

q -EXCHANGEABILITY VIA QUASI-INVARIANCE

BY ALEXANDER GNEDIN AND GRIGORI OLSHANSKI¹

*Utrecht University and Institute for Information Transmission Problems and
Independent University of Moscow*

For positive $q \neq 1$, the q -exchangeability of an infinite random word is introduced as quasi-invariance under permutations of letters, with a special cocycle which accounts for inversions in the word. This framework allows us to extend the q -analog of de Finetti's theorem for binary sequences—see Gnedin and Olshanski [*Electron. J. Combin.* **16** (2009) R78]—to general real-valued sequences. In contrast to the classical case of exchangeability ($q = 1$), the order on \mathbb{R} plays a significant role for the q -analogs. An explicit construction of ergodic q -exchangeable measures involves random shuffling of $\mathbb{N} = \{1, 2, \dots\}$ by iteration of the geometric choice. Connections are established with transient Markov chains on q -Pascal pyramids and invariant random flags over the Galois fields.

1. Introduction. A random word $w = w_1 w_2 \dots$ with letters $w_i \in \mathbb{A}$ over some alphabet $\mathbb{A} \subseteq \mathbb{R}$ is *exchangeable* if swapping the places of two neighboring letters w_i and w_{i+1} does not change the probability. We shall study the following deformation of this fundamental random symmetry property. For positive parameter $q \neq 1$, we define w to be *q -exchangeable* if, by swapping the places of two neighboring letters w_i and w_{i+1} , the probability is multiplied by factor $q^{\text{sgn}(w_{i+1} - w_i)}$. The intuitive effect of the deformation is that the arrangement of letters in the word is not completely random, as for exchangeable sequences, but rather there is a tendency for some monotonic pattern. To be definite, we shall focus temporarily on the instance $0 < q < 1$, in which case, words with smaller numbers of *inversions* are more likely, the latter defined as pairs of positions $i < j$ with $w_i > w_j$.

The same definitions apply to finite words. It is well known and easy to see that the most general exchangeable word of fixed length n can be produced by first choosing an inversion-free word $v_1 \leq \dots \leq v_n$ from an arbitrary probability distribution on the space of weakly increasing sequences over \mathbb{A} , then shuffling the letters by an independent uniformly random permutation of $\mathbb{N}_n := \{1, \dots, n\}$. We will show that the general finitely q -exchangeable word is produced similarly, with

Received August 2009; revised January 2010.

¹Supported by Utrecht University, RFBR Grant 08-01-00110 and the Project SFB 701 of Bielefeld University.

AMS 2000 subject classifications. Primary 60G09, 60C05; secondary 37A50.

Key words and phrases. q -exchangeability, ergodic decomposition, Mallows distribution.

the amendment that the random permutation should follow the Mallows distribution [15], which assigns to each particular permutation $\sigma : \mathbb{N}_n \rightarrow \mathbb{N}_n$ probability proportional to $q^{\text{inv}(\sigma)}$, with $\text{inv}(\sigma)$ being the number of inversions in σ .

The analogy between the above exchangeable and q -exchangeable representations does not extend to infinite words. According to de Finetti's theorem [1, 11], an infinite exchangeable w satisfies the strong law of large numbers: for fixed $B \subset \mathbb{A}$, the proportion of letters $w_i \in B$ in the initial subword of length n is asymptotic to $\nu(B)$, where ν is a random probability measure on \mathbb{A} . Conditionally on ν , the random word is distributed like an i.i.d. sample from ν , so the ergodic distributions for w are parametrized by probability measures on \mathbb{A} . In contrast to that, there is no principal difference between the general representations of finite and infinite q -exchangeable words. According to our main result (Theorem 4.8), every infinite q -exchangeable w can be produced by choosing a random sequence $v_1 \leq v_2 \leq \dots$ from some arbitrary distribution on the space of infinite increasing sequences over \mathbb{A} , then shuffling the letters in the order determined by an independent permutation $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ whose distribution is a properly generalized Mallows distribution on the group \mathfrak{S} of all bijections of the set \mathbb{N} . Thus, every ergodic q -exchangeable distribution for w is supported by a single orbit of the group \mathfrak{S} acting on \mathbb{A}^∞ by permutations of coordinates.

Proving the stated representation of q -exchangeable words and analysis of the Mallows distribution on \mathfrak{S} constitute the main contents of this paper. For $\mathbb{A} = \mathbb{N}_d$, we show that every q -exchangeable w can be encoded into an increasing random walk on the d -dimensional lattice with weighted edges (the q -Pascal pyramid); the ergodic measures are derived, in this case, by solving a boundary problem via path counting and asymptotics of the Gaussian multinomial coefficients.

In our recent paper [9], we observed that every homogeneous random subspace of an infinite-dimensional space V over a Galois field corresponds to a q -exchangeable sequence over $\mathbb{A} = \{0, 1\}$, for q reciprocal to the cardinality of the field. Here, homogeneity means *invariance* of the measure under the natural action on the subspaces of V by the countable group of matrices $GL(\infty) = \bigcup_{n \in \mathbb{N}} GL(n)$. In what follows, we shall extend this line of research by connecting nonstrict random flags (sequences of embedded subspaces in V) with q -exchangeable words over $\mathbb{A} = \mathbb{N}$.

2. q -exchangeability. In terms of measure theory, q -exchangeability means *quasi*-invariance of a probability measure on \mathbb{A}^∞ with respect to permutations of an arbitrary finite collection of coordinates, with a special Radon–Nikodym derivative depending on the altered number of inversions in the word. To develop this viewpoint, we first recall a general framework and some necessary facts from ergodic theory [10].

Let W be a standard Borel space and G be a countable group acting on W on the left by Borel isomorphisms $T_g : W \rightarrow W$, $g \in G$. Then, G also acts on the space of all Borel probability measures on W : namely, T_g transforms such a measure P to

$T_g P := P \circ T_g^{-1}$. We prefer to write this relation as $T_g^{-1} P = P \circ T_g$, which means that $(T_g^{-1} P)(X) = P(T_g(X))$ for every Borel set $X \subseteq W$.

A probability measure P on W is said to be *quasi-invariant* if $T_g^{-1} P$ is equivalent to P for all $g \in G$, that is, $T_g^{-1} P$ and P have the same null sets. There then exists a function $\rho(g, w)$ on $G \times W$ such that $w \mapsto \rho(g, w)$ is Borel and $T_g^{-1} P = \rho(g, \cdot) P$ for each $g \in G$, that is, $\rho(g, \cdot)$ is the Radon–Nikodym derivative $dT_g^{-1} P/dP$. The function ρ is unique modulo P -null sets and satisfies the relation

$$\rho(gh, w) = \rho(g, T_h w) \rho(h, w), \quad g, h \in G, w \in W$$

(again modulo null sets). A function ρ with this property is called a *multiplicative cocycle*.

Conversely, given a multiplicative cocycle ρ , let $\mathfrak{M}(\rho)$ denote the set of all quasi-invariant probability measures on W satisfying the relation $dT_g^{-1} P/dP = \rho(g, \cdot)$, $g \in G$. The set $\mathfrak{M}(\rho)$ has itself the structure of a standard Borel space and if $\mathfrak{M}(\rho)$ is nonempty, then it is convex and has a nonempty subset $\text{Ex } \mathfrak{M}(\rho)$ of extreme points. The set of extremes $\text{Ex } \mathfrak{M}(\rho)$ is also Borel. Moreover, every measure $M \in \mathfrak{M}(\rho)$ is uniquely representable as a mixture of the extreme measures, meaning that there exists a unique probability measure \varkappa on $\text{Ex } \mathfrak{M}(\rho)$ such that

$$M(X) = \int_{\text{Ex } \mathfrak{M}(\rho)} P(X) \varkappa(dP)$$

for every Borel subset $X \subseteq W$.

Since the generic element of $\mathfrak{M}(\rho)$ is a unique mixture of extremes, it is important to describe as explicitly as possible the set of extremes $\text{Ex } \mathfrak{M}(\rho)$. A useful criterion is that the extreme measures can be characterized as ergodic measures from $\mathfrak{M}(\rho)$. Recall that a G -quasi-invariant probability measure P on W is *ergodic* if every G -invariant Borel subset of W has P -measure 0 or 1. Since the group G is countable, the ergodicity is equivalent to the formally stronger condition that every invariant mod 0 subset has measure 0 or 1.

After these general preliminaries, we focus on a concrete instance. We shall consider the action of the group $G = \mathfrak{S}_\infty$ on the infinite product space $W = \mathbb{A}^\infty$, where \mathfrak{S}_∞ is the group of bijections $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ moving only finitely many integers and \mathbb{A} is a Borel subset of the ordered space $(\mathbb{R}, <)$. Although we assume $\mathbb{A} \subseteq \mathbb{R}$, many considerations of the present paper remain valid for an arbitrary standard Borel space endowed with a Borel-measurable linear order (e.g., \mathbb{R}^k with the lexicographic order).

Given a finite word $w = w_1 w_2 \cdots w_n \in \mathbb{A}^n$, let

$$\text{inv}(w_1 \cdots w_n) := \#\{(i, j) \mid 1 \leq i < j \leq n, w_i > w_j\}$$

denote the number of inversions in w . For an infinite word $w = w_1 w_2 \cdots \in \mathbb{A}^\infty$, let

$$\text{inv}_n(w) = \text{inv}(w_1 \cdots w_n)$$

be the number of inversions in the n -truncated word $w_1 \cdots w_n$.

For $w \in \mathbb{A}^\infty$ and $\sigma \in \mathfrak{S}_\infty$, the difference $\text{inv}_n(T_\sigma w) - \text{inv}_n(w)$ stabilizes as n becomes so large that $\sigma(i) = i$ for all $i \geq n$. We set

$$(2.1) \quad c(\sigma, w) = \text{stable value of the difference } \text{inv}_n(T_\sigma w) - \text{inv}_n(w).$$

For instance, if σ is the elementary transposition of i and $i + 1$, then $T_\sigma w$ differs from w only by transposition of the adjacent letters w_i and w_{i+1} , and then $c(\sigma, w)$ equals 1, -1 or 0, depending on whether $w_i < w_{i+1}$, $w_i > w_{i+1}$ or $w_i = w_{i+1}$, respectively.

The function $c(\sigma, w)$ is an *additive cocycle* in the sense that

$$c(\sigma\tau, w) = c(\sigma, T_\tau w) + c(\tau, w), \quad \sigma, \tau \in \mathfrak{S}_\infty.$$

Equivalently, for $q > 0$,

$$(2.2) \quad \rho_q(\sigma, w) := q^{c(\sigma, w)}$$

is a multiplicative cocycle. In accordance with the terminology of ergodic theory, the additive cocycle $c = \log_q \rho_q$ may be also called the “modular function.”

Our considerations are based on the following definition.

DEFINITION 2.1. For fixed $q > 0$, a Borel probability measure P on \mathbb{A}^∞ is called *q -exchangeable* if P is quasi-invariant with respect to the action of the group \mathfrak{S}_∞ , with the multiplicative cocycle given by (2.2).

Note that it is enough to require that (2.2) holds for the elementary transpositions because these permutations generate the group \mathfrak{S}_∞ . Thus, Definition 2.1 is equivalent to the definition of q -exchangeability given in the [Introduction](#). In the special case $q = 1$, the order on \mathbb{A} plays no role, as the cocycle ρ_q is identically equal to 1, and so our definition becomes conventional exchangeability.

It is important to understand how q -exchangeability behaves under transformations. For $f : \mathbb{A} \rightarrow \mathbb{B}$, let f^∞ denote the induced mapping $\mathbb{A}^\infty \rightarrow \mathbb{B}^\infty$ which replaces each letter w_i in a word by $f(w_i)$. First, consider the identity mapping from $(\mathbb{A}, <)$ to $(\mathbb{A}, >)$.

PROPOSITION 2.2. *If P is a q -exchangeable measure on the space of words over $(\mathbb{A}, <)$, then P is q^{-1} -exchangeable with respect to $(\mathbb{A}, >)$, that is, when the order on the basic space is reversed.*

PROOF. The claim is easily checked for the elementary transpositions which swap i and $i + 1$. \square

It is obvious that if f is an injective morphism of ordered Borel spaces, then f^∞ sends one q -exchangeable measure to another q -exchangeable measure. This applies, in particular, to $\mathbb{A} \subseteq \mathbb{R}$ and a strictly increasing function $f : \mathbb{A} \rightarrow \mathbb{R}$. It is less obvious that q -exchangeability is preserved by arbitrary monotone transformations.

PROPOSITION 2.3. *Let \mathbb{A} and \mathbb{B} be Borel subsets of \mathbb{R} . Suppose $f : \mathbb{A} \rightarrow \mathbb{B}$ is weakly increasing, that is, $a < b$ implies $f(a) \leq f(b)$. The induced Borel map $f^\infty : \mathbb{A}^\infty \rightarrow \mathbb{B}^\infty$ then preserves q -exchangeability.*

This proposition will be reduced to its restricted version involving finite random words and a finite alphabet \mathbb{A} (see Proposition 2.5 below). In the case $q = 1$, the assertion becomes a familiar property of exchangeability, one which holds for arbitrary Borel f .

Definition 2.1 has a straightforward counterpart for *finite* random words $w \in \mathbb{A}^n$. Let \mathfrak{S}_n denote the group of permutations of \mathbb{N}_n . We say that a probability measure P_n on \mathbb{A}^n is *finitely q -exchangeable* if, for each $\sigma \in \mathfrak{S}_n$, the measure $T_\sigma^{-1}P_n$ is equivalent to P_n and the Radon–Nikodym derivative $dT_\sigma^{-1}P_n/dP_n$ is given by the function $q^{\text{inv}(T_\sigma w) - \text{inv}(w)}$. If \mathbb{A} is finite or countable, then P_n is purely atomic and this condition means that, for $w = w_1 \cdots w_n \in \mathbb{A}^n$,

$$(2.3) \quad P_n(T_\sigma w) = q^{\text{inv}(T_\sigma w) - \text{inv}(w)} P_n(w), \quad \sigma \in \mathfrak{S}_n.$$

Consider the canonical projection $\mathbb{A}^\infty \rightarrow \mathbb{A}^n$ assigning to an infinite word $w = w_1 w_2 \cdots$ its n -truncation $w_1 \cdots w_n$, $n = 1, 2, \dots$. Given a probability measure P on \mathbb{A}^∞ , let P_n stand for the push-forward of P under the projection. The following result follows easily from the definitions.

LEMMA 2.4. *A probability measure P on \mathbb{A}^∞ is q -exchangeable if and only if P_n is finitely q -exchangeable for every $n = 1, 2, \dots$*

In principle, the structure of the set of finitely q -exchangeable measures on \mathbb{A}^n is clear: by finiteness of the group \mathfrak{S}_n , every such measure is a unique mixture of the extreme measures and every extreme (i.e., ergodic) measure is supported by a single \mathfrak{S}_n -orbit in \mathbb{A}^n . Moreover, every \mathfrak{S}_n -orbit carries a unique q -exchangeable probability measure, hence the extreme measures are in bijective correspondence with the set of \mathfrak{S}_n -orbits in \mathbb{A}^n . Each \mathfrak{S}_n -orbit in \mathbb{A}^n contains exactly one word $v_1 \cdots v_n \in \mathbb{A}^n$ which is *inversion-free*, that is, which satisfies $v_1 \leq \cdots \leq v_n$. Thus, the collection of inversion-free words of length n parametrizes the orbits of \mathfrak{S}_n and all extreme finitely q -exchangeable measures on \mathbb{A}^n .

We can now state a simplified version of Proposition 2.3.

PROPOSITION 2.5. *Let \mathbb{A} and \mathbb{B} be finite ordered alphabets and let $f : \mathbb{A} \rightarrow \mathbb{B}$ be a weakly increasing map. The induced map $f^n : \mathbb{A}^n \rightarrow \mathbb{B}^n$ then preserves the finite q -exchangeability of measures.*

We first show how to deduce Proposition 2.3 from Proposition 2.5. To this end, let \mathbb{A} , \mathbb{B} and f be as required in Proposition 2.3. Furthermore, let P be a q -exchangeable probability measure on \mathbb{A}^∞ and $f^\infty(P)$ be its push-forward

under f^∞ . Observe that $(f^\infty(P))_n = f^n(P_n)$ for all $n = 1, 2, \dots$. By virtue of Lemma 2.4, it suffices to prove that if a measure P_n on \mathbb{A}^n is finitely q -exchangeable, then so is its push-forward $f^n(P_n)$. This, in turn, shows that it suffices to inspect the particular case of extreme P_n . As pointed out above, every extreme measure P_n is concentrated on a single \mathfrak{S}_n -orbit so that P_n actually lives on words from a finite alphabet. This provides the desired reduction to Proposition 2.5.

PROOF OF PROPOSITION 2.5. Let P_n be a finitely q -exchangeable measure on \mathbb{A}^n and $\tilde{P}_n = f^n(P_n)$ its push-forward on \mathbb{B}^n . Since the alphabets are finite, the measures are purely atomic, supported by finite sets, so we may deal with probabilities of individual words.

It suffices to prove that for every word $u \in \mathbb{B}^n$ and every elementary transposition $\sigma = (i, i + 1)$, one has

$$\tilde{P}_n(u^*) = q^{\text{inv}(u^*) - \text{inv}(u)} \tilde{P}_n(u), \quad u^* := T_\sigma u.$$

Let us fix u and i . There are three possible cases: $u_i = u_{i+1}$, $u_i < u_{i+1}$ and $u_i > u_{i+1}$. In the first case, $u^* = u$ and the desired relation is trivial. By symmetry between the second and third cases, it suffices to examine one of them, say, the second case. Then, $\text{inv}(u^*) - \text{inv}(u) = 1$. Consider the inverse images $X = (f^n)^{-1}(u)$ and $X^* = (f^n)^{-1}(u^*)$. We then have $\tilde{P}_n(u) = P_n(X)$ and $\tilde{P}_n(u^*) = P_n(X^*)$. Thus, we are reduced to showing that

$$P_n(X^*) = q P_n(X).$$

Since f is weakly increasing, $u_i < u_{i+1}$ implies that $w_i < w_{i+1}$ for every $w \in X$, hence $P(T_\sigma w) = q P(w)$. It remains to note that the transformation $T_\sigma : \mathbb{A}^n \rightarrow \mathbb{A}^n$ maps X bijectively onto X^* . This concludes the proof. \square

Another proof will be given at the end of Section 3.

PROPOSITION 2.6. *Let $f : \mathbb{A} \rightarrow \mathbb{B}$ be as in Proposition 2.3. If a probability measure P on \mathbb{A}^∞ is q -exchangeable and extreme, then so is its push-forward $f^\infty(P)$.*

PROOF. By Proposition 2.3, $f^\infty(P)$ is q -exchangeable, hence quasi-invariant under the action of \mathfrak{S}_∞ . Obviously, the map f^∞ commutes with that action. Recall that extremality of quasi-invariant measures is equivalent to their ergodicity, so it suffices to show that $f^\infty(P)$ is ergodic if P is such, but this follows straightforwardly from the definitions. \square

3. The finite q -shuffle. We fix a positive parameter q (later, we will assume that $0 < q < 1$). For a finite permutation $\sigma \in \mathfrak{S}_n$, we denote by $\text{inv}(\sigma)$ the num-

ber of inversions, meaning the number of inversions in the permutation word $\sigma(1) \cdots \sigma(n)$. It is well known that

$$\sum_{\sigma \in \mathfrak{S}_n} q^{\text{inv}(\sigma)} = [n]_q!$$

where

$$[n]_q! := [1]_q [2]_q \cdots [n]_q, \quad [n]_q := \sum_{i=0}^{n-1} q^i$$

[this is a particular case of formula (5.4) below].

DEFINITION 3.1. For $n = 1, 2, \dots$, the *Mallows measure* \mathcal{Q}_n is the probability measure on \mathfrak{S}_n defined by

$$\mathcal{Q}_n(\sigma) = \frac{q^{\text{inv}(\sigma)}}{[n]_q!}, \quad \sigma \in \mathfrak{S}_n.$$

The Mallows measure and its relatives, introduced in [15], have been studied in statistics in the context of ranking problems; see [5, 7] for connections with card shuffling and exclusion processes, and [18] for a scaling limit of \mathcal{Q}_n .

If $q = 1$, then \mathcal{Q}_n is just the uniform measure on \mathfrak{S}_n . Thus, for general $q > 0$, \mathcal{Q}_n may be viewed as a deformation of the uniform measure.

The Mallows measure is the unique finitely q -exchangeable measure supported by the set of permutation words of length n , that is, corresponding to the inversion-free word $12 \cdots n$.

The measure \mathcal{Q}_n can be characterized by means of an important independence property partially mentioned in [15] (at the top of [15], page 125, substitute $q^{-1/2}$ for Mallows' ϕ). First, we need more notation. For $n = 1, 2, \dots$, we denote by $G_{q,n}$ the n -truncated geometric distribution on $\mathbb{N}_n = \{1, \dots, n\}$ with parameter q :

$$G_{q,n}(i) = \frac{q^{i-1}}{[n]_q}, \quad i \in \mathbb{N}_n.$$

For permutation $\sigma \in \mathfrak{S}_n$, written as the word $\sigma(1) \cdots \sigma(n)$, define *backward ranks*

$$(3.1) \quad \beta_j = \beta_j(\sigma) := \#\{i \leq j \mid \sigma(i) \leq \sigma(j)\}, \quad j = 1, \dots, n.$$

For instance, the permutation word 1324 has $\beta_1 = 1, \beta_2 = 2, \beta_3 = 2, \beta_4 = 4$. The correspondence $\sigma \mapsto (\beta_1(\sigma), \dots, \beta_n(\sigma))$ is a well-known bijection between \mathfrak{S}_n and the Cartesian product $\mathbb{N}_1 \times \cdots \times \mathbb{N}_n$.

PROPOSITION 3.2. *The Mallows measure \mathcal{Q}_n is the unique measure on \mathfrak{S}_n under which the backward ranks are independent, with each variable $j - \beta_j + 1$ distributed according to $G_{q,j}$.*

PROOF. Decompose the number of inversions as $\text{inv}(\sigma) = \sum_{j=1}^n (j - \beta_j)$ and multiply probabilities of the truncated geometric distribution to see that \mathcal{Q}_n coincides with the product measure. \square

The following shuffling algorithm is central to our construction of finitely q -exchangeable measures. The procedure is a variation of “absorption sampling” which was studied under various guises in [3, 12, 16].

DEFINITION 3.3. Given an arbitrary finite word $v_1 \cdots v_n$, its q -shuffle is the random word $w_1 \cdots w_n$ obtained by a random permutation of the letters v_1, \dots, v_n , determined by the following n -step algorithm (not to be confused with the notion of a -shuffle with integer parameter a ; see [4, 8, 17]).

Let ξ_1, \dots, ξ_n be independent random variables with ξ_j having distribution $G_{q, n-j+1}$.

At step 1, take for w_1 the ξ_1 th letter from the word $v^{(1)} := v_1 \cdots v_n$. Then, remove the letter v_{ξ_1} from $v^{(1)}$ and denote by $v^{(2)}$ the resulting word of length $n - 1$. Iterate. So, at each following step $m = 2, \dots, n$, there is a word $v^{(m)}$ which was derived from the initial word by deleting some $m - 1$ letters, a new letter $w_m = v_{\xi_m}^{(m)}$ is then chosen and, if $m < n$, the word $v^{(m+1)}$ is obtained by removing this letter from $v^{(m)}$.

PROPOSITION 3.4. Let $v = v_1 \cdots v_n$ be an inversion-free word on the ordered alphabet \mathbb{A} , so $v_1 \leq \cdots \leq v_n$. Let w be the random word obtained from v by the q -shuffle algorithm and let P_n be the distribution of w which is a probability measure concentrated on the \mathfrak{S}_n -orbit of v . Then, P_n is finitely q -exchangeable.

PROOF. First, observe that the probability $P_n(w)$ of any word w from the \mathfrak{S}_n -orbit of v is strictly positive. By the very definition of finite q -exchangeability, it suffices to prove that if σ is an elementary transposition $(i, i + 1)$, $i = 1, \dots, n - 1$, then the ratio $P_n(T_\sigma(w))/P_n(w)$ equals q, q^{-1} or 1, depending on whether $w_i < w_{i+1}$, $w_i > w_{i+1}$ or $w_i = w_{i+1}$, respectively. The latter case being trivial, we may assume, by symmetry, that $w_i < w_{i+1}$.

For $w_1 < w_2$, suppose that a word starts with $w_1 w_2$ and examine the transposition $\sigma = (1, 2)$, which swaps w_1 and w_2 . Let I and J denote the sets of indices i and j for which $v_i = w_1$ and $v_j = w_2$, respectively. If the q -shuffle algorithm results in the word w , then the first chosen letter is v_i for some $i \in I$ and the second chosen letter is v_j for some $j \in J$. Likewise, if the resulting word starts with $w_2 w_1$, then we have to choose first v_j with some $j \in J$ and afterward v_i with some $i \in I$. Let $P_{v_i v_j}$ and $P_{v_j v_i}$ stand for the corresponding probabilities.

If we fix $i \in I$ and $j \in J$, then the word $v^{(3)}$ obtained from the initial word v at the third step of the algorithm does not depend on the order in which v_i and v_j were chosen. Thus, it suffices to prove that $P_{v_i v_j}/P_{v_j v_i} = 1/q$.

The probabilities in question are easily computed. Note that $i < j$ because $v_i < v_j$. It follows that

$$P_{v_i v_j} = G_{q,n}(i)G_{q,n-1}(j - 1) = \frac{q^{i+j-3}}{[n]_q[n - 1]_q}$$

because, after the first step, the letter v_j acquires the number $j - 1$. On the other hand,

$$P_{v_j v_i} = G_{q,n}(j)G_{q,n-1}(i) = \frac{q^{i+j-2}}{[n]_q[n - 1]_q},$$

because now the position of the second letter does not change after the first step. Therefore, the ratio in question is indeed equal to $1/q$.

Finally, transpositions $\sigma = (i, i + 1)$ with $i = 2, 3, \dots$ are handled in the same way, the key point being that each of the words $v^{(2)}, v^{(3)}, \dots$ is inversion-free. \square

REMARK 3.5. Note that the claim of Proposition 3.4 fails if one drops the assumption that v is inversion-free. For instance, if $v_1 \geq \dots \geq v_n$, then the resulting probability measure on the orbit will be q^{-1} -exchangeable and hence not q -exchangeable, except the trivial cases where $v_1 = \dots = v_n$ or $q = 1$.

The connection between Definitions 3.1 and 3.3 is established by the following result.

COROLLARY 3.6. *The q -shuffle, as introduced in Definition 3.3, coincides with the action of the random permutation $\sigma \in \mathfrak{S}_n$, distributed according to the Mallows measure \mathcal{Q}_n .*

PROOF. As seen from the description of the q -shuffle, it actually acts on the positions of the letters rather than on the letters themselves. Thus, it is given by the action of the random permutation $\sigma \in \mathfrak{S}_n$, distributed according to some probability measure \mathcal{Q}'_n on \mathfrak{S}_n , which does not depend on the word to be q -shuffled. Let us identify permutations $\sigma \in \mathfrak{S}_n$ with the corresponding permutation words $\sigma(1) \cdots \sigma(n)$. Then, \mathcal{Q}'_n can be characterized as the outcome of q -shuffling the inversion-free word $v = 1 \cdot 2 \cdots n$. By Proposition 3.4, \mathcal{Q}'_n is a finitely q -exchangeable probability measure concentrated on the \mathfrak{S}_n -orbit of v . Such a measure is unique and the orbit can be identified with the group \mathfrak{S}_n itself. On the other hand, \mathcal{Q}_n is q -exchangeable, thus $\mathcal{Q}'_n = \mathcal{Q}_n$. \square

As yet another application of Proposition 3.4, we obtain an alternative proof of Proposition 2.5.

SECOND PROOF OF PROPOSITION 2.5. We will show that if P_n is an extreme q -exchangeable measure on \mathbb{A}^n , then so is $f^n(P_n)$. This will imply the claim of the proposition.

By Proposition 3.4, P_n is obtained by the q -shuffle applied to an inversion-free word $v \in \mathbb{A}^n$. Therefore, the same holds for the measure $f^n(P_n)$ and the word $f(v) := f(v_1) \cdots f(v_n)$ because the q -shuffle commutes with the map f^n . Since f is weakly increasing, the word $f(v)$ is inversion-free. Again applying Proposition 3.4, we get the desired result. \square

4. The infinite q -shuffle and statement of the main result. The above discussion of finite q -exchangeability can be summarized as follows: the extreme finitely q -exchangeable probability measures are parameterized by finite inversion-free words and can be obtained by application of the q -shuffle procedure to these words. Our aim now is to find a counterpart of this result for measures on infinite words. As in Section 2, we are dealing with an ordered alphabet $(\mathbb{A}, <)$, where \mathbb{A} is a Borel subset of \mathbb{R} . Thus far, the parameter q has been an arbitrary positive number, but:

- throughout the rest of the paper we will assume $0 < q < 1$.

By Proposition 2.2, this restriction does not lead to a loss of generality because the case $q > 1$ is reduced to the case $q < 1$ by inverting the order on the alphabet.

Let $\mathbb{N} = \{1, 2, \dots\}$ and let G_q be the geometric distribution on \mathbb{N} with parameter q :

$$G_q(i) = (1 - q)q^{i-1}, \quad i \in \mathbb{N}.$$

DEFINITION 4.1. Let $v = v_1 v_2 \cdots \in \mathbb{A}^\infty$ be an arbitrary infinite word. The *infinite q -shuffle* of v is the infinite random word $w = w_1 w_2 \cdots$ produced by the algorithm similar to that in Definition 3.3. The only changes are: (i) the independent variables with varying truncated geometric distributions should be replaced by the independent variables ξ_1, ξ_2, \dots with the same geometric distribution G_q ; (ii) the number of steps becomes infinite.

Although the infinite q -shuffle involves countably many steps, the first n letters in the output word w are specified after n steps of the algorithm. This shows, in particular, that the law of the random word w is well defined as a Borel probability measure on \mathbb{A}^∞ .

LEMMA 4.2. *The output random word w is a random permutation of the letters of the input word v . That is, all letters of v appear in w with probability 1.*

PROOF. The probability that the first letter v_1 will not be chosen in the first m steps of the algorithm is equal to q^m . As $m \rightarrow \infty$, this quantity goes to 0 so that v_1 will appear in w with probability 1. Iterating this argument, we arrive at the same conclusion for all other letters. \square

As above, we say that an infinite word $v \in \mathbb{A}^\infty$ is *inversion-free* if it has no inversions, that is, if $v_1 \leq v_2 \leq \dots$.

PROPOSITION 4.3. *If $v \in \mathbb{A}^\infty$ is an inversion-free word, then its q -shuffle produces a q -exchangeable Borel probability measure on \mathbb{A}^∞ .*

PROOF. Let $P^{(v)}$ denote the measure in question. For any $n = 1, 2, \dots$, let $P_n^{(v)}$ be the n th marginal measure of P , as in Lemma 2.4. The same argument as in the proof of Proposition 3.4 shows that each of the measures $P_n^{(v)}$ is q -exchangeable. Consequently, by virtue of Lemma 2.4, $P^{(v)}$ is also q -exchangeable. □

Let \mathfrak{S} stand for the set of *all* permutations (i.e., bijections) of the set \mathbb{N} . We will often identify permutations $\sigma \in \mathfrak{S}$ with the corresponding infinite words $\sigma(1)\sigma(2)\dots \in \mathbb{N}^\infty$. In this way, we get an embedding $\mathfrak{S} \hookrightarrow \mathbb{N}^\infty$. It is easy to check that \mathfrak{S} is a Borel subset of \mathbb{N}^∞ so that one can speak about Borel measures on \mathfrak{S} .

On the other hand, \mathfrak{S} is a group containing \mathfrak{S}_∞ as a proper subgroup. The group \mathfrak{S} acts on \mathbb{A}^∞ in the same way as \mathfrak{S}_∞ does. Namely, if $\sigma \in \mathfrak{S}$ and $w \in \mathbb{A}^\infty$, then $(T_\sigma w)_i = w_{\sigma^{-1}(i)}$.

DEFINITION 4.4. By virtue of Proposition 4.3 and Lemma 4.2, an application of the infinite q -shuffle to the inversion-free word $v = 1 \cdot 2 \cdot \dots \in \mathbb{N}^\infty$ produces a q -exchangeable Borel probability measure on \mathbb{N}^∞ , which is concentrated on the group \mathfrak{S} . We call this measure the *Mallows measure on \mathfrak{S}* and denote it \mathcal{Q} .

REMARK 4.5. In accordance with our definition of the action of permutations on words, the permutation word $\sigma(1)\sigma(2)\dots$ corresponding to an element $\sigma \in \mathfrak{S}$ coincides with $T_{\sigma^{-1}}(1 \cdot 2 \cdot \dots)$ and not with $T_\sigma(1 \cdot 2 \cdot \dots)$. It follows that the infinite q -shuffle of any infinite word coincides with the action on it by the random permutation T_σ with $\sigma \in \mathfrak{S}$ distributed according to the push-forward of \mathcal{Q} under the inversion map $\sigma \mapsto \sigma^{-1}$. However, as will be shown in the [Appendix](#), \mathcal{Q} is actually preserved by this map, so we may simply choose random σ , itself distributed according to the Mallows measure \mathcal{Q} .

Given a word $v \in \mathbb{A}^\infty$, its *support*, denoted $\text{supp}(v)$, is the subset of \mathbb{A} comprised of all distinct letters that appear in v , without regard to their multiplicities. If no assumption on v is made, then $\text{supp}(v)$ may be any finite or countable subset of \mathbb{R} and the letters from $\text{supp}(v)$ may enter v with arbitrary multiplicities, finite or infinite. This is not the case, however, if v is inversion-free, as demonstrated by the following, evident, proposition.

PROPOSITION 4.6. *The inversion-free words $v \in \mathbb{R}^\infty$ belong to one of the following two types, depending on whether the support $\text{supp}(v)$ is finite or infinite:*

(I) *The finite type: $\text{supp}(v)$ is a finite set $a_1 < \dots < a_d$. Then, for each $i = 1, \dots, d - 1$, the letter a_i enters v with a finite nonzero multiplicity l_{a_i} , while the last letter a_d has infinite multiplicity and*

$$v = \underbrace{a_1 \cdots a_1}_{l_{a_1}} \cdots \underbrace{a_{d-1} \cdots a_{d-1}}_{l_{a_{d-1}}} \underbrace{a_d a_d \cdots}_{l_{a_d} = \infty}.$$

(II) *The infinite type: $\text{supp}(v)$ is a countable set $a_1 < a_2 < \dots$. Then, for each $i = 1, 2, \dots$, the letter a_i enters v with a finite nonzero multiplicity l_{a_i} and*

$$v = \underbrace{a_1 \cdots a_1}_{l_{a_1}} \underbrace{a_2 \cdots a_2}_{l_{a_2}} \cdots.$$

For both types, the finite multiplicities l_{a_i} may take arbitrary positive integer values.

For an inversion-free word $v \in \mathbb{R}^\infty$, let $\Omega^{(v)}$ denote its \mathfrak{S} -orbit, $\Omega^{(v)} := \{T_\sigma v \mid \sigma \in \mathfrak{S}\}$, which is a Borel subset in \mathbb{R}^∞ . By the definition, the measure $P^{(v)}$ is concentrated on $\Omega^{(v)}$.

REMARK 4.7. If $\text{supp}(v)$ is finite, then $\Omega^{(v)}$ coincides with the \mathfrak{S}_∞ -orbit of v and hence is countable [except when $\text{supp}(v)$ is a singleton]. Therefore, in this case, the measure $P^{(v)}$ is purely atomic: for $w \in \Omega^{(v)}$, $P^{(v)}(w)$ is proportional to $q^{\text{inv}(w)}$. Note that, here, $\text{inv}(w)$, the total number of inversions in w , is finite. Moreover, the number

$$\mathcal{I}^{(v)}(k) := \#\{w \in \Omega^{(v)} \mid \text{inv}(w) = k\}$$

has polynomial growth in k as $k \rightarrow \infty$ so that the series $\sum_k \mathcal{I}^{(v)}(k)q^k$ converges, which explains why the measure exists. (Note that in the situation of the conventional de Finetti theorem, there are no finite invariant measures supported by a nontrivial \mathfrak{S}_∞ -orbit.) In contrast to that, if $\text{supp}(v)$ is infinite, then $\Omega^{(v)}$ has the cardinality of the continuum and the measure $P^{(v)}$ is diffuse.

We are now in a position to state the main result of the paper.

THEOREM 4.8. *Let \mathbb{A} be an arbitrary Borel subset of \mathbb{R} with order inherited from \mathbb{R} . The extreme q -exchangeable Borel probability measures on \mathbb{A}^∞ are parametrized by the infinite inversion-free words v with support contained in \mathbb{A} . The measure $P^{(v)}$ corresponding to such a word v is obtained by application of the infinite q -shuffle to v , as described in Proposition 4.3.*

Observe that the orbits $\Omega^{(v)}$ with different v 's are pairwise disjoint. It follows that Theorem 4.8 is reduced to the following, seemingly weaker, claim.

PROPOSITION 4.9. For \mathbb{A} as in Theorem 4.8, the extreme q -exchangeable measures on \mathbb{A}^∞ belong to the family of measures $\{P^{(v)}\}$, where v ranges over the set of inversion-free words in \mathbb{A}^∞ .

Indeed, combining this proposition with the above observation, we see that none of the measures in the family $\{P^{(v)}\}$ can be written as nontrivial mixtures of other measures, which implies that each $P^{(v)}$ is extreme. A proof of Proposition 4.9 will be given below.

REMARK 4.10. Given an element $\tau \in \mathfrak{S}$, let $\tilde{\tau} \in \mathbb{N}^\infty$ denote the corresponding permutation word, $\tilde{\tau} = \tau(1)\tau(2)\cdots$. The Mallows measure \mathcal{Q} (Definition 4.4) can be characterized as the only probability measure on the group \mathfrak{S} , which is quasi-invariant under the right shifts $\tau \mapsto \tau\sigma^{-1}$ by elements σ of the subgroup \mathfrak{S}_∞ , with the cocycle $\rho_q(\sigma, \tilde{\tau})$. This follows from Theorem 4.8 and the definition of \mathcal{Q} .

Next, we shall inspect the nature of the random word $w \in \mathbb{A}^\infty$ under $P^{(v)}$. The sequence of truncations $\emptyset, w_1, w_1w_2, \dots$ has transition probabilities described in the following proposition. The notation works as follows: letters a, b range over \mathbb{A} ; l_a is the multiplicity of a in v , as above; $u = w_1 \cdots w_{n-1}$ is a finite word; $\mu_a(u)$ is the multiplicity of a in u .

PROPOSITION 4.11. Let w be the infinite random word distributed according to $P^{(v)}$. The transition probabilities then have the form

$$(4.1) \quad \begin{aligned} P^{(v)}(u \rightarrow ua) &= q^{\sum_{b<a}(l_b - \mu_b(u))} (1 - q^{l_a - \mu_a(u)}) \\ &= q^{\sum_{b<a}(l_b - \mu_b(u))} - q^{\sum_{b \leq a}(l_b - \mu_b(u))}. \end{aligned}$$

PROOF. First, assume that $n = 1$, that is, $u = \emptyset$. The left-hand side of (4.1) is then the probability of $w_1 = a$, as in the first step of the q -shuffling algorithm. The string of a 's in v starts from position $i := 1 + \sum_{b<a} l_b$ and ends at position $j := \sum_{b \leq a} l_b$. Therefore, the probability in question equals

$$(1 - q)(q^{i-1} + \cdots + q^{j-1}) = q^{i-1}(1 - q^{j-i+1}).$$

The same quantity appears in the right-hand side of (4.1) when $u = \emptyset$ because then $\mu_b(u) = 0$ for all $b \in \mathbb{A}$.

For $n = 2, 3, \dots$, the argument is exactly the same, taking into account that we are dealing with the n th step of the algorithm and that the word $v^{(n)}$ is inversion-free, with letter multiplicities $l'_b = l_b - \mu_b(u)$. \square

REMARK 4.12. The following comments are relevant to formula (4.1):

1. If $\mu_a(u) = l_a$, then (4.1) shows that the transition $u \rightarrow ua$ has probability zero. This agrees with the fact that if $l_a < \infty$, then the letter a cannot enter the

random word more than l_a times. In particular, if $l_a = 0$ [which means that $a \notin \text{supp}(v)$], then a never appears.

2. The transition probability $P^{(v)}(u \rightarrow ua)$ depends on u only through the collection of multiplicities $\{\mu_a(u)\}_{a \in \mathbb{A}}$. That is, it depends only on the \mathfrak{S}_n -orbit of u .

3. Recall that the support of v is either of the form $a_1 < \dots < a_d$ or $a_1 < a_2 < \dots$. Let us set

$$x_0(u) = 1, \quad x_i(u) = q^{\sum_{j \leq i} (l_{a_j} - \mu_{a_j}(u))},$$

where $j = 1, \dots, d$ or $j = 1, 2, \dots$ for finite or infinite support, respectively. In this notation, (4.1) can be rewritten as

$$(4.2) \quad P^{(v)}(u \rightarrow ua_i) = x_{i-1}(u) - x_i(u), \quad a_i \in \text{supp}(v).$$

Now, observe that

$$1 = x_0(u) \geq x_1(u) \geq \dots \geq x_d(u) = 0$$

or

$$1 = x_0(u) \geq x_1(u) \geq x_2(u) \geq \dots \geq 0 \quad \text{with} \quad \lim_{i \rightarrow \infty} x_i(u) = 0$$

for finite or infinite support, respectively. This makes evident the fact that the transition probabilities given by (4.2) indeed sum to 1.

4. We have deduced formula (4.1) from the q -shuffling algorithm. Conversely, starting from (4.1), one can easily recover the algorithm itself.

Proposition 4.11 describes the measures $P^{(v)}$ via transition probabilities. The next proposition characterizes $P^{(v)}$ in terms of the marginal measures $P_n^{(v)}$, which are the joint distributions of the first n letters. Note that $P_n^{(v)}$ is a purely atomic measure because it is supported by the words $u = u_1 \dots u_n$ with letters u_i from the finite or countable set $\text{supp}(v)$ and the set of all such words is finite or countable. Thus, we may speak about probabilities $P_n^{(v)}(u)$ of individual words.

We recall some standard q -notation. Let

$$(x; q)_0 = 1, \quad (x; q)_k := \prod_{i=0}^{k-1} (1 - xq^i), \quad k = 1, 2, \dots$$

Likewise, we define $(x; q^{-1})_k$. Below, we use the same notation as in Proposition 4.11.

PROPOSITION 4.13. *Let $v \in \mathbb{R}^\infty$ be an inversion-free word and let u be a word of length n with letters belonging to the support of v . We have*

$$(4.3) \quad P_n^{(v)}(u) = q^{\text{inv}(u)} q^{-\sum_{b < a} \mu_b(u) \mu_a(u)} \prod_a (q^{l_a}; q^{-1})_{\mu_a(u)} q^{\mu_a(u) \sum_{b < a} l_b},$$

where a and b assume values in $\text{supp}(v)$.

Note that the product over $a \in \text{supp}(v)$ is actually finite, even if $\text{supp}(v)$ is infinite. This follows from the fact that $\mu_a(u) = 0$ implies that the corresponding factor equals 1 and that there are only finitely many a 's with $\mu_a(u) \neq 0$.

PROOF. Computing the ratio $P_{n+1}^{(v)}(ua)/P_n^{(v)}(u)$ from (4.3), one sees that the formula agrees with transition probabilities (4.1). \square

5. The case of a finite alphabet. In this section, we prove Proposition 4.9 (and hence Theorem 4.8) for a finite alphabet \mathbb{A} with cardinality $d = \#\mathbb{A} \geq 2$. The simplest case, $d = 2$, was examined in [9] and we will apply here the same method. To be definite, we take $\mathbb{A} = \mathbb{N}_d$. Following the formalism due to Kerov and Vershik [19], it is insightful to interpret q -exchangeability as a property of measures on the path space of a graded graph (Bratteli diagram) which captures the branching of orbits of \mathfrak{S}_n on \mathbb{A}^n as n varies.

Let $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$ and consider the d -dimensional lattice \mathbb{Z}_+^d . The lattice points will be denoted by λ or μ . We write lattice points as vectors $\lambda = (\lambda_1, \dots, \lambda_d)$ in the canonical basis e_1, \dots, e_d and we call $|\lambda| = \lambda_1 + \dots + \lambda_d$ the *degree* of λ . We write $\mu < \lambda$ if $\mu \neq \lambda$ and $\lambda - \mu \in \mathbb{Z}_+^d$; in this case, there is a nondecreasing lattice path connecting μ with λ .

Each λ of degree n corresponds to an inversion-free word,

$$(5.1) \quad v(\lambda) = v_1 \cdots v_n = \underbrace{1 \cdots 1}_{\lambda_1} \underbrace{2 \cdots 2}_{\lambda_2} \cdots \underbrace{d \cdots d}_{\lambda_d},$$

where the letter a does not enter if $\lambda_a = 0$. This correspondence yields a bijection between \mathfrak{S}_n -orbits in \mathbb{A}^n and vectors $\lambda \in \mathbb{Z}_+^d$ of degree n .

DEFINITION 5.1. The q -Pascal pyramid of dimension d , denoted $\Gamma(q, d)$, is the oriented graph with vertex set \mathbb{Z}_+^d and directed edges $(\lambda, \lambda + e_a)$, endowed with weights

$$(5.2) \quad \text{weight}(\lambda, \lambda + e_a) := q^{\lambda_{a+1} + \dots + \lambda_d}, \quad a \in \mathbb{N}_d.$$

Note that $\text{weight}(\lambda, \lambda + e_d) = 1$ for any λ . The n th level of the graph consists of the vertices $\lambda \in \mathbb{Z}_+^d$ with $|\lambda| = n$. Level 0 has a sole *root* vertex $\bar{0} := (0, \dots, 0)$. A *standard path* terminating at λ is a lattice path which connects $\bar{0}$ to λ and is nondecreasing in each coordinate. Similarly, we define an *infinite standard path* in $\Gamma(q, d)$ as an infinite coordinatewise nondecreasing path with initial vertex $\bar{0}$.

Observe that there is a natural bijection between \mathbb{A}^n and standard paths in $\Gamma(q, d)$ of length n . By this bijection, a word $w_1 \cdots w_n$ is mapped to the path

$$\begin{aligned} \mu(\emptyset) &= \bar{0}, & \mu(w_1) &= e_{w_1}, & \mu(w_1 w_2) &= e_{w_1} + e_{w_2}, \dots, \\ \mu(w_1 \cdots w_n) &= e_{w_1} + \dots + e_{w_n}, \end{aligned}$$

where the a th coordinate of the terminal vertex is equal to the multiplicity of the letter a in $w_1 \cdots w_n$. For $n = 1, 2, \dots$, the bijections are consistent and hence define a bijection between \mathbb{A}^∞ and the set of infinite standard paths in $\Gamma(q, d)$: under this bijection, $w_n = a$ means that the n th edge of the path connects a vertex $\mu(w_1 \cdots w_{n-1})$ of degree $n - 1$ with $\mu + e_a$. Fixing the first n vertices of a standard path corresponds to a cylinder $[w_1 \cdots w_n] \subset \mathbb{A}^\infty$. A measure P on \mathbb{A}^∞ translates as a measure on the space of infinite standard paths, with $P([w_1 \cdots w_n])$ being the probability of the corresponding initial path of length n .

DEFINITION 5.2. The *weight* of a standard path with endpoint λ is defined as the product of the weights of the edges comprising the path. Let us say that a probability measure on the path space of $\Gamma(q, d)$ is a *Gibbs measure* if, for every λ , the conditional measure of a standard path terminating at λ is proportional to the weight of this path (in the terminology of Kerov and Vershik [19], such a measure is called “central”).

PROPOSITION 5.3. For $\mathbb{A} = \mathbb{N}_d$, the q -exchangeable measures on \mathbb{A}^∞ correspond bijectively to the Gibbs measures on the space of infinite standard paths in the q -Pascal pyramid $\Gamma(q, d)$.

PROOF. Let $w \in \mathbb{A}^\infty$. Under the correspondence between words and paths, $q^{\text{inv}_n(w)}$ is equal to the weight of the standard path encoded in $w_1 \cdots w_n$, as seen by induction. Indeed, if the finite word $w_1 \cdots w_{n-1}$ corresponds to λ and $w_n = a$ is appended, then the number of inversions increases by $\text{inv}_n(w) - \text{inv}_{n-1}(w) = \lambda_{a+1} + \cdots + \lambda_d$, which is the same quantity that appears in (5.2); we then use the telescoping representation

$$\begin{aligned} \text{inv}_n(w) &= [\text{inv}_n(w) - \text{inv}_{n-1}(w)] + [\text{inv}_{n-1}(w) - \text{inv}_{n-2}(w)] + \cdots \\ &\quad + [\text{inv}_1(w) - 0]. \end{aligned}$$

On the other hand, the words in \mathbb{A}^n that correspond to standard paths with a given endpoint make up a \mathfrak{S}_n -orbit. Thus, we see that the Gibbs condition for fixed n is equivalent to finite q -exchangeability. Since this holds for every n , Lemma 2.4 allows finite q -exchangeability for $n = 1, 2, \dots$ to be translated as the Gibbs property, and conversely. \square

We shall now proceed along the lines of [14]. Denote by $\text{Path}(d)$ the space of all infinite standard paths in $\Gamma(q, d)$. With each $\lambda \in \mathbb{Z}_+^d$, we associate a unique elementary probability measure supported by the finite set of standard paths with endpoint λ . This measure corresponds to an orbital, finitely q -exchangeable probability measure on \mathbb{A}^n . We can understand this measure as a function which assigns to λ value 1 and assigns to each $\mu < \lambda$ the probability that a path passes through μ . The *Martin boundary* of $\Gamma(q, d)$ consists of probability measures on

Path(d) which are representable as weak limits of these elementary measures along a sequence of lattice points with $|\lambda| \rightarrow \infty$. We will prove that under the correspondence of Proposition 5.3, the Martin boundary is exactly the images of the measures $P^{(v)}$, with v ranging over the set of inversion-free words in \mathbb{A}^∞ . By the general theory (see [14]), the Martin boundary contains all extreme Gibbs measures, so this will imply Proposition 4.9.

To determine the boundary, we need to identify all asymptotic regimes for λ which guarantee convergence of the ratios

$$(5.3) \quad \frac{\dim(\mu, \lambda)}{\dim(\lambda)},$$

where $\dim(\lambda) = \dim(\bar{0}, \lambda)$ and $\dim(\mu, \lambda)$ is equal to the sum of weights of all nondecreasing lattice paths connecting μ and λ (the weight of each such path is defined as the product of the weights of its edges). We set $\dim(\mu, \lambda) = 0$ if $\lambda - \mu \notin \mathbb{Z}_+^d$. The ratio (5.3) is the Martin kernel for a certain Markov chain and, by analogy with the Gibbs formalism in statistical physics, $\dim \lambda$ may be called the ‘‘partition function.’’

Recall the notation

$$[0]_q! = 1, \quad [n]_q! = [1]_q [2]_q \cdots [n]_q = \frac{(q; q)_n}{(1 - q)^n}, \quad n = 1, 2, \dots$$

For nonnegative integers n_1, \dots, n_d with $n_1 + \dots + n_d = n$, the number

$$\left[\begin{matrix} n \\ n_1, \dots, n_d \end{matrix} \right]_q := \frac{[n]_q!}{[n_1]_q! \cdots [n_d]_q!} = \frac{(q; q)_n}{(q; q)_{n_1} \cdots (q; q)_{n_d}}$$

is known as the *Gaussian multinomial coefficient*.

LEMMA 5.4. *We have, for $\lambda = (\lambda_1, \dots, \lambda_d)$ and $\mu < \lambda$,*

$$\dim(\lambda) = \left[\begin{matrix} |\lambda| \\ \lambda_1, \dots, \lambda_d \end{matrix} \right]_q, \quad \dim(\mu, \lambda) = q^{N(\mu, \lambda)} \dim(\lambda - \mu),$$

where

$$N(\mu, \lambda) = \sum_{b < a} \lambda_b \mu_a - \sum_{b < a} \mu_b \mu_a.$$

PROOF. Recall that the set of finite standard paths ending at λ is encoded by the words w belonging to the $\mathfrak{S}_{|\lambda|}$ -orbit of the inversion-free word $v(\lambda)$, as defined in (5.1). Let $\{w\}$ stand for the set of these words. MacMahon’s formula for the generating function for the number of inversions in permutations of a multiset (see [2], Theorem 3.6) says, in our notation, that

$$(5.4) \quad \sum_{\{w\}} q^{\text{inv}(w)} = \left[\begin{matrix} |\lambda| \\ \lambda_1, \dots, \lambda_d \end{matrix} \right]_q.$$

This yields the formula for $\dim(\lambda)$. The formula for $\dim(\mu, \lambda)$ with

$$N(\mu, \lambda) = (\lambda_1 - \mu_1)(\mu_2 + \dots + \mu_d) + (\lambda_2 - \mu_2)(\mu_3 + \dots + \mu_d) + \dots + (\lambda_{d-1} - \mu_{d-1})\mu_d$$

follows by counting inversions in the corresponding words, which, in turn, is done by comparing the oriented subgraph rooted at μ with the whole graph $\Gamma(q, d)$. \square

A weakly increasing function $h : \mathbb{N}_d \rightarrow \{0, 1, \dots, \infty\}$ with $h(d) = \infty$ will be called a *height function* on $\mathbb{A} = \mathbb{N}_d$. We also set $h(0) := 0$, where appropriate. There is a natural bijection $h \leftrightarrow v$ between the height functions on \mathbb{N}_d and the inversion-free words in \mathbb{N}_d^∞ ,

$$(5.5) \quad v = \underbrace{1 \cdots 1}_{h(1)} \underbrace{2 \cdots 2}_{h(2)-h(1)} \underbrace{3 \cdots 3}_{h(3)-h(2)} \cdots \underbrace{r \cdots r}_{h(r)-h(r-1)} \underbrace{r+1 \cdots r+1}_{h(r+1)=\infty} \cdots$$

where, for some $0 \leq r < d$, each letter $1 \leq a \leq r$ appears $h(a) - h(a - 1) < \infty$ times (if any), and infinitely many times for $a = r + 1$.

PROPOSITION 5.5. *The Martin boundary of the graph $\Gamma(q, d)$ can be parametrized, in a natural way, by the height functions on \mathbb{N}_d .*

PROOF. Using the identity

$$\frac{(q; q)_n}{(q; q)_{n-m}} = (q^n; q^{-1})_m, \quad n \geq m \geq 0,$$

we derive, from Lemma 5.4 for $\mu \prec \lambda$, $m = |\mu|$ and $n = |\lambda|$, that

$$(5.6) \quad \frac{\dim(\mu, \lambda)}{\dim \lambda} = q^{-\sum_{b < a} \mu_b \mu_a} \frac{(q; q)_{n-m}}{(q; q)_n} \prod_{a=1}^d (q^{\lambda_a}; q^{-1})_{\mu_a} q^{\mu_a \sum_{b: b < a} \lambda_b}.$$

Observe that the constraint $\mu \prec \lambda$ can be removed; indeed, if it is not satisfied, then $\dim(\mu, \lambda) = 0$ and the right-hand side of (5.6) also vanishes because $(q^{\lambda_a}; q^{-1})_{\mu_a} = 0$ for $\lambda_a < \mu_a$.

Let us rewrite (5.6) using the notation

$$h_\lambda(a) := \lambda_1 + \dots + \lambda_a, \quad a = 1, \dots, d, \quad h_\lambda(0) := 0,$$

in the form

$$(5.7) \quad \frac{\dim(\mu, \lambda)}{\dim \lambda} = q^{-\sum_{b < a} \mu_b \mu_a} \frac{(q; q)_{n-m}}{(q; q)_n} \times \prod_{a=1}^d (q^{h_\lambda(a)-h_\lambda(a-1)}; q^{-1})_{\mu_a} q^{\mu_a h_\lambda(a-1)}.$$

It is now easy to analyze the asymptotics of this expression, assuming that μ remains fixed while λ varies so that $n = |\lambda| \rightarrow \infty$. First, note that

$$\lim_{n \rightarrow \infty} \frac{(q; q)_{n-m}}{(q; q)_n} = \frac{(q; q)_\infty}{(q; q)_\infty} = 1.$$

Next, observe that

$$0 \leq h_\lambda(1) \leq \dots \leq h_\lambda(d - 1) \leq h_\lambda(d) = n.$$

Passing to a subsequence, we may assume that there exist finite or infinite limits

$$\lim_{n \rightarrow \infty} h_\lambda(a) = h(a) \in \mathbb{Z}_+ \cup \{+\infty\}, \quad a = 1, \dots, d.$$

This means that there exists $0 \leq r < d$ such that the numbers $h_\lambda(1), \dots, h_\lambda(r)$ stabilize for n large enough, $h_\lambda(a) = h(a) < \infty$ for $1 \leq a \leq r$, while $h_\lambda(a) \rightarrow h(a) = +\infty$ for $a > r$. Note that $h_\lambda(d) = n$ always goes to infinity so that $h(d) = \infty$ in any case.

Clearly, the product in (5.7) up to $a = r$ stabilizes. Next, we have

$$(q^{h_\lambda(r+1)-h_\lambda(r)}; q^{-1})_{\mu_{r+1}} q^{\mu_{r+1}h_\lambda(r)} \rightarrow q^{\mu_{r+1}h(r)},$$

because $q^{h_\lambda(r+1)-h_\lambda(r)} \rightarrow 0$. As for the factors with $a > r + 1$, we have

$$(q^{h_\lambda(a)-h_\lambda(a-1)}; q^{-1})_{\mu_a} q^{\mu_a h_\lambda(a-1)} \rightarrow \delta_{\mu_a, 0}$$

with the Kronecker delta in the right-hand side because $h_\lambda(a - 1) \rightarrow \infty$.

We conclude that the convergence $h_\lambda \rightarrow h$ implies

$$(5.8) \quad \frac{\dim(\mu, \lambda)}{\dim \lambda} \rightarrow q^{-\sum_{b < a} \mu_b \mu_a} \prod_{a=1}^d (q^{h(a)-h(a-1)}; q^{-1})_{\mu_a} q^{\mu_a h(a-1)}$$

with the convention that $h(0) = 0$ and $h(a) - h(a - 1) = 0$ if $h(a) = h(a - 1) = +\infty$. Since, for distinct h , the limits in (5.8) are all distinct, the Martin boundary can indeed be parameterized by the height functions. \square

Observe that if $h(a) = h(a - 1)$, then the limit value (5.8) vanishes unless $\mu_a = 0$. Returning to random words $w = w_1 w_2 \dots \in \mathbb{A}^\infty$, this means that if $h(a) = h(a - 1)$, then the letter a does not occur in w , with probability 1.

PROPOSITION 5.6. *Under the correspondence $h \leftrightarrow v$, the measures on $\text{Path}(d)$ afforded by Proposition 5.5 correspond exactly to the measures $P^{(v)}$, where v ranges over the set of inversion-free words on the alphabet \mathbb{N}_d .*

PROOF. Fix a height function h and let \mathcal{P} be the corresponding Gibbs measure on $\text{Path}(d)$. Next, let P be the measure on \mathbb{N}_d^∞ which corresponds to \mathcal{P} via the bijection of Proposition 5.3. Finally, let $v \in \mathbb{N}_d^\infty$ be the inversion-free word associated with h . We have to prove that $P = P^{(v)}$. To do this, it suffices to check that

$P_n = P_n^{(v)}$ for all n . Let $u \in \mathbb{N}_d^n$. Then, $P_n(u)$ equals $q^{\text{inv}(v)}$ times the right-hand side of (5.8), where we set $\mu_a = \mu_a(u)$. Comparing with (4.3), we see that this coincides with $P_n^{(v)}(u)$. \square

This concludes the proof of Proposition 4.9 in the case of a finite alphabet \mathbb{A} .

6. The case $\mathbb{A} = \mathbb{N}$. In this section, we assume that \mathbb{A} is the countable ordered set $(\mathbb{N}, <)$ of positive integers. Our aim is to prove, for this case, Proposition 4.9 and hence Theorem 4.8.

DEFINITION 6.1. By a *height function* on \mathbb{N} , we mean a map $h: \mathbb{N} \rightarrow \mathbb{Z}_+ \cup \{+\infty\}$ which is weakly increasing [i.e., $h(a) \leq h(b)$ for $a < b$] and satisfies $\lim_{a \rightarrow \infty} h(a) = +\infty$. The set of all height functions on \mathbb{N} will be denoted $H(\mathbb{N})$.

Obviously, setting

$$l_a = h(a) - h(a - 1), \quad a \in \mathbb{N},$$

with the understanding that $h(0) = 0$ and $l_a = 0$ if $h(a) = h(a - 1) = +\infty$, we get a bijection $h \leftrightarrow v$ between $H(\mathbb{N})$ and the set of all inversion-free words $v \in \mathbb{N}^\infty$.

PROOF OF PROPOSITION 4.9 FOR $\mathbb{A} = \mathbb{N}$. Assume that P is an extreme q -exchangeable measure on \mathbb{N}^∞ . We have to show that $P = P^{(v)}$ for some v . The idea is to reduce this claim to the case $\mathbb{A} = \mathbb{N}_d$, which was examined in Section 5, by using Propositions 2.3 and 2.6.

For $d = 1, 2, \dots$ and $a \in \mathbb{N}$, set $f_d(a) = a \wedge d = \min(a, d)$. Clearly, this gives us a weakly increasing map $f_d: \mathbb{N} \rightarrow \mathbb{N}_d$. By Proposition 2.6, $f_d^\infty(P)$ is an extreme q -exchangeable measure on \mathbb{N}_d^∞ . By the results of Section 5, it coincides with some measure $P^{(v(d))}$, where $v(d) \in \mathbb{N}_d^\infty$ is an inversion-free word. Denote by h_d the corresponding height function on \mathbb{N}_d .

Let $w \in \mathbb{N}^\infty$ be the random word with law P . For each $a = 1, \dots, d - 1$, the letter a enters the random word $f_d(w)$ exactly $h_d(a) - h_d(a - 1)$ times, with probability 1. Since the map f_d does not change the letters $a = 1, \dots, d - 1$, the same holds for the initial random word w . This implies that $h_d(a) = h_{d+1}(a)$ for all $a = 1, \dots, d - 1$. Therefore, for every $a \in \mathbb{N}$, the value $h_d(a)$ stabilizes as $d \rightarrow \infty$, starting from $d = a + 1$; denote by $h(a)$ this stable value. We claim that h is a height function on \mathbb{N} . Indeed, it is obvious that h weakly increases, so we only have to check that $h(a) \rightarrow \infty$ as $a \rightarrow \infty$. If this were not the case, then $h(a)$ would assume the same (finite) value for all a large enough. However, this would mean that w contained only finitely many letters, each with a prescribed finite multiplicity $l_a = h(a) - h(a - 1)$, which is clearly impossible. Thus, h should be a height function.

Now, let $v \in \mathbb{N}^\infty$ be the inversion-free word corresponding to h . By the definition of h , we have $f_d^\infty(P) = f_d^\infty(P^{(v)})$ for all d . Clearly, this implies $P_n = P_n^{(v)}$ for all n , so $P = P^{(v)}$, as desired. \square

REMARK 6.2. An alternative proof can be based on the notion of the q -Pascal pyramid of dimension ∞ , denoted $\Gamma(q, \infty)$, which is the graph with the vertex set

$$\{\lambda \in \mathbb{Z}_+^\infty \mid \lambda_1 + \lambda_2 + \dots < +\infty\},$$

the edges $(\lambda, \lambda + e_a)$, where

$$e_a = (\underbrace{0, \dots, 0}_{a-1}, 1, 0, 0, \dots), \quad a \in \mathbb{N},$$

and the weight $q^{\sum_{b>a} \lambda_b}$ assigned to the edge $(\lambda, \lambda + e_a)$. Note that the sum in the exponent is finite because $|\lambda| := \sum_a \lambda_a$ is finite, by the definition of $\Gamma(q, \infty)$. The n th level of $\Gamma(q, \infty)$ consists of vertices with $|\lambda| = n$.

The graph $\Gamma(q, d)$ is embedded in $\Gamma(q, \infty)$ as the set of vertices with $\lambda_b = 0$ for $b > d$. Obviously, $\Gamma(q, \infty) = \bigcup_{d \geq 1} \Gamma(q, d)$. The definition of Gibbs measures on the space of standard paths in $\Gamma(q, \infty)$ and the correspondence with q -exchangeable measures on \mathbb{N}^∞ straightforwardly extend the definitions from Section 5. One can then repeat the arguments in Proposition 5.5 to show that the Martin boundary of $\Gamma(q, \infty)$ consists precisely of the Gibbs measures corresponding to measures $P^{(v)}$.

7. The case $\mathbb{A} = \mathbb{R}$. Here, we prove Proposition 4.9 and hence Theorem 4.8 for $\mathbb{A} = \mathbb{R}$. This will also cover the seemingly more general case where \mathbb{A} is an arbitrary Borel subset of $(\mathbb{R}, <)$.

Assume that the measure P on \mathbb{R}^∞ is q -exchangeable and extreme. Our aim is to show that there exists a finite or countable subset $A \subset \mathbb{R}$, of the form $a_1 < \dots < a_d$ or $a_1 < a_2 < \dots$, such that P is supported by A^∞ . The results of Sections 5 and 6 will then imply that $P = P^{(v)}$ for some inversion-free word v .

For an arbitrary word $w \in \mathbb{R}^\infty$, set $h_w(x) := \#\{j : w_j \leq x\}$. The function $h_w : \mathbb{R} \rightarrow \mathbb{Z}_+ \cup \{+\infty\}$ is weakly increasing and right-continuous, hence it is completely determined by its restriction to the set \mathbb{Q} of rational numbers.

For $x \in \mathbb{R}$, let $\phi_x : \mathbb{R}^\infty \rightarrow \{1, 2\}^\infty$ be the mapping which replaces each $w_j \in (-\infty, x]$ by 1 and each $w_j \in (x, +\infty)$ by 2. The measure $\phi_x^\infty(P)$ on $\{1, 2\}^\infty$ is q -exchangeable and extreme, by virtue of Proposition 2.6. Since $h_w(x)$ is the number of 1's in $\phi_x(w)$, the ergodicity implies that the value $h_w(x)$ is the same for P -almost all words w . Letting x run over \mathbb{Q} , we see that, outside a P -null set of words, the value $h_w(x)$ does not depend on w for each $x \in \mathbb{R}$; we denote by $h(x)$ this common value. The function $h(x)$ is again weakly increasing and right-continuous, and it assumes values in $\mathbb{Z}_+ \cup \{+\infty\}$.

Recall that in the $d = 2$ case, q -exchangeability implies the dichotomy that either 1 appears finitely many times and 2 appears infinitely often, or 2 does not appear at all. From this, $h(x) \equiv \infty$ would imply $w_j \leq x$ for all j , which is impossible. It follows that $h(x)$ cannot be identically equal to $+\infty$.

By a similar argument, $h(x)$ also cannot be identically equal to a finite constant.

Defining A to be the set of the jump points of h , we see that A is either a nonempty finite set $a_1 < \dots < a_d$ or a countably infinite set of the form $a_1 < a_2 < \dots$. In the latter case, we set $a^* = \sup\{a_i\} = \lim a_i \in \mathbb{R} \cup \{+\infty\}$. By the definition of $h(x)$, the function is constant on every interval of the form

$$(-\infty, a_1), \quad [a_{i-1}, a_i), \quad [a^*, +\infty).$$

Finally, observe that if one ignores the P -null set of words mentioned above, then any word w does not contain letters from the open intervals

$$(-\infty, a_1), \quad (a_{i-1}, a_i), \quad (a^*, +\infty).$$

We conclude that P is concentrated on A^∞ .

REMARK 7.1. We note, in passing, that this argument fails for more general ordered spaces. For instance, it cannot be applied to \mathbb{R}^k ($k > 1$) with lexicographic order because the order is not separable and h cannot be determined by its restriction to a countable set.

8. Quantization. A motivation for studying the q -exchangeability is that this property can be viewed as a quantization of conventional exchangeability. We comment briefly on this connection.

In the classical setting, each extreme exchangeable P on \mathbb{R}^∞ is of the form $\nu^{\otimes \infty}$, where ν is the limit of empirical measures, meaning that for every Borel $B \subset \mathbb{R}$, as $n \rightarrow \infty$, the random word satisfies the strong law of large numbers

$$(8.1) \quad \#\{j \leq n \mid w_j \in B\} \sim n\nu(B) \quad P\text{-a.s.}$$

Trivially, $0 < P(w_1 \in B) < 1$ if and only if $0 < \nu(B) < 1$, in which case letters from A appear in w infinitely many times for both $A = B$ and $A = B^c := \mathbb{R} \setminus B$.

In the framework of q -exchangeability (with $q < 1$), the analog of (8.1) is

$$(8.2) \quad \#\{j \leq n \mid w_j \in B\} \rightarrow \nu_q(B) \quad P\text{-a.s.},$$

where ν_q is a counting measure associated with some height function h , so the letters from B are represented in w exactly $\nu_q(B)$ times. Similarly to the above, one sees, from the formula

$$P(w_1 \in B) = \sum_{\{x \in B \mid \nu_q\{x\} > 0\}} q^{\nu_q(-\infty, x)} (1 - q^{\nu_q\{x\}}),$$

that $0 < P(w_1 \in B) < 1$ if and only if $\nu_q(B) > 0$ and $\nu_q(B^c) > 0$.

There are many ways to approach exchangeability via q -exchangeability, that is, to obtain independent sampling in the classical limit $q \rightarrow 1$. One possible explicit realization of such a limit is the following quantization of homogeneous product measures.

Let ν be a probability measure on \mathbb{R} with distribution function $F(x) := \nu(-\infty, x]$. Let $F^{-1}(p) := \inf\{x \in \mathbb{R} : F(x) \geq p\}$ be the corresponding quantile function and consider the countable collection of quantiles $\alpha_k := F^{-1}(1 - q^k)$, $k \in \mathbb{N}$, as letters of the inversion-free word $\nu := \alpha_1\alpha_2 \dots$. The idea is to create a bridge between independent sampling from ν and the q -shuffle for the counting measure $\nu_q = \sum_{j \in \mathbb{N}} \delta_{\alpha_j}$ by means of independent sampling from the measures

$$\tilde{\nu}_q = \sum_{k \in \mathbb{N}} G_q(k) \delta_{\alpha_k}.$$

PROPOSITION 8.1. *As $q \rightarrow 1$, for $\nu = \alpha_1\alpha_2 \dots$, the q -shuffle measures $P^{(\nu)}$ converge, in the sense of weak convergence of the finite-dimensional marginal measures $P_n^{(\nu)}$, $n \in \mathbb{N}$, to the product measure $\nu^{\otimes \infty}$.*

PROOF. For ξ a random variable with geometric distribution G_q , the distribution of randomized quantile α_ξ is $\tilde{\nu}_q$. It is convenient to introduce two more random variables: ζ with uniform distribution on $[0, 1]$ and ζ_q with the discrete distribution

$$(8.3) \quad \sum_{k \in \mathbb{N}} G_q(k) \delta_{1-q^k}.$$

From standard properties of the quantile function, the distribution of $F^{-1}(\zeta)$ is ν and the distribution of $F^{-1}(\zeta_q)$ is $\tilde{\nu}_q$, so we can identify $\alpha_\xi = F^{-1}(\zeta_q)$.

Now, the measure (8.3) was designed so that the mass of each interval $[0, 1 - q^k]$ is $1 - q^k$ and the largest atom has mass $1 - q$, which approaches 0 as $q \rightarrow 1$. Therefore, ζ_q converges in distribution to ζ . On the other hand, the set of discontinuities of the quantile function is at most countable and so has Lebesgue measure zero, hence F^{-1} preserves the convergence relation (see, e.g., [6], Theorem 5.1), meaning that $F^{-1}(\zeta_q) \rightarrow_d F^{-1}(\zeta)$. The latter is the same as

$$\mathbb{P}(\alpha_\xi \leq x) \rightarrow F(x) \quad \text{as } q \rightarrow 1,$$

where x is an arbitrary continuity point of F . For any nonnegative integer m , the total variation distance between ξ and the shift $\xi + m$ equals $1 - q^m$, from which the above can be strengthened as

$$\mathbb{P}(\alpha_{\xi+m} \leq x) \rightarrow F(x) \quad \text{as } q \rightarrow 1.$$

Likewise, if ξ_1, ξ_2, \dots are independent copies of ξ and m_1, \dots, m_n are arbitrary fixed nonnegative integers, then we have

$$\mathbb{P}(\alpha_{\xi_1+m_1} \leq x_1, \dots, \alpha_{\xi_n+m_n} \leq x_n) \rightarrow F(x_1) \cdots F(x_n) \quad \text{as } q \rightarrow 1,$$

where x_1, \dots, x_n are arbitrary continuity points of F .

Let $w_1 w_2 \dots$ be the q -shuffle of $1 \cdot 2 \dots$, constructed from the independent geometric ξ_1, ξ_2, \dots , as in Definition 4.1. It easily follows from the definition that $\xi_j \leq w_j < \xi_j + j$, whence the above implies

$$\mathbb{P}(\alpha_{w_1} \leq x_1, \dots, \alpha_{w_n} \leq x_n) \rightarrow F(x_1) \cdots F(x_n) \quad \text{as } q \rightarrow 1$$

for continuity points x_1, \dots, x_n , which is precisely the property of weak convergence of $P_n^{(v)}$ which we wanted to prove. \square

This construction provides quantization of homogeneous product measures on \mathbb{R}^∞ . Extension to the general exchangeable case is straightforward in the light of de Finetti’s theorem: we simply randomize ν .

9. Random flags over a Galois field. Fix $q \in (0, 1)$ and set $\tilde{q} = q^{-1}$ so that $\tilde{q} > 1$. In this section, we assume that \tilde{q} is a power of a prime number.

Let $\mathbb{F}_{\tilde{q}}$ be the Galois field with \tilde{q} elements and let V_∞ be an infinite-dimensional vector space over $\mathbb{F}_{\tilde{q}}$ with a countable basis $\{v_1, v_2, \dots\}$. Defining V_n to be the linear span of vectors v_1, \dots, v_n , we have $\bigcup_{n \geq 1} V_n = V_\infty$, so each element of V_∞ can be uniquely written in the basis as an infinite vector with finitely many nonzero components.

For $d \in \mathbb{N}$, by a *decreasing d -flag* in V_∞ , we shall mean a $(d + 1)$ -tuple $X = (X(i))$ of linear subspaces in V_∞ such that

$$V_\infty = X(0) \supseteq X(1) \supseteq \dots \supseteq X(d - 1) \supseteq X(d) = \{0\}.$$

Keep in mind that our definition disagrees with the conventional notion of a flag, in that the inclusions are not necessarily strict. In the same way, we define decreasing d -flags in each space V_n . Let $\mathbf{X}_d(V_\infty)$ and $\mathbf{X}_d(V_n)$ denote the sets of the decreasing d -flags in V_∞ and V_n , respectively.

LEMMA 9.1. *One can identify $\mathbf{X}_d(V_\infty)$ with the projective limit space $\varprojlim \mathbf{X}_d(V_n)$, where the projection $\mathbf{X}_d(V_{n+1}) \rightarrow \mathbf{X}_d(V_n)$ is determined by taking the intersection with V_n .*

PROOF. Indeed, the map $\mathbf{X}_d(V_\infty) \rightarrow \varprojlim \mathbf{X}_d(V_n)$ is defined by assigning to a flag $X = (X(i))$ in V_∞ the sequence $\{X_n \in \mathbf{X}_d(V_n)\}$ of flags with $X_n(i) = X(i) \cap V_n$. Clearly, the flags X_n are consistent with the projections $\mathbf{X}_d(V_{n+1}) \rightarrow \mathbf{X}_d(V_n)$ and hence determine an element of the projective limit space. The inverse map assigns to any such sequence $\{X_n\}$ the flag $X \in \mathbf{X}_d(V_\infty)$ with $X(i) = \bigcup X_n(i)$. \square

Using the lemma, we endow $\mathbf{X}_d(V_\infty)$ with the topology of projective limit. In other words, a small neighborhood of a flag $X = (X(i))$ is formed by the flags

$Y = (Y(i))$ such that $X(i) \cap V_n = Y(i) \cap V_n$ for all i and some fixed large n . We will consider the σ -algebra of Borel sets in $\mathbf{X}_d(V_\infty)$ relative to this topology.

Let \mathcal{G}_n be the group of all invertible linear transformations of the space V_∞ that leave V_n invariant and fix the basis vectors v_{n+1}, v_{n+2}, \dots . We then have $\{e\} = \mathcal{G}_0 \subset \mathcal{G}_1 \subset \mathcal{G}_2 \subset \dots$ and we define $\mathcal{G}_\infty := \bigcup_{n \geq 1} \mathcal{G}_n$. The group \mathcal{G}_n is finite and isomorphic to the group $GL(n, \mathbb{F}_{\tilde{q}})$ of invertible $n \times n$ matrices over $\mathbb{F}_{\tilde{q}}$. The countable group \mathcal{G}_∞ is isomorphic to the group $GL(\infty, \mathbb{F}_{\tilde{q}})$ of infinite invertible matrices (g_{ij}) , such that $g_{ij} = \delta_{ij}$ for large enough $i + j$.

The group \mathcal{G}_n acts, in a natural way, on $\mathbf{X}_d(V_n)$ and the group \mathcal{G}_∞ acts on $\mathbf{X}_d(V_\infty)$ by continuous transformations. The next proposition is an extension of [9], Lemma 5.2.

PROPOSITION 9.2. *There exists a natural bijection $P \leftrightarrow \mathcal{P}$ between q -exchangeable Borel probability measures on \mathbb{N}_d^∞ and \mathcal{G}_∞ -invariant Borel probability measures on $\mathbf{X}_d(V_\infty)$.*

PROOF. The desired bijection is constructed by understanding P as a Gibbs measure on the path space $\text{Path}(d)$ of the q -Pascal pyramid $\Gamma(q, d)$, as defined in Section 5.

We assign to P a function $\varphi(\lambda)$ on the vertices in the following way. Given a vertex $\lambda \in \Gamma(q, d)$, the probability of a finite path ending at λ equals the weight of the path times a quantity that (for given P) depends only on λ ; let us denote this quantity $\varphi(\lambda)$.

The Gibbs measure is uniquely determined by this function φ , which must satisfy the rule of addition of probabilities along the path

$$(9.1) \quad \varphi(\lambda) = \sum_{a=1}^d \text{weight}(\lambda, \lambda + e_a) \varphi(\lambda + e_a)$$

for all $\lambda \in \mathbb{Z}_+^d$, where the weight of the edge $(\lambda, \lambda + e_w)$ is specified in (5.2) as

$$(9.2) \quad \text{weight}(\lambda, \lambda + e_a) = q^k \quad \text{for } k = \lambda_{a+1} + \dots + \lambda_d.$$

One must also add the normalization condition $\varphi(\vec{0}) = 1$, which implies that

$$(9.3) \quad \sum_{\lambda \in \mathbb{Z}_+^d : |\lambda|=n} \dim(\lambda) \varphi(\lambda) = 1, \quad n = 1, 2, \dots,$$

so that $\dim(\lambda) \varphi(\lambda)$ is the probability that a random walk on $\Gamma(q, d)$ driven by P ever visits λ .

Conversely, if a nonnegative function φ satisfies (9.1) and the normalization condition, then it defines a Gibbs measure. Such functions φ play a central role in the work of Kerov and Vershik (see, e.g., [19]), who call them ‘‘harmonic.’’ However, this terminology is unfortunate as it disagrees with the conventional concept of a harmonic function in the literature on Markov processes.

We now wish to show that precisely the same functions are associated with \mathcal{G}_∞ -invariant measures. Indeed, there is a one-to-one correspondence between \mathcal{G}_∞ -invariant probability measures \mathcal{P} on $\mathbf{X}_d(V_\infty)$ and sequences $\{\mathcal{P}_n\}$ of probability measures such that each \mathcal{P}_n is a measure on $\mathbf{X}_d(V_n)$, invariant under \mathcal{G}_n , and various \mathcal{P}_n 's are consistent with respect to the projections $\mathbf{X}_d(V_{n+1}) \rightarrow \mathbf{X}_d(V_n)$. Specifically, the correspondence is established by letting \mathcal{P}_n be the push-forward of \mathcal{P} under the projection $\mathbf{X}_d(V_\infty) \rightarrow \mathbf{X}_d(V_n)$.

Observe that the \mathcal{G}_n -orbit of a d -flag $X_n = (X_n(i)) \in \mathbf{X}_d(V_n)$ is uniquely determined by the d -tuple of nonnegative integers

$$\lambda_i = \dim V_n(i - 1) - \dim V_n(i), \quad i = 1, \dots, d,$$

which determine a vector $\lambda \in \mathbb{Z}_+^d$ with $|\lambda| = n$. The reader needs to be warned that the dimension of a linear space over \mathbb{F}_q in this formula and below should not be confused with the combinatorial dimension function in the Pascal pyramid, as, for instance, in (9.3). We will say that the vertex λ is the *type* of the flag. Conversely, every such λ corresponds to an orbit. Let $\psi(\lambda)$ be the mass that \mathcal{P}_n gives to each of the flags of type λ . The consistency of the measures \mathcal{P}_n with respect to the projections means that

$$(9.4) \quad \psi(\lambda) = \sum_{a=1}^d \text{weight}'(\lambda, \lambda + e_a) \psi(\lambda + e_a), \quad \lambda \in \mathbb{Z}_+^d,$$

where $\text{weight}'(\lambda, \lambda + e_a)$ stands for the number of flags $X_{n+1} \in \mathbf{X}_d(V_{n+1})$ of type $\lambda + e_a$ projecting onto any fixed flag $X_n \in \mathbf{X}_d(V_n)$ of type λ . Conversely, each function $\psi(\lambda) \geq 0$ satisfying (9.4) and the normalization condition $\psi(\bar{0}) = 1$ determines a consistent sequence $\{\mathcal{P}_n\}$ and hence a \mathcal{G}_∞ -invariant probability measure \mathcal{P} on $\mathbf{X}_d(V_\infty)$.

We claim that

$$\text{weight}'(\lambda, \lambda + e_a) = \tilde{q}^{n-k} = q^{k-n},$$

where k is the same as in (9.2), that is, $k = \dim X_n(a)$. Indeed, if a flag X_{n+1} is projected onto X_n , then it has type $\lambda + e_a$ if and only if

$$\dim X_{n+1}(i) = \dim X_n(i) + 1 \quad \text{for } 0 \leq i \leq a - 1$$

and

$$\dim X_{n+1}(j) = \dim X_n(j) \quad \text{for } a \leq j \leq d.$$

This means that there exists a nonzero vector $v \in V_{n+1} \setminus V_n$ such that, for every $i = 0, \dots, a - 1$, the subspace $X_{n+1}(i)$ is spanned by $X_n(i)$ and v . Such a vector is uniquely defined up to a scalar multiple and addition of an arbitrary vector from $X_n(a)$. Therefore, the number of options is equal to the number of lines in $V_{n+1}/X_n(a)$ not contained in $V_n/X_n(a)$, which equals

$$\frac{\tilde{q}^{n+1-k} - 1}{\tilde{q} - 1} - \frac{\tilde{q}^{n-k} - 1}{\tilde{q} - 1} = \tilde{q}^{n-k}.$$

Viewing equations (9.1) and (9.4) as recursions on φ , respectively, ψ , we see that they are similar, with the coefficients related as

$$\text{weight}'(\lambda, \lambda + e_a) = \text{weight}(\lambda, \lambda + e_a)q^{-n}, \quad n = |\lambda|.$$

Setting

$$\varphi(\lambda) = q^{n(n-1)/2}\psi(\lambda)$$

yields an isomorphism $\{\varphi\} \leftrightarrow \{\psi\}$ between the convex compact sets of nonnegative solutions to (9.1) and (9.4), respectively. Also, note that the above relation does not affect the normalization condition. This completes the proof. \square

REMARK 9.3. By virtue of the isomorphism in Proposition 9.2, the extreme measures P correspond bijectively to extreme measures \mathcal{P} .

REMARK 9.4. Define a *decreasing* \mathbb{N} -flag in V_∞ as an infinite collection $X = (X(i))$ of subspaces such that

$$V_\infty = X(0) \supseteq X(1) \supseteq \dots, \quad \bigcap_{i \in \mathbb{N}} X(i) = \{0\}.$$

The result of Proposition 9.2 remains true when \mathbb{N}_d is replaced by \mathbb{N} . That is, q -exchangeable probability measures on \mathbb{N}^∞ correspond bijectively to \mathcal{G}_∞ -invariant probability measures on the space of decreasing \mathbb{N} -flags. The proof is identical, except with $\Gamma(q, d)$ replaced by $\Gamma(q, \infty)$.

REMARK 9.5. Let V^∞ be the dual vector space to V_∞ . We endow V^∞ with the topology of simple convergence of linear functionals; it then becomes a compact topological space. As an additive group, V^∞ is also the Pontryagin dual of V_∞ , viewed as a discrete additive group. Passing to the orthogonal complement establishes a bijection between arbitrary linear subspaces in V_∞ and *closed* linear subspaces in V^∞ . Define an *increasing* d -flag in V^∞ as a collection of closed subspaces

$$\{0\} = Y(0) \subseteq Y(1) \subseteq \dots \subseteq Y(d) = V^\infty$$

and an *increasing* \mathbb{N} -flag in V^∞ as an infinite collection of closed subspaces

$$\{0\} = Y(0) \subseteq Y(1) \subseteq \dots, \quad \overline{\bigcup_{i \in \mathbb{N}} Y(i)} = V^\infty,$$

where the horizontal line indicates closure. By duality, the increasing d -flags in V^∞ are in one-to-one correspondence with the decreasing d -flags in V_∞ . Moreover, this correspondence is consistent with the natural action of the group \mathcal{G}_∞ on V^∞ . The same also holds for \mathbb{N} -flags. Thus, instead of considering invariant measures on decreasing flags in V_∞ , one can equally well deal with invariant measures on the set of increasing flags in V^∞ .

APPENDIX: THE MALLOWS MEASURE

In this Appendix, we sketch some properties of the Mallows measures \mathcal{Q}_n and \mathcal{Q} . To state the results, we need some preparation. It is convenient to represent a generic permutation $\sigma \in \mathfrak{S}_n$ as an $n \times n$ permutation matrix $\sigma(i, j)$, where the entry $\sigma(i, j)$ equals 1 or 0, depending on whether or not $\sigma(j) = i$. Such permutation matrices are *strictly monomial*, in the sense that they have one and only one nonzero element per row and per column. Note that this realization of permutations by strictly monomial matrices takes the group multiplication into conventional matrix multiplication and the inversion map $\sigma \mapsto \sigma^{-1}$ corresponds to matrix transposition. Likewise, the group \mathfrak{S} can be realized as the group of strictly monomial matrices of infinite size.

More generally, a 0–1 matrix of finite or infinite size is *weakly monomial* if each row and each column contains *at most* one 1, the other entries being 0's. Let $M(n)$ and M denote the sets of weakly monomial 0–1 matrices of size $n \times n$ and $\infty \times \infty$, respectively. Both $M(n)$ and M are semigroups under matrix multiplication and $\mathfrak{S}_n \subset M(n)$ and $\mathfrak{S} \subset M$ are respective subgroups of invertible elements. An additional operation in $M(n)$ and M is matrix transposition, which is an involutive antiautomorphism.

For $k = 1, 2, \dots$, the *truncation operation* θ_k assigns to a matrix of size $\infty \times \infty$ or $l \times l$ with $l \geq k$ the $k \times k$ submatrix comprised of the entries (i, j) with $i, j \leq k$. Obviously, θ_k projects $M(n)$ onto $M(k)$ for any $n > k$. Likewise, θ_k projects M onto $M(k)$. Using these projections, we may identify M with the projective limit space $\varprojlim M(k)$. We endow M with the corresponding projective limit topology; M then becomes a compact topological space. By definition, a fundamental system of neighborhoods of a matrix $m \in M$ is formed by the subsets $\{m' \in M \mid \theta_k(m') = \theta_k(m)\}$, $k = 1, 2, \dots$.

It is readily checked that the restriction of $\theta_k : M \rightarrow M(k)$ to the subset $\mathfrak{S} \subset M$ is surjective for every k . It follows that \mathfrak{S} is dense in M (and even \mathfrak{S}_∞ is dense). Recall that we have endowed \mathfrak{S} with the σ -algebra of Borel sets inherited via the embedding $\mathfrak{S} \subset \mathbb{N}^\infty$. Clearly, this Borel structure coincides with that induced by the embedding $\mathfrak{S} \subset M$. Thus, any Borel probability measure on \mathfrak{S} or on $\mathfrak{S}_n \subset \mathfrak{S}$ can be viewed as a measure on M (here, we identify \mathfrak{S}_n with the subgroup in \mathfrak{S} fixing all integers from $\mathbb{N} \setminus \mathbb{N}_n$). In particular, we may view the Mallows measures \mathcal{Q}_n and \mathcal{Q} as probability measures on the compact space M . This makes sense of the following assertion.

PROPOSITION A.1. *As $n \rightarrow \infty$, \mathcal{Q}_n weakly converge to \mathcal{Q} .*

PROOF. Let $\theta_k(\mathcal{Q}_n)$ and $\theta_k(\mathcal{Q})$ denote the respective push-forwards of \mathcal{Q}_n and \mathcal{Q} under θ_k . By the definition of the topology on M and the finiteness of $M(k)$, it suffices to prove that for any k and any fixed matrix $m \in M(k)$, $\theta_k(\mathcal{Q}_n)(\{m\})$ converges to $\theta_k(\mathcal{Q})(\{m\})$.

Taking into account Remark 4.5, it is convenient to replace \mathcal{Q}_n and \mathcal{Q} by their respective push-forwards under the matrix transposition; let us denote them as \mathcal{Q}'_n and \mathcal{Q}' , respectively. Thus, we will prove the equivalent assertion that $\theta_k(\mathcal{Q}'_n)(\{m\})$ converge to $\theta_k(\mathcal{Q})(\{m\})$.

Let $w = w_1 w_2 \cdots$ be the output of the q -shuffling algorithm applied to the infinite word $1 \cdot 2 \cdots$. As usual, we identify w with the random permutation $\sigma \in \mathfrak{S}$ by writing $w = \sigma(1)\sigma(2) \cdots$. From this, one sees that the quantity $\theta_k(\mathcal{Q}')(\{m\})$ is equal to the probability of the event that for each $j = 1, \dots, k$, the letter w_j either equals some $i \in \{1, \dots, k\}$ if the matrix m has 1 in the j th column in position (i, j) , or $w_j > k$ if the j th column of m consists entirely of 0's.

For instance, if $m = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \in M(2)$, then the event in question is that the first step of the algorithm yields $w_1 > 2$ and the second step yields $w_2 = 1$.

The quantity $\theta_k(\mathcal{Q}'_n)(\{m\})$ admits exactly the same interpretation in terms of the finite q -shuffle applied to the finite word $1 \cdots n$.

Now, the desired convergence of the probabilities follows from the fact that as $n \rightarrow \infty$, the truncated geometric distributions directing the finite q -shuffle (Definition 3.3) converge to the infinite geometric distribution directing the infinite q -shuffle (Definition 4.1). \square

COROLLARY A.2. *The Mallows measures \mathcal{Q}_n and \mathcal{Q} are invariant under the group inversion map $\sigma \mapsto \sigma^{-1}$.*

PROOF. Given a matrix $m \in M(n)$, let us say that two distinct positions $\{(i_1, j_1), (i_2, j_2)\}$ occupied by 1's are *in inversion* if the two differences $i_1 - i_2$ and $j_1 - j_2$ have opposite signs (note that these differences cannot vanish) and denote by $\text{inv}(m)$ the total number of unordered pairs of positions in inversion. Clearly, $\text{inv}(m) = \text{inv}(m')$, where m' stands for the transposed matrix.

On the other hand, if $\sigma \in \mathfrak{S}_n$ and $m := [\sigma(i, j)]$ is the corresponding permutation matrix, then we obviously have $\text{inv}(\sigma) = \text{inv}(m)$. If σ is replaced by σ^{-1} , then m is replaced by m' . Therefore, $\text{inv}(\sigma) = \text{inv}(\sigma^{-1})$, which implies the desired symmetry property of \mathcal{Q}_n . The analogous property for \mathcal{Q} now follows from Proposition A.1. \square

REMARK A.3. The ‘‘absorption sampling’’ mentioned above (see [13] for history and references) seems not to have been identified with the Mallows measure on M . This connection, along with the invariance of \mathcal{Q} under matrix transposition, make obvious the unexplained symmetry in formulae like [12], equation (10) and [3], equation (2.12).

Likewise, the number of inversions is also invariant under reflection with respect to the secondary matrix diagonal, which swaps (i, j) and $(n + 1 - j, n + 1 - i)$, so \mathcal{Q}_n is also preserved by this transformation. However, this operation has no analog for the infinite group \mathfrak{S} .

REMARK A.4. Observe that the group \mathfrak{S}_∞ acts on \mathfrak{S} both by left and right shifts: an element $\sigma \in \mathfrak{S}_\infty$ maps an element $\tau \in \mathfrak{S}$ to $\sigma\tau$ or $\tau\sigma^{-1}$, respectively. Under the right action, the elementary transposition $\sigma_i := (i, i + 1) \in \mathfrak{S}_\infty$ swaps the letters of a permutation word $\tilde{\tau}$ in the i th and $(i + 1)$ th positions, while under the left action, the same element σ_i swaps the letters i and $(i + 1)$ in $\tilde{\tau}$. That is, under the right action on permutation words, we look at positions, while under the left action, we look at the letters themselves. The inversion map intertwines both actions.

We know that \mathcal{Q} is a unique probability measure on \mathfrak{S} that is quasi-invariant under the right action, with a special cocycle, (2.2). The symmetry property of the measure \mathcal{Q} implies that it is also quasi-invariant under the left action. To compute the corresponding cocycle, we return to the definition (2.1) of the additive cocycle and observe that instead of taking the n -truncated word with large n , we can equally well deal with arbitrary finite subwords, provided that they are large enough. Using this reformulation, we see that the additive cocycle is preserved under the group inversion on \mathfrak{S} , as is the corresponding multiplicative cocycle.

It follows that the cocycle corresponding to the left action remains the same. Consequently, \mathcal{Q} can also be characterized as a unique probability measure on \mathfrak{S} which is quasi-invariant under the left action of \mathfrak{S}_∞ with the same cocycle as before.

The next proposition describes the finite-dimensional distributions of the Mal'ows measure \mathcal{Q} viewed as a measure on $M = \varprojlim M(k)$. We use the following notation: m is an arbitrary matrix from $M(k)$; $I \subset \{1, \dots, k\}$ is the set of indices of the rows in m containing 1's; $J \subset \{1, \dots, k\}$ is the set of indices of the columns in m containing 1's; $r = |I| = |J|$ is the rank of m ; $\text{inv}(m)$ has the same meaning as in the proof of Corollary A.2.

PROPOSITION A.5. *Using the above notation,*

$$(A.1) \quad \theta_k(\mathcal{Q})(\{m\}) = (1 - q)^r q^{k^2 - 2kr - r + \text{inv}(m) + \sum_{i \in I} i + \sum_{j \in J} j}.$$

PROOF. We apply the same method as in Section 6, that is, reduce the alphabet \mathbb{N} to the finite alphabet \mathbb{N}_{k+1} using the monotone map $f_{k+1}(a) = a \wedge (k + 1)$. The key idea is that if $w = w_1 w_2 \dots = \sigma(1)\sigma(2)\dots$ is the random output of the infinite q -shuffle of the word $v = 1 \cdot 2 \cdot \dots$ then, as seen from the proof of Proposition A.1, the truncated matrix $\theta_k(\sigma)$ depends only on the first k letters of the word $f_{k+1}^\infty(w)$ (i.e., all of the letters $\geq k + 1$ become indistinguishable).

On the other hand, by virtue of Proposition 2.3, the random word $f_{k+1}^\infty(w)$ is the output of the infinite q -shuffle applied to the inversion-free word

$$v' := 1 \cdots k \underbrace{(k + 1)(k + 1) \cdots}_\infty \in (\mathbb{N}_{k+1})^\infty.$$

In the notation of Section 4, the law of the random word $f_{k+1}^\infty(w)$ is given by the measure $P^{(v')}$ and the distribution of the first k letters is given by the marginal $P_k^{(v')}$, for which we have an explicit expression; see (4.3). In this formula, we need to take

$$l_1 = \dots = l_k = 1, \quad l_{k+1} = \infty, \quad \mu_{k+1} = k - r,$$

$$\mu_a = \begin{cases} 1, & a \in I, \\ 0, & a \in \{1, \dots, k\} \setminus I, \end{cases}$$

and then the direct computation gives (A.1). \square

There is another way of approximating \mathcal{Q} by the \mathcal{Q}_n 's. Namely, we will see that \mathcal{Q} can be represented as the projective limit of the \mathcal{Q}_n 's. Incidentally, we will realize \mathcal{Q} as a product measure.

As usual, we will identify permutations with the corresponding permutation words. For any $n \geq 2$, we define the projection $\mathfrak{S}_n \rightarrow \mathfrak{S}_{n-1}$ as the deletion of n from a permutation word. Using these projections, we construct the projective limit space $\varprojlim \mathfrak{S}_n$, which is a compact topological space in the standard topology. We have a natural embedding

$$(A.2) \quad \mathfrak{S} \hookrightarrow \varprojlim \mathfrak{S}_n,$$

which is specified by the projection $\mathfrak{S} \rightarrow \mathfrak{S}_n$ which removes all letters larger than n from an infinite permutation word.

Note that \mathfrak{S} is a *proper* subset of $\varprojlim \mathfrak{S}_n$. Indeed, there is a natural one-to-one correspondence between elements of $\varprojlim \mathfrak{S}_n$ and all possible linear orders on the set \mathbb{N} , of which the orders induced by permutation words $\sigma(1)\sigma(2)\dots$ comprise a relatively small part. Still, \mathfrak{S} is dense in $\varprojlim \mathfrak{S}_n$.

PROPOSITION A.6. *The measures \mathcal{Q}_n are consistent with the projections $\mathfrak{S}_n \rightarrow \mathfrak{S}_{n-1}$, so we can define the projective limit $\mathcal{Q}_\infty := \varprojlim \mathcal{Q}_n$, which is a probability measure on $\varprojlim \mathfrak{S}_n$. The image of \mathfrak{S} under the embedding (A.2) has full \mathcal{Q}_∞ -measure and the restriction of \mathcal{Q}_∞ to \mathfrak{S} coincides with the Mallows measure \mathcal{Q} .*

PROOF. For a permutation $\sigma \in \mathfrak{S}_n$ (which we identify with the corresponding permutation word), set

$$\tilde{\beta}_j = \tilde{\beta}_j(\sigma) = \#\{i < j \mid i \text{ precedes } j\} + 1, \quad j = 1, \dots, n$$

[cf. (3.1)]. The link with (3.1) is the identity $\tilde{\beta}_j(\sigma) = \beta_j(\sigma^{-1})$.

The correspondence $\sigma \mapsto (\tilde{\beta}_1(\sigma), \dots, \tilde{\beta}_n(\sigma))$ is a bijection,

$$(A.3) \quad \mathfrak{S}_n \rightarrow \mathbb{N}_1 \times \dots \times \mathbb{N}_n,$$

and we have a counterpart of Proposition 3.2: under \mathcal{Q}_n , the coordinates $\tilde{\beta}_j$ are independent and $j + 1 - \tilde{\beta}_j$ is distributed according to $G_{q,j}$. This can be deduced from Proposition 3.2 taken together with the symmetry property of \mathcal{Q}_n (Proposition A.2), or can be easily checked directly.

Under the bijection (A.3), the projection $\mathfrak{S}_n \rightarrow \mathfrak{S}_{n-1}$ is simply the deletion of the last letter. This enables us to identify $\varprojlim \mathfrak{S}_n$ with the infinite product space $\prod_{n=1}^{\infty} \mathbb{N}_n$. Under this identification, the measure $\varprojlim \mathcal{Q}_n$ becomes the product of truncated geometric distributions. The image of \mathfrak{S} in $\prod_{n=1}^{\infty} \mathbb{N}_n$ consists of those sequences (i_1, i_2, \dots) for which $i_n \rightarrow \infty$. From this, it is readily checked that \mathfrak{S} has full measure.

It remains to check that the measure $\varprojlim \mathcal{Q}_n$ coincides on \mathfrak{S} with the measure \mathcal{Q} . To this end, we use the characterization of \mathcal{Q} in terms of the left action of \mathfrak{S}_{∞} , as described in Remark A.4. It is easy to see that the measure $\varprojlim \mathcal{Q}_n$ has the same transformation property with respect to the left action of elementary transpositions σ_i . Consequently, $\varprojlim \mathcal{Q}_n = \mathcal{Q}$. \square

Alternatively, one can use another chain of projections, such that the projection $\mathfrak{S}_n \rightarrow \mathfrak{S}_{n-1}$ first cuts the last letter in $\sigma(1) \cdots \sigma(n)$, then relabels the letters $\sigma(1) \cdots \sigma(n-1)$ by the increasing bijection with \mathbb{N}_{n-1} . A random element of \mathfrak{S} under \mathcal{Q} is representable by an infinite sequence of backward ranks $(\beta(1), \beta(2), \dots)$, which are independent and have distribution as in Proposition 3.2.

Acknowledgments. We are indebted to Yuliy Baryshnikov and Persi Diaconis for illuminating discussions and references.

REFERENCES

- [1] ALDOUS, D. J. (1985). Exchangeability and related topics. In *École D'été de Probabilités de Saint-Flour, XIII—1983. Lecture Notes in Math.* **1117** 1–198. Springer, Berlin. [MR883646](#)
- [2] ANDREWS, G. E. (1998). *The Theory of Partitions*. Cambridge Univ. Press, Cambridge. [MR1634067](#)
- [3] BARAKAT, R. (1985). Probabilistic aspects of particles transiting a trapping field: An exact combinatorial solution in terms of Gauss polynomials. *Z. Angew. Math. Phys.* **36** 422–432. [MR797237](#)
- [4] BAYER, D. and DIACONIS, P. (1992). Trailing the dovetail shuffle to its lair. *Ann. Appl. Probab.* **2** 294–313. [MR1161056](#)
- [5] BENJAMINI, I., BERGER, N., HOFFMAN, C. and MOSSEL, E. (2005). Mixing times of the biased card shuffling and the asymmetric exclusion process. *Trans. Amer. Math. Soc.* **357** 3013–3029 (electronic). [MR2135733](#)
- [6] BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, 2nd ed. Wiley, New York. [MR1700749](#)
- [7] DIACONIS, P. and RAM, A. (2000). Analysis of systematic scan Metropolis algorithms using Iwahori–Hecke algebra techniques. *Michigan Math. J.* **48** 157–190. [MR1786485](#)

- [8] GNEDIN, A. and OLSHANSKI, G. (2006). The boundary of the Eulerian number triangle. *Mosc. Math. J.* **6** 461–475, 587. [MR2274860](#)
- [9] GNEDIN, A. and OLSHANSKI, G. (2009). A q -analogue of de Finetti’s theorem. *Electron. J. Combin.* **16** R78. [MR2529787](#)
- [10] GRESCHONIG, G. and SCHMIDT, K. (2000). Ergodic decomposition of quasi-invariant probability measures. *Colloq. Math.* **84/85** 495–514. [MR1784210](#)
- [11] KALLENBERG, O. (2005). *Probabilistic Symmetries and Invariance Principles*. Springer, New York. [MR2161313](#)
- [12] KEMP, A. W. (1998). Absorption sampling and the absorption distribution. *J. Appl. Probab.* **35** 489–494. [MR1641849](#)
- [13] KEMP, A. W. (2001). A characterization of a distribution arising from absorption sampling. In *Probability and Statistical Models and Applications* (C. A. Charalambides et al., eds.) 239–246. Chapman and Hall/CRC Press, Boca Raton, FL.
- [14] KEROV, S., OKOUNKOV, A. and OLSHANSKI, G. (1998). The boundary of the Young graph with Jack edge multiplicities. *Int. Math. Res. Not.* **1998** 173–199. [MR1609628](#)
- [15] MALLOWS, C. L. (1957). Non-null ranking models. I. *Biometrika* **44** 114–130. [MR0087267](#)
- [16] RAWLINGS, D. (1997). Absorption processes: Models for q -identities. *Adv. in Appl. Math.* **18** 133–148. [MR1430385](#)
- [17] STANLEY, R. P. (2001). Generalized riffle shuffles and quasisymmetric functions. *Ann. Comb.* **5** 479–491. [MR1897637](#)
- [18] STARR, S. (2009). Thermodynamic limit for the Mallows model on S_n . *J. Math. Phys.* **50** 095208, 15. [MR2566888](#)
- [19] VERSHIK, A. M. and KEROV, S. V. (1987). Locally semisimple algebras: Combinatorial theory and the K_0 -functor. *J. Math. Sci. (N. Y.)* **38** 1701–1733.

MATHEMATICAL INSTITUTE
UTRECHT UNIVERSITY
POSTBUS 80010
3508 TA UTRECHT
THE NETHERLANDS
E-MAIL: A.V.Gnedin@math.uu.nl

INSTITUTE FOR INFORMATION
TRANSMISSION PROBLEMS
BOLSHOY KARETNY 19
MOSCOW 127994
AND
INDEPENDENT UNIVERSITY OF MOSCOW
RUSSIA
E-MAIL: olsh2007@gmail.com