A Hopf-Algebraic Lift of the Down-Up Markov Chain on Partitions to Permutations

C. Y. Amy Pang

Abstract

In "Card shuffling and the decomposition of tensor products", Jason Fulman explores a Markov chain on partition diagrams, where each step removes a box then adds a new box. The stationary distribution of this chain is the Plancherel measure, which suggests that it is the image under the RSK-shape map of a chain on permutations with uniform stationary distribution. Fulman proves that one possible lift, starting from the identity permutation, is the "top-to-random shuffle": remove the top card of a deck, and reinsert it at a uniformly chosen position. We construct a different lift, valid from more initial distributions, using a new and very general result regarding Markov chains from non-negative linear maps. This abstract result essentially reduces the lift construction to a well-known fact: the algebra of symmetric functions is a subquotient of the Malvenuto-Reutenauer Hopf algebra of permutations. The multistep transition probabilities of this new lift chain, starting from the identity permutation, agree with those of the top-to-random chain.

1 Introduction

In [Ful04], Jason Fulman studies tensor powers of permutation representations of a finite group G via a family of Markov chains on the irreducible representations of G. For a fixed subgroup H of G, one step of the chain proceeds as follows:

1. Restrict the current irreducible representation to H.
2. Induce this restricted representation to G.
3. Pick an irreducible constituent of this induced representation, with probability proportional to the dimension of its isotypic component - that is, proportional to the multiplicity of this irreducible in the induced representation, multiplied by the dimension of the irreducible.

[Ful04, Prop. 2.1] asserts that the stationary distribution of this chain is the well-studied *Plancherel measure*: $\pi(x) = \frac{(\dim x)^2}{|G|}$ for each irreducible representation x of G. The eigenvalues and eigenfunctions of this chain can be explicitly calculated from the irreducible characters of G [Ful04, Prop. 2.3].

This paper concentrates on the case where G, H are the symmetric groups $\mathfrak{S}_n, \mathfrak{S}_{n-1}$, for which the chain has a completely combinatorial description [Ful08, Prop. 4.2]. (This chain is an archetype for several of Fulman's frameworks: [Ful08] calculates the convergence rates for a different family of Markov chains on irreducible representations, [Ful09] studies down-up chains on branching graphs, and [Ful05] analyses related up-down chains on branching graphs.) The irreducible representations of $\mathfrak{S}_n$ are indexed by partitions of n, so restriction to $\mathfrak{S}_{n-1}$ corresponds to removing a box from the partition diagram, and the induction back to $\mathfrak{S}_n$ adds a (possibly different) box. The probability of removing or adding a certain box depends on the dimension of the irreducible representations.

Definition 1.1. One step of the *down-up chain on partitions*, starting at $\lambda \vdash n$, proceeds as follows:

1. Remove a box from λ to obtain ν, with probability $\frac{\dim \nu}{\dim \lambda}$.
2. Add a box to ν to obtain μ, with probability $\frac{1}{n}\frac{\dim \mu}{\dim \nu}$.

For an irreducible representation x of $\mathfrak{S}_n$, its dimension is the number of standard tableaux of shape x (see Section 2 for the relevant definitions). Hence the Plancherel stationary distribution $\frac{(\dim x)^2}{n!}$ counts the pairs of standard tableaux with shape x. Now the famous Robinson-Schensted-Knuth (RSK) bijection [Sta99, Sec. 7.11; Ful97, Sec. 4] sends a permutation to a pair of standard tableaux of the same shape. This prompts a natural question: is there a Markov chain on permutations, with a uniform stationary distribution, which "lumps" to the down-up chain on partitions via the RSK-shape map? That is, can one build a chain $\{X_t\}$ on permutations such that, if we observe only $\{\text{shape}(\text{RSK}\, X_t)\}$, the RSK shape of this process, then we see precisely the down-up chain on partitions? Such a lift chain $\{X_t\}$ would "explain" the Plancherel stationary distribution of the partition chain, a notion popularised by [DS05; CW07] on the PASEP chain.

[Ful04, Th. 3.1] proves that the top-to-random shuffle of a deck of cards with values $\{1, 2, \dots, n\}$ is one such lift, but only when the deck starts in ascending order (the identity permutation). Here, identify a permutation σ with the deck that has the card with value $\sigma(1)$ on top, $\sigma(2)$ second from the top, and so on, with card $\sigma(n)$ at the bottom. At each step of this chain, the top card is removed, then reinserted into a uniformly chosen position. This shuffle has been thoroughly analysed over the literature: [AD87, Sec. 1, Sec. 2] uses a strong uniform time to elegantly show that roughly $n \log n$ iterations are required to randomise the deck, and [DFP92, Cor. 2.1] finds the explicit probabilities of achieving a particular permutation after any given number of shuffles. The time-reversal of the top-to-random shuffle is the equally well-studied Tsetlin library: [Hiv+11, Sec. 4.6] describes an explicit algorithm for an eigenbasis, and [Pha91] derives the spectrum for a weighted version.

A More General Lift

Fulman finds it "surprising" that the top-to-random shuffle should be connected to the down-up chain on partitions. The primary goal of the present paper is to provide a related lift that is less mysterious: the proof falls effortlessly from two well-known maps of Hopf algebras, because of a new, extremely general result (Theorem 5.3) regarding lifts of chains from linear maps. Other advantages of this new lift are a wider range of valid starting distributions, and the existence of an intermediate lumping by RSK insertion tableau.

Here is an intuitive interpretation of the lift: you keep an electronic to-do list of n tasks. Each day, you complete the task at the top of the list, and are handed a new task, which you add to the list in a position depending on its urgency (more urgent tasks are placed closer to the top). Assume the incoming tasks are equally distributed in urgency, so they are each inserted into the list in a uniform position. Now assign a permutation (in one-line notation) to each daily list, by writing 1 for the task that's been on the list for the longest time, 2 for the next oldest task, and so on, so n denotes the task you received today. The resulting Markov chain on permutations is the "twisted-top-to-random-with-standardisation" of Definition 1.2 below. This chain also arises from performing the top-to-random shuffle and keeping track of the relative times that the cards were last touched, instead of their values.

Definition 1.2. The *twisted-top-to-random-with-standardisation* (resp. *bottom-to-random-with-standardisation*) Markov chain has as its state space the symmetric group $\mathfrak{S}_n$, viewed as words in the letters $\{1, 2, \dots, n\}$, with each letter occuring precisely once. Each timestep proceeds as follows:

1. Delete the first (resp. last) letter, and call this letter i.
2. For each letter $j > i$, relabel this letter as $j-1$. Hence the word now consists of the letters $\{1, 2, \dots, n-1\}$. (This is known as *standardisation*.)
3. Insert the letter n at a uniformly chosen position.

The *top-to-random-with-standardisation* Markov chain has the same state space, and each timestep proceeds as follows:

1. Delete the first letter, and call this letter i.
2. For each letter $j < i$, relabel this letter as $j+1$. Hence the word now consists of the letters $\{2, 3, \dots, n\}$.
3. Insert the letter 1 at a uniformly chosen position.

In order to describe the intermediate chain on standard tableaux, this paper will concentrate on bottom-to-random instead of top-to-random, but the end of Section 7 will show that all three chains defined above are lifts, with the same set of valid initial distributions. To contrast these lifts and those of [Ful04, Th. 3.1], here are a few steps of their possible trajectories in degree 5:

Top-to-random
(without standardisation)

$$\begin{tikzcd}[column sep=small]
(5,2,4,3,1) \arrow[dr] & & (2,4,3,5,1) \arrow[dr] & & (4,3,2,5,1) \arrow[dr] & & (4,3,2,5,1) \\
& (2,4,3,1) \arrow[ur] & & (4,3,5,1) \arrow[ur] & & (3,2,5,1) \arrow[ur] &
\end{tikzcd}$$

Bottom-to-random
(without standardisation)

$$\begin{tikzcd}[column sep=small]
(5,2,4,3,1) \arrow[dr] & & (5,2,4,1,3) \arrow[dr] & & (5,2,3,4,1) \arrow[dr] & & (1,5,2,3,4) \\
& (5,2,4,3) \arrow[ur] & & (5,2,4,1) \arrow[ur] & & (5,2,3,4) \arrow[ur] &
\end{tikzcd}$$

Twisted-top-to-random-
with-standardisation

$$\begin{tikzcd}[column sep=small]
(5,2,4,3,1) \arrow[dr] & & (2,4,3,5,1) \arrow[dr] & & (3,2,5,4,1) \arrow[dr] & & (5,2,4,3,1) \\
& (2,4,3,1) \arrow[ur] & & (3,2,4,1) \arrow[ur] & & (2,4,3,1) \arrow[ur] &
\end{tikzcd}$$

Botttom-to-random-
with-standardisation

$$\begin{tikzcd}[column sep=small]
(5,2,4,3,1) \arrow[dr] & & (4,1,3,5,2) \arrow[dr] & & (3,1,5,2,4) \arrow[dr] & & (5,3,1,4,2) \\
& (4,1,3,2) \arrow[ur] & & (3,1,2,4) \arrow[ur] & & (3,1,4,2) \arrow[ur] &
\end{tikzcd}$$

Top-to-random-with-standardisation

$$\begin{tikzcd}[column sep=small]
(5,2,4,3,1) \arrow[dr] & & (3,5,4,1,2) \arrow[dr] & & (5,4,1,2,3) \arrow[dr] & & (1,5,2,3,4) \\
& (3,5,4,2) \arrow[ur] & & (5,4,2,3) \arrow[ur] & & (5,2,3,4) \arrow[ur] &
\end{tikzcd}$$

Main Theorem. *Let $\{\bar{X}_t\}$ denote the down-up chain on partitions, and $\{X_t\}$ the bottom-to-random-with-standardisation chain. Then $\{X_t\}$ lumps to $\{\bar{X}_t\}$ weakly via taking the shape of the RSK tableau, from any initial distribution X_0 where any two permutations having the same RSK insertion tableau are equally probable. Furthermore, when X_0 is concentrated at the identity permutation, then, for all values of t, the distribution of X_t is equal to that of the bottom-to-random shuffle after t steps, starting also at the identity.*

Another way of phrasing the final sentence is: for any fixed t and any permutation σ, the probability of going from the identity to σ after t steps of bottom-to-random-with-standardisation is the same as the probability of going from the identity to σ after t steps of Fulman's bottom-to-random shuffle. Consequently, if starting from the identity, both chains have the same convergence rate of $n \log n$. However, the trajectories from the identity to σ are in general very different for the two chains. For example, the only way to move from 1234 to 3142 in two bottom-to-random shuffles is via 1423, whereas moving between these same two states under two steps of bottom-to-random-with-standardisation necessarily goes through 4123. So one cannot deduce from the Main Theorem that bottom-to-random is also a lift of the down-up chain on partitions.

Note that there are (many) starting distributions X_0 on permutations which lift from any given starting distribution $\bar{X}_0$ on partitions - for example, set $X_0(\sigma) = \frac{1}{(\dim \lambda)^2}\bar{X}_0(\lambda)$ whenever shape(RSK(σ)) $= \lambda$. (One way to achieve this $\bar{X}_0$ is to use [GNW79, Sec. 3(i)] to uniformly choose two standard tableaux of the required shape, then apply the inverse of RSK to obtain a permutation.)

Fulman's lift theorem holds in a more general setting: for the restriction-then-induction chain with any Young subgroup H of the symmetric group (or even any "mixture" of such chains), there is a card shuffle, starting at the identity, which lumps to it. The present construction shows that their "with standardisation" versions are also lifts, if the initial distribution again assigns equal probabilities to all permutations with the same RSK insertion tableau. In general, these two lifts do not have the same multistep transition probabilities from the identity permutation. See Corollary 7.4 and the third paragraph after it for a comparison of the two sets of lifts. It's an interesting open problem to "explain" the similarity between the two sets of lifts.

Proof Strategy

Constructing the new lift chains is a two-step process - first, lift the chain on partitions to the level of standard tableaux, then lift this chain on tableaux to bottom-to-random-with-standardisation.

As Section 3 will explain, the down-up chain on partitions arises via the Doob transform from a linear map $\mathbf{T}_n := \frac{1}{n} m \Delta_{n-1,1}$, on the Schur function basis of the Hopf algebra of symmetric functions. The first lift uses [Pan15, Th. 4.1] (a mild extension of [Pan14, Th. 4.7.1]): applying the Doob transform to a quotient $\bar{\mathcal{H}}$ of a Hopf algebra $\mathcal{H}$ (under certain conditions on their bases) creates a lumped version of the Markov chain on $\mathcal{H}$. The symmetric functions is a quotient of the Poirer-Reutenauer Hopf algebra on standard tableaux, so the chain produced by the linear map $\mathbf{T}_n$ on this tableaux algebra is a (strong) lift of the down-up chain on partitions.

The second lift relies on a new result (Theorem 5.3) - the chain created by the Doob transform on a subalgebra (under certain conditions on the basis) is also a lumping, but only for certain initial distributions (specified explicitly in the theorem statement). Since the Hopf algebra on standard tableaux is a subalgebra of the Malvenuto-Reutenauer Hopf algebra on permutations, the chain on tableaux lifts (weakly) to the level of permutations, and this lift is precisely the bottom-to-random-with-standardisation.

These two strategies of producing lifts are actually extremely flexible, applicable to Markov chains from a wide range of linear maps on various Hopf algebras. These maps include the descent operators of [Pan15], so the lift constructed here is a neat application of that big body of theory.

Organisation of the paper

Section 2 lays out all the necessary notation. Section 3 explains how the Doob transform creates Markov chains from linear maps, and casts the down-up chain on partitions in this light. Section 4 recalls that quotients of Hopf algebras lead to lumpings of their associated Markov chains, and applies this to lift the down-up chain on partitions to a chain on standard tableaux. Section 5 proves a new result relating Markov chains from subalgebras, and uses this to lift the down-up chain from tableaux to permutations. This completes the proof of the first assertion of the Main Theorem. Section 6 proves the second assertion of the Main Theorem: the new lift constructed here has the same multistep transition probabilities from the identity permutation as Fulman's lifts. Section 7 generalises these two lifting processes to the descent operator chains of [Pan15] and beyond, concluding with a more exciting lift of the down-up chain on partitions.

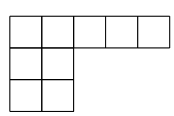
Acknowledgements

I would like to thank Nathan Williams for pointing out to me this question, Persi Diaconis, Jason Fulman and Franco Saliola for numerous helpful conversations, and Grégory Châtel, Mathieu Guay-Paquet, Aaron Lauve, Simon Rubenstein-Salzedo and Yannic Vargas for useful comments. SAGE computer software [Ste+15] was very useful.

2 Notation

Combinatorial Notation

A *partition* λ is a weakly-decreasing sequence of positive integers: $\lambda := (\lambda_1, \dots, \lambda_l)$ with $\lambda_1 \geq \cdots \geq \lambda_l > 0$. This is a *partition of* n, denoted $\lambda \vdash n$, if $\lambda_1 + \cdots + \lambda_l = n$. We will think of a partition λ as a diagram of left-justified boxes with λ_1 boxes in the topmost row, λ_2 boxes in the second row, etc. For example, (5,2,2) is a partition of 9, and below is its diagram.



A *tableau of shape* λ is a filling of each of the boxes in λ with a positive integer; for this paper, we add the condition that every row and column is strictly increasing. The *shift* of a tableaux T by an integer k, denoted $T[k]$, increases each filling of T by k. A tableau is *standard* if it is filled with $\{1, 2, \dots, n\}$, each integer occurring once. If no two boxes of a tableau T has the same filling, then its *standardisation* $\mathrm{std}(T)$ is computed by replacing the smallest filling by 1, the second smallest filling by 2, and so on. Clearly $\mathrm{std}(T)$ is a standard tableau, of the same shape as T. A box b of T

is *removable* if the difference $T \backslash b$ is a tableau. Below shows a tableau of shape $(5, 2, 2)$, its shift by 3, and its standardisation. The removable boxes in the first tableau are 11 and 13.

$$\underset{T}{\begin{array}{|c|c|c|c|c|}\hline 1 & 2 & 5 & 10 & 13 \\ \hline 4 & 8 \\ \cline{1-2} 6 & 11 \\ \cline{1-2}\end{array}} \qquad \underset{T[3]}{\begin{array}{|c|c|c|c|c|}\hline 4 & 5 & 8 & 13 & 16 \\ \hline 7 & 11 \\ \cline{1-2} 9 & 14 \\ \cline{1-2}\end{array}} \qquad \underset{\mathrm{std}(T)}{\begin{array}{|c|c|c|c|c|}\hline 1 & 2 & 4 & 7 & 9 \\ \hline 3 & 6 \\ \cline{1-2} 5 & 8 \\ \cline{1-2}\end{array}}$$

In the same vein, this paper will regard permutations as "standard words", using one-line notation: $\sigma := (\sigma(1), \dots, \sigma(n))$. The *length* of a word is its number of letters. The *shift* of a word σ by an integer k, denoted $\sigma[k]$, increases each letter of σ by k. If a word σ has all letters distinct, then its *standardisation* $\mathrm{std}(\sigma)$ is computed by replacing the smallest letter by 1, the second smallest letter by 2, and so on. Clearly $\mathrm{std}(\sigma)$ is a permutation. For example, $\sigma = (6, 1, 4, 8, 2, 11, 10, 13, 5)$ is a word of length 9. Its shift by 3 is $\sigma[3] = (9, 4, 7, 11, 5, 14, 13, 16, 8)$, and its standardisation is $\mathrm{std}(\sigma) = (5, 1, 3, 6, 2, 8, 7, 9, 4)$.

Matrix notation

Given a matrix A , let $A(x, y)$ denote its entry in row x, column y, and write A^T for the transpose of A.

Let V be a vector space with basis $\mathcal{B}$, and $\mathbf{T} : V \to V$ be a linear map. Write $[\mathbf{T}]_{\mathcal{B}}$ for the matrix of $\mathbf{T}$ with respect to $\mathcal{B}$. In other words, the entries of $[\mathbf{T}]_{\mathcal{B}}$ satisfy

$$\mathbf{T}(x) = \sum_{y \in \mathcal{B}} [\mathbf{T}]_{\mathcal{B}} (y, x) y$$

for each $x \in \mathcal{B}$.

3 The Doob Transform and a chain on Partitions

To start, here is a quick summary of the Markov chain facts required for this work. A (discrete time) Markov chain is a sequence of random variables $\{X_t\}$, where each X_t belongs to the *state space* Ω. All Markov chains here are time-independent and have a finite state space. Hence they are each described by an $|\Omega|$-by-$|\Omega|$ *transition matrix* K: for any time t,

$$P\{X_t = x_t | X_0 = x_0, X_1 = x_1, \dots, X_{t-1} = x_{t-1}\} = P\{X_t = x_t | X_{t-1} = x_{t-1}\} := K(x_{t-1}, x_t).$$

(Here, $P\{X|Y\}$ is the probability of event X given event Y.) If the probability distribution of X_t is expressed as a row vector, then taking one step of the chain is equivalent to multiplication by K on the right: $X_t = X_{t-1} K$. Note that a matrix K specifies a Markov chain in this manner if and only if $K(x, y) \geq 0$ for all $x, y \in \Omega$, and $\sum_{y \in \Omega} K(x, y) = 1$ for each $x \in \Omega$. We refer the reader to [LPW09; KS60] for the basics of Markov chain theory.

This paper concerns chains which arise from linear maps according to the Doob transform, which Theorem 3.1 below will explain. To give some motivation, recall the highly successful framework of random walks on groups [SC04] (and its extensions to left regular bands [Bro00] and R-trivial monoids [Ayy+15]): choose a measure Q on a finite group G, then define the transition probabilities as $K(x, y) := Q(yx^{-1})$. (The bottom-to-random shuffle (without standardisation) falls into this category with $G = \mathfrak{S}_n$, as do the other shuffles in Fulman's family of lifts, but the chains with standardisation do not.) The algebraic structure of the group, such as its representation theory, aids greatly in analysing the Markov chain, and one might want to replicate this philosophy on other algebraic structures. One possible generalisation starts by viewing the transition matrix as the matrix of multiplication by $\sum_{g\in G} Q(g)g$. More precisely, the set of group elements G form a basis of the group algebra $\mathbb{R}G$, on which there is a linear transformation $\mathbf{T}: \mathbb{R}G \to \mathbb{R}G$, $\mathbf{T}(x) := \sum_{g\in G} Q(g)gx$, related to the transition matrix via $K = [\mathbf{T}]_G^T$. Hence a naive generalisation is to declare new transition matrices to be $K := [\mathbf{T}]_{\mathcal{B}}^T$, for other linear transformations $\mathbf{T}$ on a vector space with basis $\mathcal{B}$. The intuition is that the transition probabilities $K(x, y)$ would represent the chance of obtaining y when applying $\mathbf{T}$ to x.

There are often natural choices of maps $\mathbf{T}$ and bases $\mathcal{B}$ for which $K = [\mathbf{T}]_{\mathcal{B}}^T$ has all entries non-negative, but these rarely satisfy $\sum_{y\in\mathcal{B}} K(x, y) = 1$. This is where the Doob transform comes in: it makes a rescaled basis $\check{\mathcal{B}}$ such that the rows of $\check{K} := [\mathbf{T}]_{\check{\mathcal{B}}}^T$ do sum to 1.

Theorem 3.1 (Doob h-transform for linear maps)**.** *[Pan14, Th. 3.1.1; KSK66, Def. 8.11, 8.12; LPW09, Sec.17.6.1] Let V be a finite-dimensional vector space with basis $\mathcal{B}$, and $\mathbf{T}: V \to V$ be a linear map for which $K := [\mathbf{T}]_{\mathcal{B}}^T$ has all entries non-negative. Suppose there is an eigenvector $\eta \in V^*$ of the dual map $\mathbf{T}^*: V^* \to V^*$, with eigenvalue 1, taking only positive values on $\mathcal{B}$. Then*

$$\check{K}(x, y) := K(x, y)\frac{\eta(y)}{\eta(x)}$$

defines a transition matrix. Equivalently, $\check{K} := [\mathbf{T}]_{\check{\mathcal{B}}}^T$, where $\check{\mathcal{B}} := \left\{\frac{x}{\eta(x)} : x \in \mathcal{B}\right\}$.

Proof. First note that $K = [\mathbf{T}^*]_{\mathcal{B}^*}$ by definition, so $\mathbf{T}^*\eta = \eta$ translates to $\sum_y K(x, y)\eta(y) = \eta(x)$. (Functions satisfying this latter condition are called *harmonic*, hence the name h-transform.) Since $\eta(x) > 0$ for all x, it is clear that $\check{K}(x, y) \geq 0$. It remains to show that the rows of $\check{K}$ sum to 1:

$$\sum_y \check{K}(x, y) = \frac{\sum_y K(x, y)\eta(y)}{\eta(x)} = \frac{\eta(x)}{\eta(x)} = 1.$$

□

As detailed in the thesis [Pan14, Sec. 3.2], one motivation for construcing Markov chains in this manner is that the eigenfunctions of the transition matrix are precisely the eigenvectors of $\mathbf{T}$ and of its dual. These eigenfunctions have a variety of applications, see [DPR14, Sec. 2.1]. As a simple example, all limiting distribution of a Markov chain must be *stationary distributions*, that is, a probability distribution π on the state space Ω satisfying $\sum_{x\in\Omega} \pi(x)K(x, y) = \pi(y)$ for each state $y \in \Omega$. It is immediate from the definition of the Doob transform that, if $\sum_{x\in\mathcal{B}} \xi_x x$ is an eigenvector of $\mathbf{T}$ of eigenvalue 1 and $\xi_x > 0$ for all $x \in \mathcal{B}$, then $\pi(x) := \xi_x\eta(x)$ is proportional to a stationary distribution. Indeed, all stationary distributions arise in this manner - see [Pan14, Prop. 3.3.1].

Down-up maps on Combinatorial Hopf Algebras

[Pan14] applies the Doob transform to the coproduct-then-product operator $m\Delta$ on combinatorial Hopf algebras. The main examples in this paper will use the variant $m\Delta_{n-1,1}$, as the resulting "down-up chains" are easier to describe. These chains fall within the recent extension [Pan15], which notates $m\Delta_{n-1,1}$ as the convolution product $\iota * \mathrm{Proj}_1$.

Loosely speaking, a combinatorial Hopf algebra $\mathcal{H} = \bigoplus \mathcal{H}_n$ is a graded vector space with a basis $\mathcal{B} = \amalg_n \mathcal{B}_n$ indexed by a family of combinatorial objects, together with a linear product map

$m : \mathcal{H} \otimes \mathcal{H} \to \mathcal{H}$ and a linear coproduct map $\Delta : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$ satisfying certain compatibility axioms. These encode respectively how the combinatorial objects combine and break. The concept was originally due to Joni and Rota [JR79], and the theory has since been expanded in [Hiv07; ABS06; BL09; AM10] and countless other works. Note that there is no universal definition of a combinatorial Hopf algebra, as each author imposes slightly different axioms. See the next section for an explicit example.

Following [AM10], write $\Delta_{i,j}$ for the composition of the coproduct with projection onto the graded subspaces $\mathcal{H}_i \otimes \mathcal{H}_j$. So $\Delta_{i,j} : \mathcal{H}_{i+j} \to \mathcal{H}_i \otimes \mathcal{H}_j$ models breaking an object of size $i + j$ into two pieces, of size i and j respectively. Note that the composition $m\Delta_{n-1,1}$ preserves the grading on $\mathcal{H}$, so we can study $\mathbf{T}_n := \frac{1}{n} m\Delta_{n-1,1} : \mathcal{H}_n \to \mathcal{H}_n$. Extend this notation to define $\Delta_{r,1,\dots,1} : \mathcal{H}_n \to \mathcal{H}_r \otimes \mathcal{H}_1^{\otimes n-r}$ as the composition $(\Delta_{r,1} \otimes \iota^{\otimes n-r-1}) \dots (\Delta_{n-2,1} \otimes \iota)\Delta_{n-1,1}$, where ι denotes the identity map. The case $r = 1$ will be particularly important.

In all Hopf algebras in this paper, the degree 1 subspace $\mathcal{H}_1$ is one-dimensional, so we can write its basis as $\mathcal{B}_1 = \{\bullet\}$. Then it follows from a direct calculation [DPR14, Th. 3.10] that the product $\bullet^n \in \mathcal{H}_n$ is an eigenvector of $\mathbf{T}_n$ with eigenvalue 1. This is in fact the unique such eigenvector up to scaling, by an argument analogous to [Pan14, Th. 2.6.2]. Applying this fact to the dual Hopf algebra shows that the function $\eta_n : \mathcal{H}_n \to \mathbb{R}$ satisfying $\Delta_{1,1,\dots,1}(x) = \eta_n(x)\bullet^{\otimes n}$ is the unique (up to scaling) eigenvector of $\mathbf{T}_n^*$ with eigenvalue 1. (When the degree is unimportant, we shall write η for η_n.) Hence, if $[\mathbf{T}_n]_{\mathcal{B}_n}^T$ has all entries non-negative and $\eta_n(x) > 0$ for all $x \in \mathcal{B}_n$, then these can be the ingredients of a Doob transform chain. (One circumstance that guarantees this conclusion for $\mathbf{T}_n$ and more general descent operator chains, is when $\mathcal{B}$ is a "state space basis" - see Definition 7.1.)

Combining these two, dual, eigenvectors proves all but the factor of $\frac{1}{n!}$ in part i of the Proposition below. Part ii comes from diagonalising $m\Delta_{n-1,1}$.

Proposition 3.2 (Stationary distribution and spectrum for down-up chains on Hopf algebras)**.** *Let $\mathcal{H} = \bigoplus \mathcal{H}_n$ be a graded Hopf algebra with basis $\mathcal{B} = \amalg \mathcal{B}_n$, such that $\mathcal{B}_1 = \{\bullet\}$. Suppose that the operators $\mathbf{T}_n := \frac{1}{n} m\Delta_{n-1,1} : \mathcal{H}_n \to \mathcal{H}_n$, and rescaling functions $\eta_n : \mathcal{H}_n \to \mathbb{R}$, defined by $\Delta_{1,1,\dots,1}(x) = \eta_n(x)\bullet^{\otimes n}$, admit the Doob transform for each n. Then*

i. [Pan15, Th. 4.5] The unique stationary distributions of these chains are

$$\pi_n(x) = \frac{1}{n!}\eta(x) \times \textit{coefficient of } x \textit{ in } \bullet^{\otimes n}.$$

ii. [Pan15, Th. 4.2] The eigenvalues of these chains are $\frac{j}{n}$ for $0 \le j \le n$, $j \neq n-1$, and their multiplicities are $\dim \mathcal{H}_{n-j} - \dim \mathcal{H}_{n-j-1}$.

□

Remark. The proof of this Proposition is considerably easier if the down-up chain is phrased in the framework of dual graded graphs [Fom94] as opposed to combinatorial Hopf algebras. Indeed, Lemma 6.3, on multiple steps of the bottom-to-random-with-standardisation chain, implicitly uses the commutation relations of a dual graded graph. However, the advantage of working in terms of combinatorial Hopf algebras is that there are many known Hopf morphisms from which we can deduce Markov chain lumpings, whereas analogous maps between dual graded graphs are relatively unexplored.

Example: the Down-Up Chain on Partitions

We express Fulman's chain as the output of a Doob transform applied to $\frac{1}{n}m\Delta_{n-1,1}$ on the algebra of symmetric functions.

Work with the algebra of symmetric functions Λ, with basis the Schur functions $\{s_\lambda\}$, which are indexed by partitions. For clarity, we will often write λ in place of s_λ. As described in [GR14, Sec. 2.5], Λ carries the following Hopf structure:

$$\begin{aligned} m(s_\nu \otimes s_\mu) = s_\nu s_\mu &= \sum_\lambda c^\lambda_{\nu\mu} s_\lambda; \\ \Delta(s_\lambda) &= \sum_{\nu,\mu} c^\lambda_{\nu\mu} s_\nu \otimes s_\mu, \end{aligned}$$

where $c^\lambda_{\nu\mu}$ are the Littlewood-Richardson coefficients.

Write $\lambda \sim \nu \cup \Box$ if the diagram of λ can be obtained by adding one box to the diagram of ν; in this case, we also have $\nu \sim \lambda \backslash \Box$. Recall that, if μ is the partition (1), then $c^\lambda_{\nu\mu} = 1$ if $\lambda \sim \nu \cup \Box$, and 0 otherwise. (This is one case of the *Pieri rule*.) Hence

$$\begin{aligned} m(\nu \otimes (1)) &= \sum_{\lambda:\lambda\sim\nu\cup\Box} \lambda; \\ \Delta_{\deg\nu-1,1}(\lambda) &= \left(\sum_{\nu:\nu\sim\lambda\backslash\Box} \nu\right) \otimes (1). \end{aligned}$$

For example,

$$m\left(\ydiagram{2,1,1} \otimes \ydiagram{1}\right) = \ydiagram{3,1,1} + \ydiagram{2,2,1} + \ydiagram{2,1,1,1};$$

$$\Delta_{4,1}\left(\ydiagram{3,1,1}\right) = \ydiagram{2,1,1} \otimes \ydiagram{1} + \ydiagram{3,1} \otimes \ydiagram{1}.$$

Since m and $\Delta_{n-1,1}$ both expand non-negatively in the $\{s_\lambda\}$ basis, the map $\mathbf{T}_n := \frac{1}{n}m\Delta_{n-1,1} : \Lambda_n \to \Lambda_n$ has a non-negative matrix with respect to the basis $\{s_\lambda : \lambda \vdash n\}$. So we can apply the Doob transform to $\mathbf{T}_n$, provided the rescaling function $\eta(\lambda)$ is positive for each partition λ. Since the standard tableaux of shape λ are in bijection with the ways to remove boxes one by one from λ, it holds that $\eta(\lambda) = \dim\lambda > 0$. Hence the Doob transform creates the following transition matrix:

$$\begin{aligned} \check{K}(\lambda,\mu) &= \sum_{\nu:\nu\sim\lambda\backslash\Box,\mu\sim\nu\cup\Box} \frac{1}{n}\frac{\dim\mu}{\dim\lambda} \\ &= \sum_{\nu:\nu\sim\lambda\backslash\Box,\mu\sim\nu\cup\Box} \frac{1}{n}\frac{\dim\mu}{\dim\nu}\frac{\dim\nu}{\dim\lambda}. \end{aligned}$$

This is indeed Fulman's down-up chain, as described in the introduction: starting at a partition λ, move to a partition ν by removing a box with probability $\frac{\dim\nu}{\dim\lambda}$, then move to μ by adding a box with probability $\frac{1}{n}\frac{\dim\mu}{\dim\nu}$. (An easy way to implement the first step is via the *hook walk* of [GNW79]: uniformly choose a box b, then uniformly choose a box in the *hook* of b - that is, to the right or below b - and continue uniformly picking from successive hooks until you reach a removable box.

Similarly, the second step can be implemented using the *complimentary hook walk* of [GNW84]: start at the box (outside the partition diagram) in row n, column n, uniformly choose a box in its *complimentary hook* - that is, to its left or above it, and outside of the partition - and continue uniformly picking from complimentary hooks until you reach an addable box.) The transition matrix $\check{K}$ for $n = 3$ is on the right below. Compare it to the matrix of the linear map $\mathbf{T}_3 = \frac{1}{3} m\Delta_{2,1}$ on the left.

$$[\mathbf{T}_3]^T_{\mathcal{B}_3} = \begin{array}{c|ccc} & (3) & (2,1) & (1,1,1) \\ \hline (3) & \frac{1}{3} & \frac{1}{3} & \\ (2,1) & \frac{1}{3} & \frac{2}{3} & \frac{1}{3} \\ (1,1,1) & & \frac{1}{3} & \frac{1}{3} \end{array} ; \qquad \check{K} = \begin{array}{c|ccc} & (3) & (2,1) & (1,1,1) \\ \hline (3) & \frac{1}{3} & \frac{2}{3} & \\ (2,1) & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \\ (1,1,1) & & \frac{2}{3} & \frac{1}{3} \end{array} .$$

As a quick aside, here is an alternative description of the down-up chain on partitions in terms of the representation theory of the symmetric groups, as in the first paragraph of the introduction. (This viewpoint is not necessary for constructing the lifted chains on tableaux or permutations.) It comes from the interpretation of multiplication and comultiplication as induction of the external product and restriction to Young subgroups - for irreducible representations corresponding to the partitions $\mu \vdash i$, $\nu \vdash j$ and $\lambda \vdash n$,

$$\mu\nu = \operatorname{Ind}_{\mathfrak{S}_i \times \mathfrak{S}_j}^{\mathfrak{S}_{i+j}} \mu \times \nu; \quad \Delta_{n-i,i}(\lambda) = \operatorname{Res}_{\mathfrak{S}_{n-i} \otimes \mathfrak{S}_i}^{\mathfrak{S}_n} \lambda.$$

So the composition $\mathbf{T}_n := \frac{1}{n} m\Delta_{n-1,1}$ is $\frac{1}{n} \operatorname{Ind}_{\mathfrak{S}_{n-1}}^{\mathfrak{S}_n} \operatorname{Res}_{\mathfrak{S}_{n-1}}^{\mathfrak{S}_n}$, and hence the transition probabilities are

$$\check{K}(\lambda, \mu) = \frac{1}{n} \frac{\dim \mu}{\dim \lambda} \times \text{coefficient of } \mu \text{ in } \operatorname{Ind}_{\mathfrak{S}_{n-1}}^{\mathfrak{S}_n} \operatorname{Res}_{\mathfrak{S}_{n-1}}^{\mathfrak{S}_n}(\lambda).$$

In other words, to take one step of this down-up chain, restrict the present irreducible representation to $\mathfrak{S}_{n-1}$, induce it back to $\mathfrak{S}_n$, and pick an irreducible constituent with probability proportional to the dimension of the isotypic component.

We conclude this section by applying Proposition 3.2 to verify that the unique stationary distribution of this down-up chain is the Plancherel measure. Note that, for $\lambda \vdash n$,

$$\text{coefficient of } \lambda \text{ in } (1)^n = |\{(\nu_2, \dots, \nu_{n-1}) : \nu_2 \sim (1) \cup \square, v_3 \sim \nu_2 \cup \square, \dots, \lambda \sim \nu_{n-1} \cup \square\}| = \dim \lambda.$$

Recall that $\eta(\lambda)$ is also $\dim \lambda$ (essentially by the same argument - this is because the Schur functions are self-dual, and does not apply to general Hopf algebras). Hence $\pi_n(\lambda) = \frac{(\dim \lambda)^2}{n!}$.

4 Strong Lumping and a lift to Standard Tableaux

Sometimes, only certain features of a Markov chain is of interest - that is, we wish to study a process $\{f(X_t)\}$ rather than $\{X_t\}$, for some function f on the state space. The process $\{f(X_t)\}$ is called a lumping (or projection), because it groups together states with the same image under f, treating them as a single state. The analysis of a lumping is easiest when it is itself a Markov chain. If this is true regardless of the initial state of the full chain $\{X_t\}$, then the lumping is strong; if it is dependent on the initial state, the lumping is weak. [KS60, Sec. 6.3, 6.4] is a very thorough exposition on these topics.

This section focuses on strong lumping; the next section will handle weak lumping.

Definition 4.1 (Strong lumping). Let $\{X_t\}, \{\bar{X}_t\}$ be Markov chains on state spaces $\Omega, \bar{\Omega}$ respectively, with transition matrices $K, \bar{K}$. Then $\{\bar{X}_t\}$ is a *strong lumping of* $\{X_t\}$ *via* θ if there is a surjection $\theta : \Omega \to \bar{\Omega}$ such that the process $\{\theta(X_t)\}$ is a Markov chain with transition matrix $\bar{K}$, irrespective of the starting distribution X_0. In this case, $\{X_t\}$ is a *strong lift of* $\{\bar{X}_t\}$ *via* θ.

A necessary and sufficient condition for strong lumping is Dynkin's criterion:

Theorem 4.2 (Strong lumping for Markov chains). *[KS60, Th. 6.3.2] Let $\{X_t\}$ be a Markov chain on a state space Ω with transition matrix K, and let $\theta : \Omega \to \bar{\Omega}$ be a surjection. Then $\{X_t\}$ has a strong lumping via θ if and only if, for every $x_1, x_2 \in \Omega$ with $\theta(x_1) = \theta(x_2)$, and every $\bar{y} \in \bar{\Omega}$, the transition probability sums satisfy*

$$\sum_{y:\theta(y)=\bar{y}} K(x_1, y) = \sum_{y:\theta(y)=\bar{y}} K(x_2, y).$$

The lumped chain has transition matrix

$$\bar{K}(\bar{x}, \bar{y}) := \sum_{y:\theta(y)=\bar{y}} K(x, y)$$

for any x with $\theta(x) = \bar{x}$. □

When the chain $\{X_t\}$ arises from linear operators via the Doob h-transform, Dynkin's criterion translates into the statement below regarding quotient operators.

Theorem 4.3 (Strong lumping for Markov chains from linear maps). *[Pan14, Th. 3.4.1] Let V be a vector space with basis $\mathcal{B}$, and $\mathbf{T} : V \to V, \eta : V \to \mathbb{R}$ be linear map allowing the Doob transform Markov chain construction of Theorem 3.1. Let $\bar{V}$ be a quotient space of V, and denote the quotient map by $\theta : V \to \bar{V}$. Suppose the distinct elements of $\{\theta(x) : x \in \mathcal{B}\}$ are linearly independent, and T, η descend to maps on $\bar{V}$ - that is, there exists $\bar{T} : \bar{V} \to \bar{V}$, $\bar{\eta} : \bar{V} \to \mathbb{R}$, such that $\theta\mathbf{T} = \bar{\mathbf{T}}\theta$ and $\bar{\eta}\theta = \eta$. Then the Markov chain defined by $\bar{\mathbf{T}}$ (on the basis $\bar{\mathcal{B}} := \{\theta(x) : x \in \mathcal{B}\}$, with rescaling function $\bar{\eta}$) is a strong lumping via θ of the Markov chain defined by $\mathbf{T}$.*

Proof. Let $K = [\mathbf{T}]_{\mathcal{B}}^T$, $\bar{K} = [\bar{\mathbf{T}}]_{\bar{\mathcal{B}}}^T$, and write $\check{K}, \check{\bar{K}}$ for the associated transition matrices. (It follows from $\bar{\eta}\theta = \eta$ that $\bar{\eta}$ is non-zero on $\bar{\mathcal{B}}$, so $\check{\bar{K}}$ is well-defined.) By Theorem 4.2 above, we need to show that, for any $x \in \mathcal{B}$ with $\theta(x) = \bar{x}$, and any $\bar{y} \in \bar{\mathcal{B}}$,

$$\check{\bar{K}}(\bar{x}, \bar{y}) = \sum_{y:\theta(y)=\bar{y}} \check{K}(x, y).$$

By definition of the Doob transform, this is equivalent to

$$\bar{K}(\bar{x}, \bar{y})\frac{\bar{\eta}(\bar{y})}{\bar{\eta}(\bar{x})} = \sum_{y:\theta(y)=\bar{y}} K(x, y)\frac{\eta(y)}{\eta(x)}.$$

Because $\bar{\eta}\theta = \eta$, the desired equality reduces to

$$\bar{K}(\bar{x}, \bar{y}) = \sum_{y:\theta(y)=\bar{y}} K(x, y).$$

(Note that this will also prove that all entries of $\bar{K}$ are non-negative, so its Doob transform $\check{\bar{K}}$ is indeed a transition matrix.)

Now expand both sides of $\bar{\mathbf{T}}\theta(x) = \theta\mathbf{T}(x)$ in the $\bar{\mathcal{B}}$ basis:

$$\sum_{\bar{y}\in\bar{\mathcal{B}}} \bar{K}(\bar{x},\bar{y})\bar{y} = \theta\left(\sum_{y\in\mathcal{B}} K(x,y)y\right) = \sum_{\bar{y}\in\bar{\mathcal{B}}}\left(\sum_{y:\theta(y)=\bar{y}} K(x,y)\right)\bar{y}.$$

Equating coefficients of $\bar{y}$ on both sides completes the proof. □

Example: the Down-Up Chain on Standard Tableaux

This section applies the Doob transform to the operator $\mathbf{T}_n := \frac{1}{n} m\Delta_{n-1,1}$ on a Hopf algebra of tableaux, and shows that the resulting Markov chain lumps strongly to the down-up chain on partitions, via taking the shape (i.e. forgetting the numbers in the tableau).

Work in the Poirer-Reutenauer Hopf algebra of standard tableaux, denoted $(\mathbb{Z}T, *, \delta)$ in [PR95] and **FSym** in [DHT02, Sec. 3.6], for "free symmetric functions". Denote its distinguished basis by $\{\mathbf{S}_T\}$, where T runs over the set of standard tableaux. As with partitions, we will often write T in place of $\mathbf{S}_T$. This algebra is graded by the number of boxes in T.

For completeness, below is an extremely brief description of the full product and coproduct of **FSym**, although the present Markov chain construction only requires the much simpler operations $\Delta_{n-1,1}$ and $m : \mathcal{H}_{n-1} \otimes \mathcal{H}_1 \to \mathcal{H}_n$. The product T_1T_2 is the sum over all standard tableaux T containing T_1 such that the difference $T \backslash T_1$ rectifies to $T_2[\deg T_1]$ (that is, T_2 with fillings shifted up by $\deg T_1$). The coproduct is

$$\Delta(T) = \sum_{T_1 \cdot T_2 = T} \operatorname{std}(T_1) \otimes \operatorname{std}(T_2),$$

where $\cdot$ denotes the tableau product of [Ful97, Sec. 2.1, proof of Cor. 2]. For example computations, see [PR95, Sec. 5c, 5d].

Following the convention in Section 3 for general Hopf algebras, let $\bullet$ be the unique tableaux with a single box (necessarily filled with 1). Then, if there are standard tableaux $T_1 \subseteq T$ with $T\backslash T_1$ rectifying to $\bullet[\deg T_1]$, then $T\backslash T_1$ must be a single box filled with the integer $\deg T_1 + 1$. So multiplication by $\bullet$ is the sum of all ways to add a new box, filled with the smallest positive integer not yet appearing in the tableau. For example,

$$m\left(\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1} 4 \\ \cline{1-1}\end{array} \otimes \bullet\right) = \begin{array}{|c|c|c|}\hline 1 & 2 & 5 \\ \hline 3 \\ \cline{1-1} 4 \\ \cline{1-1}\end{array} + \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 5 \\ \hline 4 \\ \cline{1-1}\end{array} + \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1} 4 \\ \cline{1-1} 5 \\ \cline{1-1}\end{array}.$$

To calculate $\Delta_{n-1,1}$, first note that, if T_2 is a single box, then $T_1 \cdot T_2$ is row insertion [Ful97, Sec. 1.1; Sta99, Ex. 7.11.1] of the filling of T_2 into T_1. The inverse of this procedure is unbumping [Sta99, fourth paragraph of proof of Th. 7.11.5]: for a removable box b in row i, remove b, then find the box in row $i-1$ containing the largest integer smaller than b. Call this filling b_1. Replace b_1 with b, then put b_1 in the box in row $i-2$ previously filled with the largest integer smaller than b_1, and continue this process up the rows. In the second term in the example below, these displaced fillings are $4, 3, 2$. In conclusion, each term in $\Delta_{n-1,1}(T)$ is obtained from carrying out this unbumping process, then standardising the remaining tableaux, for example

$$\begin{aligned}\Delta_{4,1}\left(\begin{array}{|c|c|c|}\hline 1&2&5\\\hline \multicolumn{1}{|c|}{3}\\\cline{1-1}\multicolumn{1}{|c|}{4}\\\cline{1-1}\end{array}\right) &= \mathrm{std}\left(\begin{array}{|c|c|}\hline 1&2\\\hline \multicolumn{1}{|c|}{3}\\\cline{1-1}\multicolumn{1}{|c|}{4}\\\cline{1-1}\end{array}\right)\otimes\bullet+\mathrm{std}\left(\begin{array}{|c|c|c|}\hline 1&3&5\\\hline \multicolumn{1}{|c|}{4}\\\cline{1-1}\end{array}\right)\otimes\bullet\\ &= \begin{array}{|c|c|}\hline 1&2\\\hline \multicolumn{1}{|c|}{3}\\\cline{1-1}\multicolumn{1}{|c|}{4}\\\cline{1-1}\end{array}\otimes\bullet+\begin{array}{|c|c|c|}\hline 1&2&4\\\hline \multicolumn{1}{|c|}{3}\\\cline{1-1}\end{array}\otimes\bullet.\end{aligned}$$

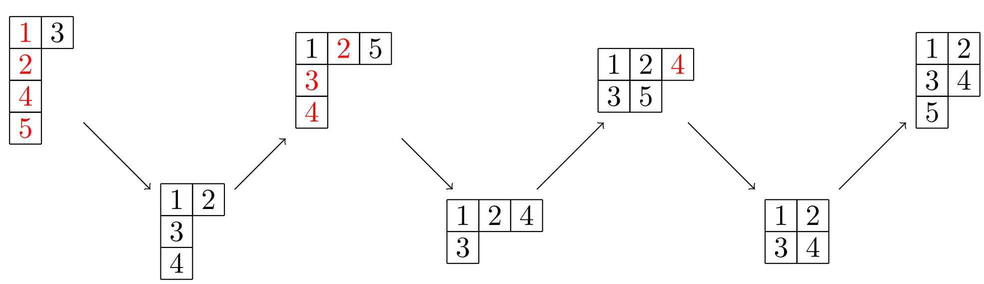
To describe the down-up chain on standard tableaux (i.e. the chain which the Doob transform fashions from the map $\mathbf{T}_n := \frac{1}{n}m\Delta_{n-1,1}$), it remains to calculate the rescaling function $\eta(T)$. This is the coefficient of $\bullet^{\otimes n}$ in $\Delta_{1,\dots,1}(T)$, which the description of $\Delta_{n-1,1}$ above rephrases as the number of ways to successively choose boxes to unbump from T. Such ways are in bijection with the standard tableaux of the same shape as T, so $\eta(T) = \dim(\operatorname{shape} T)$. Hence one step of the down-up chain on standard tableaux, starting from a tableau T of n boxes, has the following interpretation:

1. Pick a removable box b of T with probability $\frac{\dim(\operatorname{shape}(T\backslash b))}{\dim(\operatorname{shape} T)}$, and unbump b. (As for partitions, one can pick b using the hook walk of [GNW79].)

2. Standardise the remaining tableaux and call this T'.

3. Add a box labelled n to T', with probability $\frac{1}{n}\frac{\dim(\operatorname{shape}(T'\cup n))}{\dim(\operatorname{shape} T')}$. (As for partitions, one can pick where to add this box using the complimentary hook walk of [GNW84].)

Here are a few steps of a possible trajectory in degree 5 (the red marks the unbumping paths):



The transition matrix of this chain in degree 3 is

	$\begin{array}{|c|c|c|}\hline 1&2&3\\\hline\end{array}$	$\begin{array}{|c|c|}\hline 1&2\\\hline \multicolumn{1}{|c|}{3}\\\cline{1-1}\end{array}$	$\begin{array}{|c|c|}\hline 1&3\\\hline \multicolumn{1}{|c|}{2}\\\cline{1-1}\end{array}$	$\begin{array}{|c|}\hline 1\\\hline 2\\\hline 3\\\hline\end{array}$
$\begin{array}{|c|c|c|}\hline 1&2&3\\\hline\end{array}$	$\frac{1}{3}$	$\frac{2}{3}$		
$\begin{array}{|c|c|}\hline 1&2\\\hline \multicolumn{1}{|c|}{3}\\\cline{1-1}\end{array}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$
$\begin{array}{|c|c|}\hline 1&3\\\hline \multicolumn{1}{|c|}{2}\\\cline{1-1}\end{array}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$
$\begin{array}{|c|}\hline 1\\\hline 2\\\hline 3\\\hline\end{array}$			$\frac{2}{3}$	$\frac{1}{3}$

.

Recall from Proposition 3.2 that the unique stationary distribution of this chain is $\pi_n(T) = \frac{1}{n!}\eta(T)\times$coefficient of T in $\bullet^n$. Note that there is a unique way of adding outer boxes filled with $1, 2, \dots$ in succession to build a given tableau T, so each tableau of n boxes appears precisely once in the product $\bullet^n$. Hence $\pi_n(T) = \frac{1}{n!}\dim(\operatorname{shape} T)$.

Finally, we show that this chain lumps (strongly) to the down-up chain on partitions. The linear extension of the shape map, $\theta : \mathbf{FSym} \to \Lambda$, defined by $\theta(\mathbf{S}_T) = s_{\operatorname{shape}(T)}$, is a Hopf morphism [PR95, Th. 4.3.i], so $\theta(\frac{1}{n}m\Delta_{n-1,1}) = (\frac{1}{n}m\Delta_{n-1,1})\theta$. We already saw that $\eta(T) = \dim(\operatorname{shape} T) = \eta(\operatorname{shape} T)$. Hence all hypotheses of Theorem 4.3 are satisfied.

5 Weak Lumping and a lift to Permutations

To lift the down-up chain on standard tableaux to permutations, one would like to exhibit **FSym** as the quotient of a Hopf algebra on permutations. Instead, **FSym** is natually a subalgebra of **FQSym**, the Malvenuto-Reutenauer Hopf algebra on permutations. Luckily, subspaces correspond to a weaker notion of lumping, where the initial distribution matters.

Definition 5.1 (Weak lumping)**.** Let $\{X_t\}, \{X'_t\}$ be Markov chains on state spaces Ω, Ω' respectively, with transition matrices K, K'. Then $\{X'_t\}$ is a *weak lumping of* $\{X_t\}$ *via* θ, *with initial distribution* X_0, if there is a surjection $\theta : \Omega \to \Omega'$ such that the process $\{\theta(X_t)\}$, started at the specified X_0, is a Markov chain with transition matrix K'. In this case, $\{X_t\}$ is a *weak lift of* $\{X'_t\}$ *via* θ.

[KS60, Th. 6.4.1] gives a complicated necessary and sufficient condition for weak lumping. (Note that they write π for the initial distribution and α for the stationary distribution.) Their simple sufficient condition [KS60, Th. 6.4.4] has the drawback of not identifying any valid initial distribution beyond the stationary distribution - such a result would not be useful for the many descent operator chains which are absorbing. So instead we appeal to the following sufficient condition underlying [KS60, Ex. 6.4.2]:

Theorem 5.2 (Sufficient condition for weak lumping for Markov chains)**.** *Let K be the transition matrix of a Markov chain $\{X_t\}$ with state space Ω. Suppose $\Omega = \amalg\Omega^i$, and there are distributions π^i on Ω, taking non-zero values only on Ω^i, such that $\pi^i K = \sum_j K'(i,j)\pi^j$ for some constants $K'(i,j)$. Then, from any initial distribution of the form $P\{X_0 = x\} = \sum_i \alpha_i \pi^i(x)$, for constants α_i, the chain $\{X_t\}$ lumps weakly to the chain on the state space $\{\Omega^i\}$ with transition matrix K'.* □

For Markov chains arising from the Doob transform, the condition $\pi^i K = \sum_j K'(i,j)\pi^j$ translates to the existence of invariant subspaces. It may seem strange to consider the subspace spanned by $\left\{\sum_{x\in\mathcal{B}^i} x\right\}$, but this is actually very natural for combinatorial Hopf algebras.

Theorem 5.3 (Weak lumping for Markov chains from linear maps)**.** *Let V be a vector space with basis $\mathcal{B}$, and $\mathbf{T} : V \to V, \eta : V \to \mathbb{R}$ be linear maps admitting the Doob transform Markov chain construction of Theorem 3.1. Suppose $\mathcal{B} = \amalg_i \mathcal{B}^i$, and write x^i for $\sum_{x\in\mathcal{B}^i} x$. Let V' be the subspace of V spanned by the x^i, and suppose $\mathbf{T}(V') \subseteq V'$. Define a map $\theta : \mathcal{B} \to \{x^i\}$ by setting $\theta(x) := x^i$ if $x \in \mathcal{B}^i$. Then the Markov chain defined by $\mathbf{T} : V \to V$ lumps weakly to the Markov chain defined by $\mathbf{T} : V' \to V'$ (with basis $\mathcal{B}' := \{x^i\}$, and rescaling function the restriction $\eta : V' \to \mathbb{R}$) via θ, from any initial distribution of the form $P\{X_0 = x\} := \alpha_{\theta(x)}\frac{\eta(x)}{\eta(\theta(x))}$, where the αs are constants depending only on $\theta(x)$.*

Remarks.

1. When we apply this to lift the down-up chain on tableaux to permutations, it will be the case that $\eta(x) \equiv 1$ for all $x \in \mathcal{B}$. In this special case, the initial distribution condition $P\{X_0 = x\} := \alpha_{\theta(x)} \frac{\eta(x)}{\eta(\theta(x))}$ precisely mandates that X_0 is constant on each $\mathcal{B}^i$. In other words, any X_0 with $P\{X_0 = x\} = P\{X_0 = y\}$ whenever $\theta(x) = \theta(y)$ is a valid initial distribution.

2. Suppose the conditions of Theorem 5.3 hold, and let $j : V' \hookrightarrow V$ be the inclusion map. Now the dual map $j^* : V^* \twoheadrightarrow V'^*$, and $\mathbf{T}^* : V^* \to V^*$, satisfy the hypotheses of Theorem 4.3, except that there may not be suitable rescaling functions $\eta : V^* \to \mathbb{R}$ and $\bar{\eta} : V'^* \to \mathbb{R}$. Because the Doob transform chain for $\mathbf{T}^*$ is the time-reversal of the chain for $\mathbf{T}$ [Pan14, Th. 3.3.2], this is a reflection of [KS60, Th. 6.4.5].

Proof. Write T', η' for the restrictions of T, η to V'. Set $K = [\mathbf{T}]_{\mathcal{B}}^T$, $K' = [\mathbf{T}']_{\mathcal{B}'}^T$, and write $\check{K}, \check{K}'$ for the associated transition matrices. (Note that $\eta'(x^i) = \sum_{x \in \mathcal{B}^i} \eta(x)$ is positive, so $\check{K}'$ is well-defined.) In the notation of Theorem 5.2, the distribution π^i is

$$\pi^i(x) = \begin{cases} \frac{\eta(x)}{\eta(x^i)} & \text{if } x \in \mathcal{B}^i; \\ 0 & \text{otherwise.} \end{cases}$$

We need to show that, for all $y \in \mathcal{B}$,

$$\sum_{x \in \mathcal{B}^i} \pi^i(x) \check{K}(x, y) = \sum_j \check{K}'(x^i, x^j) \pi^j(y).$$

Note that $\pi^j(y)$ is zero unless $y \in \mathcal{B}^j$, so only one summand contributes to the right hand side. By substituting for $\pi^i, \check{K}$ and $\check{K}'$, the desired equality is equivalent to

$$\sum_{x \in \mathcal{B}^i} \frac{\eta(x)}{\eta(x^i)} K(x, y) \frac{\eta(y)}{\eta(x)} = K'(x^i, x^j) \frac{\eta(x^j)}{\eta(x^i)} \frac{\eta(y)}{\eta(x^j)},$$

which reduces to $\sum_{x \in \mathcal{B}^i} K(x, y) = K'(x^i, x^j)$ for $y \in \mathcal{B}^j$. (Note that this will also prove that all entries of K' are non-negative, so its Doob transform $\check{K}'$ is indeed a transition matrix.)

Now, by expanding in the $\mathcal{B}'$ basis,

$$\mathbf{T}'(x^i) = \sum_j K'(x^i, x^j) x^j = \sum_j K'(x^i, x^j) \sum_{y \in \mathcal{B}^j} y.$$

On the other hand, a $\mathcal{B}$ expansion yields

$$\mathbf{T}'(x^i) = \sum_{x \in \mathcal{B}^i} \mathbf{T}(x) = \sum_{x \in \mathcal{B}^i} \sum_{y \in \mathcal{B}} K(x, y) y.$$

So

$$\sum_{y \in \mathcal{B}} \sum_{x \in \mathcal{B}^i} K(x, y) y = \sum_j K'(x^i, x^j) \sum_{y \in \mathcal{B}^j} y = \sum_y \sum_{j : y \in \mathcal{B}^j} K'(x^i, x^j) y,$$

and since $\mathcal{B}$ is a basis, the coefficients on the two sides must be equal. $\square$

Example: the Down-Up Chain on Permutations

This section constructs the bottom-to-random-with-standardisation chain of the introduction from the linear map $\mathbf{T}_n := \frac{1}{n} m\Delta_{n-1,1}$ on a Hopf algebra of permutations, and proves that it lumps weakly to the down-up chain on standard tableaux, and hence also to the down-up chain on partitions.

Work in the Malvenuto-Reutenauer Hopf algebra of permutations, denoted $(\mathbb{Z}S, *, \Delta)$ in [MR95, Sec. 3], $(\mathbb{Z}S, *, \delta)$ in [PR95], and $\mathfrak{S}Sym$ in [AS05]. We follow the recent Parisian literature, such as [DHT02], and call this algebra **FQSym**, for "free quasisymmetric functions".

The basis of concern here is the fundamental basis $\{\mathbf{F}_\sigma\}$, as σ ranges over all permutations (of any length). As in the previous sections, we often write $\mathbf{F}_\sigma$ simply as σ. The degree of σ is its length when considered as a word. The unique permutation of degree 1 is $\bullet = (1)$.

We explain the Hopf structure on **FQSym** by example. The product $\sigma_1\sigma_2$ is $\sigma_1 \sqcup\!\sqcup\, \sigma_2[\deg \sigma_1]$, the sum of all "interleavings" or "shuffles" of σ_1 with the shift of σ_2 by $\deg(\sigma_1)$:

$$\begin{aligned}(3,1,2)(2,1) &= (3,1,2) \sqcup\!\sqcup\, (5,4) \\ &= (3,1,2,5,4) + (3,1,5,2,4) + (3,1,5,4,2) + (3,5,1,2,4) + (3,5,1,4,2) \\ &\quad + (3,5,4,1,2) + (5,3,1,2,4) + (5,3,1,4,2) + (5,3,4,1,2) + (5,4,3,1,2).\end{aligned}$$

The coproduct is

$$\Delta(\sigma) = \sum_{\sigma_1 \cdot \sigma_2 = \sigma} \operatorname{std}(\sigma_1) \otimes \operatorname{std}(\sigma_2),$$

where $\cdot$ denotes concatenation:

$$\begin{aligned}&\Delta(4,1,3,2) \\ &= () \otimes (4,1,3,2) + \operatorname{std}(4) \otimes \operatorname{std}(1,3,2) + \operatorname{std}(4,1) \otimes \operatorname{std}(3,2) + \operatorname{std}(4,1,3) \otimes \operatorname{std}(2) + (4,1,3,2) \otimes () \\ &= () \otimes (4,1,3,2) + (1) \otimes (1,3,2) + (2,1) \otimes (2,1) + (3,1,2) \otimes (1) + (4,1,3,2) \otimes ().\end{aligned}$$

Recall that we are primarily interested in $\mathbf{T}_n := \frac{1}{n} m\Delta_{n-1,1}$. Note that $\Delta_{n-1,1}$ removes the last letter of the word and standardises the result, whilst right-multiplication by $\bullet$ yields the sum of all ways to insert the letter n. Since $\Delta_{n-1,1}(\sigma)$ contains only one term, we see inductively that the rescaling function is $\eta(\sigma) \equiv 1$. Hence one step of the down-up chain on **FQSym**, starting at $\sigma \in \mathfrak{S}_n$, has the following description:

1. Remove the last letter of σ.
2. Standardise the remaining word.
3. Insert the letter n into this standardised word, in a uniformly chosen position.

Here are a few steps of a possible trajectory in degree 5:

$$\begin{tikzcd}
(5,2,4,3,1) \arrow[dr] & & (4,1,3,5,2) \arrow[dr] & & (3,1,5,2,4) \arrow[dr] & & (5,3,1,4,2) \\
& (4,1,3,2) \arrow[ur] & & (3,1,2,4) \arrow[ur] & & (3,1,4,2) \arrow[ur] &
\end{tikzcd}$$

The transition matrix of this chain in degree 3 is

	$(1,2,3)$	$(1,3,2)$	$(3,1,2)$	$(2,3,1)$	$(2,1,3)$	$(3,2,1)$
$(1,2,3)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$			
$(1,3,2)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$			
$(3,1,2)$				$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
$(2,3,1)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$			
$(2,1,3)$				$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
$(3,2,1)$				$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

Recall from Proposition 3.2 that the unique stationary distribution of this chain is $\pi_n(T) = \frac{1}{n!}\eta(T)\times$coefficient of T in $\bullet^n$. Since there is a unique way of inserting the letters $1, 2, \dots$ in that order to obtain a given permutation, each permutation of length n appears precisely once in the product $\bullet^n$. Hence $\pi_n(T) \equiv \frac{1}{n!}$.

Finally, we show that this chain lumps weakly to the down-up chain on partitions, starting from any initial distribution where permutations having the same RSK insertion tableau are equally probable. Let RSK denote the map sending a permutation to its insertion tableau under the Robinson-Schensted-Knuth algorithm (this tableau is often called P). [PR95, Th. 4.3.iii] shows that **FSym** is a subalgebra of **FQSym** under the injection $j(\mathbf{S}_T) := \sum_{\mathrm{RSK}(\sigma)=T} \mathbf{F}_\sigma$, so Theorem 5.3 applies. (Note that the rescaling function $\eta(T) = \dim(\operatorname{shape} T)$ is indeed the restriction of $\eta(\sigma) \equiv 1$, since the number of terms $\mathbf{F}_\sigma$ in the image of $\mathbf{S}_T$ is $\dim(\operatorname{shape} T)$.)

6 Multistep Transition Probabilities from the Identity

This section establishes that the bottom-to-random Markov chains, whether with or without standardisation, have the same probabilities of moving from the identity permutation to σ in t steps, for any particular value of t and σ. As warned in the introduction, this does not imply that the t-step trajectories from the identity permutation to σ are the same for both chains.

The proof idea is as follows: t iterations of the bottom-to-random-without-standardisation chain can be expressed as a linear combination of single moves in some "bottom-r-to-random-without-standardisation" chains, for $1 \leq r \leq t$. The same expression holds for t iterations of the bottom-to-random-with-standardisation chain, if we use some "bottom-r-to-random-with-standardisation" chains. For a single move starting at the identity, the bottom-r-to-random chains (for each r), with or without standardisation, gives the same result.

Definition 6.1. The *bottom-r-to-random shuffle* (without standardisation) is the following procedure: remove the bottomost r cards of the deck, then reinsert them one-by-one in a random order, each time choosing a position uniformly.

The *bottom-r-to-random-with-standardisation* chain is the following procedure: remove the bottomost r cards of the deck, standardise the remaining deck of $n-r$ cards, then insert the cards $n-r+1, n-r+2, \dots, n$ one-by-one, each time choosing a position uniformly. This chain is produced from the Doob transform for $\frac{(n-r)!}{n!} m\Delta_{n-r,1,\dots,1}$ on **FQSym**. (Recall that $\Delta_{n-r,1,\dots,1}$ is the composition $(\Delta_{n-r,1} \otimes \iota^{\otimes r-1}) \dots (\Delta_{n-2,1} \otimes \iota)\Delta_{n-1,1}$.)

Analogous to the interpretation of bottom-to-random-with-standardisation in the introduction, the bottom-r-to-random-with-standardisation chain can also be interpreted as recording the relative times that items spend on a to-do list - now one completes the bottom r tasks of the list daily, and receives r new tasks.

The exact expression for t iterations of either bottom-to-random chain is not pleasant [DFP92, Cor. 2.1]. So, examine the following recursive relation instead.

Lemma 6.2. *A bottom-r-to-random shuffle (without standardisation) followed by a bottom-to-random shuffle is the same as doing a bottom-r-to-random shuffle with probability $\frac{r}{n}$, or a bottom-$r+1$-to-random shuffle with the complementary probability $\frac{n-r}{n}$.*

Proof. After a bottom-r-to-random shuffle, there are two possibilities. With probability $\frac{r}{n}$ the bottom card is one of the r cards that were removed and reinserted, in which case composing with a bottom-to-random shuffle results in a single bottom-r-to-random shuffle. Else the bottom card was not removed during the bottom-r-to-random shuffle, but will be moved in the subsequent bottom-to-random shuffle, so $r+1$ cards were moved in total, resulting in a bottom-$r+1$-to-random shuffle. □

Next, we show the same statement is true for the chains with standardisation.

Lemma 6.3. *One step of the bottom-r-to-random-with-standardisation chain followed by one step of the bottom-to-random-with-standardisation chain is the same as taking one step of the bottom-r-to-random-with-standardisation with probability $\frac{r}{n}$, or of the bottom-$r+1$-to-random-with-standardisation with the complementary probability $\frac{n-r}{n}$.*

Proof. By definition of the bottom-r-to-random-with-standardisation chain, the lemma is equivalent to the following equality of linear maps on **FQSym**:

$$(m\Delta_{n-1,1})(m\Delta_{n-r,1,\dots,1}) = rm\Delta_{n-r,1,\dots,1} + m\Delta_{n-r-1,\dots,1}.$$

(This equality is in fact true for all graded connected Hopf algebras with $\dim \mathcal{H}_1 = 1$, by a dual graded graph commutation relation argument - see [BLL12, Th. 3.2] and [Ful09, Lem. 4.4]. So the calculation below will not use any properties specific to **FQSym**.)

Fix any $x \in \mathbf{FQSym}_n$. Since $\mathcal{B}_1 = \{\bullet\}$, the coproducts $\Delta_{n-r,1,\dots,1}(x)$ and $\Delta_{n-r-1,1,\dots,1}(x)$ have the form $y \otimes \bullet^{\otimes r}$ and $z \otimes \bullet^{\otimes r+1}$, and $\Delta_{n-r-1,1}(y) = z \otimes \bullet$ by coassociativity. Now

$$\begin{aligned}
(m\Delta_{n-1,1})(m\Delta_{n-r,1,\dots,1})(x) &= (m\Delta_{n-1,1})(y\bullet^r) \\
&= m\left[(\Delta_{n-r-1,1}(y))(\bullet^r \otimes 1) + (y \otimes 1)(\Delta_{n-r-1,1}(\bullet^r))\right] \\
&= m\left[(z \otimes \bullet)(\bullet^r \otimes 1) + (y \otimes 1)(r \bullet \otimes \bullet^{r-1})\right] \\
&= m\left[z \bullet^r \otimes \bullet + ry \bullet \otimes \bullet^{r-1}\right] \\
&= z \bullet^{r+1} + ry\bullet^r \\
&= m\Delta_{n-r-1,\dots,1}(x) + rm\Delta_{n-r,1,\dots,1}(x).
\end{aligned}$$

The second equality uses the compatibility axiom for the product and coproduct of a Hopf algebra. □

Iterating the relation in the above lemmas gives an expression for the probability, of moving from the identity permutation to σ in t steps of the bottom-to-random chains, as a weighted sum of single step transition probabilities of the bottom-r-to-random chains ($1 \leq r \leq t$). This expression is the same regardless of whether there is standardisation. And it is easy to see that, starting at the identity, a single step of the bottom-r-to-random-with-standardisation chain does not actually require any standardisation, so it agrees with a single step of the bottom-r-to-random shuffle.

7 Generalisations

The above methods of lifting Markov chains using sub-Hopf-algebras and quotient Hopf algebras generalise easily to the framework of *descent operator Markov chains* [Pan15], which model the breaking-then-combining of combinatorial objects. In this setup, there are two ingredients to each Markov chain. The first is a suitable Hopf algebra and basis:

Definition 7.1. [Pan15, Def. 2.2] A basis $\mathcal{B} = \amalg_n \mathcal{B}_n$ of a graded connected Hopf algebra $\mathcal{H} = \bigoplus_n \mathcal{H}_n$ is a *state space basis* if:

1. for all $w, z \in \mathcal{B}$, the expansion of $m(w \otimes z)$ in the $\mathcal{B}$ basis has all coefficients non-negative;
2. for all $x \in \mathcal{B}$, the expansion of $\Delta(x)$ in the $\mathcal{B} \otimes \mathcal{B}$ basis has all coefficients non-negative;
3. for $n > 1$, the basis $\mathcal{B}_n$ contains no *primitive* elements. That is, $\Delta(x) \neq 1 \otimes x + x \otimes 1$ for all $x \in \mathcal{B}_n$ with $n > 1$.

The second piece of input is a probability distribution $P(D)$ on *weak-compositions* D of a fixed n (that is, $D := (d_1, \dots, d_a)$ with $\sum d_i = n$ and $d_i \geq 0$).

Now take the linear transformation $\mathbf{T}_P : \mathcal{H}_n \to \mathcal{H}_n$ given by

$$\mathbf{T}_P := \sum_D \frac{P(D)}{\binom{n}{d_1 \dots d_a}} m \Delta_D,$$

where $\Delta_D : \mathcal{H} \to \mathcal{H}_{d_1} \otimes \dots \otimes \mathcal{H}_{d_a}$ is a projection of an iterated coproduct. (These maps are called descent operators in [Pat94] because, on a commutative or cocommutative Hopf algebra, their composition is equivalent to the multiplication in the descent algebra of the symmetric group.) An analogous argument to [Pan14, Th. 2.6.2] shows that

$$\eta(x) := \text{sum of coefficients (in the } \mathcal{B}_1 \otimes \dots \otimes \mathcal{B}_1 \text{ basis) of } \Delta_{1,\dots,1}(x)$$

is a suitable rescaling function, and the positivity hypothesis on $\mathbf{T}_P$ and η are guaranteed by the definition of a state space basis. Applying the Doob transform results in a Markov chain on the basis $\mathcal{B}_n$ with this flavour: split an object into a pieces, with $P(D)$ being the probability that the first piece has size d_1, the second has size d_2, ..., then reassemble these pieces into another object. For the down-up chains above, $P((n-1, 1)) = 1$ and $P(D) = 0$ for all other weak-compositions D.

[Pan15, Th. 4.1] is a specialisation of Theorem 4.3 above to a condition for strong lumping of descent operator chains. If $\mathbf{T}_{P,n}$ is a descent operator and θ is a Hopf morphism, then $\theta \mathbf{T}_P = \mathbf{T}_P \theta$, and $(\theta \otimes \dots \otimes \theta) \Delta_{1,\dots,1} = \Delta_{1,\dots,1} \theta$, which implies $\eta \theta = \eta$. Consequently:

Theorem 7.2 (Strong lumping for descent operator chains)**.** *[Pan15, Th. 4.1] Let $\mathcal{H}$, $\bar{\mathcal{H}}$ be graded, connected Hopf algebras with state space bases $\mathcal{B}$, $\bar{\mathcal{B}}$ respectively, and let $\mathbf{T}_P$ be the descent operator defined above. If $\theta : \mathcal{H} \to \bar{\mathcal{H}}$ is a Hopf-morphism such that $\theta(\mathcal{B}_n) = \bar{\mathcal{B}}_n$ for all n, then the Markov chain on $\mathcal{B}_n$ which the Doob transform fashions from $\mathbf{T}_P$ lumps strongly via θ to the Doob transform chain from the same map on $\bar{\mathcal{B}}_n$.*

Similarly, the last paragraph of the present Section 5 can be easily generalised to other descent operators on subalgebras of other combinatorial Hopf algebras, since any sub-Hopf-algebra must be invariant under the descent operators.

Theorem 7.3 (Weak lumping for descent operator chains)**.** *Let $\mathcal{H}$, $\mathcal{H}'$ be graded, connected Hopf algebras with state space bases $\mathcal{B}$, $\mathcal{B}'$ respectively, and let $\mathbf{T}_P$ be the descent operator defined above. Suppose $\theta : \mathcal{B}_n \to \mathcal{B}'_n$ is such that the "preimage sum" map $\theta^* : \mathcal{B}'_n \to \mathcal{H}$, defined by $\theta^*(x') := \sum_{x\in\mathcal{B},\theta(x)=x'} x$, extends to a Hopf-morphism. Then the Markov chain on $\mathcal{B}_n$ which the Doob transform fashions from $\mathbf{T}_P$ lumps weakly via θ to the Doob transform chain from the same map on $\mathcal{B}'_n$, from any starting distribution X_0 where $\frac{X_0(x)}{\eta(x)} = \frac{X_0(y)}{\eta(y)}$ whenever $\theta(x) = \theta(y)$.*

In this more general context of descent operators, the fact that Λ is a subquotient of **FQSym** proves the following:

Corollary 7.4. *Fix an integer n, and let $P(D)$ be a probability distribution on the weak-compositions of n. Consider the* P-restriction-then-induction *chain on irreducible representations of the symmetric group $\mathfrak{S}_n$, where each step goes as follows:*

1. *Choose a weak-composition D of n with probability $P(D)$.*
2. *Restrict the current irreducible representation to the chosen Young subgroup $\mathfrak{S}_{d_1} \times \cdots \times \mathfrak{S}_{d_a}$.*
3. *Induce this representation to $\mathfrak{S}_n$, then pick an irreducible constituent with probabiliy proportional to the dimension of its isotypic component.*

A weak lift of this chain is the P-shuffle-with-standardisation *chain on the permutations $\mathfrak{S}_n$ (viewed in one-line notation):*

1. *Choose a weak-composition D of n with probability $P(D)$.*
2. *Deconcatenate the current permutation into a word w_1 of the first d_1 letters, w_2 of the next d_2 letters, and so on.*
3. *Replace the smallest letter in w_1 by 1, the next smallest by 2, and so on. Then replace the smallest letter in w_2 by $d_1 + 1$, the next smallest by $d_1 + 2$, and so on for all w_i.*
4. *Interleave these words uniformly (i.e. uniformly choose a permutation where the letters $1, 2, \dots, d_1$ are in the same relative order as in the replaced w_1, where $d_1+1, d_1+2, \dots, d_1+d_2$ are in the same relative order as in the replaced w_2, etc.).*

The initial distributions for which this will be a lift are those where permutations having the same RSK insertion tableau are equally probable.

There is of course an intermediate chain on standard tableaux, coming from applying the Doob transform to $\mathbf{T}_P$ on **FSym**, but this is hard to describe for arbitrary P.

Note that [Ful04, Th. 3.1] proves that the *P-shuffle-without-standardisation* (omitting step 3 above) is also a lift of the P-restriction-then-induction chain, if started from the identity permutation. These shuffle schemes are studied in great detail in [DFP92], and [Pan15] exhibits them as the result of the Doob transform to $\mathbf{T}_P$ on the shuffle algebra. So an interesting open question is to find an appropriate sequence of Hopf-morphisms between the symmetric functions and the shuffle algebra, so that applications of Theorems 7.2 and 7.3 give a Hopf-algebraic proof of Fulman's lift result, and possibly "explain" the similarity of the two sets of lifts. (In the same, unexplained, vein, the P-shuffles with and without standardisation have the same strong lumping via the position of descents - see three paragraphs below and [Pan13] respectively.)

As an aside, let us point out some similarities and differences between the P-shuffles with and without standardisation. Sage computer calculations show that, for an arbitrary distribution

P and arbitrary permutation σ, the P-shuffles with and without standardisation have different probabilities of moving from the identity to σ in t steps. Since both are descent operator chains, by [Pan15, Th. 4.2] they have the same eigenvalues, but different multiplicities. (Strictly speaking, the case without standardisation requires a multigraded version of this theorem.) For example, the eigenvalues for both bottom-to-random chains are $\frac{j}{n}$ for $j = 0, 1, 2, \ldots, n-3, n-2, n$. Their multiplicities for the without-standardisation chain are the number of permutations with j fixed points, whilst, for the chains with standardisation, they are the number of permutations fixing $1, 2, \ldots, j$ but not $j+1$. Hence in general the smaller eigenvalues have higher multiplicities in the chain with standardisation. The P-shuffles-without-standardisation are diagonalisable (because the shuffle algebra is commutative), and so are the top-to-random-with-standardisation and bottom-to-random-with-standardisation (because of a connection with dual graded graphs), but Sage computer calculations show that the P-shuffles-with-standardisation are generally non-diagonalisable.

To see another example and non-example of lumpings induced from Hopf morphisms, consider the following commutative diagram from [PR95, Th. 4.3]:

$$\begin{tikzcd}
\mathbf{FQSym}\text{, permutations} \arrow[r, two heads, "\mathrm{Des}"] & QSym\text{, compositions} \\
\mathbf{FSym}\text{, standard tableaux} \arrow[r, two heads, "\text{shape}"] \arrow[u, hook] & \Lambda\text{, partitions} \arrow[u, hook]
\end{tikzcd}$$

The main example of this paper lifts a chain on partitions to a chain on permutations via a chain on standard tableaux, on the bottom left. Let us see why it is not possible to construct a lift via compositions, on the top right, instead. The corresponding Hopf algebra here is the algebra of quasisymmetric functions [Ges84].

There is no problem with the top Hopf morphism, which sends a permutation $\mathbf{F}_\sigma$ to the fundamental quasisymmetric function $F_{\mathrm{Des}(\sigma)}$ associated with its descent composition. Since this map sends a basis of **FQSym** to a basis of $QSym$, Theorem 7.2 applies and the descent operator chains on permutations lump strongly according to descent composition.

The problem is with the Hopf morphism on the right - this inclusion is not induced from a set map from compositions to partitions. Theorem 7.3 only applies if $s_\lambda = \sum_{\theta(I)=\lambda} F_I$ for some function θ sending compositions to partitions. This condition does not hold, as the same F_I can occur in the expansion of multiple Schur functions: $s_{(3,1)} = F_{(1,3)} + F_{(2,2)} + F_{(3,1)}$, $s_{(2,2)} = F_{(1,2,1)} + F_{(2,2)}$.

Beyond Descent Operators

This idea of lifting Markov chains using Hopf morphisms works for linear maps a little more general than the descent operators: for example, throw in a permutation of the tensorands after comultiplication and before multiplication. More precisely, for fixed n, let P now be a probability distribution on the pairs (D, σ), where D is a weak-composition of n of a parts, and $\sigma \in \mathfrak{S}_a$. Now define

$$\mathbf{T}_P := \sum_D \frac{P(D)}{\binom{n}{d_1 \ldots d_a}} m \sigma \Delta_D,$$

where σ permutes the tensorands of $\mathcal{H}^{\otimes a}$. It is straightforward to show that the rescaling function η from above continues to work for these more general maps.

As an example, take a parameter $q \in [0,1]$, and consider $\mathbf{T} := \frac{q}{n} m\sigma\Delta_{1,n-1} + \frac{1-q}{n} m\Delta_{n-1,1}$, where $\sigma \in \mathfrak{S}_2$ is the transposition. The chain that this map generates on **FQSym** is a weighted sum of the twisted-top-to-random-with-standardisation and bottom-to-random-with-standardisation:

1. With probability q, remove the first letter of the permutation; with probability $1-q$, remove the last letter of the permutation. Let i denote the letter that is removed.

2. For each letter $j > i$, relabel this letter as $j-1$. Hence the word now consists of the letters $\{1, 2, \ldots, n-1\}$.

3. Insert the letter n at a uniformly chosen position.

(This is a close cousin of the "top-or-bottom-to-random" map from [Pan15, Ex. 3.4].) Because the algebra of symmetric functions is commutative and cocommutative, this new map $\mathbf{T}$ agrees with the down-up map $\frac{1}{n} m\Delta_{n-1,1}$ on the symmetric functions. Hence this chain is also a lift of the down-up chain on the partitions.

References

[ABS06] Marcelo Aguiar, Nantel Bergeron, and Frank Sottile. "Combinatorial Hopf algebras and generalized Dehn–Sommerville relations". In: *Compos. Math.* 142.1 (2006), pp. 1–30. ISSN: 0010-437X. eprint: `10.1112/S0010437X0500165X`.

[AD87] David Aldous and Persi Diaconis. "Strong uniform times and finite random walks". In: *Adv. in Appl. Math.* 8.1 (1987), pp. 69–97. ISSN: 0196-8858. DOI: `10.1016/0196-8858(87)90006-6`. URL: `http://dx.doi.org/10.1016/0196-8858(87)90006-6`.

[AM10] Marcelo Aguiar and Swapneel Mahajan. *Monoidal functors, species and Hopf algebras.* Vol. 29. CRM Monograph Series. With forewords by Kenneth Brown and Stephen Chase and André Joyal. Providence, RI: American Mathematical Society, 2010, pp. lii+784. ISBN: 978-0-8218-4776-3.

[AS05] Marcelo Aguiar and Frank Sottile. "Structure of the Malvenuto-Reutenauer Hopf algebra of permutations". In: *Adv. Math.* 191.2 (2005), pp. 225–275. ISSN: 0001-8708. DOI: `10.1016/j.aim.2004.03.007`. URL: `http://dx.doi.org/10.1016/j.aim.2004.03.007`.

[Ayy+15] Arvind Ayyer, Anne Schilling, Benjamin Steinberg, and Nicolas M. Thiéry. "Markov chains, R-trivial monoids and representation theory". In: *Internat. J. Algebra Comput.* 25.1-2 (2015), pp. 169–231. ISSN: 0218-1967. DOI: `10.1142/S0218196715400081`. URL: `http://dx.doi.org/10.1142/S0218196715400081`.

[BL09] Nantel Bergeron and Huilan Li. "Algebraic structures on Grothendieck groups of a tower of algebras". In: *J. Algebra* 321.8 (2009), pp. 2068–2084. ISSN: 0021-8693. DOI: `10.1016/j.jalgebra.2008.12.005`. URL: `http://dx.doi.org/10.1016/j.jalgebra.2008.12.005`.

[BLL12] Nantel Bergeron, Thomas Lam, and Huilan Li. "Combinatorial Hopf algebras and towers of algebras—dimension, quantization and functorality". In: *Algebr. Represent. Theory* 15.4 (2012), pp. 675–696. ISSN: 1386-923X. DOI: `10.1007/s10468-010-9258-y`. URL: `http://dx.doi.org/10.1007/s10468-010-9258-y`.

[Bro00] Kenneth S. Brown. "Semigroups, rings, and Markov chains". In: *J. Theoret. Probab.* 13.3 (2000), pp. 871–938. ISSN: 0894-9840. DOI: 10.1023/A:1007822931408. URL: http://dx.doi.org/10.1023/A:1007822931408.

[CW07] Sylvie Corteel and Lauren K. Williams. "A Markov chain on permutations which projects to the PASEP". In: *Int. Math. Res. Not. IMRN* 17 (2007), Art. ID rnm055, 27. ISSN: 1073-7928. DOI: 10.1093/imrn/rnm055. URL: http://dx.doi.org/10.1093/imrn/rnm055.

[DFP92] Persi Diaconis, James Allen Fill, and Jim Pitman. "Analysis of top to random shuffles". In: *Combin. Probab. Comput.* 1.2 (1992), pp. 135–155. ISSN: 0963-5483. DOI: 10.1017/S0963548300000158. URL: http://dx.doi.org/10.1017/S0963548300000158.

[DHT02] Gérard Duchamp, Florent Hivert, and Jean-Yves Thibon. "Noncommutative symmetric functions. VI. Free quasi-symmetric functions and related algebras". In: *Internat. J. Algebra Comput.* 12.5 (2002), pp. 671–717. ISSN: 0218-1967. DOI: 10.1142/S0218196702001139. URL: http://dx.doi.org/10.1142/S0218196702001139.

[DPR14] Persi Diaconis, C. Y. Amy Pang, and Arun Ram. "Hopf algebras and Markov chains: two examples and a theory". In: *J. Algebraic Combin.* 39.3 (2014), pp. 527–585. ISSN: 0925-9899. DOI: 10.1007/s10801-013-0456-7. URL: http://dx.doi.org/10.1007/s10801-013-0456-7.

[DS05] Enrica Duchi and Gilles Schaeffer. "A combinatorial approach to jumping particles". In: *J. Combin. Theory Ser. A* 110.1 (2005), pp. 1–29. ISSN: 0097-3165. DOI: 10.1016/j.jcta.2004.09.006. URL: http://dx.doi.org/10.1016/j.jcta.2004.09.006.

[Fom94] Sergey Fomin. "Duality of graded graphs". In: *J. Algebraic Combin.* 3.4 (1994), pp. 357–404. ISSN: 0925-9899. DOI: 10.1023/A:1022412010826. URL: http://dx.doi.org/10.1023/A:1022412010826.

[Ful04] Jason Fulman. "Card shuffling and the decomposition of tensor products". In: *Pacific J. Math.* 217.2 (2004), pp. 247–262. ISSN: 0030-8730. DOI: 10.2140/pjm.2004.217.247. URL: http://dx.doi.org/10.2140/pjm.2004.217.247.

[Ful05] Jason Fulman. "Stein's method and Plancherel measure of the symmetric group". In: *Trans. Amer. Math. Soc.* 357.2 (2005), pp. 555–570. ISSN: 0002-9947. DOI: 10.1090/S0002-9947-04-03499-3. URL: http://dx.doi.org/10.1090/S0002-9947-04-03499-3.

[Ful08] Jason Fulman. "Convergence rates of random walk on irreducible representations of finite groups". In: *J. Theoret. Probab.* 21.1 (2008), pp. 193–211. ISSN: 0894-9840. DOI: 10.1007/s10959-007-0102-1. URL: http://dx.doi.org/10.1007/s10959-007-0102-1.

[Ful09] Jason Fulman. "Commutation relations and Markov chains". In: *Probab. Theory Related Fields* 144.1-2 (2009), pp. 99–136. ISSN: 0178-8051. DOI: 10.1007/s00440-008-0143-0. URL: http://dx.doi.org/10.1007/s00440-008-0143-0.

[Ful97] William Fulton. *Young tableaux*. Vol. 35. London Mathematical Society Student Texts. With applications to representation theory and geometry. Cambridge University Press, Cambridge, 1997, pp. x+260. ISBN: 0-521-56144-2; 0-521-56724-6.

[Ges84] Ira M. Gessel. "Multipartite P-partitions and inner products of skew Schur functions". In: *Combinatorics and algebra (Boulder, Colo., 1983)*. Vol. 34. Contemp. Math. Providence, RI: Amer. Math. Soc., 1984, pp. 289–317. DOI: 10.1090/conm/034/777705. URL: http://dx.doi.org/10.1090/conm/034/777705.

[GNW79] Curtis Greene, Albert Nijenhuis, and Herbert S. Wilf. "A probabilistic proof of a formula for the number of Young tableaux of a given shape". In: *Adv. in Math.* 31.1 (1979), pp. 104–109. ISSN: 0001-8708. DOI: 10.1016/0001-8708(79)90023-9. URL: http://dx.doi.org/10.1016/0001-8708(79)90023-9.

[GNW84] Curtis Greene, Albert Nijenhuis, and Herbert S. Wilf. "Another probabilistic method in the theory of Young tableaux". In: *J. Combin. Theory Ser. A* 37.2 (1984), pp. 127–135. ISSN: 0097-3165. DOI: 10.1016/0097-3165(84)90065-7. URL: http://dx.doi.org.proxy.bibliotheques.uqam.ca:2048/10.1016/0097-3165(84)90065-7.

[GR14] Darij Grinberg and Vic Reiner. "Hopf algebras in combinatorics". In: *ArXiv e-prints* (Sept. 2014). arXiv: 1409.8356.

[Hiv+11] Florent Hivert, Jean-Gabriel Luque, Jean-Christophe Novelli, and Jean-Yves Thibon. "The $(1-\mathbb{E})$-transform in combinatorial Hopf algebras". In: *J. Algebraic Combin.* 33.2 (2011), pp. 277–312. ISSN: 0925-9899. DOI: 10.1007/s10801-010-0245-5. URL: http://dx.doi.org/10.1007/s10801-010-0245-5.

[Hiv07] Florent Hivert. "An introduction to combinatorial Hopf algebras—examples and realizations". In: *Physics and theoretical computer science*. Vol. 7. NATO Secur. Sci. Ser. D Inf. Commun. Secur. Amsterdam: IOS, 2007, pp. 253–274.

[JR79] S. A. Joni and G.-C. Rota. "Coalgebras and bialgebras in combinatorics". In: *Stud. Appl. Math.* 61.2 (1979), pp. 93–139. ISSN: 0022-2526.

[KS60] John G. Kemeny and J. Laurie Snell. *Finite Markov chains*. The University Series in Undergraduate Mathematics. D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto-London-New York, 1960, pp. viii+210.

[KSK66] John G. Kemeny, J. Laurie Snell, and Anthony W. Knapp. *Denumerable Markov chains*. D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto, Ont.-London, 1966, pp. xi+439.

[LPW09] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. With a chapter by James G. Propp and David B. Wilson. Providence, RI: American Mathematical Society, 2009, pp. xviii+371. ISBN: 978-0-8218-4739-8.

[MR95] Clauda Malvenuto and Christophe Reutenauer. "Duality between quasi-symmetric functions and the Solomon descent algebra". In: *J. Algebra* 177.3 (1995), pp. 967–982. ISSN: 0021-8693. DOI: 10.1006/jabr.1995.1336. URL: http://dx.doi.org/10.1006/jabr.1995.1336.

[Pan13] C. Y. Amy Pang. "A Hopf-power Markov chain on compositions". In: *25th International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2013)*. Discrete Math. Theor. Comput. Sci. Proc., AS. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2013, pp. 499–510.

[Pan14] C. Y. Amy Pang. "Hopf algebras and Markov chains". In: *ArXiv e-prints* (Dec. 2014). A revised thesis. arXiv: 1412.8221.

[Pan15] C. Y. Amy Pang. "Card-Shuffling via Convolutions of Projections on Combinatorial Hopf Algebras". In: *27th International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2015)*. Discrete Math. Theor. Comput. Sci. Proc., AU. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2015, pp. 49–60.

[Pat94] F. Patras. "L'algèbre des descentes d'une bigèbre graduée". In: *J. Algebra* 170.2 (1994), pp. 547–566. ISSN: 0021-8693. DOI: 10.1006/jabr.1994.1352. URL: http://dx.doi.org/10.1006/jabr.1994.1352.

[Pha91] R. M. Phatarfod. "On the matrix occurring in a linear search problem". In: *J. Appl. Probab.* 28.2 (1991), pp. 336–346. ISSN: 0021-9002.

[PR95] Stéphane Poirier and Christophe Reutenauer. "Algèbres de Hopf de tableaux". In: *Ann. Sci. Math. Québec* 19.1 (1995), pp. 79–90. ISSN: 0707-9109.

[SC04] Laurent Saloff-Coste. "Random walks on finite groups". In: *Probability on discrete structures*. Vol. 110. Encyclopaedia Math. Sci. Springer, Berlin, 2004, pp. 263–346. DOI: 10.1007/978-3-662-09444-0_5. URL: http://dx.doi.org/10.1007/978-3-662-09444-0_5.

[Sta99] Richard P. Stanley. *Enumerative Combinatorics. Vol. 2*. Vol. 62. Cambridge Studies in Advanced Mathematics. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin. Cambridge: Cambridge University Press, 1999, pp. xii+581. ISBN: 0-521-56069-1; 0-521-78987-7.

[Ste+15] W. A. Stein et al. *Sage Mathematics Software (Version 6.6)*. http://www.sagemath.org. The Sage Development Team. 2015.