Cutoff for the Biased Random Transposition Shuffle

Alan Yan
(joint w/ Evita Nestoridi)
Harvard University

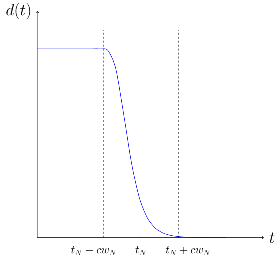
What is the Cutoff Phenomenon?

For most (discrete time) Markov chains, if we let the chain run forever it converges to a unique stationary distribution. For example, think about shuffling a deck of cards over and over again. As you shuffle the deck more and more, it becomes closer and closer to uniform.

Consider a (sequence of) Markov chain(s) $\left(X_t^{(n)}\right)_{t \geq 0}$ with transition matrix $P^{(n)}$. Let $d_n(t)$ be the distance between the chain at time t from the stationary distribution. We say that the chain has **cut-off** if there is some time t_n and window $w_n = o(t_n)$ such that

$$\lim_{c \to \infty} \limsup_{n \to \infty} d(t_n + cw_n) = 0,$$
$$\lim_{c \to -\infty} \liminf_{n \to \infty} d(t_n + cw_n) = 1.$$

If we look at the graph of $d_n(t)$ as a function in time for large n, the cutoff phenomenon can be thought of as the following behavior for the graph of $d(t)$:



Thus, the cutoff phenomenon is about the part of the graph outside of the window: it is equivalent to the outer part converging to a step function. A harder question concerns inside the window. As $n \to \infty$, what is $d(t_n + cw_n)$ as a function in c? This is the **limit profile** problem and is much less understood.

In this project, we study a generalization of the random transpositions shuffle, the first example of card shuffle studied in [1] using techniques from representation theory. We prove that it exhibits cutoff and other convergence results about the shuffle.

The Model: Biased Transposition Shuffle

Consider a deck of distinct cards $\{1, 2, \ldots, N\}$ where $N = 2n$ and half of the cards are colored red and half of the cards are colored blue. We fix parameters $0 \leq b \leq a \leq 2$ with $a + b = 2$. We can give a *weight* to the cards such that

$$\text{wt(red card)} = \frac{a}{N}, \quad \text{wt(blue card)} = \frac{b}{N}.$$

Note that wt$(\cdot)$ is a probability distribution on the cards. We shuffle our deck of cards in the following way:

- Pick cards $C_1, C_2 \in \{1, 2, \ldots, 2n\}$ independently based off of the weight wt$(\cdot)$.
- Swap the cards labeled C_1 and C_2.

We view the card shuffle as a Markov chain on the *symmetric group* $\mathfrak{S}_{2n}$. A shuffle is akin to picking a random permutation and then applying that permutation to the current state. For the random permutation, we get the probabilities:

$$\text{Pr}[(ij)] = \frac{2a^2}{N^2}, \quad \text{Pr}[(ij)] = \frac{2ab}{N^2}, \quad \text{Pr}[(ij)] = \frac{2b^2}{N^2}, \quad \text{Pr(id)} = \frac{n(a^2+b^2)}{N^2}.$$

We call this shuffle the **biased random transpositions shuffle**. When $a = b = 1$, we recover the random transposition shuffle from [1].

Main Results

Theorem. *The biased random transposition shuffle exhibits cutoff at time $t_n = \frac{1}{2b} N \log N$ with window $w_n = N$. Explicitly, we have the bounds*

$$d_N\left(\frac{1}{2b}N(\log N + c)\right) \leq C \cdot e^{-c}, \quad \textit{and}$$
$$d_N\left(\frac{1}{2b}N(\log N - c)\right) \geq 1 - e^{-\frac{1}{2}\left(\sqrt{1+\frac{1}{2}e^c} - 1\right)^2} + o(1)$$

for some $C > 0$ and all $c > 0$.

Theorem. *Let Fix_c be the number of fixed points after $\frac{1}{2b}N(\log N - c)$ shuffles. Then, if $a = b = 1$, then*

$$\text{Fix}_c \xrightarrow{\text{dist}} \text{Poiss}(1 + e^c)$$

and when $a > b$, then

$$\text{Fix}_c \xrightarrow{\text{dist}} \text{Poiss}\left(1 + \frac{e^c}{2}\right)$$

The limiting behavior for the fixed points gives us a lower bound for the total variation distance. Indeed, we immediately have

$$d_N\left(\frac{1}{2b}N(\log N - c)\right) \geq d_{\text{TV}}(\text{Fix}_c, \text{Fix})$$
$$= d_{\text{TV}}\left(\text{Poiss}(1), \text{Poiss}\left(1 + \frac{e^c}{2}\right)\right) + o(1).$$

This gives a reasonable guess for the limit profile. In the random transposition case, the number of fixed points governs the convergence. It is not a stretch to assume that the same holds in the biased random transposition case.

Representation Theory of the Shuffle

In general, we can consider the same shuffle but with an arbitrary decomposition $[N] = A \sqcup B$ into red and blue cards. In this general setting, we require that our parameters $0 \leq b \leq a$ satisfy

$$\frac{a|A| + b|B|}{N} = 1.$$

The convergence rate is governed by the eigenvalues of the transition matrix. To calculate the eigenvalues of this $N! \times N!$ matrix, we rephrase the problem into representation theory.

Fact. *The eigenvalues of P are the same as the eigenvalues of the linear operator given by left multiplication on $\mathbb{C}[\mathfrak{S}_N]$ by the element*

$$\mathscr{A} := \left(\frac{a^2|A| + b^2|B|}{N^2}\right) \cdot \text{id} + \frac{2a^2}{N^2}\mathscr{T}_A + \frac{2b^2}{N^2}\mathscr{T}_B + \frac{2ab}{N^2}\mathscr{T}_{A,B} \tag{1}$$
$$= \left(\frac{a^2|A| + b^2|B|}{N^2}\right) \cdot \text{id} + \frac{2(a^2 - ab)}{N^2}\mathscr{T}_A + \frac{2(b^2 - ab)}{N^2}\mathscr{T}_B + \frac{2ab}{N^2}\mathscr{T}_{A \cup B}. \tag{2}$$

Note that $\mathscr{A} \notin Z(\mathbb{C}[\mathfrak{S}_N])$, but $\mathscr{A} \in Z(\mathbb{C}[\mathfrak{S}_N]) + Z(\mathbb{C}[\mathfrak{S}_A]) + Z(\mathbb{C}[\mathfrak{S}_B])$. Thus we can diagonalize after restricting the module structure.

Theorem. *The transition matrix P has eigenvalues*

$$\text{Eig}^{\lambda}_{\mu,\nu} := \frac{a^2|A| + b^2|B|}{N^2} + \frac{2(a^2 - ab)}{N^2}\text{Diag}(\mu) + \frac{2(b^2 - ab)}{N^2}\text{Diag}(\nu) + \frac{2ab}{N^2}\text{Diag}(\lambda),$$

with multiplicities $f_\lambda f_\mu f_\nu c^{\lambda}_{\mu,\nu}$ for all partitions $\lambda \vdash N$, $\mu \vdash |A|$, and $\nu \vdash |B|$. Here, f_λ is the number of standard Young tableaux of shape λ and $c^{\lambda}_{\mu,\nu}$ is the Littlewood-Richardson coefficient.

For a definition of the quantity $\text{Diag}(\lambda)$, see the *Remarks*.

Remarks

Diagonalization of the transition matrix relies on decomposing $\mathfrak{S}_N$ modules into irreducible $\mathfrak{S}_A \times \mathfrak{S}_B$ modules. To this end, we can use the **Littlewood-Richardson rule**, which gives

$$\text{Res}^{\mathfrak{S}_N}_{\mathfrak{S}_A \times \mathfrak{S}_B} S^{\lambda} = \bigoplus_{\mu,\nu} (S^{\mu} \boxtimes S^{\nu})^{\oplus c^{\lambda}_{\mu,\nu}}.$$

This fact along with Schur's lemma immediately gives us the eigenvalue spectrum. The quantity $\text{Diag}(\lambda)$ is related (via Schur-Weyl duality) to characters of the symmetric group via the equation

$$\frac{\chi(\tau)}{\chi(\text{id})} = \frac{\text{Diag}(\lambda)}{\binom{N}{2}}, \quad \text{Diag}(\lambda) := \sum_{(i,j) \in \text{YD}(\lambda)} (j - i).$$

The eigenvalues of the transition matrix give us an upper bound on the cutoff time. Indeed, from standard bounds in mixing times, we have the bound

$$d(t)^2 \leq \frac{1}{4} \sum_{\substack{\lambda \vdash N \\ \lambda \neq (N)}} \sum_{\lambda : c^{\lambda}_{\mu,\nu} > 0} c^{\lambda}_{\mu,\nu} f_\lambda f_\mu f_\nu \cdot |\text{Eig}^{\lambda}_{\mu,\nu}|^{2t}.$$

Getting a suitable bound on this quantity requires understanding of the *positive cone* $\{(\lambda, \mu, \nu) : c^{\lambda}_{\mu,\nu} > 0\}$. This can be understood, for example, through the theory of **honeycombs** or **hives** [2].

Conjecture for Limit Profile

Conjecture (Limit profile)**.** *Let $a > b$, and let $d_n(c)$ be the total variation distance from uniform after $\frac{1}{2b}N(\log N - c)$ shuffles. Then,*

$$\lim_{n \to \infty} d_n(c) = d_{\text{TV}}\left(\text{Poiss}(1), \text{Poiss}\left(1 + \frac{e^c}{2}\right)\right)$$

For $a = b = 1$, we have the following theorem from [3].

Theorem (Teyssier)**.** *Let $a = b = 1$, and let $d_n(c)$ be the total variation distance from uniform after $\frac{1}{2}N(\log N - c)$ shuffles. Then,*

$$\lim_{n \to \infty} d_n(c) = d_{\text{TV}}\left(\text{Poiss}(1), \text{Poiss}(1 + e^c)\right)$$

Acknowledgements

The authors thank Ezra Edelman, Lauren Williams, Johnny Gao, Laura Colmenarejo, and Grant Barkley for enlightening conversations. The authors thank the anonymous reviewers for their helpful comments.

References

[1] P. Diaconis and M. Shahshahani. "Generating a random permutation with random transpositions". In: *Z. Wahrsch. Verw. Gebiete* 57.2 (1981), pp. 159–179. DOI: 10.1007/BF00535487.
[2] A. Knutson and T. Tao. "The honeycomb model of $\text{GL}_n(\mathbf{C})$ tensor products. I. Proof of the saturation conjecture". In: *J. Amer. Math. Soc.* 12.4 (1999), pp. 1055–1090. DOI: 10.1090/S0894-0347-99-00299-4.
[3] L. Teyssier. "Limit profile for random transpositions". In: *Ann. Probab.* 48.5 (2020), pp. 2323–2343. DOI: 10.1214/20-AOP1424.