Analysis of Top-Swap Shuffling for Genome Rearrangements

Nayantara Bhatnagar *

College of Computing, Georgia Institute of Technology
Atlanta, GA 30332
e-mail: nand@cc.gatech.edu

Pietro Caputo †

Dipartimento di Matematica, Università di Roma Tre
Largo S. Murialdo 1, 00146 Rome, Italy
e-mail: caputo@mat.uniroma3.it

Prasad Tetali‡

School of Mathematics and College of Computing
Georgia Institute of Technology
Atlanta, GA 30332
e-mail: tetali@math.gatech.edu

Eric Vigoda *

College of Computing, Georgia Institute of Technology
Atlanta, GA 30332
e-mail: vigoda@cc.gatech.edu

Abstract: We study Markov chains which model genome rearrangements. These models are useful for studying the equilibrium distribution of chromosomal lengths, and are used in methods for estimating genomic distances. The primary Markov chain studied in this paper is the top-swap Markov chain. The top-swap chain is a card-shuffling process with n cards divided over k decks, where the cards are ordered within each deck. A transition consists of choosing a random pair of cards, and if the cards lie in different decks, we cut each deck at the chosen card and exchange the tops of the two decks. We prove precise bounds on the relaxation time (inverse spectral gap) of the top-swap chain. In particular, we prove the relaxation time is $\Theta(n + k)$. This resolves an open question of Durrett.

AMS 2000 subject classifications: Primary 60J27; secondary 92D10.
Keywords and phrases: card shuffling, genome rearrangement, random transpositions, relaxation time.

1

1. Introduction

Genome rearrangements play an important role in a variety of biological studies, e.g., genomic distance [7, 8], phylogenetic studies [10], and cancer biology [6]. Rearrangements refer to chromosal fissions and fusions, reciprocal translocations between chromosomes, and inversions within a chromosome. Fissions, fusions, and reciprocal translocations are examples of translocations. In this paper we study Markov chains which model genomic rearrangements by translocations and inversions.

Stochastic models of chromosomal rearrangements by translocations were introduced by Sankoff and Feretti [9]. Such models are useful for studying the equilibrium distribution of chromosomal lengths. De, Durrett, Ferguson and Sindi [3] studied refinements of the models, including the introduction of a fitness function on chromosomal lengths, and showed that these models have a reasonable fit with data for many species. These models were used by Durrett et al [5] in a Bayesian approach for estimating genomic distance.

Various of the above models being simply Markov chains, Durrett [4] considered estimating the rate of convergence to stationarity of the corresponding Markov chains. One of the open problems raised by Durrett pertains to the analysis of, what we call henceforth, the *top-swap* chain. We study the so-called *relaxation time* of the chain, which is the inverse of the spectral gap of the transition matrix of the chain. The relaxation time is the key quantity in the rate of convergence, and hence it is of utmost importance in the efficiency of any simulations of the model. Beyond its computational significance, the convergence rate also has biological significance since it addresses the rate at which genomes reach equilibrium.

The top-swap chain has two parameters: the number of chromosomes, denoted as k, and the number of genes, denoted as n. The chain can be viewed as a card shuffling problem with k decks. More precisely, the state space of the chain is a partition of n distinct cards into k decks. The cards have some ordering within each deck, and the decks are labeled. At each transition we choose two random positions, where if the size of the i-th deck is n_i there are $n_i + 1$ positions to choose from in the i-th deck. If the positions are in the same deck we do nothing. Otherwise we cut both decks at the chosen positions. We then exchange the tops of the two decks. The figure below illustrates a sample transition (for $k = 4, n = 12$) where the chosen pair of positions is marked:

$$
\begin{array}{ccccccccc}
 & & 9 & & & & & 4 & \\
 & & & & & & & 3 & \\
4 & 7 & \underline{2} & & & & 7 & 5 & \\
3 & 1 & 11 & 8 & \implies & 9 & 1 & 11 & 8 \\
\underline{5} & 10 & 6 & 12 & & 2 & 10 & 6 & 12
\end{array}
$$

The Markov chain allows empty decks, hence fusions and fissions are modeled by this stochastic process. Translocations with non-empty decks are known as reciprocal translocations. As we note below, inversions can also be included in the above

Markov chain.

Let $\tau(n, k)$ denote the inverse of the spectral gap of the top-swap chain. This is known as the relaxation time. We prove tight bounds, up to constant factors, on the spectral gap of the top-swap Markov chain.

Theorem 1.1. *There exist constants $C_1 > C_2 > 0$ such that:*

$$C_1(n + k) \geqslant \tau(n, k) \geqslant C_2(n + k).$$

Remark 1. Durrett [4] proposes a Markov chain which is identical to the top-swap chain, except when the chosen pair of positions is in the same deck, this interval of cards is inverted. The above upper bound on the relaxation time immediately applies to this chain, since comparison of Dirichlet forms shows that extra transitions can only decrease the relaxation time.

As a byproduct of our proof, we also obtain tight bounds on the relaxation time of a k-deck random transposition chain; this chain is a natural extension of the classical (1-deck) random transposition chain to k-decks, for $k > 1$. For a precise statement of the result, see Theorem 5.2.

Before giving a high-level description of our proof we need to introduce some notation. Let Ω denote the state space of the top-swap chain, and P denote its transition matrix. We let ν denote the uniform distribution over Ω. Since P is symmetric, nu is reversible. Therefore ν is the stationary distribution of the chain. For any $f : \Omega \to \mathbb{R}$, the Dirichlet form of the process is defined as

$$\mathcal{E}(f) = \frac{1}{2} \sum_{\sigma, \eta \in \Omega} \nu(\sigma) P(\sigma, \eta)(f(\sigma) - f(\eta))^2,$$

and the variance is

$$\begin{aligned} \mathrm{Var}(f) &= \sum_{\sigma \in \Omega} \nu(\sigma)(f(\sigma) - \mathbf{E}_\nu(f))^2 \\ &= \frac{1}{2} \sum_{\sigma, \eta \in \Omega} \nu(\sigma)\nu(\eta)(f(\sigma) - f(\eta))^2. \end{aligned}$$

We then have that the relaxation time is

$$\tau(n, k) = \sup_f \frac{\mathrm{Var}(f)}{\mathcal{E}(f)},$$

where the supremum is taken over all non-constant functions f.

The lower bound on the relaxation time is easy to show by taking f as the indicator function for whether the first deck is empty. We give a high-level sketch of the proof outline in Section 2 before presenting a detailed proof in Sections 3,4 and 5.

2. High-level Proof Description

The analysis of the spectral gap of the top-swap chain has two major parts. The first part shows that it suffices to analyze the spectral gap for the 2-deck version of top-swap. (The 2-deck version of top-swap is simply the top-swap chain with $k = 2$.) We then analyze the 2-deck top-swap chain in the second part of our proof. Within both parts of the analysis, our proof relies on comparison arguments with other auxiliary chains.

2.1. Reducing to two decks

We use 2 auxiliary chains: a weighted and unweighted *deck-averaging* process. The use of a deck-averaging process is similar to the proof approach of Cancrini, Caputo and Martinelli [1] for the analysis of the so-called *L*-reversal chain.

The deck-averaging process is a continuous time Markov chain where the state space is again all the possible partitions of the n cards into k decks. For each pair of decks there is an independent Poisson clock. In the unweighted process, all clocks have the same rate, whereas for the weighted process, the clock for decks i and j has rate $n_i + n_j$ where n_ℓ denotes the number of cards in deck ℓ. The proof approach of [1] immediately yields that the unweighted deck-averaging process has a spectral gap of 1.

Our goal is to express the Dirichlet form and variance of the top-swap chain in terms of the sum of the 2-deck projections of the top-swap chain. It is straightforward to bound the Dirichlet form in terms of a sum over 2-deck Dirichlet forms, i.e. where the configuration is fixed outside the two decks. Recall that the variance is independent of the chain. Thus we can bound the variance by looking at the spectral gap of the unweighted deck-averaging process. This leaves us with the Dirichlet form for the unweighted deck-averaging process. We then use a non-trivial comparison argument to obtain the Dirichlet form for the weighted deck-averaging process. Finally, it is straightforward to bound the Dirichlet form of the weighted deck-averaging process as a sum over 2-deck variances. The end result is a bound on the spectral gap of the (k-deck) top-swap chain in terms of the spectral gap of the 2-deck top-swap chain.

2.2. Analysis of the 2–Deck Top-swap

The basic idea is to compare the 2-deck top-swap with random transpositions. However, transpositions of two cards within the same deck are a problem. Roughly speaking, we can not efficiently "simulate" these transitions by top-swap transitions (see Remark 2 below for more details). Hence, we consider a transposition chain which only allows transpositions of cards in different decks. Moreover if one of the cards is at the top of either deck, then instead of a transposition, the chain does the corresponding top-swap transition. It is straightforward to compare this modified

transposition chain with the top-swap chain. To analyze this modified transposition chain, we compare to a final chain, which we call here the balanced-swap chain. The balanced-swap chain is also defined on two decks and has two types of transitions: swapping and rearranging. The swapping transition changes the size of the decks. In particular, we choose a random position in a deck, and move the cards above it, to the top of the other deck. The rearranging transition randomizes the card order while maintaining the current deck size. However we only perform the second transition if the decks are balanced, i.e., close to the same size. (This is made precise with the notion of δ-balanced, for appropriately chosen $0 < \delta < 1$.) It turns out that this later restriction is crucial for the comparison with the modified transposition chain.

Finally, we analyze the gap of the balanced-swap chain by adapting an analytical argument given by Cancrini, Martinelli, Roberto and Toninelli in their work on kinetically constrained models [2].

In the following we first describe the solution to the 2-deck problem in Section 3, then proceed with the deck-averaging process in Section 4, and finally in Section 5 we return to the analysis of the k-deck top-swap process. The matching upper bounds on the spectral gap are briefly discussed at the end of Section 5.

3. Analysis of the 2–Deck problem

We consider two decks I_1, I_2 and we let n_1, n_2 denote the number of cards in each deck. Also, let $n = n_1 + n_2$ denote the total number of cards and let Ω_n be the set of all $(n + 1)!$ arrangements of n labeled cards in the two decks. We denote by η a generic element of Ω_n (a configuration) and call μ the uniform probability measure on Ω_n.

The 2-deck top–swap Markov chain is described as follows. At each step we choose two positions r, s, with each position drawn at random from the $n + 2$ available positions, $n_1 + 1$ from the first deck and $n_2 + 1$ from the second deck. As explained in the introduction if r and s belong to the same deck we do nothing. Otherwise we swap the tops identified by positions r, s. The extra position added to each deck allows to swap an empty top. If the current configuration is η we denote by $\mathcal{T}_{r,s}\eta$ the updated configuration. The Dirichlet form of the Markov chain can then be written as

$$\mathcal{E}_2(f) = \frac{1}{2(n+2)^2} \sum_{r,s=1}^{n+2} \mu\left[(\mathcal{T}_{r,s}f - f)^2\right], \tag{3.1}$$

where $\mathcal{T}_{r,s}f(\eta) := f(\mathcal{T}_{r,s}\eta)$ for arbitrary functions $f : \Omega_n \to \mathbb{R}$, and $\mu[f]$ stands for expectation of a function f w.r.t. the uniform probability μ. The main results in this section is the following $O(n)$ estimate on the relaxation time $\tau(n, 2)$ of the Markov chain described above.

Theorem 3.1. *There exists $C > 0$ such that, for every $n \in \mathbb{N}$,*

$$\tau(n, 2) \leqslant C\,n\,. \tag{3.2}$$

Recall that the estimate (3.2) is equivalent to showing that for every function f:

$$\mathrm{Var}_\mu(f) \leqslant C\,n\,\mathcal{E}_2(f)\,. \tag{3.3}$$

We start by introducing a convenient notation. We add a card $*$ to mark the separation between the two decks and we call $x_* = x_*(\eta)$ the position of the $*$ in η, see Figure 1.
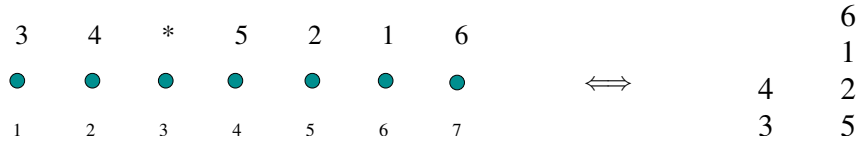

FIG 1. *Two decks with $n = 6$. The configuration $\eta = (3, 4, *, 5, 2, 1, 6)$, with $n_1 = 2, n_2 = 4$ and $x_* = 3$.*

With this representation the top–swap transformations $\mathcal{T}_{r,s}$ can be rewritten by means of the transformations $\eta \to T_{i,j}\eta$, $i, j = 1, \ldots, n+1$ described below.

3.1. Top–Swap operators $T_{i,j}$

If $i < x_* < j$ we call $T_{i,j}\eta$ the new configuration obtained by the ordinary top–swap. Namely, we collect the cards on top of j from the second deck (including j) in a deck D_j and the cards on top of i from the first deck (including i) in a deck D_i. We then swap their positions, namely D_j goes above the position $(i-1)$ in the first deck and D_i goes above the position $(j-1)$ in the second deck. Consider, for instance the configuration $\eta = (3, 4, *, 5, 2, 1, 6)$ given in Figure 1 with $x_* = 3$. If we pick $i = 1$ and $j = 5$ we obtain $T_{1,5}\eta = (2, 1, 6, *, 5, 3, 4)$. If we choose $i = 2$ and $j = 7$ then $T_{2,7}\eta = (3, 6, *, 5, 2, 1, 4)$.

If $i = x_* < j$ then we move the deck D_j above the position $(i - 1)$ in the first deck. Therefore in our example, if $i = x_* = 3$ and, say, $j = 6$, then $T_{3,6}\eta = (3, 4, 1, 6, *, 5, 2)$.

Also, if $i < x_* = j$ we move the deck D_i above the last card in the second deck. Thus, in our example if, say, $i = 1$ and $j = x_* = 3$ then $T_{1,3}\eta = (*, 5, 2, 1, 6, 3, 4)$.

Finally, if $i = j$ or $1 \leqslant i < j < x_*$ or $1 \leqslant x_* < i < j$ we do nothing, i.e. we define $T_{i,j}\eta = \eta$ in these cases. To finish the definition of $T_{i,j}$ for all $i, j = 1, \ldots, n+1$ we set $T_{j,i} := T_{i,j}$ for every $i \leqslant j$.

Note that with these definitions we may rewrite the Dirichlet form (3.1) as

$$\mathcal{E}_2(f) = \frac{1}{2(n+2)^2} \sum_{i,j=1}^{n+1} \mu\left[(T_{i,j}f - f)^2\right], \tag{3.4}$$

where $T_{i,j}f(\eta) := f(T_{i,j}\eta)$. Here we use the fact that the transformations T and $\mathcal{T}$ coincide when the chosen positions correspond to ordinary cards and are different if one of the tops to be swapped is empty. Note that the positions range from 1 to $n+2$ in (3.1) and from 1 to $n+1$ in (3.4). However, when one of the positions is the top of a deck in (3.1) then the second position must be in the other deck and there is no overcounting since in (3.4) we always perform the swap when one of the positions coincides with x_*.

3.2. Modified transpositions

A first step in the proof of Theorem 3.1 is to compare the process described above to a random transposition–type Markov chain that is defined as follows. Let $E_{i,j}$ denote the ordinary transposition operator between positions i and j. Namely, $(E_{i,j}\eta)_\ell = \eta_\ell$ for $i, j \neq \ell$ and $(E_{i,j}\eta)_i = \eta_j, (E_{i,j}\eta)_j = \eta_i$. Next, define the modified transposition $\widetilde{E}_{i,j}$, for $i < j$, by

$$
\widetilde{E}_{i,j} = \begin{cases} E_{i,j} & \text{if } i < x_* < j \\ T_{i,j} & \text{if } i = x_* < j \ \text{ or } \ i < x_* = j \\ 1 & \text{if } i < j < x_* \ \text{ or } \ x_* < i < j \end{cases} \tag{3.5}
$$

Also, set $\widetilde{E}_{i,i} = 1$ and $\widetilde{E}_{i,j} = \widetilde{E}_{j,i}$ if $i > j$. Here $T_{i,j}$ is the top–swap operation defined previously. In words, $\widetilde{E}_{i,j}$ is an ordinary transposition if the two chosen cards (in position i and j respectively) are in different decks. If they are in the same deck nothing happens. Finally, if one of the cards is the $*$ then a top–swap transformation is performed. In analogy with (3.4) we define the Dirichlet form of the modified random transpositions by

$$
\mathcal{D}_2(f) = \frac{1}{2(n+2)^2} \sum_{i,j=1}^{n+1} \mu \left[(\widetilde{E}_{i,j}f - f)^2 \right], \tag{3.6}
$$

It turns out (see Lemma 3.4 below) that the comparison $\mathcal{D}_2 \leqslant C\mathcal{E}_2$, is easy. We then have to show that $\mathcal{D}_2$ has a spectral gap not smaller than constant times n^{-1}. This is accomplished by a further comparison with a new process with Dirichlet form $\mathcal{F}_\delta(f)$ defined below.

Remark 2 We observe that comparison with ordinary random transpositions is not sufficient for our purpose. Let us call $\mathcal{E}_2^{RT}(f)$ the Dirichlet form

$$
\mathcal{E}_2^{RT}(f) = \frac{1}{2(n+2)^2} \sum_{i,j=1}^{n+1} \mu \left[(E_{i,j}f - f)^2 \right],
$$

associated to pure transpositions $E_{i,j}$. Note that cards are not required to belong to different decks. The natural normalization here, as in (3.6), would be $2(n+1)^2$ instead of $2(n+2)^2$. However, the latter has been preferred to make the analogy with (3.4) more apparent. Then it is simple to check that taking e.g. $f(\eta) = 1$ if $\eta = (1, 2, \ldots, n, *)$ and $f(\eta) = 0$ otherwise, the ratio $\mathcal{E}_2^{RT}(f)/\mathcal{E}_2(f)$ is of order n.

3.3. The balanced–swap chain

In words, this new process is described as follows. Let us fix $\delta \in (0, 1/2)$. We have two independent Poisson processes σ and τ with mean 1. At the arrival times of σ we choose uniformly at random a position i and then update the current configuration η by means of the transformation $\widetilde{E}_{i,x_*} = T_{i,x_*}$ (recall that $x_*(\eta)$ denotes the position of the $*$ in the configuration η). At the arrival times of τ we look at the current value of $x_*(\eta)$ and, if

$$\delta \leqslant \frac{x_*(\eta)}{n+1} \leqslant (1-\delta), \tag{3.7}$$

then conditioned on this value we put the system in equilibrium, i.e. we choose the new configuration η' uniformly at random among all the $n!$ configurations with the $*$ in the current position $x_*(\eta)$. If, instead, (3.7) does not hold then we do nothing. We say that η is δ–balanced if (3.7) holds. We call p_δ the μ–probability that this happens. Since μ is uniform $p_\delta = 1 - 2\delta + O(1/n)$.

We write $\mathrm{Var}_\mu(f \,|\, x_*)$ for the variance of a function $f : \Omega_n \to \mathbb{R}$ w.r.t. the conditional probability $\mu(\cdot \,|\, x_*)$ obtained from μ by conditioning on the position of the $*$. The Dirichlet form $\mathcal{F}_\delta$ of the process described above is then given by

$$\mathcal{F}_\delta(f) = \mu\left[\chi\mathrm{Var}_\mu(f \,|\, x_*)\right] + \frac{1}{2(n+1)} \sum_{i=1}^{n+1} \mu[(T_{i,x_*}f - f)^2]. \tag{3.8}$$

where

$$\chi(\eta) = \begin{cases} 1 & \text{if } \eta \text{ is } \delta\text{–balanced} \\ 0 & \text{otherwise} \end{cases} \tag{3.9}$$

We will show, see Lemma 3.3, that one has the comparison $\mathcal{F}_\delta \leqslant C\,n\,\mathcal{D}_2$. The original problem (3.3) is then reduced to estimating from below the spectral gap of $\mathcal{F}_\delta$. To this end one can use a coupling argument but the following proof seems to give a better estimate.

Lemma 3.2. *For all $n \in \mathbb{N}$, $f : \Omega_n \to \mathbb{R}$ we have*

$$\mathrm{Var}_\mu(f) \leqslant \gamma_\delta \, \mathcal{F}_\delta(f), \tag{3.10}$$

where $\gamma_\delta := (1 - \sqrt{1 - p_\delta})^{-1}$.

Proof. We have to prove that the spectral gap λ_* of the Dirichlet form $\mathcal{F}_\delta(\cdot)$ satisfies

$$\lambda_* \geqslant 1 - \sqrt{1 - p_\delta}. \tag{3.11}$$

The generator associated to $\mathcal{F}_\delta(\cdot)$ is the operator $\mathcal{L}$ acting on functions f as

$$\mathcal{L}f = \chi\left(\mu[f \,|\, x_*] - f\right) + A\,f - f, \quad A\,f := \frac{1}{n+1} \sum_{i=1}^{n+1} T_{i,x_*}\,f. \tag{3.12}$$

Note that $\mathcal{L}$ is self–adjoint and non–negative in $L^2(\mu)$ and

$$-\mu[f\mathcal{L}f] = \mathcal{F}_\delta(f). \tag{3.13}$$

We consider the eigenvalues λ and eigenfunctions f_λ of $-\mathcal{L}$:

$$-\mathcal{L}f_\lambda = \lambda f_\lambda. \tag{3.14}$$

Here $\lambda = 0$ corresponds to $f_\lambda = $ constant and any f_λ in (3.14) with $\lambda \neq 0$ satisfies $\mu[f_\lambda] = 0$. By definition, λ_* is the smallest non–zero λ such that (3.14) holds.

Observe that $\lambda = 1$ is an eigenvalue. Indeed, it suffices to consider $f(\eta) = \varphi(x_*(\eta))$ for an arbitrary $\varphi : \{1, \ldots, n+1\} \to \mathbb{R}$ such that $\frac{1}{n+1}\sum_{i=1}^{n+1} \varphi(i) = 0$. In this case, f is a function of x_* only and therefore $\mu[f \,|\, x_*] = f$ and the mean–zero property implies $Af = 0$, thus (3.14) gives $\lambda = 1$.

We may assume without loss of generality that $\lambda_* < 1$, since otherwise there is nothing to prove in (3.11). We claim that f_{λ_*} is such that

$$\mu[f_{\lambda_*} \,|\, x_*] = 0. \tag{3.15}$$

We first need the following property. For any function $g : \Omega_n \to \mathbb{R}$ we have

$$\mu[A\,g \,|\, x_*] = \mu[g], \tag{3.16}$$

i.e. $\mu[A\,g \,|\, x_*]$ does not depend on the value of x_*. To prove (3.16) we write, for every j:

$$\mu[A\,g \,|\, x_* = j] = \frac{1}{n+1}\sum_{i=1}^{n+1} \frac{\mu[(T_{i,j}g)\,1_{\{x_*=j\}}]}{\mathbb{P}_\mu[x_* = j]}.$$

Now, we show that for any pair of positions i, j:

$$\mu[(T_{i,j}g)\,1_{\{x_*=j\}}] = \mu[g\,1_{\{x_*=i\}}]. \tag{3.17}$$

Indeed, if $\Omega_n^\ell := \{\eta \in \Omega_n : x_*(\eta) = \ell\}$ then it is not hard to see that $T_{i,j} : \Omega_n^j \to \Omega_n^i$ is a bijection: namely, if e.g. $i < j$, for every $\eta \in \Omega_n^j$ and for every $\xi \in \Omega_n^i$ we have $T_{i,n+i-j+2}\xi \in \Omega_n^j$, $T_{i,j}\eta \in \Omega_n^i$ and

$$T_{i,n+i-j+2}T_{i,j}\eta = \eta, \quad T_{i,j}T_{i,n+i-j+2}\xi = \xi \tag{3.18}$$

Therefore

$$\mu[(T_{i,j}g)\,1_{\{x_*=j\}}] = \frac{1}{(n+1)!}\sum_{\eta \in \Omega_n^j} g(T_{i,j}\eta) = \frac{1}{(n+1)!}\sum_{\xi \in \Omega_n^i} g(\xi) = \mu[g\,1_{\{x_*=i\}}].$$

This proves (3.17). Using $\mathbb{P}_\mu[x_* = j] = \frac{1}{n+1}$ and (3.17), we see that

$$\mu[A\,g \,|\, x_* = j] = \sum_{i=1}^{n+1} \mu[g\,1_{\{x_*=i\}}] = \mu[g],$$

which implies (3.16).

We turn to the proof of (3.15). Taking the $\mu[\cdot \,|\, x_*]$– expectation in (3.14), using (3.16) and $\mu[f_{\lambda_*}] = 0$, we see that

$$\lambda_* \, \mu[f_{\lambda_*} \,|\, x_*] = -\mu[\mathcal{L} f_{\lambda_*} \,|\, x_*] = \mu[(f_{\lambda_*} - A f_{\lambda_*}) \,|\, x_*] = \mu[f_{\lambda_*} \,|\, x_*] \,.$$

Since $\lambda_* < 1$ this implies (3.15).

Using (3.15) in (3.14) we obtain

$$A f_{\lambda_*} = (1 - \lambda_* + \chi) \, f_{\lambda_*} =: \psi \, f_{\lambda_*} \,. \tag{3.19}$$

Note that ψ is a positive function of x_* only. In particular, taking absolute values in (3.19) we have

$$\mu[|A f_{\lambda_*}| \,|\, x_*] = \psi(x_*) \, \mu[|f_{\lambda_*}| \,|\, x_*] \,. \tag{3.20}$$

Now, using $|A f_{\lambda_*}| \leqslant A |f_{\lambda_*}|$ and (3.16)

$$\mu[|A f_{\lambda_*}| \,|\, x_*] \leqslant \mu[A |f_{\lambda_*}| \,|\, x_*] = \mu[|f_{\lambda_*}|] \,.$$

Therefore, dividing by ψ and taking μ–expectation in (3.20) we arrive at

$$\mu \left[\psi^{-1} \right] \geqslant 1 \,. \tag{3.21}$$

Inequality (3.21) can be written explicitly as

$$p_\delta \, \frac{1}{2 - \lambda_*} + (1 - p_\delta) \, \frac{1}{1 - \lambda_*} \geqslant 1 \,.$$

It is easily seen that this is satisfied iff $\lambda_*^2 - 2\lambda_* + p_\delta \leqslant 0$, or $\lambda_* \geqslant 1 - \sqrt{1 - p_\delta}$ (since $\lambda_* < 1$ by assumption). This ends the proof of Lemma 3.2. $\qquad\square$

Remark 3. The result in Lemma 3.2 shows that if $p_\delta = 1$, i.e. if there is no constraint and $\chi = 1$ (or $\delta = 0$), then the gap is equal to 1. It is however crucial for us that we are able to prove the spectral gap bound of Lemma 3.2 *with* the constraint $\chi < 1$, since we rely on this constraint to perform the comparison of Lemma 3.3 below.

3.4. From balanced–swap moves to modified transpositions

The next lemma shows how to 'simulate' balanced–swap moves by means of modified transpositions.

Lemma 3.3. *For each $\delta \in (0, 1/2)$ there exists $C_\delta > 0$ such that for every n and every function f on Ω_n:*

$$\mathcal{F}_\delta(f) \leqslant C_\delta \, n \, \mathcal{D}_2(f) \,. \tag{3.22}$$

Proof. Let us fix an arbitrary value of x_*. We want to estimate $\text{Var}(f \mid x_*)$ by using only terms of the form $\mu\left[(E_{i,\ell} - f)^2 \mid x_*\right]$ with $i < x_* < \ell$.

Since $\mu(\cdot \mid x_*)$ is nothing but the uniform probability on $n!$ permutations (arrangements of n cards in two decks of given size), recalling the spectral gap of the ordinary random transposition chain (see e.g. [1]) we know that

$$\text{Var}_\mu(f \mid x_*) \leqslant \frac{1}{n} \sum_{i \neq x_*} \sum_{j \neq x_*} \mu\left[(E_{i,j}f - f)^2 \mid x_*\right] . \tag{3.23}$$

We now assume that $\delta\, n \leqslant x_* \leqslant (1 - \delta)\, n$, for $\delta \in (0, 1/2)$ (see (3.7)). Consider the case where i, j are such that $i < x_*$ and $j < x_*$ (both cards are in the first deck). We observe that, for any $\ell > x_*$ we can write

$$E_{i,j}f - f = E_{i,\ell}E_{j,\ell}E_{i,\ell}f - f = (E_{i,\ell}f_2 - f_2) + (E_{j,\ell}f_1 - f_1) + (E_{i,\ell}f - f),$$

where $f_1 := E_{i,\ell}f$ and $f_2 := E_{j,\ell}E_{i,\ell}f$. Therefore

$$\mu\left[(E_{i,j}f - f)^2 \mid x_*\right] \leqslant$$
$$3\,\mu\left[(E_{i,\ell}f_2 - f_2)^2 \mid x_*\right] + 3\,\mu\left[(E_{j,\ell}f_1 - f_1)^2 \mid x_*\right] + 3\,\mu\left[(E_{i,\ell}f - f)^2 \mid x_*\right] .$$

From the invariance of $\mu(\cdot \mid x_*)$ under transpositions we see that

$$\mu\left[(E_{i,\ell}f_2 - f_2)^2 \mid x_*\right] = \mu\left[(E_{i,\ell}f - f)^2 \mid x_*\right] ,$$
$$\mu\left[(E_{j,\ell}f_1 - f_1)^2 \mid x_*\right] = \mu\left[(E_{j,\ell}f - f)^2 \mid x_*\right] .$$

Therefore, averaging over all $n + 1 - x_*$ positions ℓ such that $\ell > x_*$, we obtain

$$\mu\left[(E_{i,j}f - f)^2 \mid x_*\right]$$
$$\leqslant \frac{1}{(n + 1 - x_*)} \sum_{\ell > x_*} \mu\left[6\,(E_{i,\ell}f - f)^2 + 3\,(E_{j,\ell}f - f)^2 \mid x_*\right] .$$

Symmetrizing over i, j, the inequality above becomes

$$\mu\left[(E_{i,j}f - f)^2 \mid x_*\right]$$
$$\leqslant \frac{9}{2(n + 1 - x_*)} \sum_{\ell > x_*} \mu\left[(E_{i,\ell}f - f)^2 + (E_{j,\ell}f - f)^2 \mid x_*\right] .$$

Summing over all pairs i, j with $i < x_*$ and $j < x_*$ we obtain

$$\sum_{i < x_*} \sum_{j < x_*} \mu\left[(E_{i,j}f - f)^2 \mid x_*\right]$$
$$\leqslant \frac{9(x_* - 1)}{(n + 1 - x_*)} \sum_{i < x_*} \sum_{\ell > x_*} \mu\left[(E_{i,\ell}f - f)^2 \mid x_*\right] , \tag{3.24}$$

since there are $x_* - 1$ positions j such that $j < x_*$.

Repeating the reasoning leading to (3.24) for pairs i, j with $i > x_*$ and $j > x_*$ we have

$$\sum_{i > x_*} \sum_{j > x_*} \mu \left[(E_{i,j} f - f)^2 \,|\, x_* \right]$$

$$\leqslant \frac{9(n + 1 - x_*)}{(x_* - 1)} \sum_{i > x_*} \sum_{\ell < x_*} \mu \left[(E_{i,\ell} f - f)^2 \,|\, x_* \right]. \tag{3.25}$$

Inserting (3.24) and (3.25) in (3.23) we obtain

$$\mathrm{Var}(f \,|\, x_*) \leqslant \frac{2}{n} \sum_{i < x_*} \sum_{j > x_*} \mu \left[(E_{i,j} f - f)^2 \,|\, x_* \right] +$$

$$+ \frac{9}{n} \left\{ \frac{(n + 1 - x_*)}{(x_* - 1)} + \frac{(x_* - 1)}{(n + 1 - x_*)} \right\} \sum_{i < x_*} \sum_{j > x_*} \mu \left[(E_{i,j} f - f)^2 \,|\, x_* \right]$$

$$\leqslant \frac{1}{n} \left\{ 2 + \frac{9\, n^2}{(n + 1 - x_*)(x_* - 1)} \right\} \sum_{i < x_*} \sum_{j > x_*} \mu \left[(E_{i,j} f - f)^2 \,|\, x_* \right],$$

where we used $(n - t)/t + t/(n - t) \leqslant n^2/t(n - t)$, for $t \in (0, n)$. Note that if x_* satisfies (3.7) then $(n + 1 - x_*)(x_* - 1) \geqslant \delta\,(1 - \delta)\,(n + 1)^2$. In conclusion, recalling that χ forces x_* to satisfy (3.7), we see that, for any value of x_*:

$$\chi \, \mathrm{Var}(f \,|\, x_*) \leqslant \frac{1}{n} \left\{ 2 + \frac{9\, n^2}{\delta\,(1 - \delta)\,(n + 1)^2} \right\} \sum_{i < x_*} \sum_{j > x_*} \mu \left[(E_{i,j} f - f)^2 \,|\, x_* \right].$$

Taking μ–expectation and recalling the definition (3.8) of $\mathcal{F}_\delta(f)$ we see that

$$\mathcal{F}_\delta(f) \leqslant \left\{ 2 + \frac{9\, n^2}{\delta\,(1 - \delta)\,(n + 1)^2} \right\} \frac{(n + 2)^2}{n} \, \mathcal{D}_2(f) =: C_\delta \, n \, \mathcal{D}_2(f). \tag{3.26}$$

Note that $C_\delta \leqslant 2 + 9/[\delta(1 - \delta)] + O(1/n)$. $\qquad\qquad\square$

3.5. From modified transpositions to top–swaps

Here we use the technique of path–comparison to estimate the Dirichlet form $\mathcal{D}_2$ in terms of the Dirichlet form $\mathcal{E}_2$, see (3.4) and (3.6) for their definitions. The next lemma is based on the rather obvious fact that an ordinary transposition can be 'simulated' by two top–swap moves.

Lemma 3.4. *For any function f:*

$$\mathcal{D}_2(f) \leqslant 5\, \mathcal{E}_2(f) \tag{3.27}$$

Proof. Let η be a given configuration of cards and let i, j be two given positions such that $i < j$. If $i < x_*(\eta) = j$ then $\widetilde{E}_{i,j} = T_{i,j}$ so we only need to take care of the case $i < x_*(\eta) < j$. In this case $\widetilde{E}_{i,j}$ is an ordinary transposition $E_{i,j}$ and this can be written by way of two top–swap operations. Namely, suppose that $\eta \in \Omega_n^\ell$, i.e. $\eta \in \Omega_n$ is such that $x_*(\eta) = \ell$. Then it is not hard to see that

$$E_{i,j}\eta = T_{i+1,n+i-\ell+3}\, T_{i,j}\, \eta. \tag{3.28}$$

Note that $n + i - \ell + 3 \leqslant n + 2$ since we are assuming $i < x_*(\eta) = \ell$. We adopt the convention that if $n + i - \ell + 3 = n + 2$ then $T_{i+1,n+i-\ell+3}\,\xi$ changes the configuration ξ by moving the top above (and including) $i + 1$ from deck 1 to the top of deck 2.

The identity (3.28) is easily understood with the help of a picture. For instance, consider the configuration $\eta = (3, 4, *, 5, 2, 1, 6)$ of Figure 1. $E_{1,6}\eta$ (i.e. the transposition of cards labeled 3 and 1) is obtained by first applying $T_{1,6}$ to obtain $T_{1,6}\eta = (1, 6, *, 5, 2, 3, 4)$ and then applying $T_{2,7}$. To see a case where $n + i - \ell + 3 = n + 2$, consider e.g. $E_{2,5}\eta$ (the transposition of cards labeled 4 and 2). This is obtained by doing first $T_{2,5}\eta = (3, 2, 1, 6, *, 5, 4)$ and then applying $T_{3,8}$ to this last configuration, since by our convention $T_{3,8}$ is the same as $T_{3,5}$.

Using (3.28) we have, for all $i < \ell < j$:

$$\begin{aligned}
(E_{i,j}f(\eta) &- f(\eta))^2\, 1_{\{x_*=\ell\}} \\
&\leqslant 2(T_{i+1,n+i-\ell+3}f_1 - f_1)^2\, 1_{\{x_*=\ell\}} + 2(T_{i,j}f - f)^2\, 1_{\{x_*=\ell\}},
\end{aligned} \tag{3.29}$$

where $f_1(\eta) := f(T_{i,j}\eta)$ for every $\eta \in \Omega_n$. An argument similar to that used in the proof of (3.17) gives

$$\mu\left[(T_{i+1,n+i-\ell+3}f_1 - f_1)^2\, 1_{\{x_*=\ell\}}\right] = \mu\left[(T_{i+1,n+i-\ell+3}f - f)^2\, 1_{\{x_*=n+i-j+2\}}\right] \tag{3.30}$$

From (3.30) it follows that

$$\begin{aligned}
\sum_{i,j:\, i<j} \sum_{\ell=i+1}^{j-1} &\mu\left[(T_{i+1,n+i-\ell+3}f_1 - f_1)^2\, 1_{\{x_*=\ell\}}\right] \\
&= \sum_{i,j:\, i<j} \sum_{\ell=i+1}^{j-1} \mu\left[(T_{i+1,n+i-\ell+3}f - f)^2\, 1_{\{x_*=n+i-j+2\}}\right] \\
&\leqslant \sum_{i,j:\, i<j} \mu\left[(T_{i,j}f - f)^2\right]
\end{aligned} \tag{3.31}$$

Inserting (3.31) in (3.29) we have shown that

$$\sum_{i,j=1}^{n+1} \mu\left[(\widetilde{E}_{i,j}f - f)^2\, 1_{\{i<x_*<j\}}\right] \leqslant 4 \sum_{i,j:\, i<j} \mu\left[(T_{i,j}f - f)^2\right]. \tag{3.32}$$

The identity

$$\mathcal{D}_2(f) = \frac{1}{(n+2)^2} \sum_{i,j=1}^{n+1} \mu\left[(\widetilde{E}_{i,j}f - f)^2 \mathbf{1}_{\{i<x_*<j\}}\right] +$$

$$+ \frac{1}{(n+2)^2} \sum_{i=1}^{n+1} \mu\left[(T_{i,x_*}f - f)^2\right],$$

and (3.32) then imply

$$\mathcal{D}_2(f) \leqslant \frac{2}{(n+2)^2} \sum_{i,j=1}^{n+1} \mu\left[(T_{i,j}f - f)^2\right] +$$

$$+ \frac{1}{(n+2)^2} \sum_{i}^{n+1} \mu\left[(T_{i,x_*}f - f)^2\right] \leqslant 5\,\mathcal{E}_2(f).$$

$\square$

3.6. Proof of Theorem 3.1 completed

As a consequence of Lemma 3.2, Lemma 3.3 and Lemma 3.4 we know that for all n and all $f : \Omega_n \to \mathbb{R}$

$$\text{Var}_\mu(f) \leqslant 5\, C_\delta\, \gamma_\delta\, n\, \mathcal{E}_2(f). \tag{3.33}$$

We can still choose the value of $\delta \in (0, 1/2)$ (C_δ is minimized at $\delta = 1/2$ while γ_δ is minimized at $\delta = 0$).

Note that for $\delta = 0.25$ one has $p_\delta = 0.5 + O(1/n)$, $\gamma_\delta = \frac{\sqrt{2}}{\sqrt{2}-1} + O(1/n)$ and $C_\delta \leqslant 50 + O(1/n)$. Therefore (3.33) gives $\text{Var}_\mu(f) \leqslant C\, n\, \mathcal{E}_2(f)$ with $C \leqslant 875 + O(1/n)$. This ends the proof of Theorem 3.1.

3.7. Random transpositions with constraint

The analysis above can be used to obtain a spectral gap estimate for a random transposition model with the constraint that two cards need to be in two different decks to be transposed. This is the process with Dirichlet form $\mathcal{D}_2^T(f)$ as defined in (3.6) with the only difference that here the transformations $\widetilde{E}_{i,j}$ are given by

$$\widetilde{E}_{i,j}^T = \begin{cases} E_{i,j} & \text{if } i \leqslant x_* \leqslant j \\ 1 & \text{if } i < j < x_* \text{ or } x_* < i < j \end{cases} \tag{3.34}$$

with $\widetilde{E}_{i,i}^T = 1$ and $\widetilde{E}_{i,j}^T = \widetilde{E}_{j,i}^T$ if $i > j$. Namely, when $i = x_*$ we simply transpose rather than doing the top–swap as in (3.5).

Theorem 3.5. *There exists $C < \infty$ such that, for every $n \in \mathbb{N}$, for every function $f : \Omega_n \to \mathbb{R}$ we have*

$$\mathrm{Var}_\mu(f) \leqslant C \, n \, \mathcal{D}_2^T(f) \,. \tag{3.35}$$

Proof. We repeat the same arguments used in the proof of Theorem 3.1. We only need to modify the definition of the Dirichlet form $\mathcal{F}_\delta(f)$ which we replace here with

$$\mathcal{F}_\delta^T(f) = \mu\left[\chi \mathrm{Var}_\mu(f \mid x_*)\right] + \frac{1}{2(n+1)} \sum_{i=1}^{n+1} \mu[(E_{i,x_*} f - f)^2] \,. \tag{3.36}$$

It is not difficult to check that all our arguments in Lemma 3.2 apply with no modifications to this case. Moreover, the same applies to Lemma 3.3. Note that here we do not need the extra comparison of Lemma 3.4. In particular, one has that the constant C in (3.35) satisfies $C \leqslant 175 + O(1/n)$. $\qquad\square$

4. Deck–averages

We consider the following setting. There are k decks $I_1, \ldots, I_k$ and a total of n cards. The state space Ω is the set of all

$$\frac{(n+k-1)!}{(k-1)!}$$

arrangements of n cards in k decks. η will denote the general random element of Ω and ν will be the uniform probability over Ω. Also, we use $\pi_i \eta$ to denote the projection of the configuration η onto the configuration of cards in deck I_i.

4.1. Deck–averaging process

The deck–averaging process is the continuous time Markov chain obtained as follows. At each deck independently there is a rate–1 Poisson clock. When deck I_i rings we choose uniformly another deck I_j and the cards in the two decks $I_i \cup I_j$ are rearranged uniformly at random (among the two decks). The Dirichlet form of this process is

$$\bar{\mathcal{D}}(f) = \frac{1}{k} \sum_{i=1}^k \sum_{j=1}^k \nu\left[(A_{i,j}f)^2\right] \,, \tag{4.1}$$

where the averaging gradient is defined by

$$
\begin{aligned}
A_{i,j}f(\sigma) &= \sum_{\xi \in \Omega} (f(\xi) - f(\sigma)) \, \nu[\eta = \xi \mid \pi_\ell \eta = \pi_\ell \sigma \,, \; \ell \neq i,j] \\
&= \nu[f \mid \pi_\ell \eta \,, \; \ell \neq i,j](\sigma) - f(\sigma) \\
&=: \nu_{i,j}[f](\sigma) - f(\sigma) \,,
\end{aligned}
\tag{4.2}
$$

for $i \neq j$, and $\sigma \in \Omega$. We set $A_{i,i}f = 0$. Here $\nu_{i,j} = \nu[\cdot \,|\, \pi_\ell \eta, \;\; \ell \neq i, j]$ is the conditional probability obtained by freezing the configuration in all decks $I_\ell, \ell \neq i, j$. Also, note that

$$\nu\left[(A_{i,j}f)^2\right] = \nu\left[\mathrm{Var}_{i,j}(f)\right],\tag{4.3}$$

where $\mathrm{Var}_{i,j}$ stands for the variance with respect to the probability $\nu_{i,j}$, with the convention that $\mathrm{Var}_{i,i}(f) = 0$.

The result of Proposition 2.3 in [1] is that the spectral gap of the deck–average Markov chain equals 1. In particular, for every $f : \Omega \to \mathbb{R}$ we have

$$\mathrm{Var}_\nu(f) \leqslant \bar{\mathcal{D}}(f),\tag{4.4}$$

with $\mathrm{Var}_\nu(f)$ denoting the variance of f w.r.t. ν. Strictly speaking the proof of Proposition 2.3 in [1] does not apply here because the sizes of the decks are not fixed in our model. However, it is not hard to adapt that proof to show that the result (4.4) holds. This only requires small modifications, the point being that the spectrum of the operator $\mathcal{K}$ used there satisfies the right bounds in our setting, see Lemma 4.13 below and the remark following it for more details.

4.2. Weighted deck–averaging process

The weighted version of the deck–averaging chain is defined by the Dirichlet form

$$\mathcal{D}(f) = \frac{1}{k} \sum_{i=1}^{k} \sum_{j=1}^{k} \nu\left[(n_i + n_j)\mathrm{Var}_{i,j}(f)\right],\tag{4.5}$$

where $n_i = n_i(\eta)$ stands for the number of cards in deck i. A possible interpretation for (4.5) is that each *card* is equipped with a rate–1 Poisson clock; when a card c rings a deck I_j is chosen uniformly at random; if $c \in I_j$ then nothing happens, while if $c \notin I_j$, i.e. $c \in I_i$ for some $i \neq j$, then all the cards in $I_i \cup I_j$ are rearranged uniformly at random. Apart from a factor 2 the Dirichlet form of this process is given by (4.5). The main result here is the following spectral gap estimate.

Proposition 4.1. *For every $f : \Omega \to \mathbb{R}$*

$$\mathrm{Var}_\nu(f) \leqslant \frac{6\,k}{n}\,\mathcal{D}(f).\tag{4.6}$$

Proof. Since $\sum_{i=1}^{k} n_i = n$, (4.4) and (4.3) imply

$$\mathrm{Var}_\nu(f) \leqslant \frac{1}{k\,n} \sum_{i,j=1}^{k} \sum_{\ell=1}^{k} \nu\left[n_\ell\,\mathrm{Var}_{i,j}(f)\right].\tag{4.7}$$

We claim that, for every fixed triple i, j, ℓ

$$\frac{1}{2}\,\nu\left[n_\ell\,\mathrm{Var}_{i,j}(f)\right] \leqslant \nu\left[(n_i + n_j)\,\mathrm{Var}_{i,j}(f)\right] +$$
$$+ \nu\left[(n_i + n_\ell)\,\mathrm{Var}_{i,\ell}(f)\right] + \nu\left[(n_j + n_\ell)\,\mathrm{Var}_{j,\ell}(f)\right].\tag{4.8}$$

Once we have (4.8), from the estimate in (4.7) we obtain

$$\operatorname{Var}_\nu(f) \leqslant \frac{6}{n} \sum_{i,j=1}^{k} \nu\left[(n_i + n_j)\operatorname{Var}_{i,j}(f)\right],$$

which is the same as (4.6).

We turn to the proof of the claim (4.8). We can assume that i, j, ℓ are three distinct labels (i.e. $\ell \neq i$, $\ell \neq j$, $i \neq j$) since the statement is obviously true otherwise. Let $m = n_i + n_j + n_\ell$ and let μ_Δ denote the conditional probability $\nu[\cdot\,|\,\pi_v\eta,\ v \neq i, j, \ell]$ obtained by freezing all the decks I_v, $v \neq i, j, \ell$. We are going to prove that

$$\frac{m}{2}\operatorname{Var}_{\mu_\Delta}(f) \leqslant \mu_\Delta\left[(n_i + n_j)\operatorname{Var}_{i,j}(f)\right] +$$
$$+ \mu_\Delta\left[(n_i + n_\ell)\operatorname{Var}_{i,\ell}(f)\right] + \mu_\Delta\left[(n_j + n_\ell)\operatorname{Var}_{j,\ell}(f)\right]. \tag{4.9}$$

Here $\operatorname{Var}_{\mu_\Delta}(f)$ is the variance of f w.r.t. μ_Δ. This depends on the given configuration η through the projections $\pi_v\eta$ for $v \neq i, j, \ell$. Note that once the variables $\pi_v\eta$, $v \neq i, j, \ell$ are given, m becomes a constant, i.e. m is μ_Δ-a.s. constant.

Let us first show that (4.9) implies the claim (4.8). Take expectation w.r.t. ν in (4.9) and use the fact that $\nu[h] = \nu[\mu_\Delta[h]]$ for any function $h : \Omega \to \mathbb{R}$:

$$\frac{1}{2}\nu\left[m\operatorname{Var}_{\mu_\Delta}(f)\right] \leqslant \nu\left[(n_i + n_j)\operatorname{Var}_{i,j}(f)\right] +$$
$$+ \nu\left[(n_i + n_\ell)\operatorname{Var}_{i,\ell}(f)\right] + \nu\left[(n_j + n_\ell)\operatorname{Var}_{j,\ell}(f)\right]. \tag{4.10}$$

Observe that

$$\begin{aligned}
\operatorname{Var}_{\mu_\Delta}(f) &= \mu_\Delta\left[f(f - \mu_\Delta[f])\right] \\
&= \mu_\Delta\left[f(f - \mu_\Delta[f\,|\,\pi_\ell\eta])\right] + \mu_\Delta\left[f(\mu_\Delta[f\,|\,\pi_\ell\eta] - \mu_\Delta[f])\right] \\
&= \mu_\Delta\left[\operatorname{Var}_{i,j}(f)\right] + \operatorname{Var}_{\mu_\Delta}\left[\mu_\Delta[f\,|\,\pi_\ell\eta]\right] \\
&\geqslant \mu_\Delta\left[\operatorname{Var}_{i,j}(f)\right],
\end{aligned}$$

where we use the fact that $\mu_\Delta[f\,|\,\pi_\ell\eta]$ coincides with $\nu_{i,j}$.

We can now use the obvious bound $m \geqslant n_\ell$ to obtain

$$\nu\left[m\operatorname{Var}_{\mu_\Delta}(f)\right] \geqslant \nu\left[m\,\mu_\Delta\left[\operatorname{Var}_{i,j}(f)\right]\right] = \nu\left[m\operatorname{Var}_{i,j}(f)\right] \geqslant \nu\left[n_\ell\operatorname{Var}_{i,j}(f)\right].$$

With (4.10), this implies (4.8).

It remains to prove (4.9). We introduce the projection P_i defined by

$$P_i f = \mu_\Delta[f\,|\,\pi_i\eta].$$

Projectors P_j, P_ℓ are defined in a similar way. Observe that with these definitions, since $\mu_\Delta[\cdot\,|\,\pi_\ell\eta] = \nu_{i,j}$, we have

$$\mu_\Delta\left[\operatorname{Var}_{i,j}(f)\right] = \mu_\Delta\left[f(1 - P_\ell)f\right]. \tag{4.11}$$

Moreover, setting

$$\Gamma := (m - n_i)(1 - P_i) + (m - n_j)(1 - P_j) + (m - n_\ell)(1 - P_\ell) \,,$$

we see that the r.h.s. of display (4.9) coincides with the quadratic form $\mu_\Delta[f\Gamma f]$.

We call P the average

$$P = \frac{1}{3} \, (P_i + P_j + P_\ell) \,. \tag{4.12}$$

Lemma 4.2 below shows that

$$\mu_\Delta[fPf] \leqslant \frac{1}{2}\mu_\Delta[f^2] \,, \tag{4.13}$$

for any function f such that $\mu_\Delta[f] = 0$. For such f, the result (4.13) implies

$$\mu_\Delta[f\Gamma f] = 2m\mu_\Delta[f^2] - 3m\mu_\Delta[fPf] + \mu_\Delta[n_i f P_i f] + \mu_\Delta[n_j f P_j f] + \mu_\Delta[n_\ell f P_\ell f]$$
$$\geqslant \frac{m}{2} \, \mu_\Delta[f^2] \,,$$

where we use

$$\mu_\Delta[n_i f P_i f] = \mu_\Delta[n_i (P_i f)^2] \geqslant 0 \,.$$

The claim (4.9) follows since, by subtracting the mean $\mu_\Delta[f]$, we may restrict to mean zero functions to obtain

$$\mathrm{Var}_{\mu_\Delta}(f) \leqslant \frac{2}{m} \, \mu_\Delta[f\Gamma f] \,.$$

$\square$

Lemma 4.2. *Let P denote the operator defined in (4.12). For every function f such that $\mu_\Delta[f] = 0$ we have $\mu_\Delta[fPf] \leqslant \frac{1}{2} \, \mu_\Delta[f^2]$.*

Proof. In the case of three decks, the proof of Proposition 2.2 of [1] (see display (27) and display (32) there) shows that

$$\mu_\Delta[f(1 - P)f] \geqslant \ \min \left\{ \frac{1}{2}, \frac{2}{3}(1 - \lambda_2) \right\} \mu_\Delta[f^2] \,, \tag{4.14}$$

where λ_2 is the largest positive eigenvalue (other than $\lambda_0 = 1$, see below for more details) of the stochastic matrix $\mathcal{K}(\alpha, \beta)$ defined by

$$\mathcal{K}(\alpha, \beta) = \mu_\Delta[\pi_i \eta = \beta \,|\, \pi_j \eta = \alpha] \,.$$

That is, $\mathcal{K}(\alpha, \beta)$ is the μ_Δ–conditional probability of having the configuration β in I_i, given the configuration α in I_j.

The spectrum of $\mathcal{K}$ can be computed exactly as in Lemma 2.1 of [1]. The only difference here is that the size of the decks is not fixed. This does not pose any difficulty and we can proceed in the very same way. We have to compute the action

of $\mathcal{K}$ on functions of the kind $\chi_{r_1} \cdots \chi_{r_n}$ where χ_r stands for the indicator function of the event {the card labeled r belongs to deck i}. For instance, a simple computation shows that for one card r one has

$$\mathcal{K}\chi_r = \frac{1}{2}\left(1 - \chi_r\right).$$

Therefore $\chi_r - \frac{1}{3}$ is an eigenfunction with eigenvalue $\lambda_1 = -\frac{1}{2}$. Similarly, for two distinct cards r_1, r_2 one can compute

$$\mathcal{K}\chi_{r_1}\chi_{r_2} = \frac{1}{4}\left(1 - \chi_{r_1}\right)\left(1 - \chi_{r_2}\right).$$

Following the proof of Lemma 2.1 in [1] we then obtain that the spectrum of $\mathcal{K}$ consists of the eigenvalues $\lambda_\ell = \left(-\frac{1}{2}\right)^\ell$, for $\ell = 0, 1, \ldots, m$, where m is the total number of cards in the three decks. In particular, $\lambda_2 = \frac{1}{4}$ so that (4.14) implies $\mu_\Delta[f(1 - P)f] \geqslant \frac{1}{2}\mu_\Delta[f^2]$ or, equivalently

$$\mu_\Delta[fPf] \leqslant \frac{1}{2}\mu_\Delta[f^2].$$

$\square$

Remark 4. More generally, for k decks and n cards one obtains that the spectrum of $\mathcal{K}$ is given by

$$\lambda_\ell = \left(-\frac{1}{k-1}\right)^\ell,$$

$\ell = 0, 1, \ldots, n$.

5. Top–swap process with k decks

The setting is as in the previous section. In analogy with the 2–deck process defined in (3.1) we consider here the Dirichlet form

$$\mathcal{E}_k(f) = \frac{1}{2(n+k)^2} \sum_{r,s=1}^{n+k} \nu\left[(\mathcal{T}_{r,s}f - f)^2\right], \tag{5.1}$$

where we sum over all positions r, s between 1 and $n + k$, since we added to each deck I_j an extra position to allow the swap of an empty top. We shall call $\bar{I}_j$ the set of positions defined by the deck I_j and its extra position.

The main result is that the relaxation time of this chain is $O(n + k)$:

Theorem 5.1. *There exists a constant $C > 0$ such that for every n, k, every f*

$$\mathrm{Var}_\nu(f) \leqslant C\,(n + k)\,\mathcal{E}_k(f). \tag{5.2}$$

Proof. We start by decomposing $\mathcal{E}_k$ into 2–deck terms:

$$\mathcal{E}_k(f) = \frac{1}{4\,(n+k)^2} \sum_{i,j=1}^{k} \nu\left[\sum_{r,s\in\bar{I}_i\cup\bar{I}_j} (\mathcal{T}_{r,s}f - f)^2 \right]. \tag{5.3}$$

If $\nu_{i,j}$ denotes the conditional probability in (4.2) (obtained by freezing all the decks $I_\ell,\ \ell\neq i,j$) then

$$\mathcal{E}_k(f) = \frac{1}{4\,(n+k)^2} \sum_{i,j=1}^{k} \nu\left[\sum_{r,s\in\bar{I}_i\cup\bar{I}_j} \nu_{i,j}\left[(\mathcal{T}_{r,s}f - f)^2 \right] \right]$$

$$= \frac{1}{2\,(n+k)^2} \sum_{i,j=1}^{k} \nu\left[(n_i + n_j + 2)^2\, \mathcal{E}_{i,j}(f) \right], \tag{5.4}$$

where $\mathcal{E}_{i,j}(f)$ is the Dirichlet form of the top–swap process on the two decks i, j only:

$$\mathcal{E}_{i,j}(f) = \frac{1}{2\,(n_i + n_j + 2)^2}\, \nu\left[\sum_{r,s\in\bar{I}_i\cup\bar{I}_j} \nu_{i,j}\left[(\mathcal{T}_{r,s}f - f)^2 \right] \right].$$

From Theorem 3.1 we know that for all $f : \Omega \to \mathbb{R}$, for any pair of distinct decks I_i, I_j with fixed total number of cards $n_i + n_j$:

$$\mathcal{E}_{i,j}(f) \geqslant \frac{1}{C(n_i + n_j)}\, \mathrm{Var}_{i,j}(f). \tag{5.5}$$

We can then use (5.4) to obtain

$$\mathcal{E}_k(f) \geqslant \frac{1}{2\,C\,(n+k)^2} \sum_{i,j=1}^{k} \nu\left[(n_i + n_j + 2)\, \mathrm{Var}_{i,j}(f) \right]. \tag{5.6}$$

Recalling the definition (4.5) of the weighted deck–averaging Dirichlet form $\mathcal{D}(f)$ we see that

$$\mathcal{E}_k(f) \geqslant \frac{1}{2\,C\,(n+k)^2} \left[k\,\mathcal{D}(f) + 2 \sum_{i,j=1}^{k} \nu(\mathrm{Var}_{i,j}(f)) \right]. \tag{5.7}$$

Now using (4.3), (4.1), and (4.4), we get that

$$\mathcal{E}_k(f) \geqslant \frac{k}{2\,C\,(n+k)^2} \left[\mathcal{D}(f) + 2\,\mathrm{Var}_\nu(f) \right]. \tag{5.8}$$

Proposition 4.1 then implies

$$\mathcal{E}_k(f) \;\geqslant\; \frac{k}{2\,C\,(n+k)^2}\left[\frac{n}{6k}\operatorname{Var}_\nu(f) + 2\operatorname{Var}_\nu(f)\right]$$

$$= \frac{n+12k}{12\,C\,(n+k)^2}\operatorname{Var}_\nu(f)$$

$$\geqslant \frac{1}{12\,C\,(n+k)}\operatorname{Var}_\nu(f)\,.$$

This concludes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

5.1. k−deck random transpositions with constraints

Here we consider the extension to k decks of the problem discussed in Theorem 3.5. Namely, we define

$$\mathcal{E}_k^T(f) = \frac{1}{2\,(n+k)^2}\sum_{r,s=1}^{n+k}\nu\left[(E_{r,s}f - f)^2\right]\,, \tag{5.9}$$

where the transformations $E_{r,s}$ are interpreted as ordinary transpositions provided r and s belong to different extended decks $\bar{I}_i, \bar{I}_j$. If $r, s \in \bar{I}_i$ for some i then $E_{r,s} = 1$.

The computations in the proof given above can be repeated step by step. Using Theorem 3.5 we can then obtain the following analogue of Theorem 5.1.

Theorem 5.2. *There exists $C > 0$ such that for every n, k and every f*

$$\operatorname{Var}_\nu(f) \leqslant C\,(n+k)\,\mathcal{E}_k^T(f)\,. \tag{5.10}$$

5.2. Matching upper bounds on spectral gap

Consider the k-deck top-swap Dirichlet form given by (5.1). We now show that up to a constant factor, Theorem 5.1 is tight.

For $\eta \in \Omega$, let $n_1(\eta)$ denote as usual the number of cards in the first deck. Consider the indicator (test) function

$$f(\eta) = \begin{cases} 1 & \text{if } n_1(\eta) = 0 \\ 0 & \text{if } n_1(\eta) > 0\,. \end{cases} \tag{5.11}$$

Then it is easy to check that $\nu[f] = (k-1)/(n+k-1)$, and that

$$\operatorname{Var}_\nu(f) = n(k-1)/(n+k-1)^2\,.$$

On the other hand,

$$\mathcal{E}_k(f) = \frac{1}{2(n+k)^2}\left[\frac{2(k-1)n}{(n+k-1)}\right] = \frac{n(k-1)}{(n+k-1)(n+k)^2} \tag{5.12}$$

thus showing that the spectral gap is at most

$$(n + k - 1)/(n + k)^2 \leqslant 1/(n + k).$$

Note that in the first equality above, we used the fact that for η with $n_1(\eta) = 0$, the contribution to the Dirichlet form comes from choosing $r = 1$ (namely the position of the first marker) and s to be any of the n card positions (namely a non-marker position).

Finally, note that the same test function (with an identical computation) shows that the spectral gap of the k-deck random transpositions with constraints, defined using the Dirichlet form (5.9), is also at most $1/(n + k)$.

References

[1] N. Cancrini, P. Caputo, F. Martinelli, Relaxation time of L-reversal chains and other chromosome shuffles. *Ann. Appl. Probab.*, 16(3), 2006.

[2] N. Cancrini, F. Martinelli, C. Roberto, C. Toninelli, Kinetically constrained models, Preprint, 2006.

[3] A. De, M. Ferguson, S. Sindi and R. Durrett. The equilibrium distribution for a generalized Sankoff-Ferretti model accurately predicts chromosome size distributions in a wide variety of species. *J. Appl. Prob.*, 38: 324-334, 2001.

[4] R. Durrett, Genome Rearrangement: Recent Progress and Open Problems. In *Statistical Methods in Molecular Evolution*, edited by R. Nielsen, Springer, 2005.

[5] R. Durrett, R. Nielsen, and T. York. Bayesian estimation of genomic distance. *Genetics*, 166:621–62, 2004.

[6] Joe W. Gray and Colin Collins, Genome changes and gene expression in human solid tumors. *Carcinogenesis*, 21(3): 443 - 452, 2000.

[7] S. Hannenhalli and P.A. Pevzner. Transforming men into mice (polynomial algorithm for genomic distance problem). In *Proceedings of the 36th Annual IEEE Symposium on Foundations of Computer Science* (FOCS), 581-592, 1995.

[8] P. Pevzner and G. Tesler, Genome Rearrangements in Mammalian Evolution: Lessons From Human and Mouse Genomes. *Genome Res.*, 13(1):37-45, 2003.

[9] D. Sankoff and V. Ferretti, Karyotype distribution in a stochastic model of reciprocal translocation. *Genome Res.*, 6: 1-9, 1996.

[10] Li-San Wang and Tandy Warnow. Distance-Based Genome Rearrangement Phylogeny. In *Mathematics of Evolution and Phylogeny*, O. Gascuel ed., pp. 353-380, Oxford Univ. Press, 2005