_i = \beta_i f_i \text{ and } Qf_i = q_i f_i,$$

where $i = 1, \dots, |X|$ and $\beta_1 = q_1 = 1$.

1. *We have*

$$4\|P_x^t - Q_x^{t_*}\|_{T.V.}^2 \leq \sum_{i=2}^{|X|} f_i(x)^2 (\beta_i^t - q_i^{t_*})^2, \tag{1.1}$$

for every $x \in X, t, t_* \geq 0$. If P, Q are the transition matrices of transitive Markov chains then

$$4\|P_x^t - Q_x^{t_*}\|_{T.V.}^2 \leq \sum_{i=2}^{|X|} (\beta_i^t - q_i^{t_*})^2, \tag{1.2}$$

for every $x \in X, t, t_* \geq 0$.

2. *Assume P exhibits cutoff at t_n with window w_n with limit profile Φ_x and Q exhibits cutoff at $\bar{t}_n$ with window $\overline{w}_n$ with limit profile $\overline{\Phi}_x$. For $t = t_n + cw_n$ and $t_* = \bar{t}_n + c\overline{w}_n$, we have*

$$|\overline{\Phi}_x(c) - \Phi_x(c)| \leq \frac{1}{2} \lim_{n \to \infty} \left(\sum_{i=2}^{|X|} f_i(x)^2 (\beta_i^t - q_i^{t_*})^2 \right)^{1/2}, \tag{1.3}$$

for every $c \in \mathbb{R}$, $x \in X$. If P, Q are furthermore the transition matrices of transitive Markov chains then

$$|\overline{\Phi}_x(c) - \Phi_x(c)| \leq \frac{1}{2} \lim_{n \to \infty} \left(\sum_{i=2}^{|X|} (\beta_i^t - q_i^{t_*})^2 \right)^{1/2}, \tag{1.4}$$

for every $c \in \mathbb{R}$, $x \in X$.

The assumptions of Lemma 1.4 are equivalent to saying that P and Q commute, i.e. $PQ = QP$. Even though the assumptions of Lemma 1.4 sound restrictive, it is actually not rare that two reversible random walks on the same group are simultaneously diagonalizable. As explained in Section 4, it is a standard fact that random transpositions commutes with any other symmetric random walk on S_n. In general, the transition matrix of a symmetric, conjugacy invariant random walk on a group G commutes with any other symmetric transition matrix of a random walk on G (Laurent Saloff-Coste, personnal communication).

Lemma 1.4 is proven in Section 3. The proof of Theorem 1.3 is contained in Section 5. Section 4 summarizes the main tools that we need from representation theory and random transpositions. In Section 6, we discuss possible conjectures and the difficulties that occur when trying to apply Lemma 1.4 for the case of random-to-random.

2 Examples of commuting Markov chains

This section is dedicated to presenting examples of Markov chains. As Lemma 4.8 indicates, the transition matrix of a conjugacy class invariant, symmetric random walk on a group G always commutes with the transition matrix of a symmetric random walk on G. This means that in the following examples one can use Fourier analysis (just like in [29]) to get the limit profile of the conjugacy invariant random walk, and then apply Lemma 1.4 to get the limit profile of the other random walk.

2.1 The random walk on $SL_n(\mathbb{F}_q)$

Consider $G = SL_n(\mathbb{F}_q)$, the group of $n \times n$ matrices with entries on $\mathbb{F}_q$ and determinant equal to one. Hildebrand [15] proved that the random transvection random walk on G exhibits cutoff at time n with contant window. This is the random walk on the Caley graph of G with respect to the conjugacy class of $\{I_n + aE_{i,j}, a \in \mathbb{F}_q, 1 \leq i \neq j \leq n\}$, where I_n is the identity matrix and $E_{i,j}$ is the matrix whose (i, j) entry is one and the rest are zero.

Saloff-Coste and Diaconis [9] and Kassabov [17] studied the random walk X_t on G which corresponds to walking on the Caley graph of G with respect to $\{I_n + aE_{i,j}, a \in \mathbb{F}_q, 1 \leq i \neq j \leq n\}$. In other words, at time $t+1$, we pick two rows i, j of X_t and an $a \in \mathbb{F}_q$ uniformly at random. Then we add the the i–th row of X_t multiplied by a to the j–th row. This gives us X_{t+1}.

Lemma 4.8 says that the two random walks commute and therefore Lemma 1.4 can be applied, especially since the spectrum of random transvections is fully understood. Of course, determining the limit profile of random transvections and proving cutoff for the second random walk and understanding its spectrum are still open.

2.2 Kac's walk

Kac's walk is a random walk on the orthogonal group $SO(n)$. Let

$$g_{i,j}(\theta) = \begin{pmatrix} 1 & 0 & & \dots & & 0 \\ 0 & \ddots & & & & \\ & & \cos\theta & \dots & \sin\theta & \\ \vdots & & \vdots & & \vdots & \vdots \\ & & -\sin\theta & \dots & \cos\theta & \\ & & & & \ddots & 0 \\ 0 & & & \dots & 0 & 1 \end{pmatrix}$$

be the rotation by θ in the i,j plane of $\mathbb{R}^n$. Diaconis and Saloff-Coste [10] studied the random walk on $SO(n)$ generated by multiplying by $g_{i,j}(\theta)$ where i,j are chosen uniformly at random in $[n]$ and θ is a uniformly random angle in $[0,2\pi)$. Pillai and Smith [23] prove that the rate of convergence is of order at most $n^4 \log n$ and we refer to their paper for the full history of the problem.

Rosenthal [25], on the other hand, considered the random walk on $SO(n)$ generated by the conjugacy class of $g_{1,2}(\pi)$ and proved that it exhibits cutoff at $\frac{1}{4} n \log n$ with window of order n. The two walks commute and it would be interesting to compare their limit profile behavior. Hough and Jiang [16] also proved cutoff for the conjugacy invariant random walk with respect to the conjugacy class of $\{g_{1,2}(\theta) : \theta \in [0,2\pi)\}$ and this is another candidate for comparison.

2.3 Bernoulli–Laplace urn model

In the Bernoulli–Laplace model, there are two urns each one containing n balls. At time zero, the first urn contains n red balls and the second one contains n white balls. At time t, we pick one ball from each urn and swap them. Diaconis and Shahshahani [12] proved that the Bernoulli–Laplace Markov chain exhibits cutoff at $\frac{1}{4} n \log n$ with window of order n. They also prove that the orthonormal eigenbasis is given by the dual Hahn polynomials.

The dual Hahn polynomials are an eigenbasis for the generalized Bernoulli–Laplace model where k balls are picked from each urn without replacement and get swapped. This way, we get a family of commuting transition matrices. Eskenazis and Nestoridi [13] proved that for $k = o(n)$ the Bernoulli–Laplace exhibits cutoff at $\frac{n}{4k} \log n$ and window $\frac{n}{k} \log \log n$. It is very likely that the limit profile for $k = o(n)$ is the same as for the case $k = 1$.

3 The proof of Lemma 1.4

Proof of Lemma 1.4. Since P and Q are both reversible Markov chains with respect to the same stationary measure π and share the same orthonormal eigenbasis f_i, we can write

$$\frac{P_x^t(y)}{\pi(y)} = 1 + \sum_{i=2}^{|X|} f_i(x) f_i(y) \beta_i^t$$

and similarly

$$\frac{Q_x^t(y)}{\pi(y)} = 1 + \sum_{j=2}^{|X|} f_i(x) f_i(y) q_i^t,$$

as explained in Lemma 12.2 of [19]. Therefore,

$$\frac{P_x^t(\cdot) - Q_x^{t_*}(\cdot)}{\pi(\cdot)} = 1 + \sum_{i=2}^{|X|} f_i(x) f_i(\cdot) \left(\beta_i^t - q_i^{t_*} \right),$$

This way, we have expressed $\frac{P_x^t - Q_x^{t_*}}{\pi}$ as a linear combination of the elements of the orthonormal eigenbasis $\{f_i\}$. Therefore the ℓ_2 norm can be written as

$$\left\| \frac{P_x^t - Q_x^{t_*}}{\pi} \right\|_2^2 = \sum_{i=2}^{|X|} f_i(x)^2 (\beta_i^t - q_i^{t_*})^2, \tag{3.1}$$

where the ℓ^2 norm is defined as

$$\|g\|_2^2 = \sum_{x \in X} g^2(x) \pi(x).$$

Equation (3.1) and Cauchy-Schwartz give (1.1).

Note that if P and Q are transitive, then $d_x(t) = d(t)$ for every $x \in X$ and π is the uniform measure on X. Summing over $x \in X$ on both sides of (1.1), we get

$$4|X|\|P_x^t - Q_x^{t_*}\|_{T.V.}^2 \leq \sum_{x \in X} \sum_{i=2}^{|X|} f_i(x)^2 (\beta_i^t - q_i^{t_*})^2.$$

The fact that f is normal means that $\sum_{x \in X} f_i(x)^2 = |X|$, which finishes the proof of (1.2). This concludes part 1 of the lemma. Part 2 follows by taking limits on both sides of equations (1.1) and (1.2). $\square$

4 Representation theory of the symmetric group and random transpositions

In this section, we summarize features of the random transposition shuffle that are important for the analysis of the limit profile of star transpositions. The transition matrix of random transpositions is given by

$$Q(x,y) = \begin{cases} \frac{1}{n} & \text{if } y = x, \\ \frac{2}{n^2} & \text{if } y = xs, \text{ where } s \text{ is a transposition}, \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to check that Q is a symmetric matrix. Diaconis and Shahshahani [11] found all the eigenvalues of Q and their multiplicities. Their formulas are in terms of standard Young tableaux, which we now introduce.

By a *partition* $\lambda = (\lambda_1, \ldots, \lambda_k)$ of n we mean a sequence of natural numbers such that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k > 0$ and $\sum_{i=1}^k \lambda_i = n$. Every partition corresponds to a *Young diagram*, which has λ_i boxes in row i. We will also consider the transpose partition λ' of λ to be the partition occuring by transposing the Young diagram of λ.

Example 4.1. We can think of $\lambda = (3,2)$ as $\lambda = \begin{array}{|c|c|c|}\hline & & \\ \hline & \\ \cline{1-2}\end{array}$. In this case, $\lambda' = (2,2,1) = \begin{array}{|c|c|}\hline & \\ \hline & \\ \hline \\ \cline{1-1}\end{array}$.

To describe the eigenvalues of many shuffles and their multiplicities, we will need the notion of *standard Young tableaux* (SYT) of type λ, which is a filling of the Young diagram λ with all the numbers in $[n]$, so that the entries of each row and column of the diagram appear in increasing order.

Example 4.2. One SYT of $\lambda = (3,2)$ is $T_\lambda = \begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 & 5 \\ \cline{1-2}\end{array}$.

The following notion will be important when counting the multiplicities of the eigenvalues of Q.

Definition 4.3. *Let λ be a* partition *of n. We define d_λ to be the number of SYT of λ.*

The following lemma provides a useful bound on d_λ.

Lemma 4.4 (Corollary 2 of [11])**.** *Let λ be a partition of n. Let $j = n - \lambda_1$, then $d_\lambda \leq \binom{n}{j}\sqrt{j!}$.*

The following lemma presents the eigenvalues of random transpositions.

Lemma 4.5 (Lemma 7 of [11])**.** *Let $\lambda = (\lambda_1, \dots, \lambda_k)$ be a partition of n. The eigenvalues of Q are given by*

$$s_\lambda = \frac{1}{n} + \frac{n-1}{n} r_\lambda,$$

where

$$r_\lambda = \frac{1}{\binom{n}{2}} \sum_{i=1}^{k} \left[\binom{\lambda_i}{2} - \binom{\lambda_i'}{2} \right].$$

Each eigenvalue s_λ occurs with multiplicity d_λ^2.

The formula for the s_λ gives the following standard fact which is used in both [11] and [29].

Corollary 4.6. *Let λ be a partition of n. If $j = n - \lambda_1$ is constant, then*

$$s_\lambda = 1 - \frac{2j}{n} + O\left(\frac{1}{n^2}\right).$$

Similarly, if $j = n - \lambda_1'$ is constant, then

$$s_\lambda = -1 + \frac{2(j+1)}{n} + O\left(\frac{1}{n^2}\right).$$

Finally, the following lemma discusses the limit profile of random transpositions.

Lemma 4.7 (Teyssier [29])**.** *For random transpositions, we have that at $t = \frac{1}{2}n(\log n + c)$*

$$\overline{\Phi}(c) = d_{T.V.}(\text{Poiss}(1 + e^{-c}), \text{Poiss}(1)),$$

where $c \in \mathbb{R}$.

The following lemma will allow us to compare the limit profile of star transpositions to the limit profile of random transpositions. As pointed out by Laurent Saloff-Coste, it is a standard fact that the transition matrix of a symmetric random walk on a group G generated by a union of conjugacy classes commutes with the transition matrix of any symmetric random walk on G. For completeness, we include the proof of this statement for the case of Q.

Lemma 4.8. *Let P be the transition matrix of a symmetric random walk on S_n. The transition matrix Q of random transpositions commutes with P.*

Proof. Let μ be the probability measure on S_n, such that $P(x, w) = \mu(x^{-1}w)$ and define ν to be the probability measure on S_n, such that $Q(x, w) = \nu(x^{-1}w)$.

Let $x, y \in S_n$. We write

$$(PQ)(x, y) = \sum_{w \in S_n} P(x, w) Q(w, y).$$

Let S' be the set of all transpositions and let $S = S' \cup \{\text{id}\}$. We have that

$$(PQ)(x, y) = \sum_{s \in S} P(x, ys) Q(ys, y) = \sum_{s \in S} \mu(x^{-1}ys)\nu(s^{-1}).$$

We want to use the fact that ν is constant on all transpositions, which form a conjugacy class. Because of this, we write

$$(PQ)(x, y) = \sum_{s \in S} \mu((x^{-1}y)s(y^{-1}x)(x^{-1}y))\nu(s^{-1}).$$

Let $\overline{s} = (x^{-1}y)s(y^{-1}x) \in S$ and recall that $\nu(s) = \nu(s^{-1}) = \nu(\overline{s}) = \nu(\overline{s}^{-1})$. Therefore,

$$(PQ)(x, y) = \sum_{\overline{s} \in S} \mu(\overline{s}x^{-1}y)\nu(\overline{s}^{-1}) = \sum_{\overline{s} \in S} P(x\overline{s}^{-1}, y) Q(x, x\overline{s}^{-1}) = (QP)(x, y),$$

which finishes the proof. $\square$

5 Star transpositions

The goal of this section is to prove Theorem 1.3 using Lemma 1.4. In the following lemma, we recall the eigenvalues of the transition matrix P of star transpositions.

Lemma 5.1 (Theorem 3.7 of [14])**.** *Let $\lambda = (\lambda_1, \ldots, \lambda_k)$ be a partition of n. Let $\lambda^{(i)}$ be a partition of $n-1$ occurring by deleting the last box in the i-th row of λ (if this results to a partition). The eigenvalues of P corresponding to λ are given by $\overline{s}_{\lambda^{(i)}} = \frac{1}{n}(\lambda_i - i + 1)$, and they occur with multiplicity $d_\lambda d_{\lambda^{(i)}}$.*

Lemma 5.2. *Let $\lambda = (\lambda_1, \ldots, \lambda_k)$ be a partition of n and let i be a row of λ whose last box can be removed to give a partition of $n-1$. The eigenvectors of P corresponding to $\overline{s}_{\lambda^{(i)}}$ are eigenvectors for Q corresponding to s_λ.*

Proof. There are many ways to prove that there is an eigenvector for P corresponding to the eigenvalue $\overline{s}_{\lambda^{(i)}}$ that is also an eigenvector for Q corresponding to the eigenvalue s_λ. Theorem 6 of Chapter 3 of [5] says that

$$Q = \phi^* \Delta \phi \text{ and } P = \phi^* D \phi,$$

where ϕ depends only on S_n. To define Δ and D, we need to consider the Fourier transforms $\hat{Q}(\rho_i) = \sum_{x \in S_n} Q(id, x)\rho_i(x)$ and $\hat{P}(\rho_i) = \sum_{x \in S_n} P(id, x)\rho_i(x)$ at an irreducible representation $\rho_i : S_n \to GL_{d_{\rho_i}}(\mathbb{C})$. Then,

$$\Delta = \begin{pmatrix} M_Q(\rho_1) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & M_Q(\rho_k) \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} M_P(\rho_1) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & M_P(\rho_k) \end{pmatrix},$$

where $\{\rho_i : S_n \to GL_{d_{\rho_i}}(\mathbb{C})\}$ are all the irreducible representations of S_n, and $M_Q(\rho_i)$, $M_P(\rho_i) \in \mathbb{C}^{d_{\rho_i} \times d_{\rho_i}}$ are defined as

$$M_Q(\rho_i) = \begin{pmatrix} \hat{Q}(\rho_i) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \hat{Q}(\rho_i) \end{pmatrix} \quad \text{and} \quad M_P(\rho_i) = \begin{pmatrix} \hat{P}(\rho_i) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \hat{P}(\rho_i) \end{pmatrix}.$$

Diaconis and Shahshahani [11] proved that $\hat{Q}(\rho_\lambda) = s_\lambda \mathbf{I}_{d_\lambda}$ and

$$\hat{P}(\rho_\lambda) = \begin{pmatrix} \overline{s}_{\lambda^{(i_1)}} \mathbf{I}_{d_{\lambda^{(i_1)}}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \overline{s}_{\lambda^{(i_\ell)}} \mathbf{I}_{d_{\lambda^{(i_\ell)}}} \end{pmatrix},$$

where $\{i_1, \ldots, i_\ell\}$ are the rows of λ whose last box can be removed to give a valid partition of $n-1$.

To sum up, the suitable column of ϕ^* gives the common eigenvector of P and Q corresponding to the eigenvalues $\overline{s}_{\lambda^{(i)}}$ and s_λ. ☐

To prove Theorem 1.3 we will need the following lemma.

Lemma 5.3. *Let λ be a partition of n. Let $i > 1$ and denote $j = n - \lambda_1$. Then*

$$d_{\lambda^{(i)}} \leq \frac{4^j}{n} d_\lambda \quad \textit{and} \quad -\frac{j}{n} \leq \overline{s}_{\lambda^{(i)}} \leq \frac{n-j}{n}.$$

Proof. In the diagram associated to λ, the hook $h_{i,j}$ of the (i,j) box is the number of boxes which are below or on the right of our box (including our box). We call $\text{équ}(\lambda)$ the product of the hooks of the partition λ. It is a standard fact that

$$d_\lambda = \frac{n!}{\text{équ}(\lambda)}.$$

For example, the entry of each box in $\lambda = \begin{array}{|c|c|c|}\hline 4 & 3 & 1 \\ \hline 2 & 1 \\ \cline{1-2} \end{array}$ gives the corresponding hook. Then $d_\lambda = 5$.

For $\lambda^{(i)}$ to be a valid partition, we have that $h_{i,\lambda_i} = 1$. We have that

$$\frac{d_{\lambda^{(i)}}}{d_\lambda} = \frac{1}{n} \prod_{\ell=1}^{i-1} \frac{h_{\ell,\lambda_i}}{h_{\ell,\lambda_i} - 1} \prod_{k=1}^{\lambda_i - 1} \frac{h_{i,k}}{h_{i,k} - 1}.$$

Notice that each $h_{a,b}$ appearing in the above products has to be at least equal to 2, since the (i,λ_i) box and the (a,b) box contribute to $h_{a,b}$. Therefore, each ratio $\frac{h_{a,b}}{h_{a,b}-1}$ is at most 2, which gives that

$$d_{\lambda^{(i)}} \leq \frac{2^{i+\lambda_i-2}}{n} d_\lambda.$$

Using the fact that $i + \lambda_i \leq 2j + 2$ for every $i > 1$ finishes the proof of the first statement.

The final claim holds because $i \leq j+1$ and therefore

$$-\frac{j}{n} \leq \overline{s}_{\lambda^{(i)}} = \frac{1}{n}(\lambda_i - i + 1) \leq \frac{\lambda_1}{n} = \frac{n-j}{n},$$

for every $i > 1$. $\square$

The following standard fact is a consequence of the branching rule for the irreducible representations of S_n. We refer to Theorem 3.6 of [14] for the exact statement of the branching rule.

Lemma 5.4. *Let λ be a partition of n. Then $d_\lambda = \sum_i d_{\lambda^{(i)}}$ and $d_{\lambda^{(i)}} = d_{\lambda'^{(\lambda_i)}}$.*

We are now ready to prove Theorem 1.3.

Proof of Theorem 1.3. Lemma 4.8 and Lemma 5.2 allow us to use Lemma 1.4. Therefore, the goal is to prove that

$$\lim_{n\to\infty} \sum_\lambda d_\lambda \sum_i d_{\lambda^{(i)}} (s_\lambda^t - \overline{s}_{\lambda^{(i)}}^{t_*})^2 = 0,$$

where s_λ and $\overline{s}_{\lambda^{(i)}}$ are the eigenvalues of random transpositions and star transpositions respectively. Recalling Lemmas 4.5 and 5.1, we have

$$s_\lambda = \frac{1}{n} + \frac{n-1}{n} r_\lambda \text{ and } \overline{s}_{\lambda^{(i)}} = \frac{1}{n} + \frac{n-1}{n} \overline{r}_{\lambda^{(i)}},$$

where

$$r_\lambda = \frac{1}{\binom{n}{2}} \sum_{i=1}^{k} \left[\binom{\lambda_i}{2} - \binom{\lambda_i'}{2} \right] \text{ and } \overline{r}_{\lambda^{(i)}} = \frac{\lambda_i - i}{n-1}.$$

Note that

$$r_{\lambda'} = -r_\lambda \text{ and } \overline{r}_{(\lambda')^{(\lambda_i)}} = -\overline{r}_{\lambda^{(i)}}, \tag{5.1}$$

where λ' is the transpose diagram of λ.

Our strategy is to prove that for every $\varepsilon > 0$ and $c \in \mathbb{R}$, there is an $M = M(c,\varepsilon)$ such that

1. $\sum_{\lambda_1 \le n-M} d_\lambda^2 |s_\lambda|^{2t} \le \varepsilon$,
2. $\sum_{\substack{\lambda_1 \le n-M \\ \lambda_1' \le n-M}} d_\lambda \sum_{i \ge 1} d_{\lambda^{(i)}} |\overline{s}_{\lambda^{(i)}}|^{2t_*} \le \varepsilon$,
3. $\sum_{\substack{\lambda_1 \le n-M \\ \lambda_1' \le n-M}} d_\lambda |s_\lambda|^t \sum_{i \ge 1} d_{\lambda^{(i)}} |\overline{s}_{\lambda^{(i)}}|^{t_*} \le \varepsilon$, and
4. $\sum_{\lambda_1 > n-M} d_\lambda \sum_i d_{\lambda^{(i)}} |s_\lambda^t - \overline{s}_{\lambda^{(i)}}^{t_*}|^2 \le \varepsilon$, and $\sum_{\lambda_1' > n-M} d_\lambda \sum_i d_{\lambda^{(i)}} |s_\lambda^t - \overline{s}_{\lambda^{(i)}}^{t_*}|^2 \le \varepsilon$

for sufficiently large n.

Lemma 4.1 of Teyssier [29] states that for every $\varepsilon \in (0,1)$ and $c \in \mathbb{R}$, there is an $M_1 = M_1(c, \varepsilon)$ such that at $t = \frac{1}{2} n(\log n + c)$ we have that

$$\sum_{\lambda_1 \le n-M_1} d_\lambda |s_\lambda|^t \le \varepsilon,$$

where n is sufficiently large. This implies that

$$\sum_{\lambda_1 \le n-M_1} d_\lambda^2 |s_\lambda|^{2t} \le \varepsilon,$$

where n is sufficiently large. This gives 1.

Now, we want to prove 2. Let $M_2 = M_2(c, \varepsilon)$ be such that $\sum_{j \ge M_2} \frac{e^{-2cj}}{j!} \le \varepsilon/2$. Equation (5.1) implies that if $\overline{r}_{\lambda^{(i)}} \ge 0$ then $\overline{s}_{\lambda^{(i)}} = |\overline{s}_{\lambda^{(i)}}| \ge |\overline{s}_{\lambda'^{(\lambda_i)}}|$. Therefore, it suffices to consider only the cases where $\overline{r}_{\lambda^{(i)}} \ge 0$, which implies that $\overline{s}_{\lambda^{(i)}} \ge 0$. Lemmas 4.4, 5.3 and 5.4 give

$$\sum_{\substack{\lambda_1 \le n-M_2 \\ \lambda_1' \le n-M_2}} d_\lambda \sum_{\substack{i \ge 1 \\ \overline{s}_{\lambda^{(i)}} \ge 0}} d_{\lambda^{(i)}} |\overline{s}_{\lambda^{(i)}}|^{2t_*} \le \sum_{j \ge M_2} d_\lambda^2 \left(1 - \frac{j}{n}\right)^{2t_*} \le \sum_{j \ge M_2} \frac{n^{2j}}{j!} e^{-\frac{2t_* j}{n}}.$$

For $t = n \log n + cn$, we have

$$\sum_{\substack{\lambda_1 \le n-M_2 \\ \lambda_1' \le n-M_2}} d_\lambda \sum_{\substack{i \ge 1 \\ \overline{s}_{\lambda^{(i)}} \ge 0}} d_{\lambda^{(i)}} |\overline{s}_{\lambda^{(i)}}|^{2t_*} \le \sum_{j \ge M_2} \frac{e^{-2cj}}{j!} \le \varepsilon/2.$$

This gives 2.

We now focus on 3. Let $M = \max\{M_1, M_2\}$. Just like in 2, it suffices to consider only the partitions λ where $\overline{r}_{\lambda^{(i)}} \ge 0$. In combination with Lemma 5.3, this implies that $0 \le \overline{s}_{\lambda^{(i)}} \le \overline{s}_{\lambda^{(1)}} = 1 - \frac{j}{n}$, where $j = n - \lambda_1$. Therefore, Lemma 5.4 gives that

$$\sum_{\lambda_1 \le n-M} d_\lambda |s_\lambda|^t \sum_{\substack{i \ge 1 \\ \overline{s}_{\lambda^{(i)}} \ge 0}} d_{\lambda^{(i)}} |\overline{s}_{\lambda^{(i)}}|^{t_*} \le \sum_{\lambda_1 \le n-M} d_\lambda^2 |s_\lambda|^t \left(1 - \frac{j}{n}\right)^{t_*}.$$

Applying Cauchy-Schwartz on the right hand side gives

$$\left(\sum_{\lambda_1 \le n-M} d_\lambda |s_\lambda|^t \sum_{\substack{i \ge 1 \\ \overline{s}_{\lambda^{(i)}} \ge 0}} d_{\lambda^{(i)}} |\overline{s}_{\lambda^{(i)}}|^{t_*} \right)^2 \le \sum_{\lambda_1 \le n-M} d_\lambda^2 |s_\lambda|^{2t} \sum_{\lambda_1 \le n-M} d_\lambda^2 \left(1 - \frac{j}{n}\right)^{2t_*}.$$

We set $M = \max\{M_1, M_2\}$ so that both $\sum_{\lambda_1 \le n-M_3} d_\lambda^2 |s_\lambda|^{2t}$ and $\sum_{\lambda_1 \le n-M_3} d_\lambda^2 \left(1 - \frac{j}{n}\right)^{2t_*}$ are less than ε as we saw in 1 and 2. This gives 3.

We now focus on proving 4. We start by writing

$$\sum_{\lambda_1 > n-M} d_\lambda \sum_i d_{\lambda^{(i)}} |s_\lambda^t - \overline{s}_{\lambda^{(i)}}^{t_*}|^2 = \sum_{\lambda_1 > n-M} d_\lambda \sum_{i > 1} d_{\lambda^{(i)}} |s_\lambda^t - \overline{s}_{\lambda^{(i)}}^{t_*}|^2 + \sum_{\lambda_1 > n-M} d_\lambda d_{\lambda^{(1)}} |s_\lambda^t - \overline{s}_{\lambda^{(1)}}^{t_*}|^2.$$

Notice that for $j = n - \lambda_1$, with j constant Lemma 5.3 gives that $|\overline{s}_{\lambda^{(i)}}| \leq \max\{\frac{j}{n}, 1 - \frac{j}{n}\} = 1 - \frac{j}{n}$, for sufficiently large n. Therefore, Lemma 5.3 and setting $j = n - \lambda_1$ gives

$$\begin{aligned}\sum_{\lambda_1 > n-M} d_\lambda \sum_{i>1} d_{\lambda^{(i)}} |s_\lambda^t - \overline{s}_{\lambda^{(i)}}^{t_*}|^2 &\leq \frac{4^M M}{n} \sum_{\substack{j<M\\ \lambda:\lambda_1 = n-j}} d_\lambda^2 \left(|s_\lambda^t| + \left(1 - \frac{j}{n}\right)^{t_*} \right)^2 \\ &\leq \frac{4^M M}{n} \sum_{\substack{j<M\\ \lambda:\lambda_1 = n-j}} \frac{n^{2j}}{j!} \left(|s_\lambda^t| + \left(1 - \frac{j}{n}\right)^{t_*} \right)^2, \qquad (5.2)\end{aligned}$$

for sufficient large n. The last inequality occurred from applying the inequality for Lemma (4.4) for d_λ. Using Corollary 4.6, and the facts that $t = \frac{1}{2} n(\log n + c)$ and $t_* = n(\log n + c)$, we have that

$$\left(1 - \frac{j}{n}\right)^{t_*} \leq \frac{e^{-cj}}{n^j} \text{ and } s_\lambda^t = \left(1 - \frac{2j}{n} + O\left(\frac{1}{n^2}\right)\right)^t \leq 3\frac{e^{-cj}}{n^j}$$

for sufficiently large n. Therefore, we have that

$$(5.2) \leq \frac{4^{M+2} M^2}{n} M! \sum_{j<M} \frac{e^{2cj}}{j!} \leq \varepsilon/2, \qquad (5.3)$$

for sufficient large n.

Finally, Corollary 4.6 and Lemma 5.4 give that

$$\begin{aligned}\sum_{\lambda_1 > n-M} d_\lambda d_{\lambda^{(1)}} |s_\lambda^t - \overline{s}_{\lambda^{(1)}}^{t_*}|^2 &\leq \sum_{\substack{j<M\\ \lambda:\lambda_1 = n-j}} d_\lambda^2 \left(\left(1 - \frac{2j}{n} + O\left(\frac{1}{n^2}\right)\right)^t - \left(1 - \frac{j}{n}\right)^{t_*} \right)^2 \\ &\leq \sum_{\substack{j<M\\ \lambda:\lambda_1 = n-j}} d_\lambda^2 \left(\frac{e^{-cj}}{n^j} \left(1 + O\left(\frac{\log n}{n}\right)\right) - \frac{e^{-cj}}{n^j} \left(1 + O\left(\frac{j^2}{n}\right)\right) \right)^2,\end{aligned}$$

where in the last inequality, we plug in the values for t and t_* and we use the fact that $e^x\left(1 - \frac{x^2}{n}\right) \leq \left(1 + \frac{x}{n}\right)^n \leq e^x$ for $|x| \leq n$. Using Lemma 4.4, we get that

$$\sum_{\lambda_1 > n-M} d_\lambda d_{\lambda^{(1)}} |s_\lambda^t - \overline{s}_{\lambda^{(1)}}^{t_*}|^2 = O\left(\frac{\log^2 n}{n^2}\right).$$

Therefore

$$\sum_{\lambda_1 > n-M} d_\lambda d_{\lambda^{(1)}} |s_\lambda^t - \overline{s}_{\lambda^{(1)}}^{t_*}|^2 \leq \varepsilon/2, \qquad (5.4)$$

for sufficiently large n. Equations (5.3) and (5.4) give the first inequality in 4. To prove the second inequality, we write

$$\sum_{\lambda'_1 > n-M} d_\lambda \sum_i d_{\lambda^{(i)}} |s_\lambda^t - \overline{s}_{\lambda^{(i)}}^{t_*}|^2 = \sum_{\lambda'_1 > n-M} d_{\lambda'} \sum_i d_{\lambda'(\lambda_i)} \left| s_\lambda^t - \left(\frac{2}{n} - \overline{s}_{\lambda'(\lambda_i)}\right)^{t_*} \right|^2$$

and we proceed similarly as before, using the second part of Corollary 4.5 and the fact that $\overline{s}_{\lambda'(1)} = \frac{\lambda'_1}{n}$. We note that at this point it is important that t and t_* are either both odd or both even, so that the corresponding eigenvalues have the same sign. To achieve this, we could have considered t to be $\frac{1}{2} n(\log n + c) + 1$, which doesn't affect the limit profile behavior. $\square$

6 Open questions

6.1 Conjugacy class shuffles with cutoff

We are now looking at the case where $G = S_n$ and we want to apply Lemma 1.4 when Q is the transition matrix of random transpositions. Let P be the transition matrix of a lazy random walk on S_n generated by a conjugacy class (or a union of conjugacy classes).

A standard application of Schur's lemma gives that

$$D = \begin{pmatrix} M_P(\rho_1) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & M_P(\rho_k) \end{pmatrix},$$

is a diagonal matrix and in fact each Fourier transform $\hat{P}(\rho_i)$ is a multiple of the identity. Since each irreducible representation contributes only one eigenvalue, we form the following conjecture.

Conjecture 6.1. *Let P be the transition matrix of a lazy random walk on S_n generated by a conjugacy class (or a union of conjugacy classes) with number of fixed points of order $n - o(n)$. If P exhibits total variation cutoff and ℓ^2 cutoff at t_n with window w_n, then*

$$\Phi_x(c) = d_{T.V.}(\mathit{Poiss}(1 + e^{-c}), \mathit{Poiss}(1)).$$

To support this claim, we point out that the card shuffle generated by k–cycles with $k = o(n)$ satisfies this conjecture as proven in [22].

6.2 Random-to-random

One step of the random-to-random card shuffle consists of picking a card and a position of the deck uniformly and independently at random and moving that card to that position. This shuffle was introduced by Diaconis and Saloff-Coste [7], who proved that the mixing time is $O(n \log n)$. It has been studied by ([30], [27], [28], [24]), since cutoff at $\frac{3}{4} n \log n - \frac{1}{4} n \log \log n$ was conjectured by Diaconis for fifteen years [6]. Recently, Bernstein and Nestoridi [2] proved the upper bound for the mixing time, which in combination with Subag's [28] lower bound resolved the desired conjecture.

The main open question of this section is the following.

Open Question. What is the limit profile of random-to-random?

Even though Lemma 1.4 can be applied for random transpositions and random-to-random (due to Lemma 4.8), it is not clear if it can lead to determining the limit profile of random-to-random. In fact the limit profile of random-to-random seems to be different from the limit profile of random transpositions.

We can see this by looking at just at the $n - 1$ dimensional representation ρ. For random transpositions, ρ contribute the eigenvalue $1 - \frac{2}{n}$ with multiplicity n^2. For random-to-random, we have $n^{3/2}$ eigenvalues that are roughly equal to $1 - \frac{1}{n}$ and the rest are negligible. When we study the error, we get a term that is roughly equal to

$$n^{3/2} \left(\left(1 - \frac{2}{n}\right)^t - \left(1 - \frac{1}{n}\right)^{t_*} \right)^2 + n^2 \left(1 - \frac{2}{n}\right)^{2t},$$

where t, t_* are the corresponding cutoff times. This term does not converge to zero as n approches infinity, thus suggesting that the limit profiles of the two shuffles are different.

References

[1] Dave Bayer and Persi Diaconis. Trailing the dovetail shuffle to its lair. *Ann. Appl. Probab.*, 2(2):294–313, 1992. MR1161056

[2] Megan Bernstein and Evita Nestoridi. Cutoff for random to random card shuffle. *Ann. Probab.*, 47(5):3303–3320, 2019. MR4021252

[3] Alexey Bufetov and Peter Nejjar. Cutoff profile of ASEP on a segment. *Probab. Theory Related Fields*, 183(1-2):229–253, 2022. MR4421174

[4] Persi Diaconis. Applications of noncommutative Fourier analysis to probability problems. In *École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87*, volume 1362 of *Lecture Notes in Math.*, pages 51–100. Springer, Berlin, 1988. MR0983372

[5] Persi Diaconis. *Group representations in probability and statistics*, volume 11 of *Institute of Mathematical Statistics Lecture Notes—Monograph Series*. Institute of Mathematical Statistics, Hayward, CA, 1988. MR0964069

[6] Persi Diaconis. Mathematical developments from the analysis of riffle shuffling. In *Groups, combinatorics & geometry (Durham, 2001)*, pages 73–97. World Sci. Publ., River Edge, NJ, 2003. MR1994961

[7] Persi Diaconis and Laurent Saloff-Coste. Comparison techniques for random walk on finite groups. *Ann. Probab.*, 21(4):2131–2156, 1993. MR1245303

[8] Persi Diaconis and Laurent Saloff-Coste. Comparison theorems for reversible Markov chains. *Ann. Appl. Probab.*, 3(3):696–730, 1993. MR1233621

[9] Persi Diaconis and Laurent Saloff-Coste. Walks on generating sets of abelian groups. *Probab. Theory Related Fields*, 105(3):393–421, 1996. MR1425868

[10] Persi Diaconis and Laurent Saloff-Coste. Bounds for Kac's master equation. *Comm. Math. Phys.*, 209(3):729–755, 2000. MR1743614

[11] Persi Diaconis and Mehrdad Shahshahani. Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete*, 57(2):159–179, 1981. MR0626813

[12] Persi Diaconis and Mehrdad Shahshahani. Time to reach stationarity in the Bernoulli-Laplace diffusion model. *SIAM J. Math. Anal.*, 18(1):208–218, 1987. MR0871832

[13] Alexandros Eskenazis and Evita Nestoridi. Cutoff for the Bernoulli-Laplace urn model with $o(n)$ swaps. *Ann. Inst. Henri Poincaré Probab. Stat.*, 56(4):2621–2639, 2020. MR4164850

[14] Leopold Flatto, Andrew M. Odlyzko, and David B. Wales. Random shuffles and group representations. *Ann. Probab.*, 13(1):154–178, 1985. MR0770635

[15] Martin Hildebrand. Generating random elements in $\mathrm{SL}_n(\mathbf{F}_q)$ by random transvections. *J. Algebraic Combin.*, 1(2):133–150, 1992. MR1226348

[16] Bob Hough and Yunjiang Jiang. Cut-off phenomenon in the uniform plane Kac walk. *Ann. Probab.*, 45(4):2248–2308, 2017. MR3693962

[17] Martin Kassabov. Kazhdan constants for $\mathrm{SL}_n(\mathbb{Z})$. *Internat. J. Algebra Comput.*, 15(5-6):971–995, 2005. MR2197816

[18] Hubert Lacoin. The cutoff profile for the simple exclusion process on the circle. *Ann. Probab.*, 44(5):3399–3430, 2016. MR3551201

[19] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson. MR2466937

[20] Eyal Lubetzky and Yuval Peres. Cutoff on all Ramanujan graphs. *Geom. Funct. Anal.*, 26(4):1190–1216, 2016. MR3558308

[21] Peter Matthews. A strong uniform time for random transpositions. *J. Theoret. Probab.*, 1(4):411–423, 1988. MR0958246

[22] Evita Nestoridi and Sam Olesker-Taylor. Limit profiles for reversible Markov chains. *Probab. Theory Related Fields*, 182(1-2):157–188, 2022. MR4367947

[23] Natesh S. Pillai and Aaron Smith. On the mixing time of Kac's walk and other high-dimensional Gibbs samplers with constraints. *Ann. Probab.*, 46(4):2345–2399, 2018. MR3813994

[24] Chuan Qin and Ben Morris. Improved bounds for the mixing time of the random-to-random shuffle. *Electron. Commun. Probab.*, 22:Paper No. 22, 7, 2017. MR3635695

[25] Jeffrey S. Rosenthal. Random rotations: characters and random walks on $\mathrm{SO}(N)$. *Ann. Probab.*, 22(1):398–423, 1994. MR1258882

[26] Justin Salez. Cutoff for non-negatively curved Markov chains. *Journal of the European Mathematical Society*, to appear.

[27] Laurent Saloff-Coste and Jessica Zúñiga. Refined estimates for some basic random walks on the symmetric and alternating groups. *ALEA Lat. Am. J. Probab. Math. Stat.*, 4:359–392, 2008. MR2461789

[28] Eliran Subag. A lower bound for the mixing time of the random-to-random insertions shuffle. *Electron. J. Probab.*, 18:no. 20, 20, 2013. MR3035748

[29] Lucas Teyssier. Limit profile for random transpositions. *Ann. Probab.*, 48(5):2323–2343, 2020. MR4152644

[30] Jay-Calvin Uyemura-Reyes. Random walks, semi-direct products and card shuffling. *ProQuest LLC, Ann Arbor, MI, 2002, Thesis (Ph.D.)-Stanford University. MR-2703300.*

Acknowledgments. The author would like to thank Sam Olesker-Taylor, Andreas Pardon, Sarah Peluse, Laurent Saloff-Coste and Dominik Schmid for helpful discussions and comments.