Mixing of Permutations by Biased Transposition*

Shahrzad Haddadan[1] and Peter Winkler[2]

1 **Sapienza University of Rome, Dipartimento di Informatica, Rome, Italy**
shahrzad.haddadan@gmail.com

2 **Department of Mathematics, Dartmouth College, Hanover, NH, USA**
peter.winkler@Dartmouth.edu

Abstract

Markov chains defined on the set of permutations of $1, 2, \ldots, n$ have been studied widely by mathematicians and theoretical computer scientists [15, 4, 1]. We consider chains in which a position $i < n$ is chosen uniformly at random, and then $\sigma(i)$ and $\sigma(i{+}1)$ are swapped with probability depending on $\sigma(i)$ and $\sigma(i{+}1)$. Our objective is to identify some conditions that assure rapid mixing.

One case of particular interest is what we call the "gladiator chain," in which each number g is assigned a "strength" s_g and when g and g' are swapped, g comes out on top with probability $s_g/(s_g + s_{g'})$. The stationary probability of this chain is the same as that of the slow-mixing "move ahead one" chain for self-organizing lists, but an open conjecture of Jim Fill's implies that all gladiator chains mix rapidly. Here we obtain some positive partial results by considering cases where the gladiators fall into only a few strength classes.

1998 ACM Subject Classification G.3 Probability and Statistics, G.2.1 Combinatorics

Keywords and phrases Markov chains, permutations, self organizing lists, mixing time

Digital Object Identifier 10.4230/LIPIcs.STACS.2017.41

1 Introduction

For any arbitrary natural number $n \in \mathbb{N}$, we define S_n to be the set that contains all the permutations of numbers $1, 2, \ldots, n$. A natural Markov chain on S_n is the chain which picks a number $1 \leq i \leq n{-}1$ uniformly at random and operating on $\sigma \in S_n$, puts $\sigma(i{+}1)$ ahead of $\sigma(i)$ w.p. $p_{\sigma(i),\sigma(i+1)}$. We call such chains *adjacent transposition* Markov chains. These chains have been studied widely for various choices of $p_{i,j}$ [15, 5, 1, 2].

In this paper, we consider the total variation mixing time, which is defined as the time it takes until the total variation distance between the distribution of the current state and stationarity is less than ϵ (where ϵ is some fixed quantity in (0,1)). For a Markov chain $\mathcal{M}$ we denote this time by $t_\epsilon(\mathcal{M})$, or if $\epsilon = 1/4$, simply by $t(\mathcal{M})$.

Jim Fill[1] conjectured that: If the adjacent transposition Markov chain is monotone, then it is rapidly mixing (meaning the mixing time is polynomial in n). Monotonicity in this context means that for all i, j satisfying $1 \leq i < j \leq n$, $p_{i,j} \geq 1/2$ and $p_{i,j-1} \leq p_{i,j}$ and $p_{i+1,j} \leq p_{i,j}$. [5].

Here we provide a brief history of the results on adjacent transposition Markov chains. All of the chains below are monotone and rapidly mixing. Wilson and Benjamini's papers [15, 1] led to Fill's conjecture [5]; Bhakta et al. [2] verified the conjecture in two cases.

* Research supported by NSF grant DMS-1162172.

[1] Fill considered the spectral gap (another measure of mixing) in his study. Here, we are interested in total variation mixing time, which in this case is within a polynomial factor of the spectral gap.

34th Symposium on Theoretical Aspects of Computer Science (STACS 2017).
Editors: Heribert Vollmer and Brigitte Vallée; Article No. 41; pp. 41:1–41:13
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1. The simple chain. In the case where $p_{i,j} = 1/2$ for all i and j, the chain will have a simple description: Given a permutation σ, pick two adjacent elements uniformly at random, and flip a fair coin to decide whether to swap them. We call this chain, whose stationary distribution is uniform, the *simple* chain. Ironically, proving precise mixing results for this chain was not simple. Many papers targeted this problem [4, 3], and finally Wilson [15] showed the mixing time for this chain is $\Theta(n^3 \log n)$ (proving lower and upper bounds within constant factors).

2. The constant-bias chain. After Wilson's paper, Benjamini et al. [1] studied the case where $p_{i,j} = p > 1/2$ for all $i < j$, and $p_{j,i} = 1-p$, and they showed the constant-bias chain, mixes in $\Theta(n^2)$ steps.

3. "Choose your weapon" and "league hierarchy" chains. The following two special cases were studied by Bhakta et al. [2]: the *choose your weapon chain* where $p_{i,j}$ is only dependent on i, and the *league hierarchy chain* given by a binary tree T with n leaves. Each interior node v of T is labeled with some probability $1/2 \leq q_v \leq 1$, and the leaves are labeled by numbers $1 \dots n$. The probability of putting j ahead of i for $j > i$ is equal to $p_{i,j} = q_{j \wedge i}$ where $j \wedge i$ is the node that is the lowest common ancestor of i and j in T. Bhakta et al. showed that the choose your weapon chain mixes in $\mathcal{O}(n^8 \log n)$ steps and the league hierarchy chain mixes in $\mathcal{O}(n^4 \log n)$ steps.

Here we are interested in *gladiator Markov chains* which constitute a subclass of the monotone adjacent transposition chains. These chains have a connection to self organizing lists, and were introduced by Jim Fill.

Fill was interested in probabilistic analysis of algorithms for *self-organizing lists* (SOLs). Self-organizing lists are data structures that facilitate linear searching in a list of records; the objective of a self-organizing list is to sort the records in non-decreasing order of their access frequencies [13]. Since these frequencies are not known in advance, an SOL algorithm aims to move a particular record ahead in the list when access on that record is requested. There are two widely used SOL algorithms: the *move ahead one* algorithm (MA1) and the *move to front* algorithm (MTF). In MA1, if the current state of the list is $(x_1, x_2, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n)$ and the ith record is requested for access, it will go ahead in the list only one position and the list will be modified to $(x_1, x_2, \dots, x_i, x_{i-1}, x_{i+1}, \dots, x_n)$. In MTF it will go to the front and the list will be modified to $(x_i, x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$. It appears that MA1 should perform better than MTF when the list is almost sorted and worse when the low frequency records are standing in front; however, this has not been analytically studied [6]. Considering the adjacent transposition Markov chain corresponding to MA1, Fill shows [5] that there are cases in which the chain is not rapidly mixing. Hence, he poses the question of sampling from the stationary distribution of MA1, and he introduces the gladiator chain which has the same stationary distribution as MA1 and seems to be rapidly mixing for arbitrary choice of parameters.

In a gladiator chain, each element i can be thought of as a gladiator with strength $s(i)$. Every permutation of numbers $1, 2, \dots n$ can be thought of as a ranking of gladiators. In each step of the Markov chain we choose $1 \leq k < n$ uniformly at random, i.e. we choose adjacent gladiators $\sigma(k) = i$ and $\sigma(k+1) = j$. These gladiators will fight over their position in ranking. With probability $p_{j,i} = s(i)/(s(i) + s(j))$, gladiator i will be the winner of the game and will be placed ahead of j in σ if it isn't already. With probability $p_{i,j} = 1 - p_{j,i}$, j is put ahead of i. If Fill's conjecture holds the gladiator chain must mix rapidly.

Another interesting Markov chain which has received a lot of attention is the *exclusion process* ([11, 10]). In this Markov chain we have a graph $G = \langle V, E\rangle$ and $m < |V|$ particles on the vertices of G. The sample space is the set containing all the different placements of the m particles on vertices of G. At each step of the Markov chain we pick a vertex v uniformly at random with probability $1/|V|$ and one of its adjacent vertices, w with probability $1/d(v)$. If there is a particle in one of the vertices and not the other one, we swap the position of the particle with a constant probability p. We are interested in the exclusion process when the graph is a finite path with n vertices. We will see that this chain has connections with the gladiator chain. This chain was studied by Benjamini et al. [1] and is known to mix in $\Theta(n^2)$ steps[2].

Our Contribution. We study the gladiator chain when the gladiators fall into a constant number of teams, and gladiators in each team have the same strength (Definition 1). We then extend the definition of the exclusion process (studied by Benjamini et al.) by allowing particles of different types to swap their positions on a linear line. We call this new chain a *linear particle system* (Definition 2). We show that mixing results for linear particle systems can produce mixing results in gladiator Markov chains (Theorem 4).

We study the linear particle system in which there are three particle types, and in Theorem 5 we extend Benjamini et al.'s result by showing the three particle system mixes rapidly; this is our main result. Having Theorem 5 we conclude that the following adjacent transposition chains mix rapidly, and hence confirming Fill's conjecture in these cases: the gladiator chain when there are three teams of same-strength gladiators, and the league hierarchy chain for ternary trees (extending Bhakta et al.'s work [2]).

▶ Remark. We believe the linear particle systems, like the exclusion processes are interesting Markov chains that may appear as components of other Markov chains, and thus would facilitate studying mixing times of other chains (For instance see Corollary 7 in which we extend a result about binary trees to ternary trees).

Definitions and results are presented in Section 2, along with the correspondence between gladiator chains and linear particle systems. Section 3 contains the proof that the three-type system mixes rapidly under certain conditions.

2 Definitions and Results

▶ **Definition 1. Gladiator chain** (Playing in teams). Consider the Markov chain on state space S_n that has the following properties: The set $[n]$ (i.e. gladiators) can be partitioned into subsets: $T_1, T_2, \ldots, T_k$ (k teams). We have the following strength function: $s : [n] \to \mathbb{R}$, $s(g) = s_j$ iff $g \in T_j$. At each step of Markov chain, we choose $i \in [n-1]$ uniformly at random. Given that we are at state σ, and $\sigma(i) = g, \sigma(i+1) = g'$, we put g ahead of g' with probability $\frac{s(g)}{s(g)+s(g')}$. We denote a gladiator chain having n gladiators playing in k different teams by $\mathcal{G}_k(n)$.[3]

[2] Benjamini et al. use this result to prove that the constant-biased adjacent transposition chain is rapidly mixing.

[3] Although the notation $\mathcal{G}_k(n_1, n_2, \ldots, n_k)$ would be more precise (n_i being cardinality of T_i), we avoid using it for simplicity and also because our analysis is not dependent on $n_1, n_2, \ldots, n_k$.

This is a reversible Markov chain and the stationary distribution π is

$$\pi(\sigma) = \prod_{i=1}^{n} s(i)^{\sigma^{-1}(i)}/Z. \qquad (Z \text{ is a normalizing factor.}) \tag{1}$$

Note that by writing $\sigma(i) = g$ we mean gladiator g is located at position i in σ. By writing $\sigma^{-1}(g)$ we are referring to the position of gladiator g in the permutation σ. We use this notation throughout the text and for permutations presenting both gladiators and particles.

▶ **Definition 2. Linear particle systems.** Assume we have k types of particles and of each type i, we have n_i indistinguishable copies. Let $n = \sum_{i=1}^{k} n_i$. Let $\Omega_{n_1,n_1,\dots n_k}$ be the state space containing all the different linear arrangements of these n particles. If the current state of the Markov chain is σ, choose $i \in [1, n-1]$ uniformly at random. Let $\sigma(i)$ be of type t and $\sigma(i+1)$ be of type t'. If $t = t'$ do nothing. Otherwise, put $\sigma(i)$ ahead of $\sigma(i+1)$ w.p. $p_{t,t'}$ and put $\sigma(i+1)$ ahead of $\sigma(i)$ w.p. $1 - p_{t,t'}$. We denote the linear particle system having n particles of k different types by $\mathcal{X}_k(n)$.

This chain is also a reversible Markov chain. In a special case where $p_{t,t'} = \frac{s(t)}{s(t)+s(t')}$ the stationary distribution π is

$$\pi(\sigma) = \prod_{i=1}^{n} s(i)^{\sigma^{-1}(i)}/Z'. \qquad (Z' \text{ is a normalizing factor.}) \tag{2}$$

▶ **Proposition 3.** *By regarding gladiators of equal strength as indistinguishable particles, we associate to any gladiator system a linear particle system.*

Note that the state space of the gladiator system has cardinality $n!$ for n different gladiators but the linear particle system has only $n!/(n_1!n_2!\dots n_k!)$ states, since particles of the same type are indistinguishable. Thus, $Z' \ll Z$. The following theorem, whose proof will be presented later, shows the connection between the mixing times of the two chains.

▶ **Theorem 4.** *Let $t(\mathcal{X}_k)$ and $t(\mathcal{G}_k)$ be respectively the mixing times for a linear particle system and its corresponding gladiator chain. Then, $t(\mathcal{G}_k) \leq \mathcal{O}(n^8)\ t(\mathcal{X}_k)$.*

Our main result, which extends the results of Benjamini et al. [1] on exclusion processes, is the following:

▶ **Theorem 5.** *Let $\mathcal{X}_3(n)$ be a linear particle system of Definition 2, having particles of type A, B and C. Assume that we have strength functions assigned to each particle type, namely $s_A < s_B < s_C$, and thus swapping probabilities $p_{B,A} = s_A/(s_A + s_B)$, $p_{B,C} = s_C/(s_C + s_B)$ and $p_{A,C} = s_C/(s_A + s_C)$. If $s_A/s_B, s_B/s_C < 1/2$, then the mixing time of $\mathcal{X}_3(n)$ satisfies $t(\mathcal{X}_3(n)) \leq \mathcal{O}(n^{10})$.*

▶ Remark. The condition $s_A/s_B, s_B/s_C \leq 1/2$ comes from the following simple bound on q-binomials that the reader should be able to verify easily: If $q < 1/2$ then, $\binom{m}{r}_q < 2^r < (\frac{1}{q})^r$. Better bounds on q-binomials would allow the result to be improved.

We will prove Theorem 5 in Section 3. Having Theorem 5, we deduce the following case of Fill's conjecture:

▶ **Corollary 6.** *The mixing time of $\mathcal{G}_3(n)$ satisfies $t(\mathcal{G}_3(n)) \leq \mathcal{O}(n^{18})$, if $s_A/s_B < 1/2$ and $s_B/s_C < 1/2$, where C is the strongest playing team among the three, and the gladiators in team B are stronger than the gladiators in team A.*

Proof. From Theorems 5 and 4. ◀

▶ **Corollary 7.** *(Generalization of league hierarchies) Let T be a ternary tree with n leaves. The children of each interior node v are labeled with labels $A(v)$, $B(v)$, and $C(v)$, and they have strength values $s_{A(v)}$, $s_{B(v)}$, and $s_{C(v)}$. The leaves are labeled by numbers $1, 2, \ldots, n$. The probability of putting j ahead of i for $j > i$ is equal to $p_{i,j} = s_{X(v)}/(s_{X(v)} + s_{Y(v)})$ where v is the node that is the lowest common ancestor of i and j in T, and $X(v)$ is the child of v which is an ancestor of j, and $Y(v)$ is the child of v which is an ancestor of i. If for each $v \in T$, $s_{A(v)}$, $s_{B(v)}$, and $s_{C(v)}$ satisfy the conditions in Theorem 5, then the mixing time of the league hierarchy chain is $\mathcal{O}(n^{10} \log n)$.*

Proof. A correspondence between the league hierarchies for binary trees (having n leaves) and a product of $n-1$ copies of the simple exclusion processes is presented in [2]. Using this correspondence and employing Benjamini's result ([1]), Bhakta et al prove that the binary tree league hierarchies mixes rapidly. The correspondence introduced in [2] is in fact a correspondence between the k-ary league hierarchy chains and $\mathcal{X}_k$, and they prove that the mixing times of the league hierarchy and the linear particle system are related. Theorem 5 can be used to extend the results in [2] to ternary trees satisfying the restrictions. ◀

We finish this section by proving Theorem 4.

2.1 Gladiators and Particles (Proof of Theorem 4)

Consider the gladiator chain $\mathcal{G}_k(n)$ for arbitrary n being the number of gladiators and k the number of teams. Assume that we have n_i gladiators playing at team i; hence, $\sum_{i=1}^{k} n_i = n$. At each step of the chain, one of two things is happening:

1. Whisking: gladiators of the same team are fighting.
2. Sifting: gladiators of different teams are fighting.

If we were restricted to whisking steps the chain would be equivalent to a product of several simple chains analyzed by Wilson [15]. If we were restricted to sifting steps the chain would be the linear particle system chain introduced in Definition 2. In order to study the mixing time of the gladiator chain we analyze sifting and whisking steps separately, and then we employ the following decomposition theorem:

▶ **Theorem 8** (Decomposition Theorem [9]). *Let $\mathcal{M}$ be a Markov chain on state space Ω partitioned into $\Omega_1, \Omega_2, \ldots, \Omega_k$. For each i, let $\mathcal{M}_i$ be the restriction of $\mathcal{M}$ to Ω_i that rejects moves going outside of Ω. Let $\pi_i(A) = \pi(A \cap \Omega_i)/\pi(\Omega_i)$ for $A \subseteq \Omega_i$. We define the Markov chain $\bar{\mathcal{M}}$ on state space $\{1, \ldots k\}$ as follows: $Pr_{\bar{\mathcal{M}}}(i,j) = \sum_{x \in \Omega_i, y \in \Omega_j} \pi_i(x) Pr_{\mathcal{M}}(x,y)$, where $Pr_{\mathcal{M}}$ and $Pr_{\bar{\mathcal{M}}}$ are transition probabilities of $\mathcal{M}$ and $\bar{\mathcal{M}}$ respectively. Then:*

$$t(\mathcal{M}) \leq 2t(\bar{\mathcal{M}}) \max_i \{t(\mathcal{M}_i)\}.$$

To apply the decomposition theorem, we partition S_n to $S_{\sigma_1,\sigma_2,\ldots,\sigma_k}$ for all choices of $\sigma_1 \in S_{n_1}, \sigma_2 \in S_{n_2}, \ldots, \sigma_k \in S_{n_k}$; each $S_{\sigma_1,\sigma_2,\ldots,\sigma_k}$ being the set of all permutations in S_n in which all the gladiators corresponding to particle i preserve the ordering associated to them by σ_i. The restriction of $\mathcal{G}_k(n)$ to $S_{\sigma_1,\sigma_2,\ldots,\sigma_k}$ is equivalent to $\mathcal{X}_k(n)$. We define $\bar{\mathcal{G}}$ to be the Markov chain on $\prod_{i=1}^{k} S_{n_i}$ with the following transition probabilities:

$$Pr_{\bar{\mathcal{G}}}(S_{\sigma_1,\sigma_2,\ldots,\sigma_i,\ldots,\sigma_k}, S_{\sigma_1,\sigma_2,\ldots,\sigma_i',\ldots,\sigma_k}) = \frac{1}{\pi(S_{\sigma_1,\sigma_2,\ldots,\sigma_i,\ldots,\sigma_k})} \sum_{\substack{x \in S_{\sigma_1,\sigma_2,\ldots,\sigma_i,\ldots,\sigma_k}, \\ y \in S_{\sigma_1,\sigma_2,\ldots,\sigma_i',\ldots,\sigma_k}}} \pi(x) Pr_{\mathcal{G}}(x,y),$$

where σ_i and σ'_i are only different in swapping j and j+1st elements and $Pr_{\mathcal{G}}(x,y) = 1/2(n-1)$ iff j and j+1st copies of particle i are adjacent in x and swapped in y. Moreover, we observe that:

$$\frac{1}{\pi(S_{\sigma_1,\sigma_2,\ldots,\sigma_i,\ldots,\sigma_k})} \sum_{\substack{x \in S_{\sigma_1,\sigma_2,\ldots,\sigma_i,\ldots,\sigma_k},\\ y \in S_{\sigma_1,\sigma_2,\ldots,\sigma'_i,\ldots,\sigma_k}}} \pi(x) \geq 1/(n-1).$$

We can verify the above equation by the following reasoning: consider an arbitrary permutation $z \in S_{\sigma_1,\sigma_2,\ldots,\sigma_i,\ldots,\sigma_k}$ in which jth and j+1st copies of particle i are not adjacent. We can map z to two other permutations z_1 and z_2 where in z_1 we take the the jth copy of particle i down to make it adjacent to the j+1st copy, and in z_2 we take the the j+1st copy of particle i up to make it adjacent to the jth copy. We will have $\pi(z)/\pi(z_1) = \pi(z_2)/\pi(z)$, and hence one of $\pi(z_1)$ or $\pi(z_2)$ will be larger than $\pi(z)$. This mapping is in worst case $n-1$ to 1, hence the above equation holds.

Having the above observations, we realize $\bar{\mathcal{G}}$ is the product of k adjacent transposition Markov chains, and in each of these Markov chains we swap two adjacent elements with probability at least $1/2(n-1)^2$. Let these chains be $\bar{\mathcal{G}}_1, \bar{\mathcal{G}}_2, \ldots, \bar{\mathcal{G}}_k$. By comparing the conductance (for more information about conductance, see [8]) of this chain to the simple chain analyzed by Wilson [15], for each i we will have $t(\bar{\mathcal{G}}_i) \leq n_i^8$. We know by a result of Bhakta et al. [2] that if $\bar{\mathcal{G}}$ is a product of k independent Markov chains $\{\bar{\mathcal{G}}_i\}_{i=1}^k$ and it updates each $\bar{\mathcal{G}}_i$ with probability p_i, then

$$t_\epsilon(\bar{\mathcal{G}}) \leq \max_{i=1,\ldots,n} \frac{2}{p_i}\, t_{\frac{\epsilon}{2k}}(\bar{\mathcal{G}}_i).$$

Plugging in $p_i = n_i/n$, we have $t(\bar{\mathcal{G}}) \leq \max(2n/n_i)n_i^8 \leq 2n^8$. Summing up and employing the Decomposition Theorem,

$$t(\mathcal{G}_k(n)) \leq 4n^8 t(\mathcal{X}_k(n)).$$

3 Three Particle Systems (Proof of the Main Theorem)

In this section we prove Theorem 5 which states that $t(\mathcal{X}_3(n)) \leq \mathcal{O}(n^{10})$ if $s_A/s_B, s_B/s_C \leq 1/2$.

Assume that we have a copies of particle A, b copies of particle B, and c copies of particle C. We denote the set containing all the different arrangements of these particles by $\Omega_{a,b,c}$. We introduce another Markov chain $\mathcal{X}_t(n)$ on the same sample space $\Omega_{a,b,c}$. Using the comparison method (see [12]) we will show that the mixing times of $\mathcal{X}_3(n)$ and $\mathcal{X}_t(n)$ are related.

Then we will use the path congestion technique to show $\mathcal{X}_t(n)$ mixes in polynomial time, and hence we deduce Theorem 5.

$$\text{mixing time of } \mathcal{X}_3(n) \xleftarrow[\text{technique}]{\text{Comparison}} \text{mixing time of } \mathcal{X}_t(n)$$

Notation. We denote the substring $\sigma(i)\sigma(i+1)\ldots\sigma(j)$ by $\sigma[i,j]$, and by B^t we refer to a string which is t copies of particle B.

▶ **Definition 9.** Let $\mathcal{X}_t(n)$ be a Markov chain on state space $\Omega_{a,b,c}$ and $n = a+b+c$. If the current state is σ, we choose natural numbers $1 \leq i < j \leq n$ uniformly at random and swap them following these rules (Figure 1):

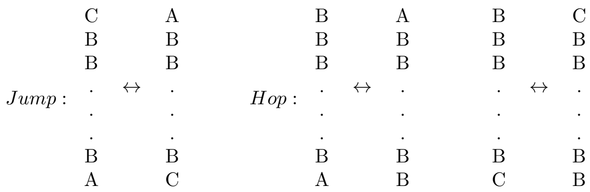

Figure 1 Jumps and Hops are the transitions in the Markov chain $\mathcal{X}_t$.

1. If $\sigma(i) = A$ and in $\sigma(j) = C$ or vice versa and $\sigma[i+1, j-1] = B^{j-i-1}$. Then, put $\sigma(i)$ and $\sigma(j)$ in increasing order of their strength w.p. $(s_C/s_A)^{(j-i)}/(1 + (s_C/s_A)^{(j-i)})$. With probability $1/(1+(s_C/s_A)^{(j-i)})$, put them in decreasing order. We call this move a *Jump* and we denote it by $\mathcal{J}_i^j(A, C)$ if $\sigma(i) = A$ and $\sigma(j) = C$; and $\mathcal{J}_i^j(C, A)$ for vice versa.
2. If $\sigma[i, j-1] = B^{j-i}$ and $\sigma(j) = A$ or if $\sigma[i+1, j] = B^{j-i}$ and $\sigma(i) = A$. Then, put $\sigma(i)$ and $\sigma(j)$ in increasing order of their strength w.p. $(s_B/s_A)^{j-i}/(1 + (s_B/s_A)^{j-i})$. With probability $1/(1 + (s_B/s_A)^{j-i})$, put them in decreasing order. We call this move a *Hop*, and we denote it by $\mathcal{H}_i^j(A, B)$ if $\sigma(i) = A$ and $\sigma(j) = B$; and $\mathcal{H}_i^j(B, A)$ for vice versa. Similar rules and notation apply when swapping B and C.
3. Else, do nothing.

It can be easily checked that $\mathcal{X}_t$ is reversible and its stationary distribution is the π in Equation 2.

▶ **Lemma 10.** $t(\mathcal{X}_3(n)) \leq 2n^4\ t(\mathcal{X}_t(n))$.

Proof. We use the comparison technique[4] in the proof of Lemma 10 (see [3, 12]). To any edge (σ, τ) in $\mathcal{X}_t$, we correspond a path from σ to τ in $\mathcal{X}_3$. Let $e_i(p, p')$ be a move in $\mathcal{X}_3$ which swaps particles p and p' located at positions i and $i + 1$ in an arrangement. To $e = (\sigma, \tau)$ making $\mathcal{J}_i^j(A, C)$ in $\mathcal{X}_t$, we correspond the following path in $\mathcal{X}_3$: $e_i(A, B), e_{i+1}(A, B), \ldots e_{j-2}(A, B), e_{j-1}(A, C), e_{j-2}(B, C), \ldots e_i(B, C)$. We denote this path by $\gamma_{\sigma\tau}$, and the set contaning all such paths by $\Gamma_{\mathcal{J}}$. Similarly, to $e = (\sigma, \tau)$ making $\mathcal{H}_i^j(A, B)$ in $\mathcal{X}_t$, we correspond the following path in $\mathcal{X}_3$: $e_i(A, B), e_{i+1}(A, B), \ldots e_{j-2}(A, B), e_{j-1}(A, B)$. We denote this path by $\gamma_{\sigma,\tau}$, and the set contaning all such paths by $\Gamma_{\mathcal{H}}$. Let $\Gamma = \{\gamma_{\sigma,\tau}\}_{\sigma,\tau\in\Omega_{a,b,c}} = \Gamma_{\mathcal{J}} \cup \Gamma_{\mathcal{H}}$.

We now bound the congestion placed by Γ on edges of $\mathcal{X}_3$. Consider an arbitrary $e = (\alpha, \beta)$ making swap $e_i(A, B)$ and assume $\alpha[i-t-1, i+d+1] = pB^tAB^dp'$ where p and p' are particles different from B. For any σ and τ in $\Omega_{a,b,c}$ if $e \in \gamma_{\sigma,\tau}$ then, there must be $i - t \leq j \leq i - 1$ and $i + 1 \leq k \leq i + d$ such that $\gamma_{\sigma\tau}$ corresponds to $\mathcal{H}_j^k(A, B)$ or it corresponds to $\mathcal{J}_j^{i+d+1}(A, p')$. Thus, the congestion placed on e only by paths in $\Gamma_{\mathcal{H}}$ is:

$$\frac{\sum_{\{\sigma,\tau|e\in\gamma_{\sigma\tau}\in\Gamma_{\mathcal{H}}\}} |\gamma_{\sigma,\tau}|\mathcal{C}(\sigma,\tau)}{\mathcal{C}(e)} = \sum_{j=i-t}^{i-1}\sum_{k=i+1}^{i+d} \frac{|\gamma_{\sigma,\tau}|(s_B/s_A)^{i-j}(1 + s_B/s_A)}{(1 + (s_B/s_A)^{k+1-j})}$$

$$\leq 2(d+t)\sum_{j'=1}^{t}\sum_{k'=1}^{d} \frac{(s_B/s_A)^{j'}(s_B/s_A)}{(1 + (s_B/s_A)^{j'+k'})} \leq 2t(d+t)\sum_{k'=1}^{d} \frac{s_B/s_A}{(s_B/s_A)^{k'}} \leq n^2.$$

[4] The comparison method was introduced by Diaconis and Saloff-Coste [3] and then Randall and Tetali extended it and employed it for analysis of Glauber dynamics [12]. For more information about this method we encourage the reader to refer to [8].

We can similarly show that the congestion placed on e by $\Gamma_{\mathcal{J}}$ is less than n^2, where n is the length of the arrangements or total number of particles. For each $e \in E(\mathcal{M})$, let $\mathcal{A}_e$ be the congestion Γ places on e. We will have $\mathcal{A}_e \le 2n^2$. We also know $\pi_{min} \ge q^{n(n+1)}$, where $q = \max\{s_A/s_B, s_B/s_C\}$ (we consider q to be a constant). Hence, employing the Comparison Theorem we have

$$t(\mathcal{X}_3(n)) \le \max_{e\in E(\mathcal{M})} \mathcal{A}_e \pi_{min} t(\mathcal{X}_t(n)) \le (2n^4) t(\mathcal{X}_t(n)). \qquad \blacktriangleleft$$

Having the above connection it suffices to bound the mixing time of $\mathcal{X}_t(n)$. In the rest of this section our goal is to bound $t(\mathcal{X}_t)$. We use the path congestion theorem, which is stated below. In particular, for any two arbitrary states $\sigma, \tau \in \Omega_{a,b,c}$, we introduce a path $\gamma_{\sigma,\tau}$. Then we show that none of the edges of the Markov chain $\mathcal{X}_t(n)$ is congested heavily by these paths. Formally, we employ Theorem 11 and in Theorem 13 we show that $t\left(\mathcal{X}_t(n)\right) \le \mathcal{O}(n^4)$.

▶ **Theorem 11** (Canonical Paths Theorem [7]). *Let $\mathcal{M}$ be a Markov chain with stationary distribution π and E the set of the edges in its underlying graph. For any two states σ and τ in the state space Ω we define a path $\gamma_{\sigma,\tau}$. The congestion factor for any edge $e \in E$ is denoted by Φ_e and is defined by $\Phi_e = \frac{1}{C(e)} \sum_{\substack{x,y\\ e\in\gamma_{x,y}}} \pi(x)\pi(y)$. We can bound the mixing time of $\mathcal{M}$ using the congestion factor: $t_\epsilon(\mathcal{M}) \le 8\Phi^2(\log \pi_{min}^{-1} + \log \epsilon)$, where $\Phi = \max_{e\in E} \phi_e$, $\pi_{min} = \min_{x\in\Omega} \pi(x)$ and ϵ is the convergence parameter.*

The Paths. For each $\sigma, \tau \in \Omega_{a,b,c}$, we introduce the following path in $\mathcal{X}_t$ from σ to τ: We partition σ and τ to $b+1$ blocks; the end points of these blocks are locations of Bs in τ. For instance if in τ, the first B is located at position i and the second B is located at position j then, the first block in both σ and τ is $[1, i]$, and the second is $[i+1, j]$. Starting from the first block, we change each block in two steps, first we use Jump moves and change the relative position of A and Cs in σ to become in the order in which they appear in τ. Then, we bring the B in that block to its location in τ. Formally, we repeat the following loop:

Notation. By saying $k = B_j(\sigma)$, we mean the jth copy of particle B is located at position k in σ.
Starting from σ, we repeat the following steps until τ is reached.
Initially, let $i, j = 1$.

1. Let $k = B_j(\tau)$. We define the jth *block* of σ and τ to be the substring starting from i and ending in k. Note that in τ, each blocks starts right after a B and ends with a B. In the jth iteration, the goal is to change $\sigma[i, k]$ until $\sigma[1, k] = \tau[1, k]$, i.e. the first j blocks equal in σ and τ.
2. Using Jumps, and starting from the lowest index i, we bring particles C or A down until A and C particles in the block $[i, k]$ have the same order in σ and τ.
3. We use Hops and bring the jth B in σ to $B_j(\tau)$. In this process, we may need to bring several copies of particle B out of the jth block in σ. In that case, we choose a random ordering of Bs and move them with respect to that order (details explained in the proof of Claim 12).
4. Set $i = B_j(\tau) + 1$.
5. Increment j.

▶ **Claim 12.** *Let $\{\gamma_{\sigma,\tau}\}_{\sigma,\tau\in\Omega_{a,b,c}}$ be the set of paths defined as above. Then, for any arbitrary edge e in the Markov chain $\mathcal{X}_t$ the congestion Φ_e, defined in Theorem 11 satisfies $\Phi_e \le n$.*

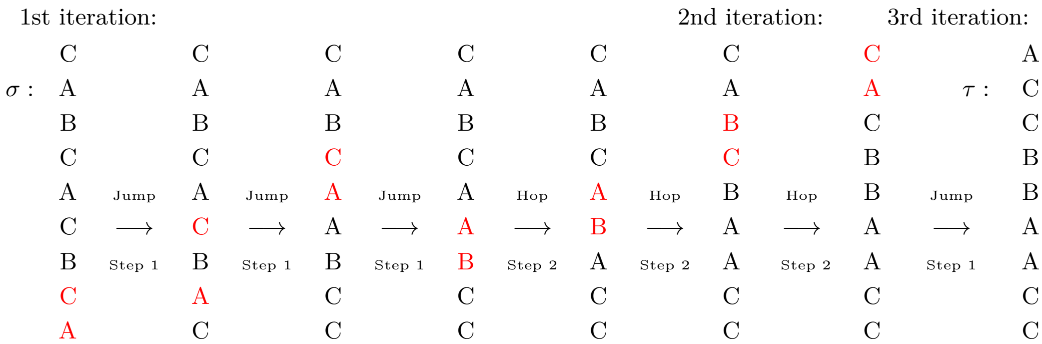

Figure 2 We use the path congestion technique to bound $t(\mathcal{X}_t)$. In each iteration we fix a block in σ until τ is reached.

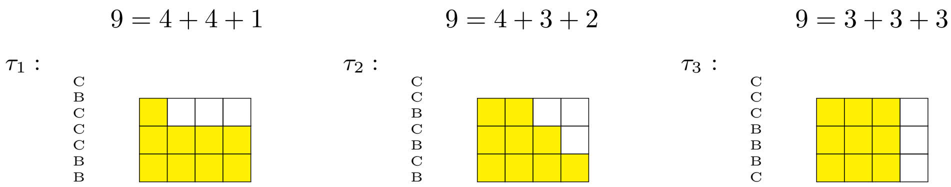

Figure 3 Correspondence of partition functions with q-binomials: There are three integer partitions of 9 that fit into a 3×4 rectangle, and there are three arrangements of gladiators in $\Omega_{0,3,4}$ with $q(\tau_1) = q(\tau_2) = q(\tau_3) = q^9$. i.e. the coefficient for q^9 in $\binom{7}{3}_q$ equals 3.

Proof. In order to verify the claim, we analyse the congestion of Jump and Hop edges separately. In both of the analyses, we consider an edge $e = (\alpha, \beta)$, and to any σ, τ such that $e \in \gamma_{\sigma,\tau}$ we assign a $\mathcal{F}_e(\sigma, \tau) \in \Omega_{a,b,c}$. The reverse image of $\mathcal{F}$ could be a subset of $\Omega_{a,b,c} \times \Omega_{a,b,c}$. However, using $q-$binomials [5] we show that $\sum_{\sigma,\tau \text{ are mapped to the same } \zeta} \pi(\sigma)\pi(\tau)$ is bounded by a polynomial function of n multiplied by $\pi(\zeta)$, and then we conclude the claim. A key factor of our analysis is the use of q-binomials. Note the following observations: Assume that we have no copies of particle A, b copies of B, and c copies of particle C. Let $M \in \Omega_{0,b,c}$ be the arrangement with maximum stationary probability, i.e. $M = B^b C^c$. Note that for each $\sigma \in \Omega_{0,b,c}$, $\pi(\sigma)/\pi(M) = (s_B/s_C)^t$, where t is the number of transpositions needed to get from M to σ. For a constant t, the number of σs requiring t transpositions is equal to the number of integer partitions of t fitting in an $b \times c$ rectangle (see Figure 3). Thus:

$$\sum_{\sigma \in \Omega_{0,b,c}} \frac{\pi(\sigma)}{\pi(M)} = \binom{b+c}{b}_q ; q = s_B/s_C.$$

Consider an edge $e = (\alpha, \beta)$ corresponding to $\mathcal{J}_k^{k+g}(C, A)$. Assume that $k = C_l(\alpha)$, $k + d = A_m(\alpha)$ (remember the notation $k = p_m(\sigma)$ meaning the mth copy of particle p is located at position k in σ), i.e. this edge is swapping the lth C with the mth A in α.

It follows from the way we set the paths that, for some j, $A_m(\alpha) \leq j < A_{m+1}(\alpha)$, $A_m(\sigma) =$

[5] More information about q-binomials can be found in Richard Stanley's course "Topics in Algebraic Combinatorics," Chapter 6 (see [14]). Here we use the the following bound which can be simply verified by expanding the formula for q-binomials: if $q < 1/2$ then, $\binom{m}{r}_q < \prod_{i=1}^{r} 1/(1-q) < 2^r < (\frac{1}{q})^r$.

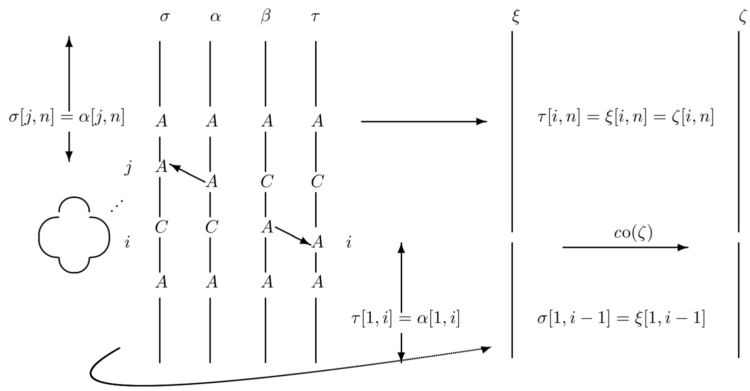

Figure 4 We define $\mathcal{F}_e(\sigma, \tau) = \zeta$. To produce ζ we first concat $\sigma[1, i-1]$ and $\tau[i, n]$ then, we substitute some particles.

j and for some i, $A_{m-1}(\beta) < i \leq A_m(\beta)$, $A_m(\tau) = i$. The preceding blocks of α have been changed in accordance with τ, and the succeeding blocks of α have not been changed yet, hence they resemble σ blocks. Therefore we have $\alpha[1, i-1] = \tau[1, i-1]$ and $\alpha[j+1, n] = \sigma[j+1, n]$ (see Figure 4).

We define the function $\mathcal{F}_e : \Omega_{a,b,c} \times \Omega_{a,b,c} \to \Omega_{a,b,c}$ as follows: For any σ, τ satisfying $e \in \gamma_{\sigma,\tau}$, let $\xi_{\sigma,\tau} := \sigma[1, i-1] | \tau[i, n]$ (the symbol $|$ denotes concatenation.) Since the arrangements of particles is changing we may have $\xi_{\sigma,\tau} \notin \Omega_{a,b,c}$. For instance we may have $\tau[i, n] \in \Omega_{x,y,z}$ and $\sigma[1, i-1] \in \Omega_{x',y',z'}$ but $x + x \neq a$ or $y + y' \neq b$ or $z + z' \neq c$. However, we know $a - (x + x') + (b - (y + y')) + (c - (z + z')) = 0$, which means there is a way to substitute the particles in $\sigma[1, i-1]$ to change ξ to ζ so that $\zeta \in \Omega_{a,b,c}$. We call this stage the substitution stage, in which we identify the particle or particles with extra copies in $\sigma[1, i-1]$, and we substitute the lowest copies of them with inadequate particles and produce $\zeta \in \Omega_{a,b,c}$. Then, we define $\mathcal{F}_e(\sigma, \tau) := \zeta$. For instance, if $a - (x + x') + (b - (y + y')) = -(c - (z + z'))$, then substitute the lowest $c - (z + z')$ copies of A and B with Cs, and produce $\mathcal{F}_e(\sigma, \tau) = \zeta$. The substitution stage will cause a *substitution cost*, we denote the substitution cost by $co(\zeta)$, and define it as: $co(\zeta) = \pi(\zeta)/\pi(\xi)$, where $\xi = \sigma[1, i-1] | \tau[i, n]$. Note that if we make t substitutions, the substitution cost is at most $(s_C/s_A)^t$. To make the analysis simpler we only analyze the worst case in which we assume we have substituted t Cs with As in $\sigma[1, i-1]$. This assumption also means that in $\sigma[i, j]$ we have t more As and t fewer Cs than in $\alpha[i, j]$.

Consider σ, τ such that $e \in \gamma_{\sigma,\tau}$. Let $\mathcal{F}_e(\sigma, \tau) = \zeta$. We have,

$$\frac{\pi(\zeta)}{\pi(\alpha)} = \left(\frac{\pi(\tau)}{\pi(\alpha)}\right)\left(\frac{\pi(\sigma)}{\pi(\alpha)}\right)\left(\frac{w^i(\alpha[i, j])}{w^i(\sigma[i, j])}\right) co(\zeta),$$

where the later term is the substitution cost, and $w^i(\sigma[i, j]) := \prod_{k=i}^{j} s(k)^{i+\sigma^{-1}(k)}$. Having $g = A_m(\alpha) - C_l(\alpha)$ we will get:

$$\Phi_e = (1 + (s_A/s_C)^g) \left(\sum_{\substack{\sigma; \\ \alpha[j+1,n]=\sigma[j+1,n]}} \frac{\pi(\sigma)}{\pi(\alpha)} \sum_{\substack{\tau; \\ \alpha[1,i-1]=\tau[1,i-1]}} \frac{\pi(\tau)}{\pi(\alpha)} \right) \pi(\alpha)$$

Let $\mathcal{S}_t$ be the set of all σs with t substitutions. We have:

$$\Phi_e \leq \sum_{\substack{\zeta \text{ needs } t \\ \text{substititions}}} \frac{1}{co(\zeta)} \sum_{\tau} \sum_{\sigma \in \mathcal{S}_t} \left(\frac{\pi(\mathcal{F}_e(\sigma, \tau))}{\pi(\alpha)}\right)\left(\frac{w^i(\sigma[i, j])}{w^i(\alpha[i, j])}\right) \pi(\alpha).$$

Let $M_t(\alpha)$ be the arrangement that we get from replacing the lowest t copies of particle C with copies of particle A in $\alpha[i,j]$. We have: $\sum_{\sigma\in S_t}\frac{w^i(\sigma[i,j])}{w^i(\alpha[i,j])}=\frac{w^i(M_t)Q^i_{\bar{B}}(M_t(\alpha))}{w^i(\alpha[i,j])}$, where $w^i(\sigma[i,j]) := \prod_{k=i}^{j} s(k)^{i+\sigma^{-1}(k)}$, and $Q^i_{\bar{B}}(M_t(\alpha)) := \sum_{\substack{\sigma:\text{fix the positions of all Bs}\\ \text{in } M_t(\alpha) \text{ and rearrange}\\ \text{the rest of particles}}} \pi(\sigma)/\pi(M_t(\alpha))$.

Note that $w^i(M_t)Q_{\bar{B}}(M_t) \le q^{t(t+1)-2t}w^i(\alpha[i,j])$, where $q = \max\{s_A/s_B, s_B/s_C\}$. This inequality holds because $Q_{\bar{B}}(M_t) \le \binom{y}{t}_{s_A/s_C} \le q^{-2t}$ and $w(M_t)/w(\alpha[i,j]) \le q^{t(t+1)}$.

Moreover, $\sum_{\substack{\zeta \text{ needs } t\\ \text{substititions}}} \frac{1}{co(\zeta)} \le \binom{t+b'}{t}_{q^2} \le q^{-2t}$, where b' is the number of Bs in $\sigma[0,i-1]$ and $q = \max\{s_A/s_B, s_B/s_C\}$.

Putting all of the above inequalities together, we will have that each edge of Move 2 is only congested by:

$$\Phi_e \le (1+q^g)\sum_t (q^{t(t+1)-4t}) \le n.$$

So far, we showed that any Jump edge is only congested by a factor of a polynomial function of n. Consider an edge corresponding to a Hop, namely e. We denote this edge by $e = (\alpha,\beta)$. Assume we are swapping A and B.

Consider a state σ traversing e to get to τ, and assume we traversed e while fixing block $[i,j]$. Since we are making a Hop, As and Cs in the block are fixed according to τ, and we are bringing the kth B to its position in τ.

Before we proceed to the proof there is a subtlety about using a Hop that needs to be explained. If A_k has to go down to reach its position in τ or if there is only one copy of it in the block there is no complication. Let's assume we have t copies of particle B in $\sigma[i,j]$. All of the t copies of B should move up and stand out of block $\sigma[i,j]$ to reach their position in τ. In order to accomplish this, we choose a subset S of $\{1_k,\dots 1_{t+k}\}$ uniformly at random and we move the elements of S in decreasing order of their index out of the block.

Assume, when going from σ to τ we used $e = (\alpha,\beta)$ and in $\alpha[i,j]$ we have t copies of particle B: $B_k,\dots,B_{k+t}$ and swapping $B_{k+l}, B_{k+l+1},\dots,B_{k+d}$ with the next A. We have, $\tau[1,i] = \alpha[1,i]$, $\sigma[j+t,n] = \alpha[j+t,n]$, and for any i if $B_{k+i}(\alpha) < B_{l+k}(\alpha)$ then, $B_{k+i}(\alpha) = B_{k+i}(\sigma)$. The following information about S can be determined by examining α and β: $B_{k+d+1},\dots B_{k+t} \notin S$ while S may contain any of $B_k,\dots B_{k+l}$. Therefore, among the random paths connecting σ to τ, there are 2^l subsets traversing through e and hence the congestion they place on e is $\pi(\tau)\pi(\sigma)/2^{t-l}$.

To bound Φ_e for each e we introduce correspondence $\mathcal{F}_e : \Omega_{a,b,c}\times\Omega_{a,b,c} \to \Omega_{a,b,c}$ satisfying:

$$\forall\zeta \in \mathcal{F}_e(\Omega_{a,b,c});\ \frac{\sum_{\substack{\sigma,\tau;\\ \mathcal{F}_e^{-1}(\zeta)=(\sigma,\tau)}} \pi(\sigma)\pi(\tau)}{\pi(\alpha)} \le 2^{t-l}\pi(\zeta); \tag{3}$$

where c is the number of Cs in $\alpha[i,j]$ and $\mathcal{F}_e(\sigma,\tau) \ne$ NULL if and only if, $e = (\alpha,\beta) \in \gamma_{\sigma,\tau}$.

Let σ and τ be two ends of a path traversing through e, we define $\mathcal{F}_e := \sigma[1,i-1]|\tau[i,n]$, to verify Equation 3 take $\zeta = \mathcal{F}_e(\sigma,\tau)$. We have, $\frac{\pi(\sigma)\pi(\tau)}{\pi(\alpha)} = \frac{\pi(\sigma)}{\pi(\alpha)}\frac{\pi(\tau)}{\pi(\alpha)}\pi(\alpha)$. Thus,

$$\begin{aligned}\frac{\pi(\zeta)}{\pi(\alpha)} &= \frac{\pi(\zeta[1,i-1])}{\pi(\alpha[1,i-1])}\frac{\pi(\zeta[i,j-1])}{\pi(\alpha[i,j-1])}\frac{\pi(\zeta[j,n])}{\pi(\alpha[j,n])} = \frac{\pi(\sigma[1,i-1])}{\pi(\alpha[1,i-1])}\frac{\pi(\tau[i,j-1])}{\pi(\alpha[i,j-1])}\frac{\pi(\tau[j,n])}{\pi(\alpha[j,n])}\\ &= \frac{\pi(\sigma')}{\pi(\sigma)}\frac{\pi(\sigma)}{\pi(\alpha)}\frac{\pi(\tau)}{\pi(\alpha)},\end{aligned}$$

where σ' is the following arrangement: $\sigma' := \alpha[1,i-1]|\sigma[i,j-1]|\alpha[j,n]$. We have $\pi(\sigma')/\pi(\alpha) = \pi(\sigma[i,j])/\pi(\alpha[i,j])$. Hence,

$$\sum_{\sigma,\tau;\mathcal{F}(\sigma,\tau)=\zeta} \frac{\pi(\sigma)\pi(\tau)}{\pi(\alpha)} = \sum_{\substack{\sigma,\tau\\ \mathcal{F}(\sigma,\tau)=\zeta}} \frac{\pi(\sigma')}{\pi(\sigma)}\pi(\zeta).$$

Since we have $t-l$ Bs with undecided position between $j{-}i$ other elements we have $\sum \frac{\pi(\sigma')}{\pi(\sigma)} \leq \binom{j-i+t-l}{t-l}_q$, where $q = \max\{s_A/s_B, s_B/s_C\}$. Thus, we have $\sum \frac{\pi(\sigma')}{\pi(\sigma)} \leq 2^{t-l}$. Hence, the congestion placed on e is:

$$\Phi_{e=(\alpha,\beta)} = (1+q^g) \sum_{\substack{\sigma,\tau \\ e\in\gamma_{\sigma,\tau}}} \frac{\pi(\sigma)\pi(\tau)}{\pi(\alpha)2^{t-l}} \leq 1.$$

Summing up, we showed the for any arbitrary edge e, $\Phi_e \leq \max\{n, 1\}$. ◀

Having the above claim, we now use the path congestion Theorem (Theorem 11) to bound $t(\mathcal{X}_t(n))$:

▶ **Theorem 13.** *If $s_A/s_B, s_B/s_C \leq 1/2$, then $t(\mathcal{X}_t(n)) \leq \mathcal{O}(n^4)$.*

Proof. Since $\pi_{min} \geq q^{n(n+1)}$, q being maximum of s_A/s_B and s_B/s_C, we can apply Theorem 11 and we will have, $t_\epsilon(\mathcal{X}_t) \leq 8n^2(n^2 + \ln(\epsilon^{-1})) \implies t(\mathcal{X}_t) \leq 8n^4$. ◀

Finally, from Lemma 10 and Theorem 13 we conclude Theorem 5.

Acknowledgements. We would like to thank Dana Randall for a very helpful conversation about the gladiator problem and Fill's conjecture, and Sergi Elizalde for his help and knowledge concerning generating functions.

References

1 I. Benjamini, N. Berger, C. Hoffman, and Mossel E. Mixing times of the biased card shuffling and the asymmetric exclusion process. *Transactions of the American Mathematical Society*, 357:3013–3029, 2005.

2 P. Bhakta, S. Miracle, D. Randall, and A. Streib. Mixing times of Markov chains for self-organized lists and biased permutations. *25th Symposium on Discrete Algorithms (SODA)*, 2014.

3 P. Diaconis and L. Saloff-Coste. Comparison techniques for random walks on finite groups. *The annals of applied probability*, 21:1149–1178, 1993.

4 P. Diaconis and M. Shahshahani. Generating a random permutation with random transpositions. *Probablity theory and related fields*, 57:159–179, 1981.

5 J. Fill. An interesting spectral gap problem, 2003. Unpublished manuscript.

6 J. H. Hester and D. S. Hirscheberg. Self-organizing linear search. *ACM Computing survey*, 19:295–311, 1985.

7 M. Jerrum and Sinclair A. Approximating the permanent. *SIAM Journal on computing*, 18:1149–1178, 1989.

8 D. A. Levin, Y. Peres, and E. L. Wilmer. *Markov chains and Mixing times.* American Mathematical Society, 2009.

9 R. Martin and D. Randall. Disjoint decomposition of Markov chains and sampling circuits in cayley graphs. *Combinatorics, Probability and Computing*, 15:411–448, 2006.

10 B. Morris. The mixing time for simple exclusion. *Annals of probabilty*, 16:63–67, 1976.

11 R. I. Oliveria. Mixing of the symmetric exclusion processes in terms of the corresponding single-particle random walk. *Annals of probablity*, 41:871–913, 2013.

12 D. Randall and P. Tatali. Analyzing glauber dynamics by comparision of Markov chains. *Journal of Mathematical Physics*, 41:1598–1615, 2000.

13 R. Rivest. On self-organizing sequential search heuristics. *Communication of the ACM*, 19:63–67, 1976.

14 R. P. Stanley. Topics in algebraic combinatorics, 2012. Course notes for Mathematics, 2012. URL: `http//www-math.mit.edu/~rstan/algcomb/algcomb.pdf`.

15 D. Wilson. Mixing times of lozenge tiling and card shuffling Markov chains. *The annals of applied probablity*, 1:274–325, 2004.