ALEA, Lat. Am. J. Probab. Math. Stat. 21, 1547–1566 (2024)
DOI: 10.30757/ALEA.v21-58

The S_k Shuffle Block Dynamics

Evita Nestoridi, Amanda Priestley and Dominik Schmid

Evita Nestoridi, Stony Brook University, Mathematics Department, 100 Nicolls Road, Stony Brook, NY 11794, USA
E-mail address: evrydiki.nestoridi@stonybrook.edu
URL: https://evitanestoridi.github.io/

Amanda Priestley, The University of Texas at Austin, Department of Computer Science, Gates Dell Complex, 2317 Speedway, Austin, Texas 78712, USA
E-mail address: amandapriestley@utexas.edu
URL: https://mndpriestley.github.io/

Dominik Schmid, University of Bonn, Institute for Applied Mathematics, Endenicher Allee 60, 53115 Bonn, Germany
E-mail address: d.schmid@uni-bonn.de
URL: https://sites.google.com/view/dominik-schmid

Abstract. We introduce and analyze the S_k shuffle on N cards, a natural generalization of the celebrated random adjacent transposition shuffle. In the S_k shuffle, we choose uniformly at random a block of k consecutive cards, and shuffle these cards according to a permutation chosen uniformly at random from the symmetric group on k elements. We study the total-variation mixing time of the S_k shuffle when the number of cards N goes to infinity, allowing also $k = k(N)$ to grow with N. In particular, we show that when $k = o(N^{\frac{2}{3}})$ the pre-cutoff phenomenon occurs. Furthermore, we show that for a suitable modification of the model, the S_k shuffle with boundaries, the cutoff phenomenon occurs when $k = o(N^{\frac{1}{6}})$.

1. Introduction

1.1. *Model and results.* When shuffling a deck of N cards, our experience suggests that shuffles that involve only "local" moves, i.e. moves that significantly affect only a small number of cards, mix slower than shuffles that involve non-local moves. Random transpositions Diaconis and Shahshahani (1981), star transpositions Nestoridi (2021), random-to-random Bernstein and Nestoridi (2019) are examples of such local card shuffles that are known to shuffle a deck of N cards in order $N \log N$ steps. Random adjacent transpositions are even slower, mixing in order $N^3 \log N$ steps Lacoin (2016); Wilson (2004). In contrast, the riffle shuffle and k-cycles for sufficiently large values of k

Received by the editors April 15th, 2024; accepted July 25th, 2024.
2020 *Mathematics Subject Classification.* Primary: 60K35; Secondary: 60K37, 60J27.
Key words and phrases. mixing times, block dynamics, card shuffling, cutoff phenomenon.
This project was initiated at the SAMSI Virtual Workshop in Combinatorial probability, which was supported by NSF Grant DMS-1929298. EN was supported by NSF Grant DMS-2052659. AP was supported by the University of Texas at Austin Dean's Strategic Fellowship. DS acknowledges the DAAD PRIME program for financial support.

mix in only order $\log N$ steps Bayer and Diaconis (1992); Berestycki et al. (2011). In this paper, we introduce the S_k shuffle, a model that interpolates between local and global moves for the shuffles, i.e. it corresponds to adjacent random transpositions when $k = 2$, and the shuffle which picks a uniform permutation in every step when $k = N$.

Our definition of the S_k shuffle is inspired by block dynamics of other well-studied models, such as the Ising model Blanca et al. (2022); Guo and Jerrum (2018); Knöpfel et al. (2020); Martinelli et al. (2003); Martinelli (1999); Yang (2023), and other non-local dynamics such as the Swendsen-Wang model Blanca et al. (2022, 2023); Long et al. (2014), the random cluster model Ganguly and Seo (2020), and the Bernoulli-Laplace model with multiple swaps Alameda et al. (2024); Eskenazis and Nestoridi (2020). In the S_k shuffle, each block of k consecutive cards is assigned an independent rate 1 Poisson clock. Whenever a clock rings, we shuffle the cards in the respective block according to a permutation chosen uniformly at random from the symmetric group on k elements. The main objective in this paper is the total-variation mixing time, $t_{\text{mix}}(\varepsilon)$, for the S_k shuffle on N cards when N goes to infinity; see Section 2 for a formal definition of the respective quantities. We have the following first result.

Theorem 1.1. *For the S_k shuffle, there exists a constant $c > 0$ such that for all $k = k(N)$ with $k = o(N^{2/3})$ and all $\varepsilon \in (0, 1)$,*

$$\frac{6}{\pi^2} \leq \liminf_{N\to\infty} \frac{k(k^2-1)\cdot t_{\text{mix}}(\varepsilon)}{N^2 \log N} \leq \limsup_{N\to\infty} \frac{k(k^2-1)\cdot t_{\text{mix}}(\varepsilon)}{N^2 \log N} \leq c\,. \tag{1.1}$$

A crucial observation is that the S_k shuffle treats cards differently depending on their positions in the deck. One might notice that cards in the first and last k positions of the deck move more slowly than those in the middle. In particular, following the first card of the deck, one sees that the mixing time must be at least of constant order; see Proposition 7.1 for a precise statement. To counter this effect, we introduce extra moves in positions 1 through $k-1$, as well as positions $N-k+1$ through N, and we refer to the first and last k positions as **boundary**. More precisely, for each $i \in \{2, \ldots, k-1\}$ we assign a rate $\delta_i^{(k)}$, respectively a rate $\delta_{N-i+1}^{(k)}$, Poisson clock to the blocks containing the first i, respectively the last i cards of the deck. Whenever a clock rings, we shuffle the cards in the respective block according to a permutation chosen uniformly at random from the symmetric group on i elements. We refer to this as the $\mathbf{S_k}$ **shuffle with boundaries**. The exact choice of the values $\delta_i^{(k)}$ is deferred to Section 2, where we formally introduce both processes; see also Section 7 for a comparison of the two processes. While the ε-mixing time $t'_{\text{mix}}(\varepsilon)$ of the S_k shuffle with boundaries still satisfies the bounds in Theorem 1.1, the upper bound can be improved for sufficiently slow growing k; see also Conjecture 7.3 when $k = o(N^{1/2})$.

Theorem 1.2. *For the S_k shuffle with boundaries, assuming $k = o(N^{1/6})$, the ε-mixing time of the S_k-shuffle with boundaries satisfies for all $\varepsilon \in (0, 1)$*

$$\lim_{N\to\infty} \frac{k(k^2-1)\cdot t'_{\text{mix}}(\varepsilon)}{N^2 \log N} = \frac{6}{\pi^2}\,. \tag{1.2}$$

The fact that the leading order of the mixing time does not depend on ε is called the **cutoff phenomenon**. When $k = 2$, Theorem 1.2 agrees with the celebrated result of Lacoin (2016) (note that our model differs from Lacoin's setup by a time change); see also earlier work by Wilson (2004) for a sharp lower bound. While we follow for Theorem 1.2 the strategy introduced by Lacoin (2016), one faces several challenges when adapting the arguments to the case $k \geq 3$. Perhaps most surprisingly, neither the spectral gap nor the other eigenvalues of the transition matrix offer a simple closed form when $k > 3$. Instead, under a suitable choice of the boundary rates, we utilize approximate eigenvalues and eigenfunctions. The idea of using approximate eigenfunctions first appeared in Nam and Nestoridi (2019) when studying a time inhomogeneous version of the adjacent transposition shuffle. It was later adapted for continuous time Markov chains in Gantert et al. (2023)

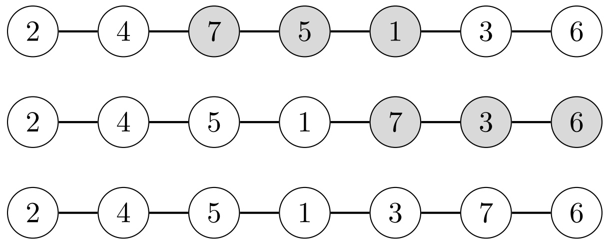

FIGURE 2.1. Example of the S_3 shuffle on a segment of length N. In each step the positions chosen to be updated are shaded gray.

to get sharp lower bounds for the symmetric exclusion process with one open boundary. In contrast to the above mentioned works, we approximate all relevant eigenvalues and eigenfunctions of the S_k shuffle with boundaries simultaneously. This allows us to perform approximate Fourier Analysis. It is a crucial ingredient in the proof of the upper bound in Theorem 1.2 as it addresses a discrete heat equation whose solution can be given via the S_k shuffle.

When providing sharp lower bounds on the mixing time, another difficulty occurs for large k as the maximal displacement by a shuffle within a block becomes comparable to the fluctuations of the S_k shuffle in equilibrium. We resolve this issue by relying on the strong Rayleigh property for the S_k shuffle with boundaries similar to Salez (2023) and Tran (2023) for the symmetric exclusion process with open boundaries. Moreover, we require a different generalized version of Wilson's Lemma, going back to the original second moment method as it was introduced in Diaconis and Shahshahani (1987, Section 4), using stricter variance bounds. Furthermore, the S_k shuffle with boundaries process requires a different interpretation of censoring than the one used previously by Gantert et al. (2023); Lacoin (2016). We introduce a generalized censoring scheme that not only restricts moves, but also alters them. Let us conclude the introduction by mentioning that adjacent transpositions can also be studied for biased card shuffling methods; see for example Benjamini et al. (2005); Zhang (2024).

1.2. *Structure of the paper.* In Section 2, we give preliminary definitions and notation which will be used throughout the paper. Section 3 discusses important properties of the S_k shuffle that are retained from the case of $k = 2$, and which play substantial roles in our proofs of the upper and lower bounds. In Section 4, we obtain lower bounds on the mixing time of the S_k shuffle using a generalized version of the second moment method. In Section 5, we introduce a coupling argument for the upper bound in Theorem 1.1. In Section 6, we adapt the argument of Lacoin from Lacoin (2016) to prove a sharp upper bound on the mixing time for the S_k shuffle with boundaries. We conclude by a discussion in Section 7 comparing the S_k shuffle with and without boundaries, and an open question.

2. Preliminaries for the S_k shuffle

In this section, we give a formal definition of the S_k shuffle with and without boundaries. For all $n \in \mathbb{N}$, we denote in the following by $\mathcal{S}_n$ the symmetric group on n elements, and refer to $\sigma \in \mathcal{S}_n$ as a permutation on $[n] \coloneqq \{1, \dots, n\}$. For integers $i, j \in [n]$ with $i < j$, and permutations $\eta \in \mathcal{S}_n$ and $\sigma \in \mathcal{S}_{j-i+1}$, we define the configuration $\eta^{\sigma,i,j}$ by

$$\eta^{\sigma,i,j}(m) \coloneqq \begin{cases} \eta(m) & \text{if } m \notin [i,j] \cap [n] \\ \eta(\sigma(m+1-i)) & \text{if } m \in [i,j] \cap [n] \end{cases} \tag{2.1}$$

for all $m \in [n]$, i.e. we permute the cards in the interval $[i, j]$ according to σ. For $N \in \mathbb{N}$ and $k \in [N]$, the $\mathbf{S_k}$ **shuffle** on a deck of size N is the continuous-time Markov chain on $\mathcal{S}_N$ whose generator is given by

$$(\mathcal{L}f)(\eta) = \sum_{i=1}^{N-k+1} \frac{1}{k!} \sum_{\sigma \in S_k} \Big(f(\eta^{\sigma,i,i+k-1}) - f(\eta) \Big) \tag{2.2}$$

for all functions $f\colon \mathcal{S}_N \to \mathbb{R}$, and $\eta \in \mathcal{S}_N$; see Figure 2.1 for a visualization. We denote the resulting dynamics by $(\eta_t)_{t\geq 0}$. For the $\mathbf{S_k}$ **shuffle with boundaries**, we set

$$\delta_i^{(k)} = \delta_{N-i+1}^{(k)} \coloneqq \frac{3i^2 + k^2 - 1}{(2i+1)(2i-1)} \tag{2.3}$$

for $i \in [2, k-1]$ and define the dynamics $(\zeta_t)_{t\geq 0}$ on $\mathcal{S}_N$ with respect to the generator

$$(\tilde{\mathcal{L}}f)(\zeta) \coloneqq (\mathcal{L}f)(\zeta) + \sum_{i=2}^{k-1} \frac{\delta_i^{(k)}}{i!} \sum_{\sigma \in S_i} \left(f(\zeta^{\sigma,1,i}) + f(\zeta^{\sigma,N-i+1,N}) - 2f(\zeta) \right) . \tag{2.4}$$

In words, we in addition apply a uniform permutation on the first and last i cards at rate $\delta_i^{(k)}$, respectively. While the choice of $\delta_i^{(k)}$ may seem slightly unnatural at first glance, we will see that this choice of rates in (2.3) allows us to reuse the eigenvalues and eigenfunctions for $k = 2$ as approximate eigenvalues and approximate eigenfunctions for $k \geq 3$.

Note that both dynamics are reversible with respect to the uniform measure on $\mathcal{S}_N$, which we denote in the following by μ_N. We let, for a probability measure ν on $\mathcal{S}_N$,

$$\|\nu - \mu_N\|_{\mathrm{TV}} \coloneqq \frac{1}{2} \sum_{\eta \in \mathcal{S}_N} |\nu(\eta) - \mu_N(\eta)| = \max_{A \subseteq \mathcal{S}_N} (\nu(A) - \mu_N(A)) \tag{2.5}$$

be the **total-variation distance** of ν and μ_N, and let the **$\boldsymbol{\varepsilon}$-mixing time** of $(\eta_t)_{t\geq 0}$ be

$$t_{\mathrm{mix}}(\varepsilon) \coloneqq \inf \left\{ t \geq 0 \,:\, \max_{\eta \in S_N} \|\mathbb{P}(\eta_t \in \cdot \mid \eta_0 = \eta) - \mu_N\|_{\mathrm{TV}} < \varepsilon \right\} \tag{2.6}$$

for all $\varepsilon \in (0,1)$. Similarly, we denote the ε-mixing time of $(\zeta_t)_{t\geq 0}$ by $t'_{\mathrm{mix}}(\varepsilon)$ for all $\varepsilon \in (0,1)$. One central tool is the **height function** of the S_k shuffle. For $\sigma \in \mathcal{S}_N$, we set

$$h_\sigma(x,y) \coloneqq \left(\sum_{z=1}^{x} \mathbb{1}_{\{\sigma(z) \leq y\}} \right) - \frac{xy}{N} , \tag{2.7}$$

with the convention that $h_\sigma(x) = h_\sigma(x, \lfloor N/2 \rfloor)$ for all $x \in [N]$ and $\sigma \in \mathcal{S}_N$. Further, with a slight abuse of notation, we write $h_t(x) = h_{\eta_t}(x, \lfloor N/2 \rfloor)$ for all $t \geq 0$ for the S_k shuffle $(\eta_t)_{t\geq 0}$, and similarly $h'_t(x) = h_{\zeta_t}(x, \lfloor N/2 \rfloor)$ for the S_k shuffle with boundaries. Observe the height functions allow one to define a partial order on the state space $\mathcal{S}_N$. We say that σ **dominates** σ', and write $\sigma \succeq \sigma'$ if, for all $x, y \in [N]$,

$$h_\sigma(x,y) \geq h_{\sigma'}(x,y) . \tag{2.8}$$

Note that the maximal element with respect to $\succeq$ is $\sigma = \mathrm{id}$, the identity on $\mathcal{S}_N$.

2.1. *An approximation of the spectrum.* In the following, we discuss the spectrum of the S_k shuffle with boundaries. Apart from the special cases $k = 2$ and $k = 3$, we shall see that the eigenfunctions and eigenvalues do not have a simple closed form, and instead we propose the following candidates as approximate eigenvalues and eigenfunctions, i.e. we set

$$\Phi_{N,y}^{(j)}(\sigma) \coloneqq \sum_{x=1}^{N-1} h_\sigma(x,y)\psi_j(x) \quad \text{where} \quad \psi_j(x) \coloneqq \sin\left(\frac{xj\pi}{N}\right) , \tag{2.9}$$

with the convention that $\Phi_N(\sigma) = \Phi^{(1)}_{N,\lfloor N/2 \rfloor}(\sigma)$. Moreover, we let for all $j \in [N]$

$$\lambda^{(j)}_{N,k} \coloneqq (k-1) - \left[2 \sum_{i=1}^{k} \frac{k-i}{k} \cos\left(\frac{ij\pi}{N}\right)\right] = \frac{kj^2\pi^2}{N^2}\left(\frac{k^2-1}{12}\right) + O\Big(\frac{k^5 j^4}{N^4}\Big). \tag{2.10}$$

The following lemma shows that for our particular choice of δ_i defined in (2.3), $\Phi^{(j)}_N$ are indeed suitable approximate eigenfunctions with respect to approximate eigenvalues $\lambda^{(j)}_{N,k}$.

Lemma 2.1. *Recall from (2.4) the definition of $\tilde{\mathcal{L}}$. Let $k = k(N)$ be such that $k = o(N)$. Then there exists some constant $C > 0$ such that for all $j \in [N]$, and for all $y \in [N-1]$*

$$|(-\tilde{\mathcal{L}}\Phi^{(j)}_{N,y})(\sigma) - \lambda^{(j)}_{N,k}\Phi^{(j)}_{N,y}(\sigma)| \leq Ck^6 j^3 N^{-3} \tag{2.11}$$

for all $\sigma \in \mathcal{S}_N$.

Proof: We consider in the following only the case $y = \frac{N}{2}$ as the remaining cases are similar. Notice that whenever we apply a permutation in $[a, b]$ chosen uniformly at random to a configuration σ, the expected height function evaluated at a position $x \in [a, b]$ is given by $(b-x)(b-a)^{-1}h_\sigma(a) + (x-a)(b-a)^{-1}h_\sigma(b)$. Therefore, by re-indexing the summation and using the definition of $\tilde{\mathcal{L}}$, we see that

$$(\tilde{\mathcal{L}}\Phi^{(j)}_N)(\sigma) = \sum_{x=1}^{N-1} (\tilde{\mathcal{L}}h)(\sigma)\psi_j(x) = \sum_{x=1}^{N-1} h_\sigma(x) a_x$$

where we set

$$a_x = \begin{cases} \sum\limits_{i=1}^{k-1} \frac{k-i}{k}(\psi_j(x-i) + \psi_j(x+i)) - (k-1)\psi_j(x) & \text{if } x \in [k, N-k] \\ \sum\limits_{i=1}^{x-1} \frac{\delta^{(k)}_x i}{x}\psi_j(i) + \sum_{i=1}^{k-1}\frac{k-i}{k}\psi_j(x+i) - \Big(x + \sum\limits_{i=x+1}^{k-1}\delta^{(k)}_i\Big)\psi_j(x) & \text{if } x < k \\ \sum\limits_{i=x}^{N-1} \frac{\delta^{(k)}_x (N-i)}{N-x+1}\psi_j(N-i) + \sum_{i=1}^{k-1}\frac{k-i}{k}\psi_j(x-i) - \Big(x + \sum\limits_{i=x+1}^{N-1}\delta^{(k)}_i\Big)\psi_j(x) & \text{if } x > N-k+1\,. \end{cases}$$

For all $x \in [N]$, a computation involving trigonometric identities shows that

$$\lambda^{(j)}_{N,k}\psi_j(x) = \sum_{i=1}^{k-1} \frac{k-i}{k}(\psi_j(x-i) + \psi_j(x+i)) - (k-1)\psi_j(x)$$

which is precisely a_x for positions in the bulk, and thus $h_\sigma(x)a_x = \lambda^{(j)}_{N,k}\psi_j(x)$ in these positions. Hence, it remains to treat the difference $|h_\sigma(x)a_x - \lambda^{(j)}_{N,k}\psi_j(x)|$ for the boundary vertices. By our choice of $\delta^{(k)}_i$ in (2.3), another computation yields that

$$\frac{\delta^{(k)}_x}{x}\left(\sum_{i=1}^{x-1} i^2\right) - \Big(x + \sum_{i=x+1}^{k-1}\delta^{(k)}_i\Big)x = \left(\sum_{i=1}^{k-1}\frac{k-i}{k}(x-i)\right) - (k-1)x \tag{2.12}$$

for all $x < k$, and similarly when $x > N-k+1$. Using Taylor approximation, we get

$$\left|\psi_j(x) - \frac{jx\pi}{N} + \frac{j^3x^3\pi}{6N^3}\right| \leq C\frac{j^5k^5}{N^5} \tag{2.13}$$

for some $C > 0$ for all N sufficiently large. By a telescopic summation, for all $x < k$

$$\delta^{(k)}_x \leq \frac{7k^2}{4x^2-1} \quad \text{and} \quad \left(\sum_{i=k-x}^{k-1}\delta^{(k)}_i\right) = \frac{x(4k-3x-1)}{4(k-x)+2} \leq \frac{x^2}{4(k-x)} + x \tag{2.14}$$

and thus, for some $c_1 > 0$, and all $j \geq 1$ and $x < k$, the error incurred by Taylor approximation is at most

$$\left| \sum_{i=1}^{x-1} \frac{\delta_x^{(k)} i^4 j^3}{x N^3} \right| + \left| \frac{x^3 j^3}{N^3} \sum_{i=x+1}^{k-1} \delta_i^{(k)} \right| \leq c_1 \frac{k^4 j^3}{N^3}. \tag{2.15}$$

Since $\max(h_\zeta(x), h_\zeta(N-x)) \leq x$ for all $\zeta \in \mathcal{S}_N$, we obtain from (2.12), (2.13) and (2.15)

$$|h_\zeta(x) a_x - \lambda_{N,k}^{(j)} \psi_j(x)| \leq c_2 \frac{j^3 k^5}{N^3} \tag{2.16}$$

for some constant $c_2 > 0$, uniformly in $x \in [k-1]$ as well as $x > N - k + 1$. Summing over all x in the boundary, we obtain the desired result. □

2.2. *Projection of the S_k shuffle.* Note that as in Lacoin (2016) and Wilson (2004), the S_k shuffle has a natural projection which can be seen as an exclusion process on a hypergraph. Let $\sigma \in \mathcal{S}_N$ and let $K \in [N-1]$ be fixed. We let $\xi_\sigma^K \in \{0,1\}^N$ be the configuration which we obtain by setting $\xi_\sigma^K(x) = 1$ if the value of the card at position x is at most K. In other words, the first K cards can be thought of as particles, while the remaining cards are given the role of empty sites. The corresponding dynamics $(\xi_t^K)_{t\geq 0}$ can then be described by the generator:

$$(\hat{\mathcal{L}} f)(\xi) = \sum_{i=1}^{N-k} \frac{1}{k!} \sum_{\xi \in \{0,1\}^N} \left(f(\xi^{\sigma,i,i+k-1}) - f(\xi) \right) , \tag{2.17}$$

where f is a function $f : \{0,1\}^N \to \mathbb{R}$. Here, $\xi^{\sigma,i,j}$ is defined as in (2.1) for $\eta^{\sigma,i,j}$.

3. Properties Preserved by the S_k Shuffle

In this section, we discuss properties that are shared by the S_k and S_2 shuffles. It is of great importance that the stationary distribution of the S_k shuffle is the uniform distribution, allowing us to transfer several properties from the case of $k = 2$ to $k \geq 3$.

3.1. *Preservation of the Censoring Inequality.* The censoring inequality is introduced by Peres and Winkler (2013) and has since been used in many contexts, such as Gantert et al. (2023) and Lacoin (2016). In this section, we define a **censoring scheme** for the S_k-shuffle and show that the censoring inequality holds. In contrast to the typical use of censoring, we also alter moves.

Formally, we define a censoring scheme $\mathcal{C} : \mathbb{R}_0^+ \to \mathcal{P}(E)$ as a càdlàg function, where $\mathcal{P}(E)$ is the power set of edges $E \coloneqq \{\{x, x+1\} \colon x \in [N-1]\}$. We obtain the **censored dynamics** $(\eta_t^{\mathcal{C}})_{t\geq 0}$ from the S_k shuffle $(\eta_t)_{t\geq 0}$ and a censoring scheme $(\mathcal{C}_t)_{t\geq 0}$ as follows: Suppose that at time t, we perform a shuffle on an interval $\mathcal{I} = [i,j]$. If $\mathcal{I}$ contains no edge from $\mathcal{C}_t$, then we perform the shuffle on $\mathcal{I}$ as in the original dynamics. However, if $\mathcal{I}$ contains at least one edge in $\mathcal{C}_t$, then we partition $\mathcal{I}$ into sub-intervals $(\mathcal{I}_m)_{m\geq 0}$ with $\mathcal{I}_m = [i_m, i_{m+1} - 1]$ such that

$$\mathcal{I} = \bigcup_{m \geq 0} [i_m, i_{m+1} - 1] \tag{3.1}$$

for some $i_0 < i_1 < i_2 < \cdots$ and that $\{i_m - 1, i_m\} \in \mathcal{C}_t$ for all m. In each interval $\mathcal{I}_m$, we perform an independent $S_{|\mathcal{I}_m|}$-shuffle of the elements.

In words, we obtain the censored dynamics by performing independent $S_\cdot$ shuffles on the sub-intervals whenever we would perform a shuffle operation along a censored edge. The censoring inequality states that the law of the censored dynamics stochastically dominates the law of the original dynamics in terms of the stochastic order $\succeq$ from (2.8) for any time $t \geq 0$. Here, recall that a measure μ **stochastically dominates** a measure ν on $\mathcal{S}_N$, denoted $\mu \succeq_{\mathrm{SD}} \nu$, whenever $\mu(A) \geq \nu(A)$ for any set $A \subseteq \mathcal{S}_N$ which is increasing with respect to $\succeq$; see Section 22.2 of Levin et al. (2009). Moreover, this stochastic domination occurs conditionally on the choice of the jump times $(\mathcal{T}_i^x)_{i\geq 1}^{x\in[N-1]}$ at which we perform the i^{th} update at the interval starting at x. Formally, we say that the **censoring inequality** holds if for all $t \geq 0$, and for a suitably defined family of open intervals $(A_i^x)_{i\in\mathbb{N}}^{x\in[N-1]}$

$$\mathbb{P}(\eta_t^{\mathcal{C}} \in \cdot \,|\, \mathcal{T}_i^x \in A_i^x) \succeq_{\text{SD}} \mathbb{P}(\eta_t \in \cdot \,|\, \mathcal{T}_i^x \in A_i^x)\,. \tag{3.2}$$

Recall that a function f is **increasing** if $f(\sigma) \geq f(\sigma')$ when $\sigma \succeq \sigma'$, and that μ_N denotes the uniform measure on $\mathcal{S}_N$. The next lemma is due to Lacoin; see Proposition 3.6 in Lacoin (2016).

Lemma 3.1. *Let ν_0 be an initial distribution for the S_2 shuffle on $\mathcal{S}_N$ such that $\sigma \mapsto \frac{\nu_0}{\mu_N}(\sigma)$ is increasing. Let $\mathcal{C}$ be a censoring scheme and let $\nu_t^{\mathcal{C}}$ be the law of the S_2 shuffle with respect to $\mathcal{C}$. Then $\sigma \mapsto \frac{\nu_t^{\mathcal{C}}}{\mu_N}(\sigma)$ is increasing and the censoring inequality holds.*

In the following, our goal to extend this result to general $k \geq 3$.

Lemma 3.2. *Let $k \geq 3$ and ν_0 be an initial distribution for the S_k shuffle on $\mathcal{S}_N$ such that $\sigma \mapsto \frac{\nu_0}{\mu_N}(\sigma)$ is increasing. Let $\mathcal{C}$ be a censoring scheme and let $\nu_t^{\mathcal{C}}$ be the law of the S_k shuffle with respect to $\mathcal{C}$. Then $\sigma \mapsto \frac{\nu_t^{\mathcal{C}}}{\mu_N}(\sigma)$ is increasing and the censoring inequality holds.*

We use the next lemma to approximate a single update in the S_k shuffle with censoring.

Lemma 3.3. *Let ν_t be the distribution of the S_2 shuffle on $\mathcal{S}_k$ at time t, then we have that*

$$\lim_{t\to\infty} \|\nu_t - \mu_k\|_{\text{TV}} = 0\,. \tag{3.3}$$

Proof: This is an immediate consequence of the fact that the S_2 shuffle is an irreducible continuous-time Markov chain, which has the uniform distribution as its unique stationary law. □

Proof of Lemma 3.2: We will in the following, only show that $\sigma \mapsto \frac{\nu_t^{\mathcal{C}}}{\mu_N}(\sigma)$ is increasing for any censoring scheme $\mathcal{C}$. The fact that the censoring inequality holds then follows from the same arguments as Theorem 22.20 in Levin et al. (2009). To do so, we proceed by a proof by contradiction. Suppose there exists a censoring scheme $\mathcal{C}$, a sequence of times $(t_i^x)_{i\geq 1}^{x\in[N-1]}$, a time $t \geq 0$, some $\delta > 0$, and permutations $\sigma \succeq \sigma'$ such that

$$\mathbb{P}(\eta_t = \sigma' \,|\, \mathcal{T}_i^x = t_i^x) - \mathbb{P}(\eta_t^{\mathcal{C}} = \sigma \,|\, \mathcal{T}_i^x = t_i^x) \geq \delta\,. \tag{3.4}$$

Let $M > 0$ which will be chosen later. Let $\mathcal{J}_t := \{(x,i) : t_x^i \leq t\}$, and let $(\tilde{\eta}_t)_{t\geq 0}$ and $(\tilde{\eta}_t^{\mathcal{C}})_{t\geq 0}$ be two processes on $\mathcal{S}_N$ defined in the following way. For all $(x,i) \in \mathcal{J}_t$, at time t_x^i in $(\tilde{\eta}_t)_{t\geq 0}$ we perform a sequence of M many (discrete time) S_2 shuffle moves on the interval $[x, x+(k-1)]$. Similarly for the process $(\tilde{\eta}_t^{\mathcal{C}})_{t\geq 0}$, we apply for all $(x,i) \in \mathcal{J}_t$ a sequence of M many S_2 shuffles, but for each interval in the decomposition $\mathcal{I}$ defined in (3.1) for the censoring scheme at t_x^i separately. By Lemma 3.3 and a standard comparison between discrete time and continuous time Markov chains – see Theorem 20.3 in Peres and Winkler (2013) – and the triangle inequality for total-variation distance, we can choose $M = M(t, \mathcal{J}_t, \delta, k, \mathcal{C})$ sufficiently large, such that

$$\begin{aligned} \|\mathbb{P}(\tilde{\eta}_t \in \cdot \,|\, \mathcal{T}_i^x = t_i^x) - \mathbb{P}(\eta_t \in \cdot \,|\, \mathcal{T}_i^x = t_i^x)\|_{\text{TV}} &\leq \frac{\delta}{4} \\ \|\mathbb{P}(\tilde{\eta}_t{}^{\mathcal{C}} \in \cdot \,|\, \mathcal{T}_i^x = t_i^x) - \mathbb{P}(\eta_t^{\mathcal{C}} \in \cdot \,|\, \mathcal{T}_i^x = t_i^x)\|_{\text{TV}} &\leq \frac{\delta}{4}\,. \end{aligned} \tag{3.5}$$

Observe that $(\tilde{\eta}_t)_{t\geq 0}$, respectively $(\tilde{\eta}_t^{\mathcal{C}})_{t\geq 0}$, is an S_2 shuffle, respectively an S_2 shuffle with respect to some censoring scheme $\tilde{\mathcal{C}}$. Thus, using Lemma 3.1, and again the triangle inequality for total-variation distance we obtain the desired contradiction to (3.4). □

Remark 3.4. Note that the same arguments as in the proof of Lemma 3.2 apply to the S_k shuffle with boundaries, establishing that the censoring inequality holds.

3.2. *Preservation of the strong Rayleigh property.* In this section we discuss the strong Rayleigh property and its relation to negative dependence. Let $n \in \mathbb{N}$, and define a function $f \in \mathbb{C}[z_1, \ldots, z_n]$ with real coefficients to be **real stable** if $f(z_1, \ldots, z_n) \neq 0$ whenever $\mathfrak{Im}(z_j) > 0$ for $1 \leq j \leq n$. Let π be a probability measure over $\{0,1\}^n$. For $(X_1, \ldots X_n) \sim \pi$ is called **strongly Rayleigh** if its generating polynomial

$$(z_1, \ldots, z_n) \longmapsto \mathbb{E}_\pi \left[\prod_{i=1}^{n} z_i^{X_i} \right] \tag{3.6}$$

is real stable. The strong Rayleigh property was introduced by Borcea, Brändén, and Liggett in Borcea et al. (2009). Recall for $K \in [N-1]$ the projection $(\xi_t^K)_{t \geq 0}$ of the S_k shuffle to the first K cards, defined in Section 2.2. The following lemma can be found as Proposition 5.1 in Borcea et al. (2009).

Lemma 3.5 (Proposition 5.1 in Borcea et al. (2009))**.** *Let $K \in [N-1]$. Let ν_t denote the law of the projection of the S_2 shuffle to the first K cards. If ν_0 is strongly Rayleigh then so is the distribution of ν_t for all $t > 0$.*

We have the following simple consequence for the S_k shuffle.

Corollary 3.6. *Let $K, k \in [N-1]$. Let ν_t denote the law of the projection of the S_k shuffle to the first K cards. If ν_0 is strongly Rayleigh then so is the distribution of ν_t for all $t > 0$. The same holds for the S_k shuffle with boundaries and censoring.*

Proof: The fact that any individual S_{i+1} update of an interval $[x, x+i]$ for any $x \in [N-i]$ and $i \geq 1$ preserves the strong Rayleigh property is a consequence of Theorem 1.2 in Borcea and Brändén (2009). Using the Trotter Product formula – Theorem 3.44 in Liggett (2010) – we obtain the desired statement for the S_k shuffle, as the generator of the S_k shuffle with boundaries and censoring can be written as the sum of generators of S_i shuffle moves for time interval in which the censoring scheme remains constant. □

Next, we say that a set of random variables $\{X_1, \ldots, X_n\}$ taking values in $\{0,1\}$ is **negatively dependent** if for all $S \subset [n]$, we have

$$\mathbb{E}\left[\prod_{i \in S} X_i\right] \leq \prod_{i \in S} \mathbb{E}[X_i]. \tag{3.7}$$

In Borcea et al. (2009) it is shown that the strong Rayleigh property implies negative dependence, and we will use the following direct consequence of negative dependence, which we state without proof.

Corollary 3.7. *Let $c_i \geq 0$ and let $Z_n := \sum_{i=1}^n c_i X_i$ be the sum of negatively dependent random variables $\{X_1, \ldots X_n\}$ for some $n \in \mathbb{N}$. Then we have*

$$\mathrm{Var}[Z_n] \leq \sum_{i=1}^{n} c_i^2 \,\mathrm{Var}[X_i].$$

4. Lower bounds on the mixing time of the S_k shuffle

4.1. *An approximate second moment method.* For the S_2 shuffle sharp lower bounds can be obtained using Wilson's Lemma as first introduced in Wilson (2004), and approximate versions of his technique can be found in Gantert et al. (2023) and Nam and Nestoridi (2019). Here we instead rely on an approximate version of the second moment method originally introduced by Diaconis and Shahshahani (1987). To state this approximate second moment method, consider a continuous-time

Markov chain $(X_t)_{t\geq 0}$ with generator $\mathcal{A}$ on a finite state space S. It is a well known result that for any function $f : S \to \mathbb{R}$ the process $(M_t)_{t\geq 0}$ with

$$M_t \coloneqq f(X_t) - f(X_0) - \int_0^t (\mathcal{A}f)(X_s)\,\mathrm{d}s \quad \text{for all } t \geq 0 \tag{4.1}$$

is a martingale. We have the following result on the mixing time of $(X_t)_{t\geq 0}$.

Lemma 4.1. *Let $\Psi : S \to \mathbb{R}$ be such that for some $\lambda, c, R > 0$, we have that*

$$|(-\mathcal{A}\Psi)(y) - \lambda\Psi(y)| \leq c \text{ for all } y \in S, \quad \text{and} \quad \mathrm{Var}[\Psi(X_t)] \leq R \text{ for all } t \geq 0\,. \tag{4.2}$$

Then for all $\varepsilon \in (0,1)$, the mixing time t_{mix} of $(X_t)_{t\geq 0}$ satisfies

$$t_{\mathrm{mix}}(1-\varepsilon) \geq \frac{1}{\lambda}\log\left(\|\Psi\|_\infty\right) - \frac{1}{2\lambda}\log\left(\frac{20\max(R, c^2\lambda^{-2})}{\varepsilon}\right). \tag{4.3}$$

Proof: Let $X_0 = \eta$ almost surely for some $\eta \in S$ with $|\Psi(\eta)| = \|\Psi\|_\infty$. Let μ denote the stationary distribution of $(X_t)_{t\geq 0}$, and $X_\infty \sim \mu$. By (4.2) and the martingale $(M_t)_{t\geq 0}$, with $f(t) \coloneqq \mathbb{E}[\Psi(X_t)]$ for all $t \geq 0$, we get

$$f'(t) = \mathbb{E}[(\mathcal{A}\Psi)(X_t)] \in [-\lambda f(t) - c, -\lambda f(t) + c] \text{ for all } t \geq 0\,.$$

Applying Gronwall's lemma yields

$$f(t) \leq f(0)e^{-\lambda t} + \int_0^t ce^{-\lambda(t-s)}\mathrm{d}s \leq f(0)e^{-\lambda t} + \frac{c}{\lambda} \text{ for all } t \geq 0\,,$$

and it follows that

$$\left|f(t) - e^{-\lambda t}f(0)\right| \leq \frac{c}{\lambda} \tag{4.4}$$

holds for all $t \geq 0$, by applying Gronwall's lemma to $-f$. Taking t equal to the right hand side of (4.3), as a lower bound on the expectation of Ψ, we have, using that $\sqrt{20} > (1 + 2\sqrt{3})$,

$$\mathbb{E}[\Psi(X_t)] \geq e^{-\lambda t}\Psi(X_0) - \frac{c}{\lambda} = e^{-\lambda t}\|\Psi\|_\infty - \frac{c}{\lambda} \geq \frac{1}{\sqrt{\varepsilon}}2\sqrt{3}\max\left(\sqrt{R}, \frac{c}{\lambda}\right).$$

Now taking $t \to \infty$ in (4.4), we see that $|\mathbb{E}[\Psi(X_\infty)]| \leq c/\lambda$. To bound the total-variation distance, let P_t^η with t from the right hand side of (4.3) be the law of X_t started from η. We get

$$\begin{aligned}\|P_t^\eta - \mu\|_{\mathrm{TV}} &\geq \mathbb{P}\left(\Psi(X_t) \geq \frac{1}{2}\mathbb{E}[\Psi(X_t)]\right) - \mathbb{P}\left(\Psi(X_\infty) \geq \frac{1}{2}\mathbb{E}[\Psi(X_t)]\right) \\ &\geq 1 - 4\frac{\mathrm{Var}(\Psi(X_t))}{\mathbb{E}[\Psi(X_t)]^2} - 4\frac{\mathrm{Var}(\Psi(X_\infty)) + \mathbb{E}[\Psi(X_\infty)]^2}{\mathbb{E}[\Psi(X_t)]^2} \geq 1 - \varepsilon\,.\end{aligned} \tag{4.5}$$

Here, the last line follows from Chebyshev's inequality and the fact that, by Markov's inequality,

$$\mathbb{P}\left(\Psi(X_\infty) \geq \frac{1}{2}\mathbb{E}[\Psi(X_t)]\right) \leq \mathbb{P}\left(\Psi(X_\infty)^2 \geq \frac{1}{4}\mathbb{E}[\Psi(X_t)]^2\right) \leq 4\frac{\mathbb{E}\left[\Psi(X_\infty)^2\right]}{\mathbb{E}[\Psi(X_t)]^2}.$$

Substituting the bounds derived above and using that $\mathrm{Var}[\Psi(X_\infty)] \leq R$ yields the desired result. □

4.2. *A lower bound from the generalized second moment method.* In the following we prove a lower bound on the mixing time for the S_k shuffle with and without boundaries, which gives the lower bounds on the mixing time in Theorems 1.1 and 1.2. Recall that

$$\Phi_N^{(j)}(\sigma) := \sum_{x=1}^{N-1} h_\sigma(x)\psi_j(x) \quad \text{where} \quad \psi_j(x) := \sin\left(\frac{xj\pi}{N}\right)$$

and, recalling the height function $(h'_t)_{t\geq 0}$ from Section 2, we set

$$\Phi_{N,t}^{(j)} := \sum_{x=1}^{N-1} h'_t(x)\psi_j(x) \tag{4.6}$$

for all $t \geq 0$. In the following, we use $a \sim b$ to denote that a is asymptotically equivalent to b.

Lemma 4.2. *Let $\tilde{\mathcal{L}}$ and $\lambda_{N,k}$ be as defined in (2.4) and (2.10) and $k = o(N)$. Then for all $\sigma \in \mathcal{S}_N$*

$$|(-\tilde{\mathcal{L}}\Phi_N^{(1)})(\sigma) - \lambda_{N,k}^{(1)}\Phi_N^{(1)}(\sigma)| \leq c \tag{4.7}$$

holds for $c \sim k^6\pi^3 N^{-3}$. Moreover, we have that $\|\Phi_N^{(1)}\|_\infty \sim N^2$ and $\mathrm{Var}(\Phi_{N,t}^{(1)}) \leq N^3$.

Proof: Claim (4.7) is immediate from Lemma 2.1, and the bound on $\|\Phi_N^{(1)}\|_\infty$ follows directly from the definition. Thus, it remains to bound the variance of the approximate eigenfunction $\Phi_{N,t}^{(1)}$. Note that the dirac mass on the projection of the identity permutation to the first $N/2$ cards is strongly Rayleigh, and thus by Corollary 3.6 so is the distribution of the projection of the S_k shuffle with boundaries on the first $N/2$ cards for all time $t > 0$. Let $(X_1^t, \ldots X_N^t)$ be the projection of the S_k shuffle with boundaries, where X_i^t is the indicator function that the card at position i has label at most $N/2$ at time t. Then by Corollary 3.7

$$\begin{aligned} \mathrm{Var}(\Phi_{N,t}^{(1)}) &= \mathrm{Var}\left(\sum_{x=1}^{N-1} h'_t(x)\psi_j(x)\right) = \mathrm{Var}\left(\sum_{m=1}^{N-1}\left(\sum_{i=m+1}^{N}\psi_j(i)\right)X_m^t\right) \\ &\leq \sum_{m=1}^{N-1}\left(\sum_{i=m+1}^{N}\psi_j(i)\right)^2 \mathrm{Var}(X_m^t) \leq \sum_{m=1}^{N-1} m^2\,\mathrm{Var}(X_m^t) \leq N^3 \end{aligned} \tag{4.8}$$

allowing us to conclude. □

Proof of the lower bounds in Theorems 1.1 and 1.2: Combining Lemma 4.1 and Lemma 4.2 with $c \sim k^6\pi^3N^{-3}$ and $\lambda = \lambda_{N,k}^{(1)} \sim \frac{\pi^2}{12}k(k^2-1)N^{-2}$ and $R = N^3$ gives the desired lower bound on the mixing time for the S_k shuffle with boundaries in Theorem 1.2. To see that the corresponding lower bound holds also for the S_k shuffle without boundaries, note that the function

$$\sigma \mapsto \sum_{x=1}^{N-1} h_\sigma(x)\psi_1(x) \tag{4.9}$$

is increasing with respect to the partial order $\succeq$ defined in (2.8). Thus, by Lemma 3.2 and Remark 3.4, treating the S_k shuffle as an S_k shuffle with boundaries and censoring

$$\mathbb{E}\left[\sum_{x=1}^{N-1} h_t(x)\psi_1(x)\right] \geq \mathbb{E}\left[\sum_{x=1}^{N-1} h'_t(x)\psi_1(x)\right] \tag{4.10}$$

for all $t \geq 0$. The lower bound on the mixing times of the S_k shuffle without boundaries follows from Chebyshev's inequality using Corollary 3.6 and the same arguments as in Lemma 4.2 to bound the variance of the function $\Phi_{N,t}^{(1)}$ for the S_k shuffle with boundaries. □

Remark 4.3. Since we only use $c/\lambda = \mathcal{O}(\sqrt{R})$ in the proof of the lower bounds in Theorems 1.1 and 1.2, observe that we in fact showed that the above lower bound on the mixing time for the S_k shuffle with and without boundaries remains valid for all $k = o(N^{5/6})$.

5. Upper bounds on the mixing time

5.1. *A general coupling for the S_k shuffle.* In this section, we provide an upper bound on the mixing time of the S_k shuffle. In contrast to our specific choice of boundary conditions in (2.3), we allow in the following for more general choices of the parameters $(\delta_i^{(k)})$.

We start by defining a coupling for the S_k shuffle with boundaries. Let $(\zeta_t)_{t\geq 0}$ and $(\zeta'_t)_{t\geq 0}$ denote the S_k shuffles started from $\zeta, \zeta' \in \mathcal{S}_N$, respectively. For both S_k shuffles, we will use the same Poisson clocks, i.e. when we update an interval $[x, x+j]$ for some x and j in $(\zeta_t)_{t\geq 0}$ at some time $s \geq 0$, we do the same in the process $(\zeta'_t)_{t\geq 0}$. Suppose that a clock associated with an interval $[x, x+j]$ rings at time s. Let $I_{x,s} \subseteq [N]$ be the set of labels for which both configurations agree at time s. For these $|I_{x,s}|$ cards, select $|I_{x,s}|$ of the $j+1$ positions in the interval $[x, x+j]$ uniformly at random, and assign the cards in both ζ_s and ζ'_s whose labels are in $I_{x,s}$ to these positions. On the remaining $(j+1) - |I_{x,s}|$ positions, we distribute the cards in both configurations ζ_s and ζ'_s uniformly at random and independently.

We refer to this as the **canonical coupling** for the S_k shuffle, and write $\mathbf{P}$ for the joint law of $(\zeta_t)_{t\geq 0}$ and $(\zeta'_t)_{t\geq 0}$ under this coupling. Let $Z_{i,t}$ and $Z'_{i,t}$ denote the positions of the cards labeled i in the configurations ζ_t and ζ'_t respectively. Moreover, let τ_i be the first time at which the cards of label i are located at the same position in both shuffles, and note that the cards of label i occupy the same position for all $s \geq \tau_i$. The next proposition states an upper bound on the mixing time of the S_k shuffle with and without boundaries.

Proposition 5.1. *Suppose that $\delta_i^{(k)} = \delta_{N-i}^{(k)} \in [0,1]$ for every $i \in [k]$, and assume that $k = o(N^{\frac{2}{3}})$. Then there exists an absolute constant $C > 0$ such that for all $\sigma, \sigma' \in \mathcal{S}_N$, and all $t \geq CN^2k^{-3}\log(N)$, we have that for all N sufficiently large*

$$\|\mathbf{P}(\zeta'_t \in \cdot \mid \zeta'_0 = \sigma') - \mathbf{P}(\zeta_t \in \cdot \mid \zeta_0 = \sigma)\|_{\mathrm{TV}} \leq N^{-1}\,. \tag{5.1}$$

For the S_k shuffle with boundary rates $(\delta_i^{(k)})$ from (2.3), the upper bound on the total-variation distance in (5.1) continues to hold for all $k = o(N)$ and $t \geq CN^2k^{-3}\log(N)$.

The proof of Proposition 5.1 will be split in two main parts. First, we investigate the time it takes for a single card to leave the sites $[4k]$, respectively $\{N-4k, \dots, N\}$, close to the boundary. In a second step, we establish tail estimates for the coalescence time τ_i for cards of label i, and obtain the desired upper bound on the mixing time by a union bound.

5.2. *An estimate on the exit time from the boundary.* Consider in the following the positions $(Z_{1,t})_{t\geq 0}$ and $(Z'_{1,t})_{t\geq 0}$ of the cards of label 1. We denote by P_x, respectively P'_x, the law of the cards of label 1 when starting the cards from position $x \in [N]$ in ζ_0 and ζ'_0, respectively. For all $y \in [N]$, let $\tilde{\tau}_{>y}$ and $\tilde{\tau}_{<y}$ be defined as

$$\tilde{\tau}_{>y} := \inf\{t \geq 0\colon Z_{1,t} > y\} \quad \text{and} \quad \tilde{\tau}_{<y} := \inf\{t \geq 0\colon Z_{1,t} < y\}\,, \tag{5.2}$$

i.e. the first time at which the card of label 1 in $(\zeta_t)_{t\geq 0}$ reaches a position larger than y, respectively smaller than y. For the following three lemmas, we assume that $k = o(N)$.

Lemma 5.2. *Let $\delta_i^{(k)} = \delta_{N-i}^{(k)} \in [0,1]$ for every $i \in [k]$. Then there exist absolute constants $c, C > 0$ such that for all $x \in [N]$, we have that for all N sufficiently large*

$$P_x\left(\tilde{\tau}_{>4k} > C\right) \leq c\,. \tag{5.3}$$

Proof: Using the canonical coupling, we can assume without loss of generality that $x = 1$. Since the interval $[k]$ is updated at rate 1, and the cards are assigned to a position chosen uniformly at random, we see that $P_1(\tilde{\tau}_{>k/2} > 2) \leq \frac{1}{4}$. Next, since all boundary rates are bounded by 1 by our assumptions, the event that the first update involving card 1, given a constant amount of time s_1 has passed, is initiated by an interval $[j, j+k-1]$ for some $j > k/4$ has positive probability uniformly in k. Thus, we get that for some absolute constants $c_1, c_2 > 0$

$$P_1(\tilde{\tau}_{>k/2} > c_1 \mid s_1 \leq 2) \leq c_2 \,. \tag{5.4}$$

Iterating this argument for $\tilde{\tau}_{>mk/4}$ with $m \in [2, 16]$, we conclude. □

We have a similar statement for the S_k shuffle with boundary rates $(\delta_i^{(k)})$ from (2.3).

Lemma 5.3. *Let $(\delta_i^{(k)})$ be defined as in (2.3). Then there exist absolute constants $c, C > 0$ such that for all $x \in [N]$, we have that for all N sufficiently large*

$$P_x\left(\tilde{\tau}_{>4k} > \frac{C}{k}\right) \leq c \,. \tag{5.5}$$

Proof: Using the canonical coupling, we can again assume without loss of generality that $x = 1$. Note that until time $\tilde{\tau}_{>k/4}$, for any update of an interval $[j]$ for some $j \geq k/2$ before time $\tilde{\tau}_{>k/4}$, we have that card 1 gets moved to some position $> k/4$ with probability at least $\frac{1}{4}$. Since $\delta_i^{(k)} \geq 1$ for all $i \in [k]$, and each interval $[j]$ is updated at rate at least 1,

$$P_1\left(\tilde{\tau}_{>k/4} > \frac{4}{k}\right) \leq \frac{1}{8} \,. \tag{5.6}$$

Since $\delta_i^{(k)} \leq 8$ for all $i \geq \frac{k}{4}$, note that the event that the first update involving card 1 after time $\tilde{\tau}_{>k/4}$ is initiated by an interval $[j, j+k-1]$ for some $j > k/8$ has positive probability, uniformly in k. Hence, we obtain that for some absolute constants $c_1, c_2 > 0$

$$P_1\left(\tilde{\tau}_{>k/2} \geq \frac{c_1}{k} \,\middle|\, \tilde{\tau}_{>k/4} \leq \frac{4}{k}\right) \geq c_2 \,. \tag{5.7}$$

Iterating this argument for $\tilde{\tau}_{>mk/4}$ with $m \in [16]$, we conclude. □

We require a final preliminary estimate on the expected return time for the S_k shuffle with boundaries when the boundary rates $(\delta_i^{(k)})$ are in $[0, 1]$.

Lemma 5.4. *Let $\delta_i^{(k)} = \delta_{N-i}^{(k)} \in [0, 1]$ for every $i \in [k]$. Then there exists an absolute constant $C > 0$ such that for all $x \in [k, 2k]$, and all N sufficiently large*

$$E_x[\tilde{\tau}_{>4k}] \leq \frac{C}{\sqrt{k}} \,, \tag{5.8}$$

where E_x denotes the expectation with respect to P_x.

Proof: From Lemma 5.2, we get that there exists some $C_1 > 0$ such that for all $x \in [4k]$, we have that $E_x[\tilde{\tau}_{>4k}] \leq C_1$. Thus, using the canonical coupling to see that $E_x[\tilde{\tau}_{>4k}]$ is decreasing in x, and iterating along $\tilde{\tau}_{>mk/4}$ for $m \in [16]$ as in (5.7), it suffices to show that

$$P_x\left(\tilde{\tau}_{<\sqrt{k}} < \tilde{\tau}_{>4k}\right) \leq \frac{c_1}{\sqrt{k}} \quad \text{and} \quad E_{\sqrt{k}}\left[\tilde{\tau}_{>k/4} \leq \frac{c_2}{\sqrt{k}} \,\middle|\, Z_{1,t} \geq \sqrt{k} \text{ for all } t \in [0, \tilde{\tau}_{>k/4}]\right] \leq \frac{c_3}{\sqrt{k}}$$

for all $x \in [k, 2k]$ and constants $c_1, c_2, c_3 > 0$. The second inequality is immediate from the fact that on the event $\{Z_{1,t} \geq \sqrt{k} \text{ for all } t \geq 0\}$, with positive probability $\tilde{\tau}_{>k/4} \leq c_4 k^{-1/2}$ holds for some constant $c_4 > 0$ by considering the first time an interval $[j, j+k-1]$ for $j < \sqrt{k}$ is updated. To see the first inequality, note that for each update of interval containing card 1, we have a positive probability, uniformly in k and the position $x \in [\sqrt{k}, k]$ of card 1, that card 1 is moved to some position $> k/4$, while the probability to move card 1 to the first $\sqrt{k}$ positions is at most $c_5 k^{-1/2}$ for some absolute constant $c_5 > 0$. □

5.3. *Proof of the upper bound on the mixing time.* We start by showing that with positive probability, the time for card 1 to exceed $\lfloor \frac{N}{2} \rfloor$ is of order at most $N^2 k^{-3}$.

Lemma 5.5. *Suppose that $\delta_i^{(k)} = \delta_{N-i}^{(k)} \in [0,1]$ for every $i \in [k]$, and $k = o(N^{\frac{2}{3}})$. Then for all $x \in [N]$, we have that for some positive constants $c_1, c_2 > 0$, and N sufficiently large,*

$$P_x\left(\tilde{\tau}_{>\lfloor N/2 \rfloor} > c_1 N^2 k^{-3}\right) \leq c_2 \,. \tag{5.9}$$

For the S_k shuffle with rates $(\delta_i^{(k)})$ from (2.3), (5.9) continues to hold for all $k = o(N)$.

Proof: In the following, we define a stopping time τ_{hit} for the process $(Z_{1,t})_{t \geq 0}$ by

$$\tau_{\text{hit}} := \min\left(\tilde{\tau}_{<2k}, \tilde{\tau}_{>N-2k}\right) \,. \tag{5.10}$$

Note that $(Z_{1,t})_{t \in [0,\tau_{\text{hit}}]}$ is a stopped martingale, and by the optional stopping theorem

$$P_{4k}\left(\tilde{\tau}_{<2k} < \tilde{\tau}_{>\lfloor N/2 \rfloor}\right) = \frac{2k}{\lfloor N/2 \rfloor - 2k} \,. \tag{5.11}$$

Let X be the amount of time $(Z_{1,t})_{t \geq 0}$ spends until time $\tilde{\tau}_{>4k}$ at sites $[2k]$. From Lemma 5.3 and Lemma 5.4, we obtain that for some constant $c > 0$, and all $x \in [4k]$

$$E_x[X] \leq \begin{cases} ck^{-1/2} & \text{if } \delta_i^{(k)} \in [0,1] \text{ for all } i \in [k] \\ ck^{-1} & \text{for rates } (\delta_i^{(k)}) \text{ from (2.3)}\,. \end{cases} \tag{5.12}$$

For the stopped martingale $(Z_{1,t})_{t \in [0,\tau_{\text{hit}}]}$, note that its quadratic variation $(\langle Z_{1,\cdot} \rangle_t)_{t \in [0,\tau_{\text{hit}}]}$ satisfies for some constant $c' > 0$, and all $t \geq 0$

$$\langle Z_{1,\cdot} \rangle_t \geq c' k^3 t \tag{5.13}$$

as card 1 moves at rate k according to an increment with a variance of order k^2. By the optional stopping theorem for the martingale $(M_t)_{t \in [0,\tau_{\text{hit}}]}$ with $M_t = (Z_{1,t})^2 - \langle Z_{1,\cdot} \rangle_t$, we see that $E_{4k}[\tau_{\text{hit}}]$ is of order $N^2 k^{-3}$. Together with (5.11) and (5.12), we conclude. □

Proof of Proposition 5.1: In the following, we argue that there exist $c_1, c_2 > 0$ independent of N and k such that

$$\mathbf{P}(\tau_1 > c_1 N^2 k^{-3}) \leq c_2 \,. \tag{5.14}$$

for any pair of starting configurations ζ_0, ζ_0'. Since the canonical coupling preserves the coalescence of cards, the upper bound on the mixing time then follows by repeating the argument to obtain a bound on the coalescence time τ_i for each card with label $i \in [N]$, and then taking a union bound over these events. By Lemma 5.5, for some $c_3, c_4 > 0$ which do not depend on N and k

$$\mathbf{P}(\tilde{\tau}_{>\lfloor N/2 \rfloor} > c_3 N^2 k^{-3}) \geq c_4 \,. \tag{5.15}$$

Since $(Z_{1,t})_{t \in [0, \tilde{\tau}_{<N/4} \wedge \tilde{\tau}_{>3N/4}]}$ is a stopped martingale with increments bounded by k,

$$\mathbf{P}\left(Z_{1,t} \in \left[\frac{N}{4}, \frac{3N}{4}\right] \text{ for all } t \in \left[\tilde{\tau}_{>\lfloor N/2 \rfloor}, \tilde{\tau}_{>\lfloor N/2 \rfloor} + CN^2 k^{-3}\right]\right) > c_5 \tag{5.16}$$

for all $C > 0$ and some $c_5 = c_5(C) > 0$. Conditioning on the event in (5.16), note that by Lemma 5.5 we can choose $C > 0$ such that with positive probability, independently of k and N, there exists some $t_* \in [\tilde{\tau}_{>\lfloor N/2 \rfloor}, \tilde{\tau}_{>\lfloor N/2 \rfloor} + CN^2 k^{-3}]$ such that $N/5 < Z'_{1,t_*} < 4N/5$. Using the Strong Markov property under the coupling $\mathbf{P}$, we see that $(Z_{1,t})_{t \geq t_*}$ and $(Z'_{1,t})_{t \geq t_*}$ coalesce with positive probability before hitting $2k$ or $N - 2k$. This gives (5.14), and hence finishes the proof. □

6. Cutoff for the S_k shuffle with boundaries

6.1. *Approximate Fourier Analysis.* Recall from Section 2 the height function $(h_{\zeta_t})_{t\geq 0}$ of the S_k shuffle with boundaries $(\zeta_t)_{t\geq 0}$ and, with a slight abuse of notation, set

$$h'_t(x,y) \coloneqq h_{\zeta_t}(x,y)\,. \tag{6.1}$$

In Lacoin (2016), a key observation is that the expected height function $(x,y,t)\mapsto \mathbb{E}[h'_t(x,y)]$ of the S_2 shuffle is a solution $f:\{0,\dots,N\}^2\times\mathbb{R}_0^+\to\mathbb{R}$ to the discrete heat equation

$$\begin{cases} \partial_t f = \Delta_x f \\ f(0,y,t) = f(N,y,t) = 0 \\ f(x,y,0) = \mathbb{E}[h'_0(x,y)] \end{cases} \tag{6.2}$$

where Δ_x denotes the discrete Laplace operator

$$\Delta_x(f)(x) \coloneqq \frac{1}{2}f(x-1,y,t) + \frac{1}{2}f(x+1,y,t) - f(x)\,. \tag{6.3}$$

This allows for sharp estimates on $\mathbb{E}[h'_t(x,y)]$ for the S_2 shuffle; see Lemma 4.1 in Lacoin (2016). Then next lemma provides a similar result for the S_k shuffle with boundaries for $k\geq 3$.

Lemma 6.1. *Let $k\geq 3$. There exists a constant $C>0$ such that for any initial configuration $\sigma\in\mathcal{S}_N$ of the S_k shuffle with boundaries, for all $t\geq 0$, and all $y\in[N]$,*

$$\max_{x\in\{0,\dots,N\}} \mathbb{E}[h'_t(x,y)] \leq 8\min(y,N-y)e^{-t\cdot\lambda_{1,N,k}} + Ck^3. \tag{6.4}$$

Proof: By the standard Fourier decomposition, we obtain that

$$\mathbb{E}[h'_t(x,y)] = \frac{2}{N}\sum_{i=1}^{N}\sin\left(\frac{i\pi x}{N}\right)\sum_{j=1}^{N}\mathbb{E}\left[h'_t(j,y)\sin\left(\frac{ij\pi}{N}\right)\right]. \tag{6.5}$$

Recall $\Phi^{(j)}_{N,t}$ from (4.6) and the approximate eigenvalues $\lambda_{j,N,k}$ from (2.10). By Lemma 2.1 and the same arguments as for (4.4) in Lemma 4.1, we see that for all $y\in[N-1]$ and $t\geq 0$

$$\left|\mathbb{E}\left[\Phi^{(i)}_{N,t}\right] - e^{-t\cdot\lambda_{i,N,k}}\Phi^{(i)}_{N,0}(y)\right| \leq \frac{c_i}{\lambda_{i,N,k}}\,, \tag{6.6}$$

where we set $c_i = Ci^3k^6N^{-3}$ and take C from Lemma 2.1. Together with (6.5) we have

$$|\mathbb{E}[h'_t(x,y)]| \leq \left|\frac{2}{N}\sum_{i=1}^{N}\sin\left(\frac{i\pi x}{N}\right)e^{-t\cdot\lambda_{i,N,k}}\Phi^{(i)}_{N,0}(y)\right| + \left|\frac{2}{N}\sum_{j=1}^{N}\left|\sin\left(\frac{j\pi x}{N}\right)\right|\frac{c_j}{\lambda_{j,N,k}}\right|. \tag{6.7}$$

Note that the second summand in (6.7) is bounded from above by $2Ck^3$. For the first summand, note that $|\Phi^{(i)}_N(y)|\leq \min(y,N-y)N$ for all $y\in[N]$ and $i\in[N-1]$. Moreover, $\lambda_{j,N,k}\geq j\cdot\lambda_{1,N,k}$ holds for all $j\geq 1$. Using the fact that $|\sin(z)|<|z|$ for all $z\in\mathbb{R}$, we get

$$\left|\mathbb{E}[h'_t(x,y)]\right| \leq 2Ck^3 + 8y\sum_{j=1}^{N-1}e^{-jt\lambda_{1,N,k}} \leq 2Ck^3 + \frac{8ye^{-t\lambda_{1,N,k}}}{1-e^{-t\lambda_{1,N,k}}}\,. \tag{6.8}$$

When $e^{-t\lambda_{1,N,k}}\leq\frac{1}{2}$, this allows us to conclude (6.4). For $e^{-t\lambda_{1,N,k}}>\frac{1}{2}$, (6.4) is immediate from the fact that $\max_{x\in\{0,\dots,N\}} h_\sigma(x,y)\leq\min(y,N-y)$ for all $y\in[N]$ and $\sigma\in\mathcal{S}_N$. □

6.2. *Proof of the upper bound in Theorem 1.2.* As we follow in large parts the arguments of Lacoin (2016), we give the necessary adjustments, rather than the arguments in full detail. Let us start by introducing the main objects and outlining the main strategy for the proof. Fix some $K \in \mathbb{N}$ chosen later, and recall the height function representation $h_\sigma(x,y)$ from (2.7) for a permutation $\sigma \in \mathcal{S}_N$. Moreover, we recall the following definitions from Lacoin (2016).

Let $x_i := \lceil iN/K \rceil$ for all $i \geq 0$. For a permutation $\sigma \in \mathcal{S}_N$, we define in the following two projections $\widehat{\sigma}$ and $\bar{\sigma}$. The **semi-skeleton projection** $\widehat{\sigma} = (\widehat{\sigma}_{x,i})_{x\in[N], i\in[K]}$ and **skeleton projection** $\bar{\sigma} = (\bar{\sigma}(m,i))_{i,m\in[K]}$ are given by

$$\widehat{\sigma}(x,i) = h_\sigma(x, x_i) \quad \text{ and } \quad \bar{\sigma}(m,i) = h_\sigma(x_m, x_i) \tag{6.9}$$

respectively. We denote by $\widehat{\mathcal{S}}_N$ and $\bar{\mathcal{S}}_N$ the corresponding image spaces of $\mathcal{S}_N$ under these projections. Given a probability measure ν on $\mathcal{S}_N$, we use $\widehat{\nu}$, respectively $\bar{\nu}$, to denote the image measures of ν on $\widehat{\mathcal{S}}_N$, respectively $\bar{\mathcal{S}}_N$, under the semi-skeleton and skeleton projection, i.e. we set for all $\widehat{\sigma} \in \widehat{\mathcal{S}}_N$ and $\bar{\sigma} \in \bar{\mathcal{S}}_N$

$$\widehat{\nu}(\widehat{\sigma}) := \nu(\{\sigma \in \mathcal{S}_N : \sigma \mapsto \widehat{\sigma}\}) \quad \text{ and } \quad \bar{\nu}(\bar{\sigma}) := \nu(\{\sigma \in \mathcal{S}_N : \sigma \mapsto \bar{\sigma}\}) \,. \tag{6.10}$$

Let $\Delta x_i := x_i - x_{i-1}$, and let $\widetilde{\mathcal{S}}_N$ be the largest subgroup of $\mathcal{S}_N$ which is for all $i \in [K]$ invariant under permuting the cards of labels between $x_{i-1}+1$ and x_i. Note that $\widetilde{\mathcal{S}}_N$ is isomorphic to the product space $\bigotimes_{i=1}^K \mathcal{S}_{\Delta x_i}$. For a probability measure ν on $\mathcal{S}_N$, we define the measure $\tilde{\nu}$ on $\mathcal{S}_N$ by setting, for all $\sigma \in \mathcal{S}_N$,

$$\widetilde{\nu}(\sigma) := \frac{1}{\Pi_{i=1}^K (\Delta x_i)} \sum_{\widetilde{\sigma} \in \widetilde{\mathcal{S}}_N} \nu(\widetilde{\sigma} \circ \sigma) \,. \tag{6.11}$$

In words, to obtain the measure $\widetilde{\nu}$ from ν, we apply a uniformly chosen permutation which only shuffles for all $i \in [K]$ the cards of labels $x_{i-1}+1$ to x_i among each other.

To show to upper bound on the mixing time in Theorem 1.2, let $\delta > 0$ and set

$$t_1 = \frac{\delta N^2}{3k^3} \log(N) \qquad t_2 = \left(\frac{2\delta}{3} + \frac{4}{\pi^2}\right) \frac{N^2}{k^3} \log(N) \qquad t_3 = \left(\delta + \frac{4}{\pi^2}\right) \frac{N^2}{k^3} \log(N) \,.$$

We consider the censoring scheme $\mathcal{C} = (\mathcal{C}_t)_{t\geq 0}$ for the S_k shuffle with boundaries given by

$$\mathcal{C}_t := \begin{cases} \{\{x_i, x_i+1\} \colon i \in [K]\} & \text{if } t \in [0,t_1) \cup [t_2,t_3) \\ \emptyset & \text{if } t \notin [0,t_1) \cup [t_2,t_3) \,. \end{cases} \tag{6.12}$$

In the following, let $(\nu_t)_{t\geq 0}$ be the law of the S_k shuffle with boundaries under the censoring scheme $\mathcal{C}$ from (6.12) started from the identity, and let $\mu = \mu_N$ denote the uniform distribution on $\mathcal{S}_N$. The proof proceeds now in two steps. First, we argue that by time t_1, for all $i \in [K]$, the cards of labels $x_{i-1}+1$ to x_i in the S_k shuffle with boundaries and censoring scheme $\mathcal{C}$ are well mixed among each other. Second, we argue that by time t_3, the semi-skeleton has well mixed. To do so, the key task is to verify that the skeleton of the S_k shuffle with boundaries and censoring scheme $\mathcal{C}$ has well mixed by time t_2. This strategy is summarized and made precise in the following two propositions.

Proposition 6.2 (Proposition 5.1 in Lacoin (2016))**.** *For all $\varepsilon > 0$, there exists some $N_0 = N_0(\varepsilon)$ such that for all $N \geq N_0$ and $t \geq t_1$*

$$\|\widetilde{\nu}_t - \nu_t\|_{\mathrm{TV}} \leq \varepsilon/3 \,. \tag{6.13}$$

The proof of Proposition 6.2 is deferred to Section 6.3.

Proposition 6.3 (Proposition 5.3 in Lacoin (2016))**.** *For all $\varepsilon > 0$ and $k = o(N^{1/6})$, there exists some $N_1 = N_1(\varepsilon)$ such that for all $N \geq N_1$ and $t \geq t_3$*

$$\|\widehat{\nu}_t - \widehat{\mu}\|_{\mathrm{TV}} \leq 2\varepsilon/3 \,. \tag{6.14}$$

The proof of Proposition 6.3 is deferred to Section 6.4.

Proof of the upper bound in Theorem 1.2: Note that since the S_k shuffle with boundaries $(\zeta_t)_{t\geq 0}$ is a transitive Markov chain – see Section 2.6.2 in Levin et al. (2009) – it suffices to bound the distance from the stationary distribution when starting from the identity. Since the Dirac measure on the identity is increasing with respect to the partial order $\succeq$ from (2.8), using the censoring inequality Lemma 3.2 for the first step, and Lemma 4.3 in Lacoin (2016) for the second step, we obtain that

$$\|\mathbb{P}(\zeta_{t_3} \in \cdot \,|\, \zeta_0 = \mathrm{id}) - \mu\|_{\mathrm{TV}} \leq \|\nu_{t_3} - \mu\|_{\mathrm{TV}} \leq \|\widehat{\nu}_{t_3} - \widehat{\mu}\|_{\mathrm{TV}} + \|\widetilde{\nu}_{t_3} - \nu_{t_3}\|_{\mathrm{TV}} \,. \tag{6.15}$$

As $\delta > 0$ for t_3 was arbitrary, we combine Propositions 6.2 and 6.3 to conclude. □

6.3. *Proof of Proposition 6.2.* In the following, we will only describe the necessary changes in the proof of Proposition 5.1 in Lacoin (2016) in order to obtain Proposition 6.2 for the S_k shuffle with boundaries, and refer to Section 5.3 of Lacoin (2016) for a detailed proof of the corresponding result for the S_2 shuffle. Note that as we start from the Dirac measure on the identity δ_{id}, the measure $\widetilde{\delta_{\mathrm{id}}}$ can be identified with the product measure $\bigotimes_{i=1}^{K} \mu_{\Delta x_i}$ on $\bigotimes_{i=1}^{K} \mathcal{S}_{\Delta x_i}$, where we recall (6.11) and that μ_i denotes the uniform distribution on $\mathcal{S}_i$. By our choice of the censoring scheme $\mathcal{C}$, note that the measure ν_t for $t < t_1$ corresponds to the law of K many S_k shuffles on $\bigotimes_{i=1}^{K} \mathcal{S}_{\Delta x_i}$ with laws $(\nu_t^i)_{t\geq 0}$ and suitable boundary parameters $(\delta_i^{(k)})$, satisfying the assumptions of Proposition 5.1, i.e. we have $\delta_i^{(k)} \in [0,1]$ for all $\mathcal{S}_{\Delta x_j}$ with $j \in \{2, \dots, K-1\}$. Using the canonical coupling from Section 5 and Proposition 5.1, we choose $K = K(\delta)$ large enough such that for all N sufficiently large

$$\|\nu_t - \widetilde{\nu}_{t_1}\|_{\mathrm{TV}} \leq \sum_{i=1}^{K} \|\nu_t^i - \mu_i\|_{\mathrm{TV}} \leq \sum_{i=1}^{K} \|\nu_{t_1}^i - \mu_i\|_{\mathrm{TV}} \leq \frac{\varepsilon}{3} \,. \tag{6.16}$$

6.4. *Proof of Proposition 6.3.* In order to show Proposition 6.3, we first state a bound on the skeleton projection $\bar{\nu}_t$.

Proposition 6.4 (Proposition 5.2 in Lacoin (2016))**.** *For all $\varepsilon > 0$ and $k = o(N^{1/6})$, there exists some $N_2 = N_2(\varepsilon)$ such that for all $N \geq N_2$ and $t \geq t_2$*

$$\|\bar{\nu}_t - \bar{\mu}_t\|_{\mathrm{TV}} \leq \varepsilon/3 \,. \tag{6.17}$$

Proof: Let ν be a probability measure on $\mathcal{S}_N$, which is increasing with respect to the partial order $\succeq$. Lemma 5.5 in Lacoin (2016) states that (6.17) for some $\varepsilon > 0$ follows whenever, for some sufficiently small $\kappa(\varepsilon, K)$, one can show that

$$\mathbb{E}_\nu \left[\sum_{i=1}^{K-1} \sum_{j=i}^{K-1} h_\sigma(x_i, x_j) \right] < \kappa \sqrt{N} \,, \tag{6.18}$$

Here $\mathbb{E}_\nu$ denotes the expectation with respect to ν, and we let $\sigma \sim \nu$. Choose $\nu = \nu_{t_2}$ and note that ν_{t_2} is increasing by Lemma 3.2. As $k = o(N^{1/6})$, we get from Lemma 6.1 that starting the S_k shuffle with boundaries from ν_{t_1}, for any $\kappa > 0$ and all N sufficiently large

$$\mathbb{E}_\nu \left[\sum_{i=1}^{K-1} \sum_{j=i}^{K-1} h_\sigma(x_i, x_j) \right] \leq (K-1)^2 (2N e^{-(t_2 - t_1)\lambda_{N,k}} + ck^3)) \leq \kappa \sqrt{N} \,, \tag{6.19}$$

allowing us to conclude. □

It remains to deduce Proposition 6.3 from Proposition 6.4. As this follows along the same lines as the proof of Proposition 5.3 in Lacoin (2016), we will highlight only the required adjustments. Let $\sigma_t \sim \nu_t$ be the configuration of the S_k shuffle at time t, and observe that the skeleton $\bar{\sigma}_{t_2}$ remains unchanged for times $t \in [t_2, t_3]$ by our choice of the censoring scheme $\mathcal{C}$. Conditioning on the event

$\{\sigma_{t_2} = \xi\}$ for some $\xi \in \mathcal{S}_N$, let $\mu_{\bar{\xi}}$ be the uniform distribution on set of permutations with skeleton $\bar{\xi}$. From the definition of the semi-skeleton for the first step, and Proposition 5.1 together with the same reasoning as in Proposition 6.2 of decomposing the S_k shuffle into K independent $S_{\Delta x_i}$ shuffles for $i \in [K]$ in the second step,

$$\|\hat{\nu}_{t_3}(\,\cdot\,|\bar{\sigma}_{t_2} = \bar{\zeta}) - \hat{\mu}(\,\cdot\,|\bar{\sigma}_{t_2} = \bar{\zeta})\|_{\mathrm{TV}} \leq \max_{\xi \in \mathcal{S}_N} \|\mathbb{P}(\sigma_{t_3} \in \cdot \,|\, \sigma_{t_2} = \xi) - \mu_{\bar{\xi}}\|_{\mathrm{TV}} \leq \frac{\varepsilon}{3} \tag{6.20}$$

for all $\zeta \in \mathcal{S}_N$ and N sufficiently large; see also equation (5.38) in Lacoin (2016). Since we have that

$$2\|\hat{\nu}_{t_3} - \hat{\mu}\|_{\mathrm{TV}} \leq 2\left(\|\bar{\nu}_t - \bar{\mu}_t\|_{\mathrm{TV}} + \sum_{\bar{\zeta} \in \bar{\mathcal{S}}_N} \bar{\nu}_{t_3}(\bar{\zeta})\|\hat{\nu}_{t_3}(\,\cdot\,|\bar{\sigma}_{t_2} = \bar{\zeta}) - \hat{\mu}(\,\cdot\,|\bar{\sigma}_{t_2} = \bar{\zeta})\|_{\mathrm{TV}}\right) \tag{6.21}$$

by equation (5.39) in Lacoin (2016), we get Proposition 6.3 by combining Proposition 6.4 and (6.20).

7. Comparison between the S_k shuffle with and without boundaries

In Theorem 1.1, we saw for $k = o(N^{2/3})$ that the S_k shuffle exhibits **pre-cutoff**, i.e. the ratio between the ε-mixing time and $N^2k^{-3}\log(N)$ is bounded for $N \to \infty$ from below and above by positive constants, which do not depend on $\varepsilon \in (0,1)$. While the S_k shuffle with boundaries also exhibits pre-cutoff when $k = o(N^{2/3})$, we argue that for $k = N^\alpha$ with $\alpha \in (\frac{2}{3}, \frac{3}{4})$ the behavior of the S_k shuffle and the S_k shuffle with boundaries is fundamentally different due to a different treatment of the cards near the boundaries.

Proposition 7.1. *For all $k = o(N^{3/4})$ and $\varepsilon \in (0,1)$, the mixing time $t'_{\mathrm{mix}}(\varepsilon)$ of the S_k shuffle with boundaries satisfies*

$$\frac{6}{\pi^2} \leq \liminf_{N\to\infty} \frac{k(k^2-1)\cdot t'_{\mathrm{mix}}(\varepsilon)}{N^2\log(N)} \leq \limsup_{N\to\infty} \frac{k(k^2-1)\cdot t'_{\mathrm{mix}}(\varepsilon)}{N^2\log(N)} \leq c \tag{7.1}$$

for some constant $c > 0$. In particular, pre-cutoff occurs. For $k = N^\alpha$ with $\alpha \in (\frac{2}{3}, 1)$, the mixing time $t_{\mathrm{mix}}(\varepsilon)$ of the S_k shuffle is of constant order where

$$\lim_{\varepsilon\to 0}\liminf_{N\to\infty} t_{\mathrm{mix}}(\varepsilon) = \infty\,, \tag{7.2}$$

and for any fixed $\varepsilon > 0$, there exist some $C = C(\varepsilon) > 0$ such that

$$\limsup_{N\to\infty} t_{\mathrm{mix}}(\varepsilon) \leq C\,. \tag{7.3}$$

In particular, pre-cutoff does not occur.

In order to show Proposition 7.1, we require some setup. Set $\alpha' = \frac{1}{24}(3\alpha - 2) > 0$, and recall from Section 5 that we denote by $(Z_{i,t})_{t\geq 0}$ and $(Z'_{i,t})_{t\geq 0}$ the position of the cards of label i in two S_k shuffles under the canonical coupling $\mathbf{P}$. For $i \in [N]$ and $T \geq 0$, we define

$$B_{i,T} := \left\{ Z_{i,t} \in \left[N^{\frac{1}{3}+\alpha'}, N - N^{\frac{1}{3}+\alpha'}\right] \text{ for all } t \in \left[T, T + N^{-\alpha'}\right] \right\}\,, \tag{7.4}$$

and similarly define $B'_{i,T}$ with respect to $(Z'_{i,t})_{t\geq 0}$. We have the following result on the coalescence time τ_i of the cards of label $i \in [N]$ under the canonical coupling $\mathbf{P}$.

Lemma 7.2. *Let $\alpha > \frac{2}{3}$ and $i \in [N]$. Then for all $T \geq 0$ and N sufficiently large*

$$\mathbf{P}\big(B_{i,T} \,\big|\, Z_{i,T} \in \left[N^{\frac{1}{3}+\alpha'}, N - N^{\frac{1}{3}+\alpha'}\right]\big) \geq 1 - N^{-3\alpha'}\,, \tag{7.5}$$

and similarly for the events $B'_{i,T}$. Moreover, we have that for all N sufficiently large

$$\mathbf{P}\big(\tau_i > T + N^{-\alpha'} \,\big|\, B_{i,T} \cap B'_{i,T}\big) \leq N^{-2}\,. \tag{7.6}$$

Proof: For the first statement (7.5), note that $Z_{i,t} < N^{\frac{1}{3}+\alpha'}$ for some $t \geq T$ can only occur by an update of an interval $[j, j+k-1]$ for some $j < N^{\frac{1}{3}+\alpha'}$. Let X_T be the total number of updates of these intervals between time T and $T+N^{-\alpha'}$, and note that each such update places the card of label i in the first $N^{\frac{1}{3}+\alpha'}$ positions independently with probability at most $N^{\frac{1}{3}+\alpha'-\alpha}$. As X_T is Poisson-($N^{1/3}$)-distributed, we have that $\mathbf{P}(X_T \geq N^{\frac{1}{3}+\alpha'}) \leq \frac{1}{4}N^{-3\alpha'}$ uniformly in $T>0$, and for all N large enough. Hence

$$\begin{aligned}\mathbf{P}\big(Z_{i,T} > N^{\frac{1}{3}+\alpha'} \text{ for all } t \in [T, T+N^{-\alpha'}] \,\big|\, Z_{i,T} \in \big[N^{\frac{1}{3}+\alpha'}, N - N^{\frac{1}{3}+\alpha'}\big]\big) \\ \geq \mathbf{P}(X_T \geq N^{\frac{1}{3}+\alpha'}) + (1 - N^{\frac{1}{3}+\alpha'-\alpha})^{N^{\frac{1}{3}+\alpha'}} \geq 1 - \frac{1}{2}N^{-3\alpha'}\end{aligned}$$

for all N sufficiently large. A similar statement for $Z_{i,t} < N - N^{\frac{1}{3}+\alpha'}$ gives the desired lower bound on the probability of $B_{i,T}$. For the second statement (7.6), we recall (5.12) in the proof of Lemma 5.5 which implies that for any starting position $> N^{\frac{1}{3}+\alpha'}$, the expected time to reach some position $> k$ is of order at most $N^{-\frac{1}{3}+2\alpha'}$. Using the same arguments as in the proof of Proposition 5.1, together with the fact that

$$\max\left(N^{-\frac{1}{3}+\alpha'}\frac{N}{k}, \frac{N^2}{k^3}\right) \leq N^{-2\alpha'} \tag{7.7}$$

for all N large enough, we see that there exist constants $c_1, c_2 > 0$ such that

$$\mathbf{P}\big(\tau_i > T + c_1 N^{-2\alpha'} \,\big|\, B_{i,T} \cap B'_{i,T}\big) \leq 1 - c_2\,. \tag{7.8}$$

Iterating (7.8) now $N^{\alpha'}$ many times gives the desired result. □

Proof of Proposition 7.1: Note that we obtain (7.1) by combining Remark 4.3 and Proposition 5.1. The lower bound in (7.2) on the mixing time of the S_k shuffle follows from observing that the ε-mixing time is bounded from below by the time T it takes such that at least one of the cards initially at positions 1 or N has moved with probability at least $1-\varepsilon$ until time T. Hence, it remains to prove (7.3). Using (7.6) in Lemma 7.2, it suffices to show that with probability at least $1-\varepsilon/2$, there exists some $n \in \mathbb{N}$, some constant $C = C(\varepsilon, n) > 0$, and a sequence of non-negative times $(T_1, T_2, \dots, T_n)$ such that

$$T_{j+1} - T_j > N^{-\alpha'} \text{ for all } j \in [n-1] \quad \text{ and } \quad T_n \leq C\,, \tag{7.9}$$

and with the property that for every $i \in [N]$, at least one of events B_{i,T_j} occurs for some $j = j(i) \in [n]$. In order to define these times $(T_j)_{j\in[n]}$, consider for all $m \in \mathbb{N}$ the event D_m that during the time interval $[2m-1, 2m]$, both intervals $[k]$ and $\{N-k+1, \dots, N\}$ receive an update. Let $(\mathcal{F}_t)_{t\geq 0}$ denote the natural filtration under the coupling $\mathbf{P}$. Then

$$\mathbf{P}(D_m \,|\, \mathcal{F}_{2m-1}) \geq \frac{1}{4} \quad \text{ and } \quad \mathbf{P}(B_{i,2m} \cap B'_{i,2m} \,|\, D_m, \mathcal{F}_{2m-1}) \geq 1 - 2N^{-\alpha'} \tag{7.10}$$

for all $m \in \mathbb{N}$, uniformly in $i \in [N]$, iterating (7.5) of Lemma 7.2 for the second statement. Let T_j be the j^{th} time that the event D_m occurs. Choosing $C > 0$ sufficiently large, we see that for any fixed $n \in \mathbb{N}$, the times $(T_i)_{i\in[n]}$ satisfy (7.9) with probability at least $1-\varepsilon/4$. To ensure that with probability at least $1-\varepsilon/4$, for every $i \in [N]$ at least one of events B_{i,T_j} occurs for some $j \in [n]$, we set $n = 4/\alpha'$. From (7.10), we see that for every fixed $i \in [N]$, with probability at least $1 - N^{-2}$, the event $B_{i,T_j} \cap B'_{i,T_j}$ holds for some $j \in [n]$, and all N large enough. Together with a union bound over all $i \in [N]$, this finishes the proof. □

We conclude with a few conjectures on the mixing time of the S_k shuffle.

Conjecture 7.3. *Let $k = o(N^{1/2})$. Then for all $\varepsilon \in (0,1)$, we have that*

$$\lim_{N\to\infty} \frac{k(k^2-1)\cdot t_{\mathrm{mix}}(\varepsilon)}{N^2\log(N)} = \lim_{N\to\infty} \frac{k(k^2-1)\cdot t'_{\mathrm{mix}}(\varepsilon)}{N^2\log(N)} = \frac{6}{\pi^2}\,, \tag{7.11}$$

i.e. the S_k shuffle with and without boundary exhibits cutoff.

In Lacoin (2016), the author also proves a separation cutoff result, where the separation distance of a measure μ on $\mathcal{S}_N$ from the uniform measure μ_N is defined as

$$\max_{A\in\mathcal{S}_N}\left[1 - \frac{\nu(A)}{\mu_N(A)}\right] \tag{7.12}$$

Thus, defining $t_{\mathrm{sep}}(\varepsilon)$ analogously to 2.6, we conjecture the following separation cutoff result should hold.

Conjecture 7.4. *Let $k = o(N^{1/2})$. Then for all $\varepsilon \in (0,1)$, we have that*

$$\lim_{N\to\infty} \frac{k(k^2-1)\cdot t_{\mathrm{sep}}(\varepsilon)}{N^2\log(N)} = \lim_{N\to\infty} \frac{k(k^2-1)\cdot t'_{\mathrm{sep}}(\varepsilon)}{N^2\log(N)} = \frac{12}{\pi^2}\,, \tag{7.13}$$

i.e. the S_k shuffle with and without boundary exhibits separation cutoff at twice the total variation mixing time.

Acknowledgements

We thank Jonathan Hermon for helpful comments. Moreover, we thank the anonymous referee for valuable suggestions which significantly helped to improve the paper.

References

Alameda, J. S., Bang, C., Brennan, Z., Herzog, D. P., Kritschgau, J., and Sprangel, E. Cutoff in the Bernoulli-Laplace urn model with swaps of order n. *Electron. Commun. Probab.*, **29**, Paper No. 2, 13 (2024). MR4696437.

Bayer, D. and Diaconis, P. Trailing the dovetail shuffle to its lair. *Ann. Appl. Probab.*, **2** (2), 294–313 (1992). MR1161056.

Benjamini, I., Berger, N., Hoffman, C., and Mossel, E. Mixing times of the biased card shuffling and the asymmetric exclusion process. *Trans. Amer. Math. Soc.*, **357** (8), 3013–3029 (2005). MR2135733.

Berestycki, N., Schramm, O., and Zeitouni, O. Mixing times for random k-cycles and coalescence-fragmentation chains. *Ann. Probab.*, **39** (5), 1815–1843 (2011). MR2884874.

Bernstein, M. and Nestoridi, E. Cutoff for random to random card shuffle. *Ann. Probab.*, **47** (5), 3303–3320 (2019). MR4021252.

Blanca, A., Caputo, P., Chen, Z., Parisi, D., Štefankovič, D., and Vigoda, E. On mixing of Markov chains: coupling, spectral independence, and entropy factorization. *Electron. J. Probab.*, **27**, Paper No. 142, 42 (2022). MR4505379.

Blanca, A., Chen, Z., Štefankovič, D., and Vigoda, E. The Swendsen-Wang dynamics on trees. *Random Structures Algorithms*, **62** (4), 791–831 (2023). MR4597159.

Borcea, J. and Brändén, P. The Lee-Yang and Pólya-Schur programs. I. Linear operators preserving stability. *Invent. Math.*, **177** (3), 541–569 (2009). MR2534100.

Borcea, J., Brändén, P., and Liggett, T. M. Negative dependence and the geometry of polynomials. *J. Amer. Math. Soc.*, **22** (2), 521–567 (2009). MR2476782.

Diaconis, P. and Shahshahani, M. Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete*, **57** (2), 159–179 (1981). MR626813.

Diaconis, P. and Shahshahani, M. Time to reach stationarity in the Bernoulli-Laplace diffusion model. *SIAM J. Math. Anal.*, **18** (1), 208–218 (1987). MR871832.

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

Ganguly, S. and Seo, I. Information percolation and cutoff for the random-cluster model. *Random Structures Algorithms*, **57** (3), 770–822 (2020). MR4144084.

Gantert, N., Nestoridi, E., and Schmid, D. Mixing times for the simple exclusion process with open boundaries. *Ann. Appl. Probab.*, **33** (2), 972–1012 (2023). MR4564424.

Guo, H. and Jerrum, M. Random cluster dynamics for the Ising model is rapidly mixing. *Ann. Appl. Probab.*, **28** (2), 1292–1313 (2018). MR3784500.

Knöpfel, H., Löwe, M., Schubert, K., and Sinulis, A. Fluctuation results for general block spin Ising models. *J. Stat. Phys.*, **178** (5), 1175–1200 (2020). MR4081224.

Lacoin, H. Mixing time and cutoff for the adjacent transposition shuffle and the simple exclusion. *Ann. Probab.*, **44** (2), 1426–1487 (2016). MR3474475.

Levin, D. A., Peres, Y., and Wilmer, E. L. *Markov chains and mixing times.* American Mathematical Society, Providence, RI (2009). ISBN 978-0-8218-4739-8. MR2466937.

Liggett, T. M. *Continuous time Markov processes. An introduction*, volume 113 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI (2010). ISBN 978-0-8218-4949-1. MR2574430.

Long, Y., Nachmias, A., Ning, W., and Peres, Y. A power law of order 1/4 for critical mean field Swendsen-Wang dynamics. *Mem. Amer. Math. Soc.*, **232** (1092), vi+84 (2014). MR3243141.

Martinelli, F. *Lectures on Glauber Dynamics for Discrete Spin Models*, pp. 93–191. Springer Berlin Heidelberg, Berlin, Heidelberg (1999). ISBN 978-3-540-48115-7. DOI: 10.1007/978-3-540-48115-7_2.

Martinelli, F., Sinclair, A., and Weitz, D. The Ising model on trees: boundary conditions and mixing time. In *44th Annual IEEE Symposium on Foundations of Computer Science, 2003. Proceedings.*, pp. 628–639 (2003). ISBN 0-7695-2040-5. DOI: 10.1109/SFCS.2003.1238235.

Nam, D. and Nestoridi, E. Cutoff for the cyclic adjacent transposition shuffle. *Ann. Appl. Probab.*, **29** (6), 3861–3892 (2019). MR4047994.

Nestoridi, E. The Limit Profile of Star Transpositions. *ArXiv Mathematics e-prints* (2021). arXiv: 2111.03622.

Peres, Y. and Winkler, P. Can extra updates delay mixing? *Comm. Math. Phys.*, **323** (3), 1007–1016 (2013). MR3106501.

Salez, J. Universality of cutoff for exclusion with reservoirs. *Ann. Probab.*, **51** (2), 478 – 494 (2023). DOI: 10.1214/22-AOP1600.

Tran, H.-Q. Cutoff for the non reversible SSEP with reservoirs. *Electron. J. Probab.*, **28**, 1 – 24 (2023). DOI: 10.1214/23-EJP1044.

Wilson, D. B. Mixing times of Lozenge tiling and card shuffling Markov chains. *Ann. Appl. Probab.*, **14** (1), 274–325 (2004). MR2023023.

Yang, S. Cutoff and dynamical phase transition for the general multi-component Ising model. *J. Stat. Phys.*, **190** (9), Paper No. 151, 43 (2023). MR4635907.

Zhang, L. Cutoff profile of the Metropolis biased card shuffling. *Ann. Probab.*, **52** (2), 713–736 (2024). MR4718404.