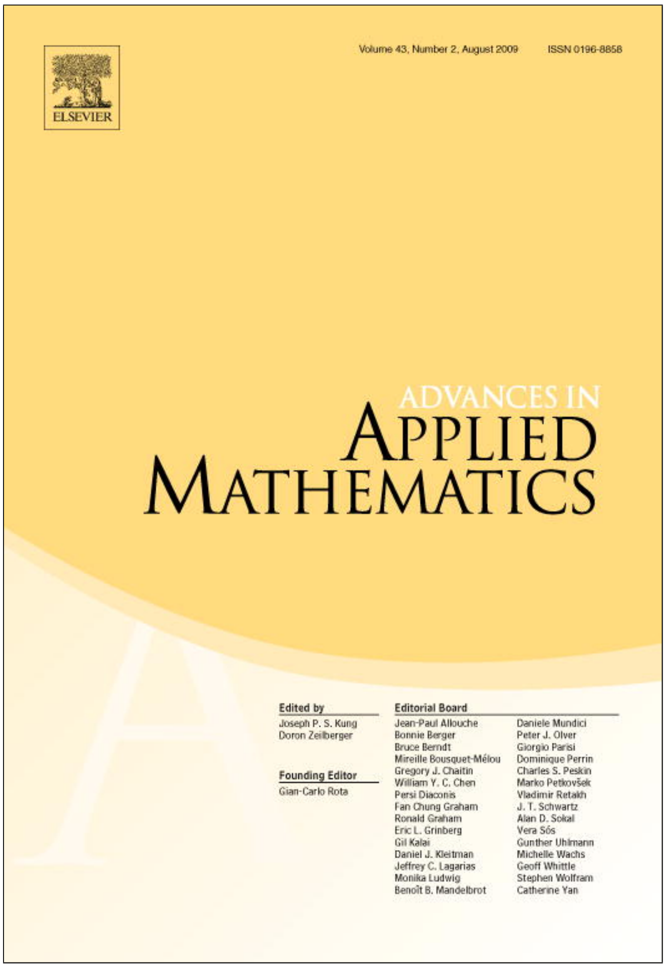
ELSEVIER
Volume 43, Number 2, August 2009
ISSN 0196-8858
ADVANCES IN
APPLIED
MATHEMATICS
Edited by
Joseph P. S. Kung
Doron Zeilberger
Founding Editor
Gian-Carlo Rota
Editorial Board
Jean-Paul Allouche
Bonnie Berger
Bruce Berndt
Mireille Bousquet-Mélou
Gregory J. Chaitin
William Y. C. Chen
Persi Diaconis
Fan Chung Graham
Ronald Graham
Eric L. Grinberg
Gil Kalai
Daniel J. Kleitman
Jeffrey C. Lagarias
Monika Ludwig
Benoît B. Mandelbrot
Daniele Mundici
Peter J. Olver
Giorgio Parisi
Dominique Perrin
Charles S. Peskin
Marko Petkovšek
Vladimir Retakh
J. T. Schwartz
Alan D. Sokal
Vera Sós
Gunther Uhlmann
Michelle Wachs
Geoff Whittle
Stephen Wolfram
Catherine Yan

Advances in Applied Mathematics 43 (2009) 176–196

Contents lists available at ScienceDirect

Advances in Applied Mathematics

www.elsevier.com/locate/yaama

# Carries, shuffling, and symmetric functions

Persi Diaconis [a], Jason Fulman [b,*]

[a] *Department of Mathematics and Statistics, Stanford University, Stanford, CA 94305, USA*
[b] *Department of Mathematics, University of Southern California, Los Angeles, CA 90089, USA*

ARTICLE INFO

*Article history:*
Received 1 February 2009
Accepted 11 February 2009
Available online 15 April 2009

*MSC:*
60C05
60J10
05E05

*Keywords:*
Carries
Shuffling
Symmetric function
Autoregressive process
Veronese map

ABSTRACT

The “carries” when $n$ random numbers are added base $b$ form a Markov chain with an “amazing” transition matrix determined in a 1997 paper of Holte. This same Markov chain occurs in following the number of descents when $n$ cards are repeatedly riffle shuffled. We give generating and symmetric function proofs and determine the rate of convergence of this Markov chain to stationarity. Similar results are given for type $B$ shuffles. We also develop connections with Gaussian autoregressive processes and the Veronese mapping of commutative algebra.



## 1. Introduction

We use generating functions and symmetric function theory to explain a surprising coincidence: when $n$ long integers are added base-$b$, the distribution of “carries” is the same as the distribution of descents when $n$ cards are repeatedly riffled shuffled. The explanation yields a sharp analysis of convergence to stationarity of the associated Markov chains. A similar analysis goes through for “type $B$” shuffles. In this introduction, we first explain the carries process, then riffle shuffling and finally the connection.

* Corresponding author. Address for correspondence: USC Department of Mathematics, KAP 108-MC 2532, 3620 South Vermont Ave, Los Angeles, CA 90089-2532, USA. Fax: +1 213 740 2424.
*E-mail address:* fulman@usc.edu (J. Fulman).

0196-8858/$ – see front matter 
doi:10.1016/j.aam.2009.02.002

### 1.1. Carries

Consider adding three 50-digit numbers base 10 (in the top row, italics are used to indicate the carries):

$$
\begin{array}{cllllllllll}
\mathit{1} & \mathit{12021} & \mathit{01111} & \mathit{11111} & \mathit{11111} & \mathit{11011} & \mathit{10111} & \mathit{01111} & \mathit{11111} & \mathit{21011} & \mathit{1112} \\
 & 43935 & 23749 & 58561 & 74916 & 62215 & 47448 & 33196 & 51990 & 19807 & 27075 \\
 & 48537 & 53642 & 77448 & 32760 & 14421 & 72142 & 82116 & 37225 & 43300 & 51498 \\
 & 33618 & 41327 & 41561 & 16257 & 43616 & 55134 & 82714 & 63369 & 87142 & 45607 \\
\hline
1 & 26091 & 18719 & 77571 & 23934 & 20253 & 74725 & 98027 & 52585 & 50250 & 24180
\end{array}
$$

For this example, $6/50 = 12\%$ of the columns have a carry of zero, $40/50 = 80\%$ have a carry of one and $4/50 = 8\%$ have a carry of two.

If $n$ integers (base $b$) are produced by choosing their digits uniformly at random in $\{0, 1, 2, \ldots, b-1\}$, the sequence of carries $\kappa_0 = 0, \kappa_1, \kappa_2, \ldots$ forms a Markov chain taking values in $\{0, 1, 2, \ldots, n-1\}$. Holte [23] studied this Markov chain and found fascinating structure in its "amazing" transition matrix $(P(i, j))$. Here $P(i, j)$ is the probability that the next carry is $j$ given that the last carry was $i$, and he showed, for $0 \leqslant i, j \leqslant n-1$, that

$$P(i, j) = \frac{1}{b^n} \sum_{l=0}^{j-\lfloor i/b \rfloor} (-1)^l \binom{n+1}{l} \binom{n-1-i+(j+1-l)b}{n}. \tag{1.1}$$

For example, when $n = 3$ the matrix becomes

$$\frac{1}{6b^2} \begin{pmatrix} b^2+3b+2 & 4b^2-4 & b^2-3b+2 \\ b^2-1 & 4b^2+2 & b^2-1 \\ b^2-3b+2 & 4b^2-4 & b^2+3b+2 \end{pmatrix}.$$

Among many other things, Holte shows that the $j$th entry of the left eigenvector with eigenvalue 1 is $A(n, j)/n!$, with $A(n, j)$ the Eulerian number: the number of permutations in the symmetric group $S_n$ with $j$ descents. Here $\sigma \in S_n$ is said to have a descent at $i$ if $\sigma(i+1) < \sigma(i)$. So $\mathbf{5}\,1\,\mathbf{3}\,2\,4$ has two descents. The fundamental theorem of Markov chain theory gives that $A(n, j)/n!$ is the long term frequency of carries of $j$ when long random numbers are added. Note that this is independent of the base $b$. When $n = 3$, $A(3, 0)/6 = 1/6$, $A(3, 1)/6 = 2/3$, $A(3, 2)/6 = 1/6$ very roughly matching the example above. We give alternative derivations of this at the end of Section 2.

We will not detail the many nice properties Holte found but warmly recommend his paper [23]. Some further properties are in [8,14], which give appearances of this same matrix in card shuffling and in the Veronese construction for graded algebras. This is developed briefly in Section 5 below.

### 1.2. Shuffling

The usual method of shuffling cards proceeds by cutting a deck of $n$ cards into two approximately equal piles and then riffling the two piles together into one pile. A realistic mathematical model was created by Gilbert–Shannon–Reeds: cut off $c$ cards with probability $\binom{n}{c}/2^n$. Drop cards sequentially as follows: if the left pile has $A$ cards and the right pile has $B$ cards, drop the next card from the bottom of the left pile with probability $A/(A+B)$ and from the right pile with probability $B/(A+B)$. This is continued until all cards are dropped.

A careful analysis of riffle shuffles is carried out in [3] using a generalization to $b$-shuffles. There, a deck of cards is cut into $b$ packets of size $c_1, c_2, \ldots, c_b$ with probability $\binom{n}{c_1 \ldots c_b}/b^n$. The packets are riffled together by dropping the next card with probability proportional to packet size. Thus the original Gilbert–Shannon–Reeds model corresponds to a 2-shuffle. Two basic facts established in [3] are:

- The probability of the permutation $\sigma$ arising after a $b$-shuffle is

$$\frac{\binom{n+b-d(\sigma^{-1})-1}{n}}{b^n} \tag{1.2}$$

with $d(\sigma^{-1})$ the number of descents in $\sigma^{-1}$.
- An $a$-shuffle followed by a $b$-shuffle is the same as an $ab$-shuffle.

Thus the result of $r$ 2-shuffles is the same as a single $2^r$ shuffle and so formula (1.2) gives a closed form expression for the probability of any permutation after $r$ 2-shuffles. This and some calculus allow a sharp analysis of the rate of convergence: roughly $\frac{3}{2}\log_2 n + c$ shuffles suffice to make the distribution within $2^{-c}$ of the uniform distribution. Further details are in [3].

The combinatorics of riffle shuffles has expanded. An enumerative theory of cycle and other properties under the $b$-shuffle measure (1.2) is equivalent to the Gessel–Reutenauer enumeration jointly by cycles and descents [17,22]. The combinatorics of riffle shuffling is essentially the same as quasi-symmetric function theory [21,32]. There are extensions to other types (see Section 4 below) and to random walk on the chambers of hyperplane arrangements [7,10] and buildings [9]. Much of this development is surveyed in [13]. Interesting new developments are in [1].

### 1.3. The connection

Carries and riffle shuffling seem like different subjects. However, if $P_b$ denotes the matrix (1.1), Holte [23] showed that

$$P_a P_b = P_{ab}. \tag{1.3}$$

The eigenvalues of the matrix $P_b$ turn out to be the same as the eigenvalues of the $b$-shuffle transition matrix (the multiplicities are different). This, and the appearance of descents in both subjects, led us to suspect and then prove an intimate connection. In Section 2 we prove the following.

**Theorem 1.1.** *The probability that the base-$b$ carries chain goes from* 0 *to $j$ in $r$ steps is equal to the probability that the permutation in $S_n$ obtained by performing $r$ successive $b$-shuffles* (*started at the identity*) *has $j$ descents.*

We give a generating function proof which also yields a similar statement for the inverse permutation along with enumerative results of Gessel in Section 2. We have subsequently found a bijective proof of the theorem which shows that the transition matrices of carries (1.1) and the Markov chain generated by the number of descents after successive $b$-shuffles are the same [14].

The more analytic proof given here allows us to use the Robinson–Schensted–Knuth (RSK) correspondence and symmetric function theory to show that the number of descents (and in fact any function of the descent set) after $r$ 2-shuffles is close to stationarity when $r = \log_2 n + c$. (Note from [3] that $\frac{3}{2}\log_2 n + c$ steps are required for *all* aspects of the permutation to be close to stationarity.) The correspondence with carries shows that the carries chain "settles down" after $\log_2 n + c$ steps. Refining this, we show that for large $n$, $\frac{1}{2}\log_b(n) + c$ steps of the carries chain are necessary and sufficient for convergence to stationarity. Details are in Section 3.

The discussion so far has all been on the permutation group. There are well-established "type $B$" (hyperoctahedral)-shuffles [3,5,20]. In Section 4 we develop a parallel "carries process" and show that theorems about type $B$ shuffles translate into theorems about adding numbers. We also point out a connection with the theory of rounding. Section 5 shows that for large $n$, the carries process is well approximated by a Gaussian autoregressive process, and develops the connection with the Veronese mapping of commutative algebra.

## 2. Two Markov chains

In this section we show that two processes derived from the Markov chain of repeated $b$-shuffles on the symmetric group are Markov chains with transition probabilities from 0 to $j$, the same as the carries chain. As background, note that usually a function of a Markov chain is *not* a Markov chain. A simple example is nearest neighbor random walk on the integers mod $n$, with $n$ odd, $n \geqslant 7$. Let the walk start at 0 and move left or right with probability 1/2. Let $f(j) = 1$ for $0 \leqslant j \leqslant (n-1)/2$, $f(j) = -1$ otherwise. If steps of the original walk are denoted $X_0 = 0, X_1, X_2, \dots$ and $Y_j = f(X_j)$, then $\{X_j\}_{j=0}^{\infty}$ is a Markov chain but $\{Y_j\}_{j=0}^{\infty}$ is not: $\mathbb{P}\{Y_3 = + | Y_2 = +\} = 2/3$, $\mathbb{P}\{Y_3 = + | Y_2 = +, Y_1 = +\} = 3/4$. The literature on conditions for the Markov property is often called "lumping of Markov chains." A useful introduction is [25] with [28] a sophisticated extension.

To begin, we show that the two basic facts about riffle shuffles give a generating function identity of Gessel (unpublished).

**Proposition 2.1.** *Let $\sigma$ be a permutation with $d$ descents. Let $c_{ij}^d$ be the number of ordered pairs $(\tau, \mu)$ of permutations in $S_n$ such that $\tau$ has $i$ descents, $\mu$ has $j$ descents, and $\tau\mu = \sigma$. Then*

$$\sum_{i,j \geqslant 0} \frac{c_{ij}^d s^{i+1} t^{j+1}}{(1-s)^{n+1}(1-t)^{n+1}} = \sum_{a,b \geqslant 0} \binom{n+ab-d-1}{n} s^a t^b.$$

**Proof.** Since an $a$-shuffle followed by a $b$-shuffle is an $ab$-shuffle, the formula (1.2) implies that

$$\sum_{\mu \in S_n} \binom{n+a-d(\mu)-1}{n} \mu^{-1} \cdot \sum_{\tau \in S_n} \binom{n+b-d(\tau)-1}{n} \tau^{-1} = \sum_{\sigma \in S_n} \binom{n+ab-d(\sigma)-1}{n} \sigma^{-1}.$$

Multiplying both sides by $s^a t^b$, summing over all $a, b \geqslant 0$, and then taking the coefficient of $\sigma^{-1}$ on both sides yields that

$$\begin{aligned}
\sum_{a,b \geqslant 0} \binom{n+ab-d-1}{n} s^a t^b &= \sum_{\substack{(\tau,\mu) \\ \tau\mu=\sigma}} \left[ \sum_{a \geqslant 0} s^a \binom{n+a-d(\mu)-1}{n} \cdot \sum_{b \geqslant 0} t^b \binom{n+b-d(\tau)-1}{n} \right] \\
&= \sum_{\substack{(\tau,\mu) \\ \tau\mu=\sigma}} \frac{s^{d(\mu)+1}}{(1-s)^{n+1}} \frac{t^{d(\tau)+1}}{(1-t)^{n+1}} \\
&= \sum_{i,j \geqslant 0} \frac{c_{ij}^d s^{i+1} t^{j+1}}{(1-s)^{n+1}(1-t)^{n+1}}. \qquad \square
\end{aligned}$$

Recall that if a Markov chain has transition probabilities $P(i, j)$, its formal time reversal with respect to a stationary measure $\pi$ is defined to have transition probabilities $P^*(i, j) = P(j, i)\pi(j)/\pi(i)$. This $P^*$ is a Markov transition matrix which also has $\pi$ as stationary measure. A Markov chain $P$ is reversible with respect to $\pi$ if and only if $P = P^*$.

Theorem 2.2 identifies the carries Markov chain with the formal time reversal of a chain arising in the theory of riffle shuffles. As in the introduction, $\pi$ denotes the distribution on $\{0, 1, \dots, n-1\}$ defined by $\pi(j) = A(n, j)/n!$, where $A(n, j)$ is the number of permutations in $S_n$ with $j$ descents.

**Theorem 2.2.** *Let a Markov chain on the symmetric group $S_n$ begin at the identity and proceed by successive independent $b$-shuffles. Then the number of descents of $\tau^{-1}$ forms a Markov chain with stationary distribution $\pi(j) = A(n, j)/n!$, and its formal time reversal with respect to $\pi$ is identical with the carries Markov chain.*

**Proof.** Let $d(\tau_r^{-1})$ denote the number of descents of the inverse of the permutation $\tau_r$ obtained after $r$ independent $b$-shuffles. Corollary 2 of [3] showed that $d(\tau_r^{-1})$ forms a Markov chain. Note that the stationary distribution of this chain is given by $\pi(j) = A(n, j)/n!$, since $\tau_r^{-1}$ tends to a uniform element of $S_n$ as $r \to \infty$.

We compute the transition probabilities of the Markov chain formed by $d(\tau_r^{-1})$. By (1.2), $\mathbb{P}(d(\tau_{r-1}^{-1}) = i) = A(n, i)\binom{n+b^{r-1}-i-1}{n}/b^{(r-1)n}$. Clearly

$$\mathbb{P}\bigl(d\bigl(\tau_{r-1}^{-1}\bigr) = i, d\bigl(\tau_r^{-1}\bigr) = j\bigr) = \sum_{\sigma:\ d(\sigma^{-1})=i} \frac{\binom{n+b^{r-1}-i-1}{n}}{b^{(r-1)n}} \sum_{k \geqslant 0} \sum_{\substack{\mu:\ d(\mu^{-1})=k \\ d(\sigma^{-1}\mu^{-1})=j}} \frac{\binom{n+b-k-1}{n}}{b^n}.$$

Thus

$$\begin{aligned}\mathbb{P}\bigl(d\bigl(\tau_r^{-1}\bigr) = j \big| d\bigl(\tau_{r-1}^{-1}\bigr) = i\bigr) &= \frac{\mathbb{P}(d(\tau_{r-1}^{-1}) = i, d(\tau_r^{-1}) = j)}{\mathbb{P}(d(\tau_{r-1}^{-1}) = i)} \\ &= \frac{1}{A(n,i)} \sum_{\sigma:\ d(\sigma^{-1})=i} \sum_{k \geqslant 0} \sum_{\substack{\mu:\ d(\mu^{-1})=k \\ d(\sigma^{-1}\mu^{-1})=j}} \frac{\binom{n+b-k-1}{n}}{b^n}.\end{aligned}$$

In the notation of Proposition 2.1, this is

$$\frac{A(n,j)}{A(n,i)} \frac{1}{b^n} \sum_{k \geqslant 0} c_{ik}^j \binom{n+b-k-1}{n}.$$

Letting $[x^h]f(x)$ denote the coefficient of $x^h$ in a series $f(x)$, this can be rewritten as

$$\bigl[t^b\bigr] \frac{A(n,j)}{A(n,i)} \frac{1}{b^n} \sum_{k \geqslant 0} c_{ik}^j \frac{t^{k+1}}{(1-t)^{n+1}} = \bigl[t^b s^{i+1}\bigr] \frac{A(n,j)}{A(n,i)} \frac{(1-s)^{n+1}}{b^n} \sum_{i,k \geqslant 0} c_{ik}^j \frac{s^{i+1} t^{k+1}}{(1-s)^{n+1}(1-t)^{n+1}}.$$

By Proposition 2.1, this is equal to

$$\begin{aligned}&\bigl[t^b s^{i+1}\bigr] \frac{A(n,j)}{A(n,i)} \frac{(1-s)^{n+1}}{b^n} \sum_{a,d \geqslant 0} \binom{n+ad-j-1}{n} s^a t^d \\ &= \bigl[s^{i+1}\bigr] \frac{A(n,j)}{A(n,i)} \frac{(1-s)^{n+1}}{b^n} \sum_{a \geqslant 0} \binom{n+ab-j-1}{n} s^a \\ &= \frac{A(n,j)}{A(n,i)} \frac{1}{b^n} \sum_{l \geqslant 0} (-1)^l \binom{n+1}{l} \binom{n-1-j+(i+1-l)b}{n}.\end{aligned}$$

This is equal to $\pi(j)P(j,i)/\pi(i)$ where $P$ is the transition probability of the carries chain (1.1). $\square$

The next result gives a second, more direct, interpretation of the transition probabilities of the carries chain.

**Theorem 2.3.** *The probability that the base-$b$ carries chain goes from* 0 *to $j$ in $r$ steps is equal to the probability that a permutation in $S_n$ obtained by performing $r$ successive $b$-shuffles (started at the identity) has $j$ descents.*

**Proof.** By (1.2) and the fact that an $a$-shuffle followed by a $b$-shuffle is an $ab$-shuffle, the probability that $r$ successive $b$-shuffles (started at the identity) lead to a permutation with $j$ descents is

$$\sum_{i\geqslant 0}\frac{1}{b^{rn}}\binom{n+b^r-i-1}{n}c^0_{ij}, \tag{2.1}$$

where as in Proposition 2.1, $c^0_{ij}$ denotes the number of $\sigma\in S_n$ such that $d(\sigma^{-1})=i$ and $d(\sigma)=j$.

Proposition 2.1 gives that

$$\sum_{i,k\geqslant 0}\frac{c^0_{ik}s^{i+1}t^{k+1}}{(1-s)^{n+1}(1-t)^{n+1}}=\sum_{a,d\geqslant 0}\binom{n+ad-1}{n}s^a t^d.$$

Taking the coefficient of $s^{b^r}$ on both sides gives that

$$\sum_{i,k\geqslant 0}\frac{c^0_{ik}\binom{n+b^r-i-1}{n}t^{k+1}}{(1-t)^{n+1}}=\sum_{d\geqslant 0}\binom{n+b^r d-1}{n}t^d.$$

Comparing with Eq. (2.1) gives that the probability that a permutation obtained after $r$ successive $b$-shuffles has $j$ descents is

$$\begin{aligned}&\frac{1}{b^{rn}}\left[t^{j+1}\right](1-t)^{n+1}\sum_{d\geqslant 0}\binom{n+b^r d-1}{n}t^d\\&\quad=\frac{1}{b^{rn}}\sum_{l\geqslant 0}(-1)^l\binom{n+1}{l}\binom{n-1+(j+1-l)b^r}{n}.\end{aligned}$$

From (1.1), this is equal to the carries transition probability $P_{b^r}(0,j)$. By Eq. (1.3), this is $P^r_b(0,j)$, as claimed. □

An immediate corollary of Theorem 2.3 is that the carries chain also has $\pi(j)=A(n,j)/n!$ as its stationary distribution. We conclude this section with two alternative derivations of the stationary distribution of the carries chain. The following lemma will be helpful. Stanley [30] and Pitman [27] give bijective proofs.

**Lemma 2.4.** *Let $X_1,\dots,X_n$ be independent uniform $[0,1]$ random variables. Then for all integers $j$, $\mathbb{P}(j\leqslant\sum_{i=1}^n X_i<j+1)$ is equal to the probability that a uniformly chosen random permutation on $n$ symbols has $j$ descents.*

As usual we let $P^r(0,j)$ denote the distribution on $\{0,1,\dots,n-1\}$ after $r$ steps of the carries chain (for the base $b$ addition of $n$ numbers) started from 0.

**Theorem 2.5.** *(See [23].) The stationary distribution $\pi$ of the carries chain satisfies $\pi(j)=A(n,j)/n!$, where $A(n,j)$ is the number of permutations in $S_n$ with $j$ descents.*

**Proof.** By Holte [23], $r$ steps of the base $b$ carries chain is equivalent to one step of the base $b^r$ carries chain. Letting $Y_1,\dots,Y_n$ be independent discrete uniform random variables on $\{0,1,\dots,b^r-1\}$, it follows that

$$P^r(0, j) = \mathbb{P}\left(jb^r \leqslant \sum_{i=1}^{n} Y_i < (j+1)b^r\right).$$

Letting $U_1, \ldots, U_n$ be independent continuous uniforms on $[0, b^r]$, this implies that

$$\begin{aligned} P^r(0, j) &= \mathbb{P}\left(jb^r \leqslant \sum_{i=1}^{n} \lfloor U_i \rfloor < (j+1)b^r\right) \\ &= \mathbb{P}\left(jb^r \leqslant \sum_{i=1}^{n} U_i - \sum_{i=1}^{n} \big(U_i - \lfloor U_i \rfloor\big) < (j+1)b^r\right) \\ &= \mathbb{P}\left(j \leqslant \sum_{i=1}^{n} X_i - E < j+1\right). \end{aligned}$$

Here the $X_i = U_i/b^r$ are independent uniforms on $[0, 1]$ and $E = 1/b^r \cdot \sum_{i=1}^{n} (U_i - \lfloor U_i \rfloor)$.

Although $E$ is not independent of the $X_i$'s, note that when $n$ is fixed and $r \to \infty$, $E$ converges in probability to 0. Indeed, this follows since $|E| \leqslant n/b^r$ with probability 1. Thus Slutsky's theorem implies that

$$\lim_{r\to\infty} P^r(0, j) = \mathbb{P}\left(j \leqslant \sum_{i=1}^{n} X_i < j+1\right),$$

and the result follows from Lemma 2.4. □

A simple analytic way to find the stationary distribution uses the closed form for $P^r(0, j)$. As $r$ tends to infinity,

$$\frac{1}{b^{rn}} \binom{n-1+(j+1-l)b^r}{n} \to \frac{(j+1-l)^n}{n!}.$$

Thus by (1.1) and (1.3),

$$\begin{aligned} P^r(0, j) &= \frac{1}{b^{rn}} \sum_{l=0}^{j} (-1)^l \binom{n+1}{l} \binom{n-1+(j+1-l)b^r}{n} \\ &\to \frac{1}{n!} \sum_{l=0}^{j} (-1)^l \binom{n+1}{l} (j+1-l)^n = \frac{A(n, j)}{n!}. \end{aligned}$$

The last equality is an identity, due to Euler, for the $A(n, j)$ [12].

## 3. Rates of convergence

This section presents both upper and lower bounds on convergence to stationarity for the equivalent Markov chains of Section 2. Theorem 3.2 shows that the descent set of a permutation (not just the number of descents) is close to its stationary distribution after $r$ $b$-shuffles if $r = \log_b(n) + c$. This uses symmetric function theory. Theorem 3.3 uses stochastic monotonicity to bound convergence of the carries chain: it shows that at least $r = \frac{1}{2}\log_b(n) + c$ steps are needed and that $r = \log_b(n) + c$ steps suffice. Theorem 3.4 shows that for large $n$, $\frac{1}{2}\log_b(n) + c$ steps are sufficient.

All of our results involve the total variation distance between probability measures $P$ and $Q$ on a finite set $\mathcal{X}$, defined as

$$\|P - Q\|_{\mathrm{TV}} = \frac{1}{2}\sum_x \big|P(x) - Q(x)\big| = \max_{A \subseteq \mathcal{X}} \big|P(A) - Q(A)\big|.$$

**Theorem 3.1.** *Consider the carries chain for base b addition of n numbers. Let* $r = \lceil \log_b(cn) \rceil$ *with* $c > 0$*. Let* $P_0^r$ *denote the distribution on* $\{0, 1, \ldots, n-1\}$ *given by taking r steps in the carries chain, started from 0. Let* $\pi$ *be the stationary distribution of the carries chain. Then*

$$\big\| P_0^r - \pi \big\|_{\mathrm{TV}} \leqslant \frac{1}{2}\sqrt{e^{1/(2c^2)} - 1}.$$

In fact, we prove a stronger result. This uses the notion of the descent set of a permutation $\sigma$, defined as the set of $i$, $1 \leqslant i \leqslant n-1$, such that $\sigma(i) > \sigma(i+1)$. For instance **5**1**3**24 has descent set $\{1, 3\}$. Let $\widetilde{P^r}(S)$ denote the probability that a permutation obtained after the iteration of $r$ $b$-shuffles (or equivalently a single $b^r$-shuffle) has descent set $S$, and let $\widetilde{\pi}(S)$ denote the probability that a uniformly chosen random permutation has descent set $S$. Theorem 3.2 uses symmetric function theory to upper bound the total variation distance between $\widetilde{P^r}$ and $\widetilde{\pi}$. Chapter 7 of the text [31] provides background on the concepts used in the proof of Theorem 3.2 (i.e. Young tableaux, the RSK correspondence, and symmetric functions).

**Theorem 3.2.** *Let* $r = \lceil \log_b(cn) \rceil$ *with* $c > 0$*. Then*

$$\big\| \widetilde{P^r} - \widetilde{\pi} \big\|_{\mathrm{TV}} \leqslant \frac{1}{2}\sqrt{e^{1/(2c^2)} - 1}.$$

**Proof.** We use the RSK correspondence which associates to a permutation $\sigma$ a pair of standard Young tableaux $(P, Q)$ called the insertion and recording tableau of $\sigma$ respectively. One says that a standard Young tableau $T$ has a descent at $i$ $(1 \leqslant i \leqslant n-1)$ if $i+1$ is in a row lower than $i$ in $T$. We let $d(T)$ denote the number of descents of $T$. By Lemma 7.23.1 of [31], the descent set of $\sigma$ is equal to the descent set of $Q(\sigma)$. This implies that

$$\widetilde{\pi}(S) = \sum_{|\lambda|=n} \frac{f_\lambda(S) f_\lambda}{n!},$$

where $f_\lambda$ is the number of standard Young tableaux of shape $\lambda$, and $f_\lambda(S)$ is the number of standard Young tableaux of shape $\lambda$ with descent set $S$.

From Theorem 3 of [21], the probability that $Q(\sigma) = T$ (for $\sigma$ obtained from a $b^r$ shuffle) is $s_\lambda(1/b^r, \ldots, 1/b^r)$ for any standard Young tableau $T$ of shape $\lambda$. Here $b^r$ coordinates of the Schur function $s_\lambda$ are equal to $1/b^r$ and the rest are 0. Thus,

$$\begin{aligned}
\big\| \widetilde{P^r} - \widetilde{\pi} \big\|_{\mathrm{TV}} &= \frac{1}{2} \sum_{S \subseteq \{1,\ldots,n-1\}} \big| \widetilde{P^r}(S) - \widetilde{\pi}(S) \big| \\
&= \frac{1}{2} \sum_S \left| \sum_{|\lambda|=n} \left[ f_\lambda(S) s_\lambda\left(\frac{1}{b^r}, \ldots, \frac{1}{b^r}\right) - \frac{f_\lambda(S) f_\lambda}{n!} \right] \right| \\
&\leqslant \frac{1}{2} \sum_S \sum_{|\lambda|=n} \left| f_\lambda(S) s_\lambda\left(\frac{1}{b^r}, \ldots, \frac{1}{b^r}\right) - \frac{f_\lambda(S) f_\lambda}{n!} \right|
\end{aligned}$$

$$= \frac{1}{2} \sum_{|\lambda|=n} \left| s_\lambda\left(\frac{1}{b^r}, \ldots, \frac{1}{b^r}\right) - \frac{f_\lambda}{n!} \right| \sum_S f_\lambda(S)$$

$$= \frac{1}{2} \sum_{|\lambda|=n} \left| f_\lambda s_\lambda\left(\frac{1}{b^r}, \ldots, \frac{1}{b^r}\right) - \frac{f_\lambda^2}{n!} \right|.$$

By the Cauchy–Schwarz inequality, this is at most

$$\frac{1}{2}\sqrt{\sum_{|\lambda|=n} \left[ s_\lambda\left(\frac{1}{b^r}, \ldots, \frac{1}{b^r}\right) - \frac{f_\lambda}{n!} \right]^2 \sum_{|\lambda|=n} f_\lambda^2} = \frac{1}{2}\sqrt{n! \sum_{|\lambda|=n} \left[ s_\lambda\left(\frac{1}{b^r}, \ldots, \frac{1}{b^r}\right) - \frac{f_\lambda}{n!} \right]^2}.$$

The functions $f_\lambda s_\lambda(1/b^r, \ldots, 1/b^r)$ and $f_\lambda^2/n!$ both define probability measures on the set of partitions of size $n$; the first is the distribution on RSK shapes after a $b^r$ riffle shuffle [32], and the second is known as Plancherel measure. Hence the previous expression simplifies to

$$\frac{1}{2}\sqrt{n! \sum_{|\lambda|=n} s_\lambda\left(\frac{1}{b^r}, \ldots, \frac{1}{b^r}\right)^2 - 1}.$$

Let $[u^n]f(u)$ denote the coefficient of $u^n$ in a series $f(u)$. By the Cauchy identity for Schur functions [31, p. 322],

$$\sum_{|\lambda|=n} s_\lambda\left(\frac{1}{b^r}, \ldots, \frac{1}{b^r}\right)^2 = [u^n] \sum_{|\lambda|\geqslant 0} s_\lambda\left(\frac{u}{b^r}, \ldots, \frac{u}{b^r}\right) s_\lambda\left(\frac{1}{b^r}, \ldots, \frac{1}{b^r}\right)$$

$$= [u^n]\left(1 - \frac{u}{b^{2r}}\right)^{-b^{2r}} = b^{-2rn}\binom{b^{2r}+n-1}{n}.$$

Thus

$$n! \sum_{|\lambda|=n} s_\lambda\left(\frac{1}{b^r}, \ldots, \frac{1}{b^r}\right)^2 - 1 = \prod_{i=1}^{n-1}\left(1 + \frac{i}{b^{2r}}\right) - 1.$$

Since $\log(1+x) \leqslant x$ for $x > 0$, it follows that

$$\log\left(\prod_{i=1}^{n-1}\left(1 + \frac{i}{b^{2r}}\right)\right) = \sum_{i=1}^{n-1} \log\left(1 + \frac{i}{b^{2r}}\right) \leqslant \binom{n}{2}/b^{2r}.$$

Thus

$$\prod_{i=1}^{n-1}\left(1 + \frac{i}{b^{2r}}\right) - 1 \leqslant \exp\left(\binom{n}{2}/b^{2r}\right) - 1.$$

Summarizing, it has been shown that

$$\|\widetilde{P}^r - \widetilde{\pi}\|_{\mathrm{TV}} \leqslant \frac{1}{2}\sqrt{\exp\left(\binom{n}{2}/b^{2r}\right) - 1}.$$

If $b^r = cn$ with $c > 0$, then $\binom{n}{2}/b^{2r} \leqslant 1/(2c^2)$, which proves the result. $\square$

**Proof of Theorem 3.1.** Theorem 2.3 showed that the base-$b$ carries chain started from 0 is the same as the chain for the number of descents after successive $b$-shuffles started from the identity. Thus Theorem 3.2 also upper bounds the total variation distance between $r$ iterations of the base-$b$ carries chain (started from 0) and its stationary distribution. $\square$

Next we give a different approach to proving convergence using stochastic monotonicity and also give a lower bound. The arguments show that $\log_b n + c$ steps suffice for convergence and that $\frac{1}{2}\log_b n$ steps are not enough.

**Theorem 3.3.** *For $n \geqslant 3$, any starting state $i$, and any $r \geqslant 0$, the Markov chain $P$ of* (1.1) *satisfies*

$$\left\| P^r(i,\cdot) - \pi \right\|_{\mathrm{TV}} \leqslant \left(\frac{n-1}{2} + i\right)/b^r.$$

*Conversely, for any $\epsilon$, $0 < \epsilon < 1$, if $1 \leqslant r \leqslant \log_b[\frac{\epsilon|i-\frac{n-1}{2}|}{\sqrt{n}}]$, then*

$$\left\| P^r(i,\cdot) - \pi \right\|_{\mathrm{TV}} \geqslant 1 - \epsilon.$$

**Proof.** Recall that a Markov chain on $\{0,1,\ldots,n-1\}$ is *stochastically monotone* if for all $i \leqslant i'$, $P(i,\{0,\ldots,j\}) \geqslant P(i',\{0,\ldots,j\})$ for all $j$. We show that $P$ is stochastically monotone by coupling. Consider two copies of the carries chain, one at $i$ and one at $i'$ with $i \leqslant i'$. Each chain proceeds by adding $n$ random base-$b$ digits. Couple them by adding the *same* digits to both. If the first process results in a carry of $k$, the second process results in a carry of $k$ or $k+1$. This implies stochastic monotonicity.

From Holte [23, Theorem 4] and the fact that $n \geqslant 3$, the right eigenfunctions for eigenvalues $1/b$, $1/b^2$ can be taken as

$$f_1(i) = i - \frac{n-1}{2}, \qquad f_2(i) = i^2 - (n-1)i + \frac{(n-2)(3n-1)}{12}.$$

The upper bound follows from stochastic monotonicity and the first eigenvector via [15, Theorem 2.1]. For the lower bound, note that $f_1^2 = f_2 + A$, with $A = (n+1)/12$. This, and a simple computation show that

$$\int \bigl(f_1(x) - f_1(y)\bigr)^2 P(x,dy) = \left(1 - \frac{1}{b}\right)^2 f_1^2(x) + A\left(1 - \frac{1}{b^2}\right).$$

This is the required input for the lower bound, using [15, Theorem 2.3]. One obtains that $\|P^r(i,\cdot) - \pi\|_{\mathrm{TV}} \geqslant 1-\epsilon$ for $r \leqslant \log_b[\frac{\epsilon|i-\frac{n-1}{2}|}{\sqrt{8(n+1)/12}}]$, and the result follows since $8(n+1)/12 \leqslant n$ when $n \geqslant 3$. $\square$

**Remark.** The argument for stochastic monotonicity does not depend on the assumption that the digits are uniform and independently distributed. Any joint distribution within a column (with columns independent) leads to a stochastically monotone Markov chain. In [14] it is shown that the transition matrix $P$ is totally positive of order 2. This implies stochastic monotonicity via [24, Proposition 1.3.1, p. 22].

To close this section, we prove that $\frac{1}{2}\log_b(n) + c$ steps are sufficient for total variation convergence when $n$ is large.

**Theorem 3.4.** *With the notation of Theorem* 3.1*, there is a constant $B > 0$ (independent of $n, b, c > 1$) such that for $r = \frac{1}{2}\log_b(nc)$,*

$$\| P_0^r - \pi \|_{TV} \leqslant \frac{B}{\sqrt{n}} + \frac{B}{\sqrt{c}}.$$

**Proof.** From Theorem 4.3 of [14], there is $j^* \in \{0, 1, \ldots, n-1\}$ such that $P^r(0, j) \geqslant \pi(j)$ for $0 \leqslant j \leqslant j^*$, and that $P^r(0, j) \leqslant \pi(j)$ for $j^* + 1 \leqslant j \leqslant n-1$. Thus

$$\| P_0^r - \pi \|_{TV} = P_0^r\{0, 1, \ldots, j^*\} - \pi\{0, 1, \ldots, j^*\}. \tag{3.1}$$

From the proof of Theorem 2.5, $P^r(0, j) = \mathbb{P}(jb^r \leqslant \sum_{i=1}^n Y_i < (j+1)b^r)$ with $Y_i$ independent discrete uniform random variables on $\{0, 1, \ldots, b^r - 1\}$. From Lemma 2.4, one has that $\pi(j) = \mathbb{P}(j \leqslant \sum_{i=1}^n U_i < j+1)$ with $U_i$ independent uniform random variables on $[0, 1]$. Thus

$$P_0^r(j) = \mathbb{P}\left(\left\lfloor \frac{1}{b^r}\sum_{i=1}^n Y_i \right\rfloor = j\right) \quad \text{and} \quad P_0^r\{0, 1, \ldots, j^*\} = \mathbb{P}\left(\frac{1}{b^r}\sum_{i=1}^n Y_i < j^* + 1\right),$$
$$\pi(j) = \mathbb{P}\left(\left\lfloor \sum_{i=1}^n U_i \right\rfloor = j\right) \quad \text{and} \quad \pi\{0, 1, \ldots, j^*\} = \mathbb{P}\left(\sum_{i=1}^n U_i < j^* + 1\right).$$

From the above considerations, we have

$$\| P_0^r - \pi \|_{TV} \leqslant \sup_x \left| \mathbb{P}\left(\frac{1}{b^r}\sum_{i=1}^n Y_i < x\right) - \mathbb{P}\left(\sum_{i=1}^n U_i < x\right) \right|. \tag{3.2}$$

Let $\mu_n = n/2, \sigma_n^2 = n/12$ and $\nu_n = (n/2) \cdot (1 - 1/b^r), \tau_n^2 = (n/12) \cdot (1 - 1/b^{2r})$. The right-hand side of (3.2) is

$$\begin{aligned}
&\sup_x \left| \mathbb{P}\left[\frac{(\frac{1}{b^r}\sum_{i=1}^n Y_i - \mu_n)}{\sigma_n} < \frac{(x - \mu_n)}{\sigma_n}\right] - \mathbb{P}\left[\frac{(\sum_{i=1}^n U_i - \mu_n)}{\sigma_n} < \frac{(x - \mu_n)}{\sigma_n}\right] \right| \\
&\leqslant \sup_y \left| \mathbb{P}\left[\frac{(\frac{1}{b^r}\sum_{i=1}^n Y_i - \mu_n)}{\sigma_n} < y\right] - \Phi(y) \right| + \sup_y \left| \mathbb{P}\left[\frac{(\sum_{i=1}^n U_i - \mu_n)}{\sigma_n} < y\right] - \Phi(y) \right| \\
&= I + II.
\end{aligned}$$

Here $\Phi(y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^y e^{-t^2/2}\,dt$ denotes the cumulative distribution function of the normal distribution.

From the usual Berry–Esseen bound, $II \leqslant B_1/\sqrt{n}$ with $B_1$ involving the second and third moments of the uniform on $[0, 1]$, uniformly bounded. Rewrite $I$ as

$$\begin{aligned}
&\sup_y \left| \mathbb{P}\left[\frac{(\frac{1}{b^r}\sum_{i=1}^n Y_i - \nu_n)}{\tau_n} < \frac{(\sigma_n y + \mu_n - \nu_n)}{\tau_n}\right] - \Phi(y) \right| \\
&\leqslant \sup_z \left| \mathbb{P}\left[\frac{(\frac{1}{b^r}\sum_i Y_i - \nu_n)}{\tau_n} \leqslant z\right] - \Phi(z) \right| + \sup_z \left| \Phi(z) - \Phi(a_1 z + a_2) \right|
\end{aligned}$$

with $a_1 = \tau_n/\sigma_n$, $a_2 = (\nu_n - \mu_n)/\sigma_n$. Using the Berry–Esseen bound again, the first term is bounded above by $B_2/\sqrt{n}$ with $B_2$ involving the ratio

$$\mathbb{E}\left|\frac{Y_1}{b^r} - \frac{b^r - 1}{2b^r}\right|^3 \Big/ \left(\mathbb{E}\left|\frac{Y_1}{b^r} - \frac{b^r - 1}{2b^r}\right|^2\right)^{3/2}.$$

This is uniformly bounded in $b, n, c$. To bound the final term, we use the following inequality: for any $\mu \in \mathbb{R}$, $\sigma^2 \in \mathbb{R}^+$,

$$\sup_z \left|\Phi(z) - \Phi(\sigma z + \mu)\right| \leqslant \left|\sigma^2 - 1\right| + |\mu| \frac{\sqrt{2\pi}}{4}. \tag{3.3}$$

An elegant proof of (3.3) using Stein's identity was communicated by Sourav Chatterjee. Let $W$ be Normal$(\mu, \sigma^2)$ and $Z$ be Normal$(0, 1)$. For any bounded $f$ with a bounded, piecewise continuous derivative, $\mathbb{E}(Wf(W)) = \mu\mathbb{E}(f(W)) + \sigma^2\mathbb{E}(f'(W))$ (Stein's identity being used). Thus

$$\mathbb{E}\big(Wf(W) - f'(W)\big) = \mu\mathbb{E}\big(f(W)\big) + \big(\sigma^2 - 1\big)\mathbb{E}\big(f'(W)\big).$$

As in [33, p. 22], choose $f_{w_0}$ so that for all $w$, one has

$$wf_{w_0}(w) - f'_{w_0}(w) = \delta_{w \leqslant w_0} - \Phi(w_0). \tag{3.4}$$

Here $w_0$ is fixed. Stein shows that $|f_{w_0}(w)| \leqslant \sqrt{2\pi}/4$ for all $w$, and that $|f'_{w_0}(w)| \leqslant 1$ for all $w$. Taking expectations in (3.4) proves (3.3). Taking $\sigma^2 = 1 - 1/b^{2r}$, $\mu = -\sqrt{3n}/b^r = -\sqrt{3/c}$, it follows that

$$\sup_z \left|\Phi(z) - \Phi(a_1 z + a_2)\right| \leqslant B_3/\sqrt{c}$$

with $B_3$ independent of $n, b, c > 1$. $\square$

**Remark.** If $W$ is Normal$(\nu, \tau^2)$ and $Z$ is Normal$(\mu, \sigma^2)$, the bound (3.3) shows that the Kolmogorov distance between their distributions is at most

$$\min\left(\frac{|\mu - \nu|}{\tau}\frac{\sqrt{2\pi}}{4} + \left|\frac{\sigma^2}{\tau^2} - 1\right|, \frac{|\mu - \nu|}{\sigma}\frac{\sqrt{2\pi}}{4} + \left|\frac{\tau^2}{\sigma^2} - 1\right|\right).$$

## 4. Signed permutations

Let $B_n$, the hyperoctahedral group, be represented as signed permutations. Thus $B_n$ has $2^n n!$ elements. We associate elements of $B_n$ to arrangements of a deck of $n$ cards with cards allowed to be face up or face down. A natural analog of the Gilbert–Shannon–Reeds shuffling model was studied in [3]; the deck is cut approximately in half, the top half turned face up, and the two halves are riffled together according to the G–S–R prescription. These shuffles have similarly neat combinatorial properties which allow sharp analysis of mixing times. Of course, shuffling is a natural algebraic operation and type $B$ shuffles have been studied from an algebraic viewpoint (with applications to Hochschild homology) in [5,6,20]. This section develops a corresponding carries process in rough parallel with Section 2. We also give an application to the theory of rounding.

From the previous sections, we see that the key idea is to use the fact that an $a$-shuffle followed by a $b$-shuffle is equivalent to an $ab$-shuffle. A hyperoctahedral analog of $(2a+1)$-shuffles was considered in [5] (see also [20] for connections with the affine Weyl group). A $(2a+1)$-shuffle is defined by multinomially cutting the deck into $2a+1$ piles, then flipping over the even numbered piles, and riffling them together.

View $B_n$ as the signed permutations on $n$ symbols, using the linear ordering

$$1 < 2 < \cdots < n < -n < \cdots < -2 < -1.$$

Say that

(1) $\sigma$ has a descent at position $i$ $(1 \leqslant i \leqslant n-1)$ if $\sigma(i) > \sigma(i+1)$.
(2) $\sigma$ has a descent at position $n$ if $\sigma(n) < 0$.

For example, $-1\ -2\ -3 \in B_3$ has three descents. Let $\overline{A(n,j)}$ denote the number of elements of $B_n$ with $j$ descents. The Bergerons [5] give analogs of basic properties of riffle shuffles. More precisely, they show that if a Markov chain on the hyperoctahedral group begins at the identity and proceeds by successive independent $(2b+1)$-shuffles, then

- The probability of obtaining the signed permutation $\tau$ after $r$ steps is

$$\frac{\binom{n+\frac{(2b+1)^r-1}{2}-d(\tau^{-1})}{n}}{(2b+1)^{rn}}. \tag{4.1}$$

- A $(2a+1)$-shuffle followed by a $(2b+1)$-shuffle is equivalent to a $(2a+1)(2b+1)$-shuffle.

Using these gives a type $B$ analog of Proposition 2.1. Gessel also has an unpublished proof of Proposition 4.1 using $P$-partitions.

**Proposition 4.1.** *Let $\sigma \in B_n$ have d descents. Let $c_{ij}^d$ be the number of ordered pairs $(\tau,\mu)$ of elements of $B_n$ such that $\tau$ has i descents, $\mu$ has j descents, and $\tau\mu = \sigma$. Then*

$$\sum_{i,j\geqslant 0} \frac{c_{ij}^d s^i t^j}{(1-s)^{n+1}(1-t)^{n+1}} = \sum_{a,b\geqslant 0} \binom{n+2ab+a+b-d}{n} s^a t^b.$$

**Proof.** Since a $(2a+1)$-shuffle followed by a $(2b+1)$-shuffle is equivalent to a $(2a+1)(2b+1)$-shuffle, the $r=1$ case of (4.1) gives that

$$\sum_{\mu\in B_n} \binom{n+a-d(\mu)}{n}\mu^{-1} \cdot \sum_{\tau\in B_n} \binom{n+b-d(\tau)}{n}\tau^{-1} = \sum_{\sigma\in B_n} \binom{n+2ab+a+b-d(\sigma)}{n}\sigma^{-1}.$$

As in the proof of Proposition 2.1, one multiplies both sides by $s^a t^b$, sums over all $a, b \geqslant 0$, and then takes the coefficient of $\sigma^{-1}$ on both sides to obtain the result. □

Next we define a "type $B$" carries process, to which we will relate the type $B$ hyperoctahedral shuffle. This is defined as the usual carries process, where one adds $n$ length $m$ numbers base $2b+1$, and to these adds the length $m$ number $(b, b, \ldots, b)$. Note that the state space of the type $B$ carries chain is $\{0, 1, \ldots, n\}$ (for usual carries, the most one can carry is $n-1$). For example when $b = 1$ (so $2b+1 = 3$), adding 222 and 201 followed by appending 111 gives

$$\begin{array}{cccc} \mathit{2} & \mathit{1} & \mathit{1} & \\ & 2 & 2 & 2 \\ & 2 & 0 & 1 \\ & 1 & 1 & 1 \\ \hline 2 & 0 & 1 & 1 \end{array}$$

with carries $\kappa_0 = 0$, $\kappa_1 = 1$, $\kappa_2 = 1$, $\kappa_3 = 2$.

**Theorem 4.2.** *For $0 \leqslant i, j \leqslant n$,*

(1) *The transition probabilities of the type $B$ carries chain are*

$$P(i,j) = \frac{1}{(2b+1)^n} \sum_{l \geqslant 0} (-1)^l \binom{n+1}{l} \binom{n+(j-l)(2b+1)+b-i}{n}.$$

(2) *The $r$-step transition probabilities of the type $B$ carries chain are*

$$P^r(i,j) = \frac{1}{(2b+1)^{rn}} \sum_{l \geqslant 0} (-1)^l \binom{n+1}{l} \binom{n+(j-l)(2b+1)^r + \frac{(2b+1)^r-1}{2} - i}{n}$$

(*i.e. one replaces $2b+1$ by $(2b+1)^r$ in part* (1)).

**Proof.** From the definition of the type $B$ carries chain,

$$P(i,j) = \mathbb{P}\big(j(2b+1) - b \leqslant i + X_1 + \cdots + X_n \leqslant j(2b+1) + b\big)$$

where $X_1, \ldots, X_n$ are independent discrete uniform random variables in $\{0, 1, \ldots, 2b\}$. Equivalently,

$$P(i,j) = (2b+1) \cdot \mathbb{P}\big(i + X_1 + \cdots + X_n + Y = j(2b+1) + b\big),$$

where $X_1, \ldots, X_n, Y$ are independent discrete uniform random variables in $\{0, 1, \ldots, 2b\}$. Letting $[x^h]f(x)$ denote the coefficient of $x^h$ in a series $f(x)$, it follows that

$$\begin{aligned}
P(i,j) &= \frac{1}{(2b+1)^n} \big[x^{j(2b+1)+b-i}\big] \left(\frac{1-x^{2b+1}}{1-x}\right)^{n+1} \\
&= \frac{1}{(2b+1)^n} \sum_{l \geqslant 0} (-1)^l \binom{n+1}{l} \big[x^{(j-l)(2b+1)+b-i}\big] \left(\frac{1}{1-x}\right)^{n+1} \\
&= \frac{1}{(2b+1)^n} \sum_{l \geqslant 0} (-1)^l \binom{n+1}{l} \binom{n+(j-l)(2b+1)+b-i}{n}.
\end{aligned}$$

Thus the first part is proved.

To prove the second half of the theorem, we show that $r$ steps of the base-$(2b+1)$ carries chain is equivalent to one step of the base $(2b+1)^r$ carries chain. To compute the carry after $r$ steps of the type $B_n$ carries chain, add $b(1 + (2b+1) + \cdots + (2b+1)^{r-1})$ to the sum of $n$ length $r$ numbers base $2b+1$. To compute the carry after one step of the type $B_n$ base $(2b+1)^r$ carries chain, add $[(2b+1)^r - 1]/2$ to the sum of $n$ length 1 numbers base $(2b+1)^r$. These computations are equivalent, so the result follows by replacing $2b+1$ by $(2b+1)^r$ in part (1). $\square$

Now we relate hyperoctahedral shuffles to type $B$ carries. In what follows, $\pi$ denotes the distribution on $\{0, 1, \ldots, n\}$ defined by $\pi(j) = \overline{A(n,j)}/(2^n n!)$.

**Theorem 4.3.** *Let a Markov chain on the hyperoctahedral group $B_n$ begin at the identity and proceed by successive independent $(2b+1)$-shuffles. Then the number of descents of $\tau^{-1}$ forms a Markov chain, and its formal time reversal with respect to its stationary distribution $\pi$ is identical with the carries Markov chain.*

**Proof.** Let $\tau_r$ be the element of $B_n$ obtained after $r$ independent $b$-shuffles (started at the identity). Arguing as in the proof of Theorem 2.2 gives that $d(\tau_r^{-1})$ forms a Markov chain with transition probabilities

$$\mathbb{P}\big(d\big(\tau_r^{-1}\big)=j|d\big(\tau_{r-1}^{-1}\big)=i\big)=\frac{\overline{A(n,j)}}{\overline{A(n,i)}(2b+1)^n}\sum_{k\geqslant 0}c_{ik}^j\binom{n+b-k}{n}.$$

Here $c_{ik}^j$ is as in the statement of Proposition 4.1.

Letting $[x^h]f(x)$ denote the coefficient of $x^h$ in a series $f(x)$, the transition probability in the previous paragraph can be written as

$$\begin{aligned}&\big[t^b\big]\frac{\overline{A(n,j)}}{\overline{A(n,i)}(2b+1)^n}\sum_{k\geqslant 0}c_{ik}^j\frac{t^k}{(1-t)^{n+1}}\\ &\quad=\big[t^bs^i\big]\frac{\overline{A(n,j)}}{\overline{A(n,i)}(2b+1)^n}(1-s)^{n+1}\sum_{i,k\geqslant 0}c_{ik}^j\frac{s^it^k}{(1-s)^{n+1}(1-t)^{n+1}}.\end{aligned}$$

By Proposition 4.1 this is equal to

$$\begin{aligned}&\big[t^bs^i\big]\frac{\overline{A(n,j)}}{\overline{A(n,i)}(2b+1)^n}(1-s)^{n+1}\sum_{a,c\geqslant 0}\binom{n+2ac+a+c-j}{n}s^at^c\\ &\quad=\frac{\overline{A(n,j)}}{\overline{A(n,i)}(2b+1)^n}\sum_{l\geqslant 0}(-1)^l\binom{n+1}{l}\big[t^b\big]\sum_{c\geqslant 0}\binom{n+(i-l)(2c+1)+c-j}{n}t^c\\ &\quad=\frac{\overline{A(n,j)}}{\overline{A(n,i)}(2b+1)^n}\sum_{l\geqslant 0}(-1)^l\binom{n+1}{l}\binom{n+(i-l)(2b+1)+b-j}{n}.\end{aligned}$$

Comparing with Theorem 4.2, this is equal to $\pi(j)P(j,i)/\pi(i)$, as needed. $\square$

The next theorem is easily proved by the technique used to prove Theorem 2.3 (using Proposition 4.1 instead of Proposition 2.1).

**Theorem 4.4.** *The probability that the type B carries chain goes from* 0 *to $j$ in $r$ steps is equal to the probability that an element of $B_n$ obtained by performing $r$ successive* $(2b+1)$*-shuffles* (*started at the identity*) *has $j$ descents.*

The following corollary is immediate from Theorem 4.4.

**Corollary 4.5.** *The stationary distribution of the type B carries chain is given by* $\pi(j)=\overline{A(n,j)}/(2^nn!)$, *where $\overline{A(n,j)}$ is the number of signed permutations on $n$ symbols with $j$ descents.*

Corollary 4.6 gives a closed formula for $\overline{A(n,j)}$. (This can also be obtained by combining Proposition 4.7 below with Eq. (19) of [11].)

**Corollary 4.6.**

$$\overline{A(n,j)}=\sum_{l=0}^{j}(-1)^l\binom{n+1}{l}(2j-2l+1)^n.$$

**Proof.** Let $r \to \infty$ in part (2) of Theorem 4.2, and apply Corollary 4.5. $\square$

As an application of the above results, we give a new proof of the following lovely fact from [29] (see also Section 9 of [11] for closely related results). Note that it can be interpreted as computing the probability that the sum of $n$ independent uniform random variables on $[0,1]$, when rounded to the nearest integer, is equal to $j$.

**Proposition 4.7.** *Let $U_1,\dots,U_n$ be independent, identically distributed continuous uniform random variables in* $[0,1]$. *Then*

$$\mathbb{P}\left(j-\frac{1}{2}\leqslant U_1+\cdots+U_n\leqslant j+\frac{1}{2}\right)=\frac{\overline{A(n,j)}}{2^n n!}.$$

**Proof.** Let $X_1,\dots,X_n$ be independent discrete uniform random variables on $\{0,1,\dots,2b\}$. From the definition of the type $B_n$ base-$(2b+1)$ carries chain,

$$P(0,j)=\mathbb{P}\left(j(2b+1)-b\leqslant\sum_{i=1}^{n}X_i<j(2b+1)+b+1\right).$$

Let $Y_1,\dots,Y_n$ be independent continuous uniform random variables on $[0,2b+1]$. Then it follows that

$$\begin{aligned}P(0,j)&=\mathbb{P}\left(j(2b+1)-b\leqslant\sum_{i=1}^{n}\lfloor Y_i\rfloor<j(2b+1)+b+1\right)\\&=\mathbb{P}\left((2b+1)\left(j-\frac{1}{2}\right)\leqslant\sum_{i=1}^{n}Y_i-\sum_{i=1}^{n}\bigl(Y_i-\lfloor Y_i\rfloor\bigr)-\frac{1}{2}<(2b+1)\left(j+\frac{1}{2}\right)\right)\\&=\mathbb{P}\left(j-\frac{1}{2}\leqslant\sum_{i=1}^{n}U_i-E<j+\frac{1}{2}\right).\end{aligned}$$

Here the $U_i=Y_i/(2b+1)$ are independent continuous uniform random variables on $[0,1]$, and

$$E=\frac{1}{2b+1}\sum_{i=1}^{n}\bigl(Y_i-\lfloor Y_i\rfloor\bigr)+\frac{1}{2(2b+1)}.$$

Note that when $n$ is fixed and $b\to\infty$, $E$ converges to 0 with probability 1. Thus Slutsky's theorem implies that

$$\lim_{b\to\infty}P(0,j)=\mathbb{P}\left(j-\frac{1}{2}\leqslant\sum_{i=1}^{n}U_i<j+\frac{1}{2}\right).$$

However by Theorem 4.2 and Corollary 4.5,

$$\lim_{b\to\infty}P(0,j)=\pi(j)=\frac{\overline{A(n,j)}}{2^n n!},$$

which completes the proof. $\square$

## 5. Two final topics

The carries matrix also comes up in studying sections of generating functions via the Veronese map. The large $n$ limit of the carries *process* is well approximated by a classical auto-regressive process.

### *5.1. Eulerian polynomials and Hilbert series of Veronese subrings*

Some natural sequences $a_k$, $0 \leqslant k < \infty$, have generating functions

$$\sum_{k=0}^{\infty} a_k x^k = \frac{h(x)}{(1-x)^{n+1}}$$

with $h(x) = h_0 + h_1 x + \cdots + h_{n+1} x^{n+1}$ a polynomial of degree at most $n+1$. Suppose we are interested in every $b$th term $\{a_{bk}\}$, $0 \leqslant k < \infty$. It is not hard to see that

$$\sum_{k=0}^{\infty} a_{bk} x^k = \frac{h^{\langle b \rangle}(x)}{(1-x)^{n+1}}$$

for another polynomial $h^{\langle b \rangle}(x)$ of degree at most $n+1$. The study of these generating functions arises naturally in algebraic geometry [18] and lattice point enumeration [4].

Brenti and Welker [8] show that the $i$th coefficient of $h^{\langle b \rangle}(x)$ satisfies

$$h_i^{\langle b \rangle} = \sum_{j=0}^{n+1} C(i,j) h_j$$

with $C$ an $(n+2) \times (n+2)$ matrix with $(i,j)$ entry $(0 \leqslant i, j \leqslant n+1)$ equal to the number of solutions to $a_1 + \cdots + a_{n+1} = ib - j$ where $0 \leqslant a_l \leqslant b-1$ are integers. In [14] we show that the $n \times n$ matrix given by deleting the first and last rows and columns of $C$, then multiplying by $b^{-n}$ and taking the transpose is precisely the carries matrix $(P(i,j))$ of (1.1).

Since iterates of the carries chain converge, the matrix $(C(i,j))$ has nice limiting behavior. Brenti and Welker [8] show that the zeros of $h^{\langle b \rangle}$ converge and Beck and Stapledon [4] show that the zeros converge to the zeros of the $n$th Eulerian polynomial $p_n(x)$, defined as $\sum_{j \geqslant 0} A(n,j) x^{j+1}$, where $A(n,j)$ is the number of permutations in $S_n$ with $j$ descents. The following is a refinement.

**Theorem 5.1.** *Suppose that $h(1) \neq 0$ and let $p_n(x)$ be the nth Eulerian polynomial. Then as $b \to \infty$ with $n$ fixed,*

$$\frac{h^{\langle b \rangle}(x)}{b^n \cdot h(1)} \to \frac{p_n(x)}{n!}.$$

**Proof.** Let $[y^k] f(y)$ denote the coefficient of $y^k$ in a power series $f(y)$. Then the definition of $C(i,j)$ gives that

$$C(i,j) = \left[y^{ib-j}\right] \frac{(1-y^b)^{n+1}}{(1-y)^{n+1}} = \sum_{l \geqslant 0} (-1)^l \binom{n+1}{l} \left[y^{(i-l)b-j}\right] \frac{1}{(1-y)^{n+1}}$$
$$= \sum_{l \geqslant 0} (-1)^l \binom{n+1}{l} \binom{n+(i-l)b-j}{n}.$$

Supposing that $1 \leqslant i \leqslant n$, it follows that

$$\lim_{b\to\infty} \frac{C(i,j)}{b^n} = \frac{1}{n!}\sum_{l=0}^{i}(-1)^l \binom{n+1}{l}(i-l)^n = \frac{A(n,i-1)}{n!} \tag{5.1}$$

where the second equality uses a well-known formula for Eulerian numbers [12]. Since $C(0,j) = \delta_{0,j}$ and $C(n+1,j) = \delta_{n+1,j}$, clearly

$$\lim_{b\to\infty} \frac{C(0,j)}{b^n} = \lim_{b\to\infty} \frac{C(n+1,j)}{b^n} = 0. \tag{5.2}$$

Combining Eqs. (5.1) and (5.2) yields that

$$\begin{aligned}
\lim_{b\to\infty} \frac{h^{\langle b\rangle}(x)}{b^n \cdot h(1)} &= \lim_{b\to\infty} \frac{\sum_{i=0}^{n+1}[\sum_{j=0}^{n+1} C(i,j)h_j]x^i}{b^n \cdot h(1)} \\
&= \frac{\sum_{i=1}^{n}[A(n,i-1)\sum_{j=0}^{n+1} h_j]x^i}{n! \cdot h(1)} \\
&= \frac{p_n(x)}{n!}. \qquad \square
\end{aligned}$$

Here is an example. The coordinate ring $R$ of a projective variety in $n+1$ variables decomposes into its graded pieces $R_k$, $0 \leqslant k \leqslant \infty$ and the Hilbert series has the form [2, Theorem 11.1]

$$\sum_{k=0}^{\infty} \dim(R_k)x^k = \frac{h(x)}{(1-x)^{n+1}}.$$

The *bth Veronese embedding* replaces the variables by all degree $b$ monomials in these variables. (If $b=3$, $\{x,y\}$ are replaced by $x^3, x^2y, xy^2, y^3$.) The image of the coordinate ring has Hilbert series $h^{\langle b\rangle}(x)/(1-x)^{n+1}$. As a simple special case, the full projective space has coordinate ring $\mathbb{C}[x_1 \ldots x_{n+1}]$. The degree $k$ homogeneous polynomials have dimension $\binom{n+k}{n}$ and

$$\sum_{k=0}^{\infty} \binom{n+k}{n} x^k = \frac{1}{(1-x)^{n+1}}. \tag{5.3}$$

When $n+1=2$,

$$\sum_{k=0}^{\infty}(k+1)x^k = \frac{1}{(1-x)^2} \quad \text{and} \quad \sum_{k=0}^{\infty}(bk+1)x^k = \frac{(b-1)x+1}{(1-x)^2}.$$

When $n+1=3$,

$$\sum_{k=0}^{\infty}\binom{k+2}{2}x^k = \frac{1}{(1-x)^3} \quad \text{and} \quad \sum_{k=0}^{\infty}\binom{bk+2}{2}x^k = \frac{x^2(b(b-3)+2)+x(b(b+3)-4)+2}{2(1-x)^3}.$$

Dividing the right-hand sides of these expressions by $b^n \cdot h(1)$ (here $b$ and $b^2$ respectively), then multiplying by $(1-x)^{n+1}$ and letting $b \to \infty$, gives $p_n(x)/n!$ (here $x$ and $(x^2+x)/2$ respectively).

*5.2. Autoregressive approximation*

This section studies the large $n$ limit of the carries process and shows it is well approximated by a classical autoregressive process. Throughout, we work with a general base $b$ and let $n$ be the number of summands. Let $\kappa_0 = 0, \kappa_1, \kappa_2, \ldots$ be the carries process on $\{0, 1, \ldots, n-1\}$. Let $Y_t = (\kappa_t - n/2)/\sqrt{n/12}$, $t = 0, 1, 2, \ldots$. Theorem 5.2 relates $Y_t$ to a Gaussian autoregressive process $W_0, W_1, W_2, \ldots$ defined by $W_0 = -\sqrt{3n}$, $W_{t+1} = W_t/b + \epsilon_t$, with the $\epsilon_t$ independent Normal$(0, 1-1/b^2)$ random variables.

**Theorem 5.2.** *Let $P_n$ be the law of the process $Y_t$, $0 \leqslant t < \infty$, on $\mathbb{R}^\infty$. Let $Q$ be the law of the process $W_0, W_1, \ldots$ on $\mathbb{R}^\infty$. Then $P_n \Rightarrow Q$ as $n \to \infty$.*

The following lemma will be helpful for proving Theorem 5.2.

**Lemma 5.3.** *The base b carries process can be represented as follows*:

$$\textit{If } \kappa_t = r \bmod b, \quad \textit{let } \kappa_{t+1} = \frac{\kappa_t - r}{b} + \epsilon_{t+1} \tag{5.4}$$

*where $\mathbb{P}(\epsilon_{t+1} = k)$ is the probability that the sum of $n+1$ independent discrete uniform random variables on $\{0, 1, \ldots, b-1\}$ is equal to $bk + b - r - 1$, given that the sum is congruent to $b - r - 1$ mod $b$.*

**Proof.** From page 140 of Holte [23] one can write the carries transition probability as

$$P(i, j) = \frac{1}{b^n}\left[x^{(j+1)b-i-1}\right]\left(1 + x + \cdots + x^{b-1}\right)^{n+1} \tag{5.5}$$

where $[x^k]f(x)$ denotes the coefficient of $x^k$ in a polynomial $f(x)$. If $i = r \bmod b$, write $j = (i-r)/b + \epsilon_{t+1}$. Then (5.5) becomes

$$\frac{1}{b^n}\left[x^{b-r-1+(\epsilon_{t+1})b}\right]\left(1 + x + \cdots + x^{b-1}\right)^{n+1}.$$

To see that this implies the lemma, note that the sum of $n+1$ discrete uniform random variables on $\{0, 1, \ldots, b-1\}$ is equidistributed mod $b$, and so is congruent to $b - r - 1$ mod $b$ with probability $1/b$. $\square$

**Proof of Theorem 5.2.** We show convergence by showing that $\{P_n\}_{n=1}^{\infty}$ is tight and that the finite dimensional distributions of $P_n$ converge to the finite dimensional distributions of $Q$. This is enough from [19, 2.2, 4.3, 4.5]. From [19, 2.4], $P_n$ is tight if and only if the family $P_n^h$ of $h$th marginal distributions is tight. Thus it is enough to show that the finite dimensional distributions converge.

By Lemma 5.3 the carries process can be represented as:

$$\text{If } \kappa_t = r \bmod b, \quad \text{let } \kappa_{t+1} = \frac{\kappa_t - r}{b} + \epsilon_{t+1} \tag{5.6}$$

where $\mathbb{P}(\epsilon_{t+1} = k)$ is the probability that the sum of $n+1$ independent uniform random variables on $\{0, 1, \ldots, b-1\}$ is equal to $bk + b - r - 1$, given that the sum is congruent to $b - r - 1$ mod $b$. Hence $b\epsilon_{t+1}$ (for $\kappa_t = b - 1 \bmod b$) has the distribution of the sum of $n+1$ uniform random variables on $\{0, 1, \ldots, b-1\}$ given that the sum is congruent to 0 mod $b$. A generating function argument then shows that if $n \geqslant 2$ then $\epsilon_{t+1}$ has mean $\frac{n+1}{2}(1 - 1/b)$ and variance $\frac{n+1}{12}(1 - 1/b^2)$. By the local central limit theorem for sums of independent, identically distributed random variables,

$$\frac{\epsilon_{t+1} - \frac{n}{2}(1 - \frac{1}{b})}{\sqrt{\frac{n}{12}(1 - \frac{1}{b^2})}} \implies \mathcal{N}(0, 1)$$

as $n \to \infty$, and a similar argument gives the same conclusion for $\kappa_t$ congruent to any $r$ mod $b$, with error term $O(n^{-1/2})$ since $b$ is fixed.

From these considerations, the joint distribution of $(\epsilon_1, \epsilon_2, \ldots, \epsilon_h)$, normalized as above, converges to the product of $h$ independent standard normal variables ($h$ fixed, $n$ large).

Next, write $\kappa_{t+1} = (\kappa_t - \delta_t)/b + \epsilon_{t+1}$ with $\delta_t = \kappa_t$ mod $b$. Thus, for $t = 1, 2, 3, \ldots, h-1$, with $\kappa_0 = \delta_0 = 0$,

$$\kappa_{t+1} = \frac{\kappa_0}{b^{t+1}} + \left(\epsilon_{t+1} + \frac{\epsilon_t}{b} + \cdots + \frac{\epsilon_1}{b^t}\right) - \left(\frac{\delta_t}{b} + \frac{\delta_{t-1}}{b^2} + \cdots + \frac{\delta_0}{b^{t+1}}\right).$$

Since $\kappa_{t+1} = \sqrt{n/12} \cdot Y_{t+1} + n/2$ and $\kappa_0 = \sqrt{n/12} \cdot Y_0 + n/2$, it follows that

$$\begin{aligned} Y_{t+1} &= \frac{Y_0}{b^{t+1}} - \frac{n/2}{\sqrt{n/12}}\left(1 - \frac{1}{b^{t+1}}\right) + \frac{1}{\sqrt{n/12}}\left(\epsilon_{t+1} + \frac{\epsilon_t}{b} + \cdots + \frac{\epsilon_1}{b^t}\right) \\ &\quad - \frac{1}{\sqrt{n/12}}\left(\frac{\delta_t}{b} + \frac{\delta_{t-1}}{b^2} + \cdots + \frac{\delta_0}{b^{t+1}}\right) \\ &= \frac{Y_0}{b^{t+1}} + \frac{1}{\sqrt{n/12}}\left[\left(\epsilon_{t+1} - \frac{n}{2}\left(1 - \frac{1}{b}\right)\right) + \cdots + \frac{(\epsilon_1 - \frac{n}{2}(1 - \frac{1}{b}))}{b^t}\right] \\ &\quad - \frac{1}{\sqrt{n/12}}\left(\frac{\delta_t}{b} + \frac{\delta_{t-1}}{b^2} + \cdots + \frac{\delta_0}{b^{t+1}}\right). \end{aligned}$$

As noted earlier, the $[\epsilon_i - \frac{n}{2}(1 - 1/b)]/\sqrt{n/12}$ converge to independent $\mathcal{N}(0, 1 - 1/b^2)$'s. Since $|\delta_i| \leqslant b$ for all $i$, the term involving the $\delta$'s converges to 0 with probability 1, so by Slutsky's theorem, it can be disregarded. We thus have that the joint distribution of $(Y_0, Y_1, \ldots, Y_h)$ converges to the joint distribution of $(W_0, W_1, \ldots, W_h)$, and the proof is complete. □

**Remark.** The Gaussian autoregressive process $X_{n+1} = \frac{1}{b} X_n + \epsilon_{n+1}$ (with $X_0 = x$) is carefully studied in [16]. It has eigenvalues $1, 1/b, 1/b^2, \ldots$ and takes order $\log_b |x| + c$ steps to converge [16, Proposition 4.9]. Taking $x = -\sqrt{3n}$ as in Theorem 5.2, this is consistent with our result in Section 3 that $\frac{1}{2}\log_b(n) + c$ steps is the right answer for the carries chain.

**Remark.** Theorem 5.2 implies that many properties of Gaussian autoregressive processes (here the discrete time Ornstein–Uhlenbeck process) apply to carries—at least in the limit. For example Corollary 2 of Lai [26] implies, in the notation above, that

$$\mathbb{P}(W_t \geqslant b_t \text{ i.o.}) = 1 \text{ or } 0 \quad \text{according as } \sum_{t=0}^{\infty} b_t^{-1} e^{-b_t^2/2} = \infty \text{ or } < \infty.$$

It follows for carries that

$$\mathbb{P}\left(\kappa_t \geqslant \frac{n}{2} + \sqrt{\frac{n}{12}\big(\log(t)\big)^{1+\epsilon}} \text{ i.o.}\right) = 1 \text{ or } 0 \quad \text{according as } \epsilon \leqslant 1 \text{ or } \epsilon > 1.$$

## Acknowledgments

We thank Sourav Chatterjee and Tze Lai for their help. Diaconis was partially supported by NSF grant DMS 0505673. Fulman was partially supported by NSF grant DMS 0802082 and NSA grant H98230-08-1-0133.

## References

[1] S. Assaf, P. Diaconis, K. Soundararajan, A rule of thumb for riffle shuffling, preprint, Department of Statistics, Stanford University, Stanford, CA, 2008.
[2] M. Atiyah, I.G. Macdonald, Introduction to Commutative Algebra, Addison–Wesley, Reading, MA, 1969.
[3] D. Bayer, P. Diaconis, Trailing the dovetail shuffle to its lair, Ann. Appl. Probab. 2 (1992) 294–313.
[4] M. Beck, A. Stapledon, On the log-concavity of Hilbert series of Veronese subrings and Ehrhart series, arXiv:0804.3639, 2008.
[5] F. Bergeron, N. Bergeron, Orthogonal idempotents in the descent algebra of $B_n$ and applications, J. Pure Appl. Algebra 79 (1992) 109–129.
[6] N. Bergeron, Hyperoctahedral operations on Hochschild homology, Adv. Math. 110 (1995) 255–276.
[7] P. Bidigare, P. Hanlon, D. Rockmore, A combinatorial description of the spectrum for the Tsetlin library and its generalization to hyperplane arrangements, Duke Math. J. 99 (1999) 135–174.
[8] F. Brenti, V. Welker, The Veronese construction for formal power series and graded algebras, Adv. in Appl. Math. 42 (2009) 545–556.
[9] K.S. Brown, Semigroups, rings, and Markov chains, J. Theoret. Probab. 13 (2000) 871–938.
[10] K.S. Brown, P. Diaconis, Random walks and hyperplane arrangements, Ann. Probab. 26 (1998) 1813–1854.
[11] D. Chakerian, D. Logothetti, Cube slices, pictorial triangles, and probability, Math. Mag. 64 (1991) 219–241.
[12] L. Comtet, Advanced Combinatorics, Reidel, Dordrecht, 1974.
[13] P. Diaconis, Mathematical developments from the analysis of riffle-shuffling, in: A. Ivanov, M. Liebeck, J. Saxl (Eds.), Groups, Combinatorics and Geometry, World Scientific, New Jersey, 2003, pp. 73–97.
[14] P. Diaconis, J. Fulman, Carries, shuffling, and an amazing matrix, Amer. Math. Monthly (2008), in press, available at http://www-rcf.usc.edu/~fulman/carries.pdf.
[15] P. Diaconis, K. Khare, L. Saloff-Coste, Gibbs sampling, conjugate priors, and coupling, Technical report, Department of Statistics, Stanford University, 2008.
[16] P. Diaconis, K. Khare, L. Saloff-Coste, Gibbs sampling, conjugate priors, and orthogonal polynomials, with discussion, Statist. Sci. 23 (2008) 151–200.
[17] P. Diaconis, M. McGrath, J. Pitman, Riffle shuffles, cycles and descents, Combinatorica 15 (1995) 11–29.
[18] D. Eisenbud, A. Reeves, B. Totaro, Initial ideals, Veronese subrings, and rates of algebras, Adv. Math. 109 (1994) 168–187.
[19] S. Ethier, T. Kurtz, Markov Processes, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, Wiley, New York, 1986.
[20] J. Fulman, Applications of the Brauer complex: Card shuffling, permutation statistics, and dynamical systems, J. Algebra 243 (2001) 96–122.
[21] J. Fulman, Applications of symmetric functions to cycle and increasing subsequence structure after shuffles, J. Algebraic Combin. 16 (2002) 165–194.
[22] I. Gessel, C. Reutenauer, Counting permutations with given cycle structure and descent set, J. Combin. Theory Ser. A 64 (1993) 189–215.
[23] J. Holte, Carries, combinatorics, and an amazing matrix, Amer. Math. Monthly 104 (1997) 138–149.
[24] S. Karlin, Total Positivity, vol. I, Stanford University Press, Stanford, CA, 1968.
[25] J.G. Kemeny, J.L. Snell, Finite Markov Chains, Univ. Ser. Undergrad. Math., D. van Nostrand, Princeton, NJ, 1960.
[26] T.L. Lai, Gaussian processes, moving averages and quick detection problems, Ann. Probab. 1 (1973) 825–837.
[27] J.W. Pitman, Probabilistic bounds on the coefficients of polynomials with only real zeros, J. Combin. Theory Ser. A 77 (1997) 279–303.
[28] L.C.G. Rogers, J.W. Pitman, Markov functions, Ann. Probab. 9 (1981) 573–582.
[29] F. Schmidt, R. Simion, Some geometric probability problems involving the Eulerian numbers, Electron. J. Combin. 4 (1997), Research Paper 18, approx. 13 pp. (electronic).
[30] R. Stanley, Eulerian partitions of a unit hypercube, in: M. Aigner (Ed.), Higher Combinatorics, Reidel, Dordrecht/Boston, 1977, p. 49.
[31] R. Stanley, Enumerative Combinatorics II, Cambridge University Press, Cambridge, 1999.
[32] R. Stanley, Generalized riffle shuffles and quasisymmetric functions, Ann. Comb. 5 (2001) 479–491.
[33] C. Stein, Approximate Computation of Expectations, IMS Lecture Notes Monogr. Ser., vol. 7, Institute of Mathematical Statistics, Hayward, CA, 1986.