Combinatorics, Probability and Computing (2011) 20, 651–682.
doi:10.1017/S0963548311000307

Concave Majorants of Random Walks and Related Poisson Processes

JOSH ABRAMSON and JIM PITMAN†

Department of Statistics, University of California at Berkeley, CA 94720, USA
(e-mail: josh@stat.berkeley.edu, pitman@stat.berkeley.edu)

Received 14 November 2010; revised 4 July 2011

We offer a unified approach to the theory of concave majorants of random walks, by providing a path transformation for a walk of finite length that leaves the law of the walk unchanged whilst providing complete information about the concave majorant. This leads to a description of a walk of random geometric length as a Poisson point process of excursions away from its concave majorant, which is then used to find a complete description of the concave majorant of a walk of infinite length. In the case where subsets of increments may have the same arithmetic mean, we investigate three nested compositions that naturally arise from our construction of the concave majorant.

1. Introduction

Let $S_0 = 0$ and $S_j = \sum_{i=1}^{j} X_i$ for $1 \leqslant j \leqslant n$, where $X_1, \dots, X_n$ are exchangeable random variables. Let **A** be the assumption that almost surely no two subsets of $X_1, \dots, X_n$ have the same arithmetic mean, and assume for now that **A** holds. Let $S^{[0,n]} := \{(j, S_j) : 0 \leqslant j \leqslant n\}$, so that $S^{[0,n]}$ is the random walk of length n with increments distributed like $X_1, \dots, X_n$. Let

$$0 < N_{n,1} < N_{n,1} + N_{n,2} < \cdots < N_{n,1} + \cdots + N_{n,F_n} = n$$

be the successive times j with $0 \leqslant j \leqslant n$ such that $S_j = \bar{C}^{[0,n]}(j)$, where $\bar{C}^{[0,n]}$ is the *concave majorant* of the walk $S^{[0,n]}$, *i.e.*, the least concave function C on $[0,n]$ such that $C(j) \geqslant S_j$ for $1 \leqslant j \leqslant n$. The random variable F_n is the *number of faces* of the concave majorant. Without assumption **A**, more care needs to be taken in defining the faces of the concave majorant; this will be discussed further in Section 6.

The ith face of the concave majorant is a chord from $(N_{n,1} + \cdots + N_{n,i-1}, S_{N_{n,1}+\cdots+N_{n,i-1}})$ to $(N_{n,1} + \cdots + N_{n,i}, S_{N_{n,1}+\cdots+N_{n,i}})$. We define the *length*, *increment* and *slope* of the ith face to be N_i, $\Delta_{n,i}$ and $\frac{\Delta_{n,i}}{N_i}$ respectively, where

$$\Delta_{n,i} := (S_{N_{n,1}+\cdots+N_{n,i}} - S_{N_{n,1}+\cdots+N_{n,i-1}}), \quad \text{for } 1 \leqslant i \leqslant F_n.$$

† Research supported in part by National Science Foundation Award 0806118.

In the 1950s, E. Sparre Andersen [17] discovered the following remarkable result: for any exchangeable $X_1, \ldots, X_n$ satisfying assumption **A**, there is the equality in distribution

$$F_n \stackrel{d}{=} K_n = \sum_{j=1}^{n} I_j, \tag{1.1}$$

where K_n is the number of cycles in a uniformly distributed random permutation of the set $[n] := \{1, \ldots, n\}$, and I_j, $j = 1, 2, \ldots$ is a sequence of independent Bernoulli variables with $\mathbb{P}(I_j = 1) = 1/j$ and $\mathbb{P}(I_j = 0) = 1 - 1/j$ for each j. The second equality in (1.1) is an elementary and well-known representation of K_n which holds for a number of natural constructions of uniform random permutations of n simultaneously for all n, including both the construction from records of the X_i [7], and the Chinese Restaurant Process [12].

A further result that seems to have been known by Spitzer [19], and shown explicitly by Goldie [7] using a generalization by Brunk of Spitzer's lemma [3], is that under assumption **A** the distribution of the *partition of n* generated by the lengths of the faces of the concave majorant on $[0, n]$, which may be encoded by these lengths in non-increasing order, has the same distribution as the partition of n generated by the cycles of a uniform random permutation; we will prove this result as a corollary of our main theorem. Thus the partition generated by the lengths of the faces of the concave majorant may be generated by a discrete *uniform stick-breaking process* on $[0, n]$ [12]. The result raises the following problem.

The rearrangement problem. Conditionally given that the partition of n generated by the lengths of the faces of the concave majorant of the random walk $S^{[0,n]}$ has segment lengths $n_1, \ldots, n_k$ with $n_1 \geqslant n_2 \geqslant \ldots \geqslant n_k > 0$, we ask the following.

- In what order and with what increments should the faces $f_1, \ldots, f_k$ of the concave majorant with lengths $n_1, \ldots, n_k$, respectively, be arranged to recreate the concave majorant of the random walk $S^{[0,n]}$?
- Given the concave majorant, what is the distribution of values of the random walk $S^{[0,n]}$ between vertices of the concave majorant?

We answer this question by giving in Theorem 1.1 a simultaneous construction of the walk and its concave majorant, conditional on the partition generated by the lengths of the faces of the concave majorant. The theorem will be proved under assumption **A** in Section 2, and in the general case in Section 6, with the key idea of both proofs being that it is enough to show that the theorem is true when $X_1, \ldots, X_n$ are samples without replacement from a set of n real numbers. Since the construction given in the theorem applies to general exchangeable $X_1, \ldots, X_n$, it allows us to investigate in Section 6 the structure of the concave majorant in the general case. The statement of the theorem is complicated, but easy to describe informally, particularly under assumption **A**, in which case the construction is as follows. Conditional on the lengths of the blocks of the partition generated by the concave majorant being $(n_1, \ldots, n_k)$:

- Split $X_1, \ldots, X_n$ into k blocks,

$$(X_1, \ldots, X_{n_1})(X_{n_1+1}, \ldots, X_{n_1+n_1}) \cdots (X_{\sum_{i=1}^{k-1} n_i + 1}, \ldots, X_{\sum_{i=1}^{k} n_i}).$$

- Arrange the blocks in order of decreasing arithmetic mean.

https://doi.org/10.1017/S0963548311000307 Published online by Cambridge University Press

- Perform the unique cyclic permutations of the increments within each block such that the walk with those cyclically permuted increments remains below the line joining its start and end points.

This process defines a permutation of the original increments which leaves the distribution of the walk $S^{[0,n]}$ unchanged and at the same time provides us with information about the concave majorant. In the case where $X_1, \ldots, X_n$ are independent, we may just generate independent walks of length $n_1, \ldots, n_k$, cyclically permute the increments of each walk appropriately, and then arrange the walks in order of decreasing slope. The idea of using cyclic permutations to transform random walk bridges into excursions is due to Vervaat [21].

When assumption **A** is not satisfied there are two more complications. Some of the blocks may have the same arithmetic mean, in which case their ordering is chosen uniformly, and within a block there may be more than one cyclic permutation of increments that leaves the walk with those increments below the line joining its start and end points, in which case the cyclic permutation is chosen uniformly from the possible options. By exchangeability, it would also work to take the blocks with the same arithmetic mean in order of appearance rather than randomly ordering them, but this makes the statement of the theorem harder and in fact does not make the proof any easier.

To facilitate the statement of the theorem, it is necessary to define the set of all permutations that cyclically permute increments within certain blocks and then arrange those blocks in some order.

Definition. Let Σ_n be the set of permutations of $[n]$, and let $\mathcal{P}_n$ be the set of partitions of n, encoded in non-increasing order. For $(n_1, \ldots, n_k) \in \mathcal{P}_n$, let $\Sigma_{(n_1,\ldots,n_k)} \subseteq \Sigma_n$ be such that $\sigma \in \Sigma_{(n_1,\ldots,n_k)}$ if and only if, for some $\tau \in \Sigma_k$ and $(r_1, \ldots, r_k) \in \mathbb{Z}^k$, we have

$$\sigma\left(\sum_{l=1}^{i-1} n_{\tau(l)} + j\right) = \textstyle\sum_{l=1}^{\tau(i)-1} n_l + ((j-1+r_i) \mod n_{\tau(i)}) + 1,$$

for $1 \leqslant j \leqslant n_{\tau(i)}$, $1 \leqslant i \leqslant k$.

In the definition of $\Sigma_{(n_1,\ldots,n_k)}$ just given, the cyclic shift chosen for the $\tau(i)$th block is given by r_i and the ordering of the k blocks is given by τ.

Theorem 1.1. *Let $S_0 = 0$ and $S_j = \sum_{\ell=1}^{j} X_\ell$ for $1 \leqslant j \leqslant n$, where $X_1, \ldots, X_n$ are random variables with any exchangeable joint distribution. Let $S^{[0,n]} = \{(j, S_j) : 0 \leqslant j \leqslant n\}$. Independent of $X_1, \ldots, X_n$, let $L_{n,1}, L_{n,2}, \ldots, L_{n,K_n}$ be a sequence of random variables distributed like the lengths of cycles of a random permutation of $[n]$ arranged in non-increasing order. Conditionally given $\{K_n = k\}$ and $\{L_{n,i} = n_i : 1 \leqslant i \leqslant k\}$, let B be the random subset of Σ_n defined by the following relation: σ is in B if and only if $\sigma \in \Sigma_{(n_1,\ldots,n_k)}$, and there exists $\tau \in \Sigma_k$ such that the function defined on $[k]$ by*

$$i \mapsto \Delta_{n,i}^{\sigma,\tau} := \frac{1}{n_{\tau(i)}}\left(\sum_{\ell=n_{\tau(1)}+\cdots+n_{\tau(i-1)}+1}^{n_{\tau(1)}+\cdots+n_{\tau(i)}} X_{\sigma(\ell)}\right) \tag{1.2}$$

is non-increasing in i, and for each $1 \leqslant i \leqslant k$ *we have*

$$\frac{1}{m}\left(\sum_{\ell=n_{\tau(1)}+\cdots+n_{\tau(i-1)}+1}^{n_{\tau(1)}+\cdots+n_{\tau(i-1)}+m} X_{\sigma(\ell)}\right) \leqslant \Delta_{n,i}^{\sigma,\tau} \quad \text{for } 1 \leqslant m \leqslant n_{\tau(i)}. \tag{1.3}$$

Conditionally given B*, let* ρ *be a uniform random element of* B*, independently of all previously introduced random variables. For* $1 \leqslant j \leqslant n$*, let* $S_j^\rho = \sum_{\ell=1}^{j} X_{\rho(\ell)}$ *and let*

$$S_\rho^{[0,n]} = \{(j, S_j^\rho) : 0 \leqslant j \leqslant n\}.$$

Then $S_\rho^{[0,n]} \stackrel{d}{=} S^{[0,n]}$.

The condition involving (1.2) ensures that the permutation that we end up choosing puts the blocks of increments in non-increasing order of arithmetic mean, *i.e.*, in non-increasing order of slope, and the condition involving (1.3) ensures that the cyclic permutation chosen for each block makes the walk stay below the line joining the start and end points of the increments of that block. In the case where $X_1, \ldots, X_n$ satisfy assumption **A**, the random set B almost surely only consists of one element and thus the additional random variable ρ is not needed.

Some of the ideas of our construction are contained within the work of Spitzer [19], who observed that if $\Delta_{n,i}$ is the increment of the walk over the ith face of the concave majorant, then for the maximum

$$M_n := \max_{0 \leqslant k \leqslant n} S_k$$

there is the almost sure representation

$$M_n = \sum_{i=1}^{F_n} \Delta_{n,i} 1(\Delta_{n,i} \geqslant 0). \tag{1.4}$$

Spitzer showed the much simpler representation in distribution

$$M_n \stackrel{d}{=} \sum_{i=1}^{K_n} \Delta_{n,i}^* 1(\Delta_{n,i}^* \geqslant 0), \tag{1.5}$$

where K_n is the number of cycles of a random permutation independent of the random walk $S^{[0,n]} = \{(j, S_j) : 0 \leqslant j \leqslant n\}$, and given $K_n = k$ and that the permutation has cycles of length say $L_{n,1}, \ldots, L_{n,k}$, the $\Delta_{n,i}^*$ are conditionally independent, with

$$(\Delta_{n,i}^* \mid K_n = k, L_{n,i} = \ell) \stackrel{d}{=} S_\ell, \quad \text{for } 1 \leqslant i \leqslant k, \text{ and } 1 \leqslant \ell \leqslant n.$$

This is an immediate corollary of our theorem, and something we investigate further in Section 5.3. Some consequences of this result lead to other ideas which arise in this paper. Let $S_\ell^+ = S_\ell \vee 0$. As pointed out by Spitzer, Hunt's remarkable identity [11, Theorem 4.1],

$$\mathbb{E}(M_n) = \sum_{\ell=1}^{n} \frac{\mathbb{E}(S_\ell^+)}{\ell}, \tag{1.6}$$

follows easily from (1.5), along with the following complete description of the distribution of M_n for every $n = 1, 2, \ldots$ (this description is known as Spitzer's identity). For $|q| < 1$,

$$\sum_{n=0}^{\infty} q^n \mathbb{E} e^{itM_n} = \exp\left(\sum_{k=1}^{\infty} \frac{q^k}{k} \mathbb{E} e^{itS_k^+}\right). \tag{1.7}$$

To indicate how (1.6) follows from (1.5), recall that the expected number of cycles of length ℓ in a random permutation of $[n]$ is ℓ^{-1}. So (1.6) decomposes the expectation of the sum in (1.5) according to the contributions from cycles of various sizes ℓ. To provide a similar interpretation of (1.7), let $n(q)$ denote a random variable with geometric distribution with parameter $1-q$, so $\mathbb{P}(n(q) \geqslant n) = q^n$ for $n = 0, 1, \ldots$, and assume $n(q)$ is independent of the random walk. Then, multiplying (1.7) by $1-q$ and using the expansion $-\log(1-q) = \sum_{k=1}^{\infty} q^k/k$ allows (1.7) to be rewritten [9] as

$$\mathbb{E} e^{itM_{n(q)}} = \exp\left(\sum_{k=1}^{\infty} \frac{q^k}{k} (\mathbb{E} e^{itS_k^+} - 1)\right). \tag{1.8}$$

Otherwise stated, the maximum $M_{n(q)}$ of the walk up to the independent geometric time $n(q)$ has a compound Poisson distribution,

$$M_{n(q)} \stackrel{d}{=} \sum_{k=1}^{\infty} \sum_{i=1}^{N(q^k/k)} S_{k,i}^+, \tag{1.9}$$

where for fixed q the $N(q^k/k)$ are independent Poisson variables with parameters q^k/k for $k = 1, 2, \ldots$, and given these variables the $S_{k,i}$ for $1 \leqslant i \leqslant N(q^k/k)$ are independent with $S_{k,i} \stackrel{d}{=} S_k$. As observed by Greenwood and Pitman [9], the identity in distribution (1.9), and the companion result which determines the common distribution of $S_n - M_n$ and $\min_{0 \leqslant k \leqslant n} S_k$ for every n, can be derived, along with other results of fluctuation theory for the distribution of ladder heights and ladder times, from the decomposition

$$S_{n(q)} = M_{n(q)} + (S_{n(q)} - M_{n(q)}), \tag{1.10}$$

which expresses the compound Poisson variable $S_{n(q)}$ as the sum of two independent compound Poisson variables, with positive and negative ranges respectively. Moreover, as shown in [8], this discussion can be passed to a continuous time limit to derive the companion circle of fluctuation identities for maxima, minima and ladder processes associated with Lévy processes. In Section 5.3 we give new explanations for the compound Poisson distributions mentioned above.

The rest of this article is structured as follows. In Section 2 we will prove Theorem 1.1 under assumption **A** and give corollaries relating to the partition and composition induced by the concave majorant. In Section 3 we will analyse some specific examples of composition probabilities, including the Cauchy increment case, which turns out to be particularly simple. In Section 4 we extend the description to the case where n is replaced by $n(q)$, a geometric random variable with parameter $1-q$, which results in a description of the concave majorant and the excursions under each face as a Poisson point process. In Section 5 we apply the Poissonian theory. First, by letting $q \to 1$ we find a description of the concave majorant for the random walk on $[0, \infty)$, and the associated excursions under each face. Then we analyse the behaviour of the concave majorant as n grows. As a final application we investigate the pre- and post-maximum parts of the walk. In Section 5.3 we investigate the two concave majorants that result from decomposing the random walk at its maximum, and their associated partitions. In Section 6 we extend the theory to $X_1, \ldots, X_n$ not satisfying assumption **A**. Also in Section 6 we investigate three nested compositions of integers that arise naturally. At the end of this section we give some examples of how the general theory can be applied. In Section 7 we finish answering the rearrangement problem mentioned above by describing the law of a random walk conditional on the value of its concave majorant. Finally, in Section 8, we describe an important path transformation that provides Pitman and Uribe Bravo with the basis for a full investigation into the concave majorant of a Lévy process [13].

2. Proof of Theorem 1.1 under assumption A and the partition and composition laws

We begin with a simple lemma due to Spitzer relating to cyclic permutations of increments of walks, which shows that under assumption **A** the appropriate cyclic permutations discussed in the Introduction are almost surely unique.

Lemma 2.1 ([19], Theorem 2.1). *Let $x = (x_1, \ldots, x_n)$ be a vector such that no two subsets of the coordinates have the same arithmetic mean. For $1 \leqslant k \leqslant n$ let $x_{k+n} = x_k$, and let $x(k) = (x_k, x_{k+1}, \ldots, x_{k+n})$. Then there is a unique $1 \leqslant k^* \leqslant n$ such that the walk with increments $x(k^*) = (x_{k^*}, x_{k^*+1}, \ldots, x_{k^*+n})$ lies below the chord joining its start and end points.*

Proof of Theorem 1.1 under assumption A. By conditioning on the set of values taken by $X_1, \ldots, X_n$, it is enough to show that $S_\rho^{[0,n]} \overset{d}{=} S^{[0,n]}$ in the case where $X_1, \ldots, X_n$ are samples without replacement from n real numbers $x_1, \ldots, x_n$ such that no two subsets of $x_1, \ldots, x_n$ have the same arithmetic mean. Thus it is enough to show that, for every permutation $\sigma \in \Sigma_n$, we have

$$\mathbb{P}(X_{\rho(1)} = x_{\sigma(1)}, \ldots, X_{\rho(n)} = x_{\sigma(n)}) = \frac{1}{n!},$$

and, without loss of generality, it is enough to show this for σ, the identity permutation. Suppose the concave majorant of the deterministic walk with increments $(x_1, \ldots, x_n)$ has k faces, whose lengths *in order of appearance* are $(m_1, \ldots, m_k)$, so that the composition induced by the lengths of the faces of the concave majorant is $(m_1, \ldots, m_k)$. Let $\tau \in \Sigma_k$ be such that

$$(n_1, \ldots, n_k) := (m_{\tau(1)}, \ldots, m_{\tau(k)})$$

are the lengths of the k faces in *non-increasing order*, so that the partition induced by the lengths of the faces of the concave majorant is $(n_1, \ldots, n_k)$.

First suppose that each element of $(n_1, \ldots, n_k)$ is distinct. Then the event $\{X_{\rho(\ell)} = x_\ell : 1 \leqslant \ell \leqslant n\}$ occurs if and only if:

(i) the partition chosen according to the lengths of the cycles of a random permutation is $(n_1, \ldots, n_k)$;

(ii) for each $1 \leqslant i \leqslant k$, the ordered list $(X_{n_1+\cdots+n_{i-1}+1}, \ldots, X_{n_1+\cdots+n_i})$ is one of the n_i cyclic permutations of the ordered list

$$(x_{m_1+m_2+\cdots+m_{\tau(i)-1}+1}, \ldots, x_{m_1+m_2+\cdots+m_{\tau(i)}}).$$

https://doi.org/10.1017/S0963548311000307 Published online by Cambridge University Press

According to the Ewens Sampling Formula, the event in (i) has probability $\prod_{i=1}^{k} \frac{1}{n_i}$. The event in (ii) is independent of the event in (i), and has probability $\frac{1}{n!}\prod_{i=1}^{k} n_i$.

Now suppose that the elements of $(n_1, \ldots, n_k)$ are not distinct. For $1 \leqslant j \leqslant n$, let $I_j = \{i : n_i = j\}$ and let $a_j = |I_j|$. The event $\{X_{\rho(\ell)} = x_\ell : 1 \leqslant \ell \leqslant n\}$ occurs if and only if:

(i) the partition chosen according to the lengths of the cycles of a random permutation is $(n_1, \ldots, n_k)$;

(ii) for each $1 \leqslant j \leqslant n$, for each $i \in I_j$ the ordered list $(X_{n_1+\cdots+n_{i-1}+1}, \ldots, X_{n_1+\cdots+n_i})$ is one of the $n_i = j$ cyclic permutations of the ordered list

$$(x_{m_1+m_2+\cdots+m_{\tau(i')-1}+1}, \ldots, x_{m_1+m_2+\cdots+m_{\tau(i')}})$$

for some $i' \in I_j$.

By the Ewens Sampling Formula, the event in (i) has probability

$$\left(\prod_{i=1}^{k} \frac{1}{n_i}\right)\left(\prod_{j=1}^{n} \frac{1}{a_j!}\right).$$

The event in (ii) is independent of the event in (i), and has probability

$$\frac{1}{n!}\left(\prod_{i=1}^{k} n_i\right)\left(\prod_{j=1}^{n} a_j!\right).$$

Hence $\mathbb{P}(X_{\rho(\ell)} = x_\ell : 1 \leqslant \ell \leqslant n) = \frac{1}{n!}$. □

As a direct consequence of Theorem 1.1 we have the result of Goldie [7] mentioned in the Introduction.

Corollary 2.2. *Let $M_{n,1}, \ldots, M_{n,F_n}$ be the lengths of the faces of the concave majorant of $S^{[0,n]}$ arranged in non-increasing order. Then, under assumption* **A** *the joint distribution of $M_{n,1}, \ldots, M_{n,F_n}$ is given by the formula*

$$\mathbb{P}(F_n = k, M_{n,i} = n_i, 1 \leqslant i \leqslant k) = \prod_{j=1}^{n} \frac{1}{j^{a_j} a_j!}$$

for all $(n_1, \ldots, n_k) \in \mathcal{P}_n$, where $a_j = \#\{i : 1 \leqslant i \leqslant k, n_i = j\}$ for $1 \leqslant j \leqslant n$. That is, the partition of n induced by the lengths of the faces of the concave majorant of $S^{[0,n]}$ has the law of a partition of n induced by the cycle lengths of a random permutation.

Proof. Following the construction in Theorem 1.1, the lengths $L_{n,1}, \ldots, L_{n,K_n}$ are exactly the lengths of the faces of the concave majorant of $S_\rho^{[0,n]}$, and the conclusion follows since $S^{[0,n]} \stackrel{d}{=} S_\rho^{[0,n]}$. □

Further, Theorem 1.1 allows us to describe the law of the composition induced by the lengths of the faces of the concave majorant.

https://doi.org/10.1017/S0963548311000307 Published online by Cambridge University Press

Corollary 2.3. *Let $(N_{n,1},\ldots,N_{n,F_n})$ be the composition of n induced by the lengths of the faces of the concave majorant of $S^{[0,n]}$. Then, under assumption* **A** *the joint distribution of $N_{n,1},\ldots,N_{n,F_n}$ is given by the formula*

$$\mathbb{P}(F_n = k, N_{n,i} = n_i, 1 \leqslant i \leqslant k) = \mathbb{P}\left(\frac{S^{(1)}_{n_1}}{n_1} > \frac{S^{(2)}_{n_2}}{n_2} > \cdots > \frac{S^{(k)}_{n_k}}{n_k}\right)\prod_{i=1}^{k}\frac{1}{n_i}$$

for all compositions $(n_1,\ldots,n_k)$ of $[n]$ into k parts, where for $1 \leqslant i \leqslant k$

$$S^{(i)}_{n_i} := S_{n_1+\cdots+n_i} - S_{n_1+\cdots+n_{i-1}} \overset{d}{=} S_{n_i}.$$

In particular, if the X_i are independent, then so are the $S^{(i)}_{n_i}$ for $1 \leqslant i \leqslant k$.

Proof. Fix a composition $(n_1,\ldots,n_k)$ and let $(\overrightarrow{n}_{\tau(1)},\ldots,\overrightarrow{n}_{\tau(k)})$ be $(n_1,\ldots,n_k)$ in non-increasing order. Let T be the set of $\tau \in \Sigma_k$ such that $(\overrightarrow{n}_{\tau(1)},\ldots,\overrightarrow{n}_{\tau(k)}) = (n_1,\ldots,n_k)$. Then $|T| = \prod_{j=1}^{n} a_j$, where $a_j = \#\{i : 1 \leqslant i \leqslant k, n_i = j\}$ for $1 \leqslant j \leqslant n$. We are interested in comparing the slopes of the faces of the concave majorant that result from the construction in Theorem 1.1. In this direction, for $1 \leqslant i \leqslant k$ let

$$S^{(\tau(i))}_{\overrightarrow{n}_{\tau(i)}} = S_{\overrightarrow{n}_1+\cdots+\overrightarrow{n}_{\tau(i)}} - S_{\overrightarrow{n}_1+\cdots+\overrightarrow{n}_{\tau(i)-1}} \overset{d}{=} S_{\overrightarrow{n}_{\tau(i)}} = S_{n_i}.$$

Under the construction in Theorem 1.1, the events $\{F_n = k\}$ and $\{N_{n,i} = n_i : 1 \leqslant i \leqslant k\}$ occur if and only if:

(i)

$$(L_{n,1},\ldots,L_{n,K_n}) = (\overrightarrow{n}_1,\ldots,\overrightarrow{n}_k);$$

(ii)

$$\frac{S^{(\tau(1))}_{\overrightarrow{n}_{\tau(1)}}}{n_1} > \frac{S^{(\tau(2))}_{\overrightarrow{n}_{\tau(2)}}}{n_2} > \cdots > \frac{S^{(\tau(k))}_{\overrightarrow{n}_{\tau(k)}}}{n_k}, \quad \text{for some } \tau \in T.$$

As before, the event in (i) has probability

$$\left(\prod_{i=1}^{k}\frac{1}{n_i}\right)\left(\prod_{j=1}^{n}\frac{1}{a_j!}\right).$$

The event in (ii) is independent of the event in (i), and by exchangeability the probability that it occurs for one particular element of T is

$$\mathbb{P}\left(\frac{S^{(1)}_{n_1}}{n_1} > \frac{S^{(2)}_{n_2}}{n_2} > \cdots > \frac{S^{(k)}_{n_k}}{n_k}\right).$$

Recalling that $|T| = \prod_{j=1}^{n} a_j$ completes the proof. □

3. Examples of composition probabilities

The special case of Cauchy increments gives rise to the following appealing version of Corollary 2.3.

Corollary 3.1. *Suppose that the X_i are independent and such that S_k/k has the same distribution for every k, as when the X_i have a Cauchy distribution. Then*

$$\mathbb{P}(F_n = k; N_{n,i} = n_i, 1 \leqslant i \leqslant k) = \frac{1}{k!}\prod_{i=1}^{k}\frac{1}{n_i},$$

and hence $\{N_{n,i} : 1 \leqslant i \leqslant F_n\}$ has the same distribution as the composition of n created by first choosing a random permutation of n and then putting the cycle lengths in uniform random order.

Proof. Since $\frac{S^{(1)}_{n_1}}{n_1}, \ldots, \frac{S^{(k)}_{n_k}}{n_k}$ is an i.i.d. sequence, each of the $k!$ orderings is equally likely, and hence

$$\mathbb{P}\left(\frac{S^{(1)}_{n_1}}{n_1} > \cdots > \frac{S^{(k)}_{n_k}}{n_k}\right) = \frac{1}{k!}. \qquad \square$$

Note that the continuum limit of this result can be read from Bertoin's work [2]. The above result shows that the Cauchy discrete model is the same as that derived by random sampling from the continuum Cauchy model, as per Gnedin's theory of sampling consistent compositions of positive integers [6]. That is, let $U_1, \ldots, U_n$ be independent identically distributed uniform random variables on $[0,1]$ and let X be a Cauchy process on $[0,1]$. Generate a composition of n by putting i in the same block as j if and only if U_i and U_j fall in the same segment of the composition of $[0,1]$ induced by the lengths of the faces of the concave majorant of X, and then ordering blocks according to the ordering of the faces of the concave majorant of X. Then the composition of n that is generated will have the same distribution as $(N_{n,1}, \ldots, N_{n,F_n})$ in Corollary 3.1. This does not seem at all obvious *a priori*, and according to simulation is not true in the Brownian case, suggesting that it is not true in general.

Now let $X_1, \ldots, X_n$ be any exchangeable sequence of random variables satisfying assumption **A**, as in Corollary 2.3. We now give some numerical examples of composition probabilities when n is small. Let

$$p(n_1, \ldots, n_k) := \mathbb{P}(F_n = k, N_{n,i} = n_i, 1 \leqslant i \leqslant k).$$

Using symmetry and the partition probabilities given in Corollary 2.2, universal values are

$$p(1,1) = 1/2, \quad p(2) = 1/2,$$
$$p(3) = 1/3, \quad p(2,1) = p(1,2) = 1/4, \quad p(1,1,1) = 1/6,$$
$$p(4) = 1/4, \quad p(1,3) = p(3,1) = 1/6, \quad p(2,2) = 1/8, \quad p(1,1,1,1) = 1/24.$$

As n increases, the first values that depend on the particular choice of increment distributions are

$$p(1,1,2) = p(2,1,1) = \frac{1}{2}\mathbb{P}(X_1 > X_2 > \tfrac{1}{2}(X_3 + X_4)),$$
$$p(1,2,1) = \frac{1}{2}\mathbb{P}(X_1 > \tfrac{1}{2}(X_2 + X_3) > X_4),$$

where, according to the partition probabilities, we must have

$$p(1,1,2) + p(2,1,1) + p(1,2,1) = 1/4.$$

https://doi.org/10.1017/S0963548311000307 Published online by Cambridge University Press

We consider two special cases: independent Cauchy increments and independent Gaussian increments. When the increments are independent and Cauchy, the three probabilities above are equal, with

$$2p(1,2,1) = \mathbb{P}(X_1 > \tfrac{1}{2}(X_2+X_3) > X_4) = 1/6 = 0.1666666\ldots.$$

Note that

$$\mathbb{P}(X_1 > \tfrac{1}{2}(X_2+X_3) > X_4) = \mathbb{P}\big(\tfrac{1}{2}(X_2+X_3) - X_1 < 0 \quad \text{and} \quad X_4 - \tfrac{1}{2}(X_2+X_3) < 0\big).$$

In the centred Gaussian case with $\mathrm{Var}(X_1) = 1$, this is the probability of the negative quadrant for a centred bivariate normal with equal variances $3/2$ and covariance $-1/2$ and thus correlation $\rho = -1/3$. That probability is given by

$$\frac{1}{4} + \frac{\arcsin(-1/3)}{2\pi} = 0.195913276.$$

The difference from the Cauchy case is quite small. The fact that it is larger is consistent with the known differences in behaviour of the limit partitions for large n after scaling; it is known that the concave majorant of Brownian motion is more likely to have longer faces in its central region than the concave majorant of a Cauchy process. We conclude this section by conjecturing that $p(1,2,1)$ is a monotonic function of the stability index α for symmetric stable laws.

4. A Poisson point process description

The concave majorant of $S^{[0,n]}$ can be viewed as a random point process on $\{1,\ldots,n\} \times \mathbb{R}$, where a point at (j,s) means that one of the faces of the concave majorant has length j and increment s. Let $A_n(j)$ be the number of faces of the concave majorant of $S^{[0,n]}$ that have length j for $1 \leqslant j \leqslant n$, and let $\Sigma_j^{(1)},\ldots,\Sigma_j^{(A_n(j))}$ be the increments of the faces with length j in uniform random order. Thus, if $X_1,\ldots,X_n$ are independent then for each $1 \leqslant j \leqslant n$, conditionally given $A_n(j) = a_j$, $\Sigma_j^{(\ell)}$ is an independent copy of S_j for each $1 \leqslant \ell \leqslant a_j$. Figure 1 shows an example of such a point process. To construct the concave majorant from this point process, the faces with lengths and increments indicated by the points are arranged in decreasing order of slope.

Now suppose we have an infinite sequence of exchangeable random variables $X_1, X_2, \ldots$, such that almost surely no two subsets have the same arithmetic mean. As before, let $S_0 = 0$ and $S_j = \sum_{i=1}^{j} X_i$ for $j \geqslant 1$. Following ideas from the fluctuation theory of Greenwood and Pitman [9], we now randomize the length of the walk by setting the number of steps of the random walk equal to $n(q)$, where $n(q)$ is a geometric random variable with parameter $1-q$, so that

$$\mathbb{P}(n(q) \geqslant n) = q^n \quad \text{for } n = 0,1,2,\ldots.$$

Let $S^{[0,n(q)]} = \{(j,S_j) : 0 \leqslant j \leqslant n(q)\}$, and let

$$0 < N_{n(q),1} < N_{n(q),1} + N_{n(q),2} < \cdots < N_{n(q),1} + \cdots + N_{n(q),F_{n(q)}} = n(q)$$

be the successive times that $S^{[0,n(q)]}$ meets its concave majorant, where $F_{n(q)}$ is the number of faces of the concave majorant of $S^{[0,n(q)]}$. The following lemma, which involves a fundamental Poisson representation of the geometric distribution, is due to Shepp and Lloyd [15], who were just working with partitions generated by random permutations, not concave majorants.

https://doi.org/10.1017/S0963548311000307 Published online by Cambridge University Press

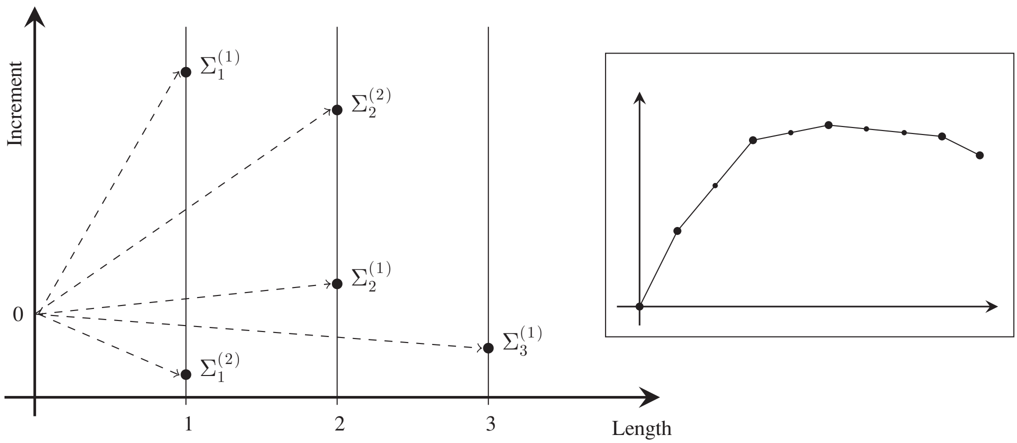

Figure 1. An example point process and the resulting concave majorant. The dashed lines show the slope of each face, and these faces are arranged in decreasing order of slope.

Lemma 4.1. *Let $A_j = \#\{i : 1 \leqslant i \leqslant F_{n(q)}, N_{n(q),i} = j\}$ for $j \geqslant 1$. Then A_j has the Poisson distribution with mean q^j/j, independently for each $j \geqslant 1$.*

Proof. Noting that $\log(1-q) = -\sum_j q^j/j$, we have that

$$\begin{aligned}\mathbb{P}(A_j = a_j, j \geqslant 1) &= \mathbb{P}(n(q) = \textstyle\sum_{j\geqslant 1} j a_j)\mathbb{P}(A_j = a_j, j \geqslant 1 | n(q) = \textstyle\sum_{j\geqslant 1} j a_j) \\ &= (1-q)q^{\sum_j j a_j} \frac{1}{\prod_j j^{a_j} a_j!} \\ &= \prod_j \frac{(\frac{q^j}{j})^{a_j} e^{-\frac{q^j}{j}}}{a_j!},\end{aligned}$$

where the second equality comes from Corollary 2.2. □

For the next theorem, and in fact the rest of this section, it is important that we assume $X_1, X_2, \ldots$ are independent with common continuous distribution. The theorem asserts that the point process discussed above is a Poisson point process under this assumption.

Theorem 4.2. *If $X_1, X_2, \ldots$ are independent with common continuous distribution, then the point process of lengths and increments of faces of the concave majorant of $S^{[0,n(q)]}$ is a Poisson point process on $\{1, 2, \ldots\} \times \mathbb{R}$ with intensity $j^{-1} q^j \mathbb{P}(S_j \in dx)$ for $j = 1, 2, \ldots, x \in \mathbb{R}$. Moreover, let $T_i = \sum_{l=1}^{i} N_{n(q),l}$, $0 \leqslant i \leqslant F_{n(q)}$, be the consecutive times at which $S^{[0,n(q)]}$ meets its concave majorant, so that $T_0 = 0$ and $T_{F_{n(q)}} = n(q)$. Then the sequence of path segments,*

$$\{(S_{T_i+k} - S_{T_i}, 0 \leqslant k \leqslant N_{n(q),i}), i = 0, \ldots, F_{n(q)} - 1\},$$

is a list of the points of a Poisson point process in the space of finite random walk segments

$$\{(s_1, \ldots, s_j) \text{ for some } j = 1, 2, \ldots\},$$

https://doi.org/10.1017/S0963548311000307 Published online by Cambridge University Press

whose intensity measure on paths of length j is $q^j j^{-1}$ times the conditional distribution of $(S_1, \ldots, S_j)$, given that $S_k < (k/j)S_j$ for all $1 \leqslant k \leqslant j-1$.

Proof. Conditionally given $A_j = a_j$, the increment for each face of length j is an independent copy of S_j by Theorem 1.1. Combined with Lemma 4.1 this proves the first statement.

Conditional on the concave majorant of $S^{[0,n(q)]}$ having a face of length j and increment s, the increments of $S^{[0,n(q)]}$ over that face of the concave majorant have the distribution of $(X_1, \ldots, X_j)$ given that $\sum_{\ell=1}^{k} X_\ell < (k/j)s$ for all $1 \leqslant k \leqslant j-1$ and $\sum_{\ell=1}^{j} X_\ell = s$, and this law is independent for each face of $S^{[0,n(q)]}$. This implies the second statement. □

A simple but important corollary of Theorem 4.2 is the following.

Corollary 4.3. $(n(q), S_{n(q)})$ *has a compound Poisson distribution, and the total number of faces* $F_{n(q)}$ *of the concave majorant of* $S^{[0,n(q)]}$ *has Poisson distribution with mean*

$$\sum_{j=1}^{\infty} j^{-1} q^j = -\log(1-q).$$

The first assertion of Corollary 4.3 can in fact be seen directly since

$$(n(q), S_{n(q)}) = \sum_{i=1}^{n(q)} (1, X_i)$$

and $n(q)$ is itself compound Poisson. Explicitly, $n(q)$ is a Poisson compound of a log-series law: $n(q)$ has probability generating function $\mathbb{E}z^{n(q)} = (1-q)/(1-qz)$, which can be expressed as $e^{-\lambda(1-\phi(z))}$, where $\lambda = -\ln(1-q)$ and ϕ is the probability generating function of the log-series law with parameter q. This well-known decomposition of a geometric random variable reappears later in Lemma 6.7.

5. Applications of the Poissonian description

5.1. The random walk on $[0, \infty)$

By letting $q \to 1$ it is possible to deduce the structure of the concave majorant of the random walk on $[0, \infty)$ using Theorem 4.2. Groeneboom [10] gave a Poissonian description of the concave majorant of BM on $[0, \infty)$; that there is a closely parallel description for random walks does not seem to have been pointed out before. The case of Lévy processes will be covered in the forthcoming paper by Pitman and Uribe Bravo [13].

Suppose $\mathbb{E}(X_1) = \mu \in [-\infty, \infty)$. Informally, as $q \to 1$ the intensity measure of the Poisson point process of face lengths and increments approaches $j^{-1}\mathbb{P}(S_j \in dx)$, but since the slope of the concave majorant converges downwards to μ but does not reach it, only the faces with slope greater than μ will contribute to the concave majorant in the limit. Therefore, by Poisson thinning we get a new intensity measure $j^{-1}\mathbb{P}(S_j \in dx)1(x > j\mu)$. Moreover, we can also describe path segments of the walk below each face of the concave majorant as a Poisson point process.

Theorem 5.1. *Let $S_0 = 0$ and $S_j = \sum_{i=1}^{j} X_i$ for $j \geqslant 1$, where $X_1, X_2, \ldots$ are independent random variables with common continuous distribution that has a well-defined mean $\mu := \mathbb{E}(X_1) \in [-\infty, \infty)$. Let $S^{[0,\infty)} = \{(j, S_j) : j \geqslant 0\}$. Let $0 = T_0 < T_1 < T_2 < \cdots$ be the successive times that $S^{[0,\infty)}$ meets its concave majorant, and let $N_i = T_i - T_{i-1}$ for $i \geqslant 1$. Then the sequence of path segments*

$$\{(S_{T_i+k} - S_{T_i}, 0 \leqslant k \leqslant N_i), i = 0, 2, \ldots\}$$

is a list of the points of a Poisson point process in the space of finite random walk segments

$$\{(s_1, \ldots, s_j) \text{ for some } j = 1, 2, \ldots\},$$

whose intensity measure on paths of length j is j^{-1} times the restriction to $S_j \in (j\mu, \infty)$ of the conditional distribution of $(S_1, \ldots, S_j)$ given that $S_k < (k/j)S_j$ for all $1 \leqslant k < j$.

Proof. The combination of the following four facts is enough to prove the theorem.

(i) The number of faces of length j has a Poisson distribution with mean $j^{-1}\mathbb{P}(S_j > j\mu)$.
(ii) These numbers are independent as j varies.
(iii) Given all of these numbers, and with n faces of length j, the n walks on the associated faces, when listed in a uniform random order independently of the walks on the faces, are n independent processes each distributed according to $(S_1, \ldots, S_j)$, given that $S_k < (k/j)S_j$ for all $1 \leqslant k < j$ and $S_j > j\mu$.
(iv) Given n faces of length j, the increments of these faces, when listed in uniform random order, are distributed like n independent copies of S_j given $S_j > j\mu$.

The main thing to check is that (i) and (ii) are true, *i.e.*, that the counts

$$A_\infty(j) := \#\{j : N_i = j\}$$

are independent Poisson variables with mean $j^{-1}\mathbb{P}(S_j \geqslant j\mu)$. Once we have shown this, (iii) and (iv) follow from Poisson thinning and previous discussions relating to the independence of the walks below each segment.

Let $n(q)$ be a geometric random variable with parameter $1-q$. Let $S^{[0,n(q)]} = \{(j, S_j) : 0 \leqslant j \leqslant n(q)\}$, so that the concave majorant of $S^{[0,n(q)]}$ and $S^{[0,\infty)}$ agree up until some random time $T^*_{n(q)}$.

Lemma 5.2. *$T^*_{n(q)}$ is the maximal T_i with $T_i \leqslant n(q)$.*

Proof. To see this, let i be such that $T_i \leqslant n(q)$. Since the concave majorant of $S^{[0,n(q)]}$ is everywhere less than or equal to the concave majorant of $S^{[0,\infty)}$, if they did not agree at time T_i then the concave majorant of $S^{[0,n(q))}$ would go beneath the point (T_i, S_{T_i}), but this is a contradiction since (T_i, S_{T_i}) is in $S^{[0,n(q)]}$. □

Let

$$A_{n(q)}(j) := \#\{i : N_{n(q),i} = j\},$$

https://doi.org/10.1017/S0963548311000307 Published online by Cambridge University Press

where $N_{n(q),1}, \ldots, N_{n(q),F_{n(q)}}$ are the lengths of faces of the concave majorant of $S^{[0,n(q)]}$. There are the obvious decompositions

$$A_\infty(j) = A_\infty(j)(0, T^*_{n(q)}] + A_\infty(j)(T^*_{n(q)}, \infty], \tag{5.1}$$

$$A_{n(q)}(j) = A_{n(q)}(j)(0, T^*_{n(q)}] + A_{n(q)}(j)(T^*_{n(q)}, \infty], \tag{5.2}$$

where, *e.g.*, $A_\infty(j)(0, T^*_{n(q)}]$ is the number of faces of the concave majorant of $S^{[0,\infty)}$ of length j up to and including the face ending at time $T^*_{n(q)}$, and the other terms are defined similarly. Moreover, since $T^*_{n(q)}$ is by definition the maximal common vertex of the concave majorants of $S^{[0,n(q)]}$ and $S^{[0,\infty)}$, it is clear that

$$\begin{aligned} A_\infty(j)(0, T^*_{n(q)}] &= A_{n(q)}(j)(0, T^*_{n(q)}] \\ &= \#\{i : N_{n(q),i} = j,\ S_{T_i} - S_{T_{i-1}} > j\alpha_{n(q)}\}, \end{aligned} \tag{5.3}$$

where $\alpha_{n(q)}$ is the right derivative of the concave majorant of $S^{[0,\infty)}$ at time $T^*_{n(q)}$. Conditionally given $\alpha_{n(q)}$, by Poisson thinning and Theorem 4.2 the distribution of the right-hand side of (5.3) is Poisson with mean $q^j j^{-1}\mathbb{P}(S_j > j\alpha_{n(q)})$, independently for each j. The strategy at this point is to let $q \to 1$, so that $T_{n(q)} \to \infty$ and $\alpha_{n(q)} \to \mu$, resulting in $A_\infty(j)$ having Poisson distribution with mean $j^{-1}\mathbb{P}(S_j > j\mu)$, independently for each j, *i.e.*, resulting in (i) and (ii).

Let $\{q_m\}_{m\geqslant 1}$ be any sequence such that if $\{n(q_m)\}_{m\geqslant 1}$ is a sequence of independent geometric random variables with parameters $1 - q_m$, then $n(q_m) \to \infty$ almost surely as $m \to \infty$ (so that necessarily $q_m \to 1$). Suppose that $T_{(n(q_m))} \to \infty$ and $\alpha_{n(q_m)} \to \mu$ almost surely, so that

$$\begin{aligned} A_\infty(j) &= \lim_{m\to\infty} A_\infty(j)(0, T_{(n(q_m))}] \\ &= \lim_{m\to\infty} \#\{i : N_{n(q_m),i} = j,\ S_{T_i} - S_{T_{i-1}} > j\alpha_{n(q_m)}\}, \end{aligned} \tag{5.4}$$

where the first equality is from (5.1) and the second is from (5.3). Since $\alpha_{n(q_m)} \to \mu$ almost surely, by continuity of the function $x \mapsto \mathbb{P}(S_j > jx)$ the distribution of the right-hand side of (5.4) is Poisson with parameter $j^{-1}\mathbb{P}(S_j > j\mu)$, independently for each j. This proves (i) and (ii).

It remains to prove that $T_{(n(q_m))} \to \infty$ and $\alpha_{n(q_m)} \to \mu$ almost surely as $m \to \infty$. For every $i \geqslant 1$, since $T_i < \infty$ we will have $n(q_m) > T_i$ eventually, and hence by Lemma 5.2 for every $i \geqslant 1$ we will have $T_{(n(q_m))} \geqslant T_i$ eventually. Since $T_i \to \infty$ this implies that $T_{(n(q_m))} \to \infty$ almost surely.

Lemma 5.3. *Almost surely no face of the concave majorant of $S^{[0,\infty)}$ can have slope less than μ.*

Proof. If $\mu = -\infty$ then the conclusion is clear. Suppose $\mu \in (-\infty, \infty)$; then since $S_n - n\mu$ is a mean zero random walk and hence recurrent, for every $i \geqslant 1$ there will almost surely be some $n_i > T_i$ such that $S_{n_i} > S_{T_i} + (n_i - T_i)\mu$, and hence, for any vertex of the concave majorant, the slope of the face to the right must be greater than μ. □

Lemma 5.4. *For every $\epsilon > 0$ there will almost surely be a face of the concave majorant with slope x such that $\mu < x < \mu + \epsilon$.*

Proof. For any $\mu \in [-\infty, \infty)$, by the strong law of large numbers $S_n/n \to \mu$ almost surely as $n \to \infty$. But if there was no slope of the concave majorant on $[0, \infty)$ with slope $x < \mu + \epsilon$ then we would have $\limsup_n S_n/n > \mu$. Combined with Lemma 5.3, this gives the conclusion. □

https://doi.org/10.1017/S0963548311000307 Published online by Cambridge University Press

We already have that $T_{(n(q_m))} \to \infty$ almost surely. Since $\alpha_{n(q_m)}$ is the right derivative of the concave majorant of $S^{[0,\infty)}$ at $T_{(n(q_m))}$, Lemma 5.4 implies that $\alpha_{n(q_m)} \to \mu$ almost surely as $m \to \infty$. This concludes the proof of Theorem 5.1. □

5.2. The structure of the concave majorant of $S^{[0,n]}$ as n varies

Theorem 1.1 relates to the structure of the concave majorant of a random walk of fixed length, and Theorems 4.2 and 5.1 allow randomized lengths or infinite length. So far, though, we have not discussed how the structure changes as the number of steps of the walk increases, but Theorem 5.1 and its proof now allow us to make some comments. Recall that F_n is the number of faces of the concave majorant of $S^{[0,n]} = \{(j, S_j) : 0 \leqslant j \leqslant n)\}$, and in the case where $X_1, \ldots, X_n$ are independent with common continuous distribution, we know from (1.1) that for each fixed n there is the equality in distribution

$$F_n \stackrel{d}{=} K_n := \sum_{j=1}^{n} I_j,$$

where the I_j are independent Bernoulli variables with $\mathbb{P}(I_j = 1) = 1/j$. However, as observed by Steele [20], the identity in law between F_n and K_n does not hold jointly as n varies, and as pointed out by Qiao and Steele [14], the asymptotic behaviour of F_n and K_n as $n \to \infty$ may be quite different. They provide an example of a continuous distribution of X_i such that, for each $m = 1, 2, \ldots,$

$$\mathbb{P}(F_n = m \text{ infinitely often}) = 1.$$

It is an easy consequence of Theorem 5.1 that

$$\mathbb{P}(F_n = 1 \text{ infinitely often}) = 1$$

if and only if $\mathbb{E}(X^+) = \infty$. It appears that the Poisson analysis of $F_{n(q)}$ can be used to provide a more thorough description of the possible asymptotic behaviours of F_n as n varies. In particular, as a consequence of the argument of the proof of Lemma 5.2, if $\mathbb{E}(X^+) < \infty$ then F_n is bounded below by the number of faces of the majorant on $[0,n]$ which are part of the majorant on $[0,\infty)$, and this number is increasing in n, with limit ∞.

5.3. Decomposition at the maximum

Theorem 4.2 provides tools for analysing the behaviour of the random walk $S^{[0,n(q)]}$ before and after the time it achieves its maximum. By conditioning on $n(q) = n$, we can then do the same for $S^{[0,n]}$. The key idea is that by taking the faces of the concave majorant that have positive slope we get only those faces that lie in the region up to where the random walk achieves its maximum, and by taking the faces with negative slope we get only those faces that lie in the region after the time when the random walk achieves its maximum. This approach was used by Spitzer to find identities involving the maximum of a random walk [19], as indicated in Section 1.

Let $X_1, X_2, \ldots$ be a sequence of independent random variables with common continuous distribution, and let $S_0 = 0$ and $S_j = \sum_{i=1}^{j} X_i$ for $j \geqslant 1$. Let $S^{[0,n]} = \{(j, S_j) : 0 \leqslant j \leqslant n\}$ and $S^{[0,n(q)]} = \{(j, S_j) : 0 \leqslant j \leqslant n(q)\}$. Let L_n be the almost surely unique time at which $S^{[0,n]}$ achieves its maximum, and let the value of the maximum be M_n. Let F_n denote the number of faces of the concave majorant of the walk $S^{[0,n]}$, with the convention $F_0 = 0$, and let $(N_{n,i}, \Delta_{n,i})$

https://doi.org/10.1017/S0963548311000307 Published online by Cambridge University Press

denote the length and increment associated with the ith of these faces. We make similar definitions when n is randomized to $n(q)$.

Theorem 5.5. *$(L_{n(q)}, M_{n(q)})$ and $(n(q) - L_{n(q)}, S_{n(q)} - M_{n(q)})$ are independent and both have compound Poisson distributions.*

As discussed in Section 1, the compound Poisson nature of $M_{n(q)}$ and $S_{n(q)} - M_{n(q)}$, and their independence, was discovered by Greenwood and Pitman [9], but this section gives a more explicit explanation of their distribution.

Proof. By construction,

$$\Delta_{n,i} = S_{N_{n,1}+\cdots N_{n,i-1}+N_{n,i}} - S_{N_{n,1}+\cdots N_{n,i-1}}$$

and

$$(L_n, M_n) = \sum_{i=1}^{K_n} (N_{n,i}, \Delta_{n,i}) 1(\Delta_{n,i} > 0),$$
$$(n - L_n, S_n - M_n) = \sum_{i=1}^{K_n} (N_{n,i}, \Delta_{n,i}) 1(\Delta_{n,i} \leqslant 0).$$

From Theorem 4.2 the $(N_{n(q),i}, \Delta_{n(q),i})$ are the points of a Poisson point process on $\{1, 2 \ldots\} \times \mathbb{R}$ with intensity $j^{-1} q^j \mathbb{P}(S_j \in dx), j \in \{1, 2, \ldots\}, x \in \mathbb{R}$, and thus the conclusion follows. □

In the special case where $\mathbb{P}(S_j > 0)$ is constant for $1 \leqslant j \leqslant n$, by conditioning on the event $n(q) = n$ and $L_{n(q)} = \ell$ we can deduce results about the concave majorant of $S^{[0,n]}$ either side of its maximum.

Theorem 5.6. *Let $X_1, \ldots, X_n$ be independent with common continuous distribution. Let $S_0 = 0$ and $S_j = \sum_{i=1}^{j} X_i$ for $1 \leqslant j \leqslant n$, and let $S^{[0,n]} = \{(j, S_j) : 0 \leqslant j \leqslant n\}$. Suppose that $\mathbb{P}(S_j > 0) = p_+$ for $1 \leqslant j \leqslant n$. Then conditionally given $L_n := Theorem \max_{0 \leqslant j \leqslant n} S_j = \ell$, the partition generated by the lengths of the faces of the concave majorant of $S^{[0,n]}$ on the interval $[0, \ell]$ is distributed according to the Ewens sampling formula with parameter p_+. That is, if A_j^+ is the number of faces of the concave majorant with positive slope of length j, then for any $\{a_j : j \geqslant 1\}$ such that $\sum_j j a_j = \ell \leqslant n$,*

$$\mathbb{P}(A_j^+ = a_j, j \geqslant 1 | L_n = \ell) = \frac{\Gamma(p_+) \ell!}{\Gamma(p_+ + \ell)} \prod_{j=1}^{\ell} \frac{(p_+)^{a_j}}{j^{a_j} a_j!}. \tag{5.5}$$

The partition generated by the lengths of the faces of the concave majorant of $S^{[0,n]}$ on the interval $[\ell, n]$ is also distributed according to the Ewens sampling formula, but with parameter $p_- = 1 - p_+$.

Proof. Let $A_{n(q),j}^+$ be the number of faces of the concave majorant of $S^{[0,n(q)]}$ with positive slope of length j. From the proof of Theorem 5.5 it is easy to see that $A_{n(q),j}^+$ has a Poisson distribution with parameter $j^{-1}q^jp_-$, independently for each j, and independently of $S^{[0,n(q)]}$ after time $L_{n(q)}$. Thus, for any $\{a_j : j \geqslant 1\}$ such that $\sum_j ja_j = \ell$,

$$\begin{aligned}\mathbb{P}(A_j^+ = a_j, j \geqslant 1 | L_n = \ell) &= \mathbb{P}(A_{n(q),j}^+ = a_j, j \geqslant 1 | L_{n(q)} = \ell, n(q) = n) \\ &= \mathbb{P}(A_{n(q),j}^+ = a_j, j \geqslant 1 | L_{n(q)} = \ell) \\ &= \frac{\mathbb{P}(A_{n(q),j}^+ = a_j, j \geqslant 1)}{\mathbb{P}(L_{n(q)} = \ell)} \\ &= \frac{\prod_j \frac{(p_+)^{a_j} q^{ja_j}}{j^{a_j} a_j!} \exp\{-\frac{p_+ q^j}{j}\}}{\mathbb{P}(L_{n(q)} = \ell)}. \end{aligned} \tag{5.6}$$

Under the assumption $\mathbb{P}(S_j > 0) = p_+$ for $1 \leqslant j \leqslant n$, it is known [5, Chapter XII, (8.12)] that for the random walk $S^{[0,n]}$, the almost surely unique index L_n such that $S_{L_n} = \max_{0 \leqslant j \leqslant n} S_j$ has the beta-binomial distribution

$$\mathbb{P}(L_n = \ell) = (-1)^n \binom{p_- - 1}{\ell} \binom{p_+ - 1}{n - \ell} \quad (0 \leqslant \ell \leqslant n),$$

which is the mixture of binomial(n, p) distributions for p with beta(p_+, p_-) distribution on $[0, 1]$. Thus

$$\mathbb{P}(L_{n(q)} = \ell) = \frac{\Gamma(p_+ + \ell) q^\ell (1-q)^{p_+}}{\Gamma(p_+) \ell!}.$$

Thus (5.6) reduces to (5.5). The partition after the maximum is proved similarly. □

6. The general case

Let $S_j = \sum_{i=1}^j X_i$ for $1 \leqslant j \leqslant n$, where $X_1, X_2, \ldots$ is a sequence of exchangeable random variables. Let $S^{[0,n]} = \{(j, S_j) : 1 \leqslant j \leqslant n\}$, and let $\bar{C}^{[0,n]}$ be the concave majorant of $S^{[0,n]}$. The concave majorant in this case, where there may some subsets of $X_1, \ldots, X_n$ that have the same arithmetic mean, is less well studied. However, the literature does contain some results for the case where $X_1, X_2, \ldots$ are also assumed to be independent.

Sparre Andersen [18] introduced the random variable H_n, the number of $1 \leqslant j \leqslant n$ such that $S_j = \bar{C}^{[0,n]}(j)$, and F_n, the number of faces of the concave majorant, *i.e.*, the number of distinct slopes in the concave majorant (note that Andersen uses K_n instead of F_n, but we will always use K_n to represent the number of cycles in a random permutation of $[n]$). Figure 2 shows an example of a random walk with $F_n = 3$ and $H_n = 8$. Clearly, $F_n \leqslant H_n$, and in the case of continuous distributions we have $F_n = H_n$ almost surely. Sparre Andersen derived the generating function

$$H(s,t) := \sum_{n=0}^{\infty} \sum_{m=0}^{n} \mathbb{P}(H_n = m) s^n t^m \tag{6.1}$$

for all distributions of X_1. As will be shown in Theorem 6.5, the theory presented in this section provides a powerful new method of deriving this formula, and in addition a formula for a similar generating function involving F_n.

Sherman [16] introduced a further variable J_n relating to the concave majorant with $H_n \leqslant J_n \leqslant F_n$. Sherman deduces a Spitzer identity which relates the generating functions of J_n and

https://doi.org/10.1017/S0963548311000307 Published online by Cambridge University Press

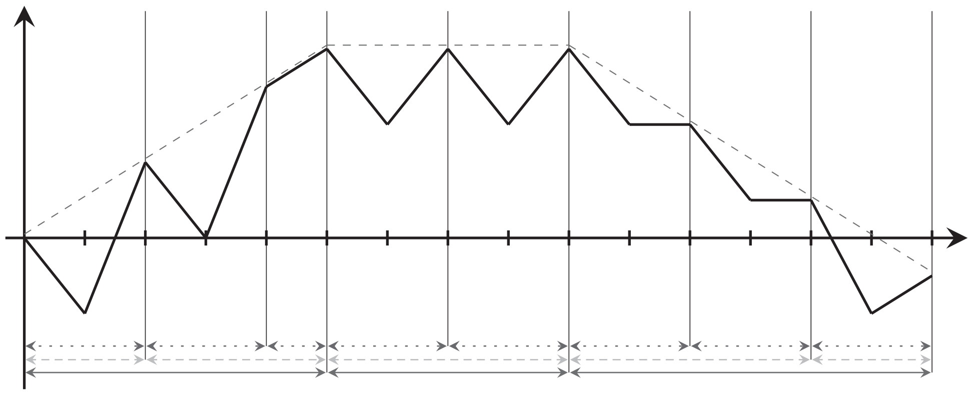

Figure 2. An example of a random walk with non-continuous increment distribution for $n = 15$, with $F_n = 3$ and $H_n = 8$. The concave majorant is shown with a dashed line. The compositions induced by the excursion lengths and face lengths are fixed by the values of the walk, and an example of a possible composition induced by the lengths of the chords associated with the partition segments is shown. The compositions going from top to bottom are $\Xi^H_{[0,15]} = (2,2,1,2,2,2,2,2,)$, $\Xi^K_{[0,15]} = (2,3,4,4,2)$ and $\Xi^F_{[0,15]} = (5,4,6)$.

Φ_n, the periodicity of $(X_1,\dots,X_n)$, that is, the maximal number ϕ such that $(X_1,\dots,X_n) = (X_1,\dots,X_{n/\phi}, \dots ,X_1,\dots,X_{n/\phi})$.

In this section it will be important to make a distinction between *excursions*, *segments* and *faces*, and between their associated compositions of n. The following definitions are illustrated in Figure 2.

- An *excursion* is a section of a walk between two integer-valued times with the property that the walk touches its concave majorant at the end points of the excursion but lies strictly below it between the end points. The number of distinct excursions of $S^{[0,n]}$ is equal to H_n. Let $\Xi^H_{[0,n]}$ be the composition of n induced by the lengths of the excursions of $S^{[0,n]}_\rho$, the transformed walk of Theorem 1.1. Although this has the same distribution as the composition induced by the lengths of the excursions of $S^{[0,n]}$, the forthcoming discussion about *segment* compositions only makes sense for $S^{[0,n]}_\rho$. We say that the *slope* of an excursion is the slope of the line joining its start and end points.
- A *segment* will always refer to one segment of a partition. That is, if $(n_1,\dots,n_k)$ is a partition of n, then we say it has k segments with associated lengths $n_1,\dots,n_k$. As we described in the Introduction, to generate a walk with the law of $S^{[0,n]}$ whilst simultaneously getting information about its concave majorant, *i.e.*, to generate $S^{[0,n]}_\rho$, we first choose a random partition induced by the cycle lengths of a uniform random permutation. If we are just interested in the concave majorant of $S^{[0,n]}_\rho$, then we only need to associate a slope with each segment of that partition and then arrange the segments in order of non-increasing slope, where the ordering of any segments with the same slope is chosen uniformly randomly. Keeping track of the end points of the segments results in another induced composition of n, which we call $\Xi^K_{[0,n]}$. This composition arises from our construction and cannot be read off from a given random walk.
- A *face* will mean one face of the concave majorant. The number of distinct faces is equal to F_n. Let $\Xi^F_{[0,n]}$ be the composition of n induced by the lengths of the faces of $S^{[0,n]}_\rho$. Again,

https://doi.org/10.1017/S0963548311000307 Published online by Cambridge University Press

this has the same distribution as the composition of n induced by the lengths of the faces of $S^{[0,n]}$.

- The terms *excursion block*, *segment block* and *face block* will mean blocks of the compositions $\Xi^H_{[0,n]}$, $\Xi^K_{[0,n]}$ and $\Xi^F_{[0,n]}$ respectively, where for example the *blocks* of the composition $(3,4,1)$ of 8 in order are defined to be $[0,3]$, $[3,7]$ and $[7,8]$. The slope associated with any block $[a,b]$ is defined by $(S^\rho_b - S^\rho_a)/(b-a)$.

Since the values of any walk on $[0,n]$ between two vertices of its concave majorant, *i.e.*, between the start and end points of some face, are composed of one or many consecutive excursions, $\Xi^H_{[n]}$ is some refinement of $\Xi^F_{[n]}$, which we write as $\Xi^H_{[n]} \preceq \Xi^F_{[n]}$. For $S^{[0,n]}_\rho$ constructed as in Theorem 1.1, define H^ρ_n and F^ρ_n similarly to H_n and F_n, and note that $H_n \stackrel{d}{=} H^\rho_n$ and $F_n \stackrel{d}{=} F^\rho_n$. Recall that K_n is the number of segments in the partition chosen at the beginning of the construction. We will have $H^\rho_n \leqslant K_n \leqslant F^\rho_n$, and moreover $\Xi^K_{[0,n]}$ will be such that $\Xi^H_{[0,n]} \preceq \Xi^K_{[0,n]} \preceq \Xi^F_{[0,n]}$. We will discuss these nested compositions further after proving Theorem 1.1 in the general case.

Proof of Theorem 1.1. As in the proof of Theorem 1.1 under assumption **A**, it is enough to show that if $X_1,\dots,X_n$ are samples without replacement from a list $x_1,\dots,x_n$ of real numbers, where now each number is labelled but no longer necessarily distinct in value, then

$$\mathbb{P}(X_{\rho(1)} = x_1,\dots,X_{\rho(n)} = x_n) = \frac{1}{n!}.$$

Let $x = (x_1,\dots,x_n)$, and suppose this is fixed throughout the proof of the theorem. Let $\bar{c}^{[0,n]}$ be the concave majorant of the deterministic walk with increments $x_1,\dots,x_n$. Some notation and a couple of combinatorial lemmas are needed before continuing.

For any $n \in \mathbb{N}$, let $\mathcal{N}_n$ be the set of all compositions of n. Let $f \in \mathbb{N}$, $h \in \mathbb{N}$ and $(v_1,\dots,v_f) \in \mathcal{N}_h$. Let $\mathcal{N}_{(v_1,\dots,v_f),(k_1,\dots,k_f)}$ be the set

$$\Big\{ \big(h_1,\dots,h_{\sum_{i=1}^f k_i}\big) \in \mathcal{N}_h : \big(h_{\sum_{i=1}^{j-1} k_i},\dots,h_{\sum_{i=1}^{j} k_i}\big) \in \mathcal{N}_{v_j} \ \text{ for } 1 \leqslant j \leqslant f \Big\}.$$

Thus an element of $\mathcal{N}_{(v_1,\dots,v_f),(k_1,\dots,k_f)}$ is a composition of h formed by joining together compositions of $v_1,\dots,v_f$ which contain $k_1,\dots,k_f$ blocks respectively (so $\mathcal{N}_{(v_1,\dots,v_f),(k_1,\dots,k_f)}$ may be an empty set for some values of $(k_1,\dots,k_f)$).

Lemma 6.1. *Let $f \in \mathbb{N}$, $h \in \mathbb{N}$ and $(v_1,\dots,v_f) \in \mathcal{N}_h$. Then*

$$\sum_{k=f}^{h} \sum_{(k_1,\dots,k_f)\in\mathcal{N}_k} \sum_{(h_1,\dots,h_k)\in\mathcal{N}_{(v_1,\dots,v_f),(k_1,\dots,k_f)}} \prod_{i=1}^{k} \frac{1}{k_1!\cdots k_f!}\frac{1}{h_1\cdots h_k} = 1. \tag{6.2}$$

Proof. The numbers that are being summed over bear a strong resemblance to the unsigned Stirling numbers of the first kind $|S(n,k)|$, which enumerate the number of permutations of n with k cycles. Using this as a guide, consider a set A consisting of permutations of $v_1,\dots,v_f$, where permutations corresponding to v_i and v_j with $i \neq j$ are considered distinct even if they are identical. The number of such sets where, for each $1 \leqslant j \leqslant f$, the permutation of v_j has k_j cycles

https://doi.org/10.1017/S0963548311000307 Published online by Cambridge University Press

of sizes $h_{\sum_{i=1}^{j-1} k_i}, \dots, h_{\sum_{i=1}^{j} k_i}$ is

$$\frac{v_1! \cdots v_f!}{k_1! \cdots k_f! \cdot h_1 \cdots h_k}.$$

Since the total number of elements of A is $v_1! \cdots v_f!$, and the summation in (6.2) simplifies to be the sum over the subsets of A such that, for each $1 \leqslant j \leqslant f$, the permutation of v_j has k_j cycles of size $h_{\sum_{i=1}^{j-1} k_i}, \dots, h_{\sum_{i=1}^{j} k_i}$, the value of the sum must be 1. □

Let $f(\bar{c}^{[0,n]})$ be the number of faces of $\bar{c}^{[0,n]}$, and let $\ell_1(\bar{c}^{[0,n]}), \dots, \ell_{f(\bar{c}^{[0,n]})}(\bar{c}^{[0,n]})$ be the lengths of those faces, arranged in the order those faces appear in $\bar{c}^{[0,n]}$. Let $\mathcal{N}(\bar{c}^{[0,n]})$ be the set

$$\left\{(n_1, \dots, n_k) \in \mathcal{N}_n : \exists\, k_1 < \cdots < k_{f(\bar{c}^{[0,n]})} \quad \text{s.t.} \sum_{i=k_{j-1}}^{k_j} n_i = \ell_j(\bar{c}^{[0,n]}),\ 1 \leqslant j \leqslant f(\bar{c}^{[0,n]})\right\}.$$

Loosely, $\mathcal{N}(\bar{c}^{[0,n]})$ is the set of possible values for $\Xi^K_{[0,n]}$ conditionally given that the concave majorant of $S^{[0,n]}_\rho$ is $\bar{c}^{[0,n]}$. For $(n_1, \dots, n_k) \in \mathcal{N}(\bar{c}^{[0,n]})$, let

$$\{k_j(n_1, \dots, n_k), 1 \leqslant j \leqslant f(\bar{c}^{[0,n]})\} = \left\{(k_1, \dots, k_{f(\bar{c}^{[0,n]})}) : \sum_{i=k_{j-1}}^{k_j} n_i = \ell_j(\bar{c}^{[0,n]})\right\}.$$

Then $k_j(\Xi^K_{[0,n]})$ represents the number of blocks of $\Xi_{[0,n]}$ that lie in the jth face block, *i.e.*, in the jth block of $\Xi^F_{[0,n]}$. Finally, let

$$\mathcal{N}_x(\bar{c}^{[0,n]}) = \left\{(n_1, \dots, n_k) \in \mathcal{N}(\bar{c}^{[0,n]}) : \sum_{j=1}^{n_i} x_j = \bar{c}^{[0,n]}(n_i) \ \text{ for } 1 \leqslant i \leqslant k\right\}.$$

Then $\mathcal{N}_x(\bar{c}^{[0,n]})$ is the set of possible values for $\Xi^K_{[0,n]}$ conditionally given that $\{X_{\rho(i)} = x_i : 1 \leqslant i \leqslant n\}$.

Lemma 6.2. *For every composition $(n_1, \dots, n_k) \in \mathcal{N}_x(\bar{c}^{[0,n]})$, for $1 \leqslant i \leqslant k$ let*

$$h_i(x, n_1, \dots, n_k) = \#\left\{j : n_1 + \cdots + n_{i-1} < j \leqslant n_1 + \cdots + n_i,\ \sum_{l=1}^{j} x_l = \bar{c}^{[0,n]}(j)\right\}.$$

Then

$$\sum_{k=1}^{n} \sum_{(n_1,\dots,n_k) \in \mathcal{N}_x(\bar{c}^{[0,n]})} \left(\prod_{i=1}^{k} \frac{1}{h_i(x, n_1, \dots, n_k)}\right)\left(\prod_{j=1}^{f(\bar{c}^{[0,n]})} \frac{1}{k_j(n_1, \dots, n_k)!}\right) = 1. \tag{6.3}$$

Proof. Let $h = \#\{j : 1 \leqslant j \leqslant n, \sum_{l=1}^{j} = \bar{c}^{[0,n]}(j)\}$, and for $1 \leqslant i \leqslant f(\bar{c}^{[0,n]})$ let

$v_i(x) =$

$$\#\left\{j : \ell_1(\bar{c}^{[0,n]}) + \cdots + \ell_{i-1}(\bar{c}^{[0,n]}) < j \leqslant \ell_1(\bar{c}^{[0,n]}) + \cdots + \ell_i(\bar{c}^{[0,n]}),\ \sum_{l=1}^{j} x_l = \bar{c}^{[0,n]}(j)\right\}.$$

https://doi.org/10.1017/S0963548311000307 Published online by Cambridge University Press

Associate with each composition $(n_1, \ldots, n_k) \in \mathcal{N}_x(\bar{c}^{[0,n]})$ of length k a composition of h,

$$(h_1(x, n_1, \ldots, n_k), h_2(x, n_1, \ldots, n_k), \ldots, h_k(x, n_1, \ldots, n_k)),$$

so that there is a bijection between the elements of $\mathcal{N}_x(\bar{c}^{[0,n]})$ with k blocks and the set of compositions $(h_1, \ldots, h_k)$ of h with k blocks that are formed by joining together in order compositions of $v_1, \ldots, v_{f(\bar{c}^{[0,n]})}$ which have $k_1, \ldots, k_{f(\bar{c}^{[0,n]})}$ blocks respectively. Thus the term on the left-hand side of (6.3) is

$$\sum_{k=f}^{h} \sum_{(k_1,\ldots,k_{f(\bar{c}^{[0,n]})}) \in \mathcal{N}_k} \sum_{(h_1,\ldots h_k) \in \mathcal{N}_{(v_1,\ldots,v_f),(k_1,\ldots,k_f)}} \prod_{i=1}^{k} \frac{1}{k_1! \cdots k_f!} \frac{1}{h_1 \cdots h_k},$$

which by Lemma 6.1 is 1. □

Fix a composition $(n_1, \ldots, n_k)$ of n. For $1 \leqslant j \leqslant n$, let $I_j = \{i : n_i = j\}$ and let $a_j = |I_j|$. Following the construction of $S_\rho^{[0,n]}$ described in the Introduction, we see that the event $\{\Xi_{[0,n]}^K = (n_1, \ldots, n_k)$ and $X_{\rho(\ell)} = x_\ell, 1 \leqslant \ell \leqslant n\}$ occurs if and only if:

(i) $L_{n,1}, \ldots, L_{n,K_n}$ is $(n_1, \ldots, n_k)$ in non-increasing order;

(ii) for each $1 \leqslant j \leqslant n$, for each $i \in I_j$ the ordered list $(X_{n_1+\cdots+n_{i-1}+1}, \ldots, X_{n_1+\cdots+n_i})$ is one of the $n_i = j$ cyclic permutations of the ordered list

$$(x_{m_1+m_2+\cdots+m_{\tau(i')-1}+1}, \ldots, x_{m_1+m_2+\cdots+m_{\tau(i')}})$$

for some $i' \in I_j$;

(iii) for each $1 \leqslant j \leqslant n$, for each $i \in I_j$ the cyclic permutation that is chosen for the ordered list of increments $(X_{n_1+\cdots+n_{i-1}+1}, \ldots, X_{n_1+\cdots+n_i})$ is the unique cyclic permutation that results in the ordered list becoming exactly

$$(x_{m_1+m_2+\cdots+m_{\tau(i')-1}+1}, \ldots, x_{m_1+m_2+\cdots+m_{\tau(i')}});$$

(iv) for each $1 \leqslant j \leqslant f(\bar{c})^{[0,n]})$, the ordering of the $k_j(n_1, \ldots, n_k)$ segments within the jth face is chosen correctly out of the $k_j!$ possible orderings.

Recall that for $1 \leqslant i \leqslant k$ we have

$$h_i(x, n_1, \ldots, n_k) = \#\left\{ j : n_1 + \cdots + n_{i-1} < j \leqslant n_1 + \cdots + n_i, \sum_{l=1}^{j} x_l = \bar{c}^{[0,n]}(j) \right\}$$

so that in (iii) there are $\prod_{i=1}^k h_i(x, n_1, \ldots, n_k)$ possible choices of combinations of cyclic permutations. Then the probability of the event

$$\{\Xi_{[0,n]}^K = (n_1, \ldots, n_k) \text{ and } X_{\rho(\ell)} = x_\ell, 1 \leqslant \ell \leqslant n\}$$

is

$$\left(\prod_{j=1}^{n} \frac{1}{a_j!} \prod_{i=1}^{k} \frac{1}{n_i}\right) \left(\frac{1}{n!} \prod_{i=1}^{k} n_i \prod_{j=1}^{n} a_j!\right) \left(\prod_{i=1}^{k} \frac{1}{h_i(x, n_1, \ldots, n_k)}\right) \left(\prod_{j=1}^{f(\bar{c}^{[0,n]})} \frac{1}{k_j(n_1, \ldots, n_k)!}\right),$$

where the first two terms should be familiar from the proof of Theorem 1.1 under assumption **A**. Finally, by summing this probability over all possible compositions, we have that the probability of the event $\{X_{\rho(\ell)} = x_\ell, 1 \leqslant \ell \leqslant n\}$ is

$$\frac{1}{n!}\sum_{k=1}^{n}\sum_{(n_1,\dots,n_k)\in\mathcal{N}_x(\bar{c}^{[0,n]})}\left(\prod_{i=1}^{k}\frac{1}{h_i(x,n_1,\dots,n_k)}\right)\left(\prod_{j=1}^{f(\bar{c}^{[0,n]})}\frac{1}{k_j(n_1,\dots,n_k)!}\right) = \frac{1}{n!},$$

where the equality is by Lemma 6.2. This completes the proof of Theorem 1.1. □

In the case where $X_1, X_2, \dots$ are independent, the Poisson point process ideas of Section 4 lead to a simpler description of the concave majorant. For the rest of this section it is assumed that $X_1, X_2, \dots$ is a sequence of independent and identically distributed random variables and $n(q)$ is a geometric variable with parameter $1-q$. Let $S^{[0,n(q)]} = \{(j, S_j) : 0 \leqslant j \leqslant n(q)\}$, where $S_0 = 0$ and $S_j = \sum_{i=1}^{j} X_i$ for $j \geqslant 1$. Let $\bar{C}^{[0,n]}$ be the concave majorant of $S^{[0,n(q)]}$. The following theorem is the extension to the non-continuous increment case of Theorem 4.2.

Theorem 6.3. *If $X_1, X_2, \dots$ are independent with common distribution and $n(q)$ a geometric variable with parameter $1-q$, then the lengths and increments of the faces of the concave majorant of the random walk $S^{[0,n(q)]}$ have the following law. Let $\mathfrak{P}$ be a Poisson point process on $\{1,2,\dots\}\times\mathbb{R}$ with intensity $j^{-1}q^j\mathbb{P}(S_j \in dx)$ for $j = 1,2,\dots, x \in \mathbb{R}$. Note that this process may result in multiple points at the same location. Each point of $\mathfrak{P}$ represents the length and increment of a chord associated with some segment of a partition of $n(q)$. Chords with the same slope are joined together in uniform random order, independently of their lengths, to form the faces of the concave majorant. Moreover, let $K_{n(q)}$ be the total number of chords associated with partition segments, and for $1 \leqslant i \leqslant K_{n(q)}$ let $N_{n(q),i}$ be the length of the ith of these chords once they have been ordered by decreasing slope and uniform randomization of ties. Then the sequence of path segments*

$$\left\{\left(S_{\sum_{l=1}^{i-1}N_{n(q),l}+k} - S_{\sum_{l=1}^{i-1}N_{n(q),l}}, 0 \leqslant k \leqslant N_{n(q),i}\right),\ i = 1,\dots,K_{n(q)}\right\}$$

is a list of the points of a Poisson point process in the space of finite random walk segments

$$\{(s_1,\dots,s_j)\ \textit{for some}\ j = 1,2,\dots\},$$

whose intensity measure on paths of length j, given that $S_k < (k/j)S_j$ for all $1 \leqslant k < j$, is j^{-1} times the conditional distribution of $(S_1,\dots,S_j)$. Again, this Poisson point process may result in multiple points at the same location.

Proof. For any $n \in \mathbb{N}$, conditionally given $n(q) = n$, the projection of the points of $\mathfrak{P}$ onto $\{1,\dots,n\}$ has the law of a partition of n generated by the cycle lengths of a random permutation of $[n]$ by Lemma 4.1. Hence we know from Theorem 1.1 that for every $n \in \mathbb{N}$, conditionally given $n(q) = n$, the process described in the theorem gives the correct law for the concave majorant of $S^{[0,n]}$ and gives the correct law for $\Xi^K_{[0,n]}$, the composition induced by the lengths of the partition segments involved in creating $S^{[0,n]}_\rho$. The remaining assertions follow by independence of the walks associated with each partition segment. □

We now move towards describing the joint law of the nested compositions

$$\Xi^H_{[0,n(q)]} \preceq \Xi^K_{[0,n(q)]} \preceq \Xi^F_{[0,n(q)]}$$

https://doi.org/10.1017/S0963548311000307 Published online by Cambridge University Press

in the case where $X_1, X_2, \ldots$ are independent and the walk has geometric length. The full description of this law will be given in Theorem 6.8 at the end of this section, along with some applications of the theory. Let $S_\rho^{[0,n(q)]}$ be such that, conditionally given $n(q) = n$, $S_\rho^{[0,n(q)]}$ is constructed in the same way as $S_\rho^{[0,n]}$ in Theorem 1.1, and let $\bar{C}_\rho^{[0,n(q)]}$ be the concave majorant of $S_\rho^{[0,n]}$. We begin by describing the laws of $H_{n(q)}$, $K_{n(q)}$ and $F_{n(q)}$, which are defined to be the number of excursions, segments and faces, respectively, of $\bar{C}_\rho^{[0,n(q)]}$.

We need some new notation, some of which is taken from Sparre Andersen [18]. Let $x_1, x_2, \ldots$ be an enumeration of the set of real numbers x for which $\mathbb{P}(S_k = kx)$ is positive for some $k > 0$, and let

$$\mu_j(q) = \sum_{k=1}^{\infty} k^{-1} q^k \mathbb{P}(S_k = kx_j), \quad \text{for } j = 1, 2, \ldots$$

$$\mu_0(q) = \sum_{k=1}^{\infty} k^{-1} q^k \mathbb{P}(S_k \neq kx_j \ \text{ for } j = 1, 2, \ldots)$$

$$= -\log(1-q) - \sum_{j=1}^{\infty} \mu_j(q).$$

Proposition 6.4. *Let $H_{q,j}$, $K_{q,j}$ and $F_{q,j}$ be the number of excursions, segments and faces in $\bar{C}_\rho^{[0,n(q)]}$ of slope x_j for $j \geqslant 1$. Then for each $j \geqslant 1$:*

(i) *$H_{q,j}$ is a geometric random variable with parameter $\exp(-\mu_j(q))$, independently of $\{H_{q,i} : i \neq j\}$;*
(ii) *$K_{q,j}$ is a Poisson random variable with parameter $\mu_j(q)$, independently of $\{K_{q,i} : i \neq j\}$;*
(iii) *$F_{q,j}$ is a Bernoulli random variable with parameter $1 - \exp(-\mu_j(q))$, independently of $\{F_{q,i} : i \neq j\}$.*

Let $H_{q,0}$, $K_{q,0}$ and $F_{q,0}$ be the number of excursions, segments and faces with slope not equal to x_j for any $j \geqslant 1$. Then:

(iv) *$H_{q,0} = K_{q,0} = F_{q,0}$ almost surely, and their common distribution is Poisson with parameter $\mu_0(q)$, independently of $\{H_{q,j}, K_{q,j}, F_{q,j} : j \geqslant 1\}$.*

Proof. Here (ii) follows from Theorem 6.3, (iii) is implied by (ii), since a face of slope x exists if and only if there is at least one segment of slope x, and (iv) is also implied by Theorem 6.3 since it concerns the restriction of the Poisson point process to slopes which have zero probability, as in the case of continuous increment distributions.

Fix $j \geqslant 1$. (ii) implies that $\mathbb{P}(H_{q,j} \geqslant 1) = \mathbb{P}(K_{q,j} \geqslant 1) = 1 - \exp(-\mu_j(q))$. Given that there are at least n excursions of slope x_j, by the memoryless property of the geometric distribution of $n(q)$, the law of the remaining values of the walk $S_\rho^{[0,n(q)]}$ is the same as the law of a walk generated by the Poisson process of path segments in Theorem 6.3 but thinned to only include segments with slope $x \geqslant x_j$. Thus

$$\mathbb{P}(H_{q,j} \geqslant n+1 | H_{q,j} \geqslant n) = \mathbb{P}(K_{q,j} \geqslant 1) = 1 - \exp(-\mu_j(q)),$$

which proves (i). □

https://doi.org/10.1017/S0963548311000307 Published online by Cambridge University Press

Theorem 6.5. *Let H_n and F_n be the number of excursions and faces for $S^{[0,n]}$, and let K_n be the number of segments for $S_\rho^{[0,n]}$. Then, for $0 \leqslant s,t \leqslant 1$,*

$$H(s,t) = e^{t\mu_0(s)} \prod_{j=1}^{\infty} \frac{1}{1-t+te^{-\mu_j(s)}},$$

$$K(s,t) = e^{t\mu_0(s)} \prod_{j=1}^{\infty} e^{t\mu_j(s)} = (1-s)^{-t},$$

$$F(s,t) = e^{t\mu_0(s)} \prod_{j=1}^{\infty} (1-t+te^{\mu_j(s)}).$$

The generating function of $G_{K_n}(z) = \sum_{m=1}^{\infty} z^m \mathbb{P}(K_n = m)$ is well-known from the equality in (1.1). $H(s,t)$ is as in (6.1) and agrees with Sparre Andersen's formula [18, Theorem 2].

Proof. Recall first that $H_n^\rho \stackrel{d}{=} H_n$ and $F_n^\rho \stackrel{d}{=} F_n$. Let $n(s)$ be a geometric random variable with parameter $1-s$ and consider the walk of $n(s)$ steps. We have by definition

$$H_{n(s)} = H_{s,0} + \sum_{j=1}^{\infty} H_{s,j}.$$

Thus the generating function of $H_{n(s)}$ is the product of the generating functions of $H_{s,0}$ and $H_{s,j}$, $j \geqslant 1$. These are known from Proposition 6.4, thus

$$\begin{aligned}\sum_{m=0}^{\infty} t^m \mathbb{P}(H_{n(s)} = m) &= e^{(t-1)\mu_0(s)} \prod_{j=1}^{\infty} \frac{e^{-\mu_j(s)}}{1-t+te^{-\mu_j(s)}} \\ &= (1-s)e^{t\mu_0(s)} \prod_{j=1}^{\infty} \frac{1}{1-t+te^{-\mu_j(s)}}.\end{aligned}$$

We can conclude that

$$\begin{aligned}H(s,t) &= \sum_{n=0}^{\infty} \sum_{m=0}^{n} \mathbb{P}(H_n = m) s^n t^m \\ &= (1-s)^{-1} \sum_{m=0}^{\infty} t^m \sum_{n=m}^{\infty} (1-s) s^n \mathbb{P}(H_n = m) \\ &= (1-s)^{-1} \sum_{m=0}^{\infty} t^m \mathbb{P}(H_{n(s)} = m) \\ &= e^{t\mu_0(s)} \prod_{j=1}^{\infty} \frac{1}{1-t+te^{-\mu_j(s)}}.\end{aligned}$$

The deduction for $F(s,t)$ is similar, and, as already mentioned, $K(s,t)$ is well known. □

In order to fully describe the joint law of the nested compositions, two more lemmas are necessary. The first contains information about the lengths of each segment or excursion, and the

https://doi.org/10.1017/S0963548311000307 Published online by Cambridge University Press

second describes how many excursions there are in each segment. We already know from the Poissonian description of the concave majorant the distribution of the number of segments with a given slope, and thus we already know the distribution of the number of segments within each face (see Theorem 6.8 for the full description).

Lemma 6.6. *Consider the walk of $n(q)$ steps. For $j \geqslant 1$, conditionally given $K_{q,j} = k_{q,j}$, let $L^K_{q,j,1}, \ldots, L^K_{q,j,k_{q,j}}$ be the lengths of the $k_{q,j}$ segments of $S_\rho^{[0,n(q)]}$ of slope x_j. Then the $L^K_{q,j,1}, \ldots, L^K_{q,j,k_{q,j}}$ are independent from each other and the lengths of all other segments. Moreover, they are identically distributed with common probability generating function $G_{L^K_{q,j}}(z) = \mu_j(zq)/\mu_j(q)$.*

For $j \geqslant 1$, conditionally given $H_{q,j} = h_{q,j}$, let $L^H_{q,j,1}, \ldots, L^H_{q,j,h_{q,j}}$ be the lengths of the $h_{q,j}$ excursions of $S_\rho^{[0,n(q)]}$ of slope x_j. Then $L^K_{q,j,1}, \ldots, L^K_{q,j,h_{q,j}}$ are independent from each other and the lengths of all other segments. Moreover they are identically distributed with common probability generating function $G_{L^H_{q,j}}(z) = (1 - e^{-\mu_j(zq)})/(1 - e^{-\mu_j(q)})$.

Furthermore, each excursion in the face of slope x_j is independent and has the law of a random walk with increment distribution X_1 conditioned on making its first return to the line through the origin with slope x_j before $n(q)$, an independent geometric random variable with parameter $1 - q$, and remaining below that line before its first return time; the excursion is taken to be that walk up to the time of its first return to the line with slope x_j.

Proof. By Poisson process properties, each $L^K_{q,j,1}, \ldots, L^K_{q,j,h_{q,j}}$ are independent from each other and the lengths of all other segments. By Poisson thinning,

$$\mathbb{P}(L^K_{q,j,1} = l) = l^{-1} q^l \mathbb{P}(S_k = kx_j),$$

which gives the claimed generating function.

By the memoryless property of the geometric distribution of $n(q)$, each excursion of slope x_j is independent, and is clearly independent from all excursions of other slopes. This gives the final assertion of the lemma. By considering the total lengths of the face with slope x_j, we see that

$$\sum_{i=1}^{H_{q,j}} L^H_{q,j,i} = \sum_{i=1}^{K_{q,j}} L^K_{q,j,i}.$$

By comparing the generating functions of both sides and using Proposition 6.4 we can deduce the claimed generating function $G_{L^H_{q,j}}(z)$. □

Lemma 6.7. *Conditionally given there are $k_{q,j}$ segments of $S_\rho^{[0,n(q)]}$ of slope x_j, let*

$$E_{q,j,1}, \ldots, E_{q,j,k_{q,j}}$$

be the number of excursions in each of those $k_{n(q)}$ segments. Then $E_{q,j,1}, \ldots, E_{q,j,k_{q,j}}$ are independent of each other and all other excursions, and are identically distributed. Their common distribution is the log-series distribution with parameter $1 - e^{-\mu_j(q)}$, that is,

$$\mathbb{P}(E_{q,j,1} = i) = \frac{(1 - e^{-\mu_j(q)})^i}{i\mu_j(q)}, \quad i = 1, 2, \ldots.$$

https://doi.org/10.1017/S0963548311000307 Published online by Cambridge University Press

Proof. By Theorem 6.3 the values of the walk $S_\rho^{[0,n(q)]}$ over each segment are independent, which gives the independence of $E_{q,j,1},\dots,E_{q,j,k_{q,j}}$. By the independence of the excursions in the face of slope x_j and the independence of the walks over each segment of slope x_j, $L^H_{q,j,1},\dots,L^H_{q,j,E_{q,j}}$ are independent and identically distributed. By considering the total length of each segment of slope x_j, we have the identity in distribution

$$L^K_{q,j,1} \stackrel{d}{=} \sum_{i=1}^{E_{q,j,1}} L^H_{q,j,1},$$

which after applying generating function analysis reveals that

$$G_{E_{q,j,1}}(z) := \sum_{l=1}^{\infty} z^i \mathbb{P}(E_{q,j,1}=i) = \sum_{i=1}^{\infty} z^i \frac{(1-e^{-\mu_j(q)})^i}{i\mu_j(q)}. \qquad \square$$

We are now ready to describe the joint law of the three nested compositions $\Xi^H_{[0,n(q)]} \preceq \Xi^K_{[0,n(q)]} \preceq \Xi^F_{[0,n(q)]}$. The following theorem is a summary of most of the information from Theorem 6.3 to Lemma 6.7.

Theorem 6.8. *Let $n(q)$ be a geometric random variable with parameter $1-q$. Let $X_1,X_2,\dots$ be independent and identically distributed. Let $S_j=\sum_{i=1}^j X_i$ for $j \geqslant 1$. Let $x_1,x_2,\dots$ be an enumeration of the set of real numbers x for which $\mathbb{P}(S_k=kx)$ is positive for some $k>0$, and for $j\geqslant 1$ let*

$$\mu_j(q) = \sum_{k=1}^{\infty} k^{-1} q^k \mathbb{P}(S_k = kx_j).$$

Let $S_\rho^{[0,n(q)]}$ be such that, conditionally given $n(q)=n$, $S_\rho^{[0,n(q)]}$ is constructed in the same way as $S_\rho^{[0,n]}$ in Theorem 1.1. Let $\bar{C}_\rho^{[0,n(q)]}$ be the concave majorant of $S_\rho^{[0,n(q)]}$. Then, independently for each $j\geqslant 1$, we have the following.

- *There is a face of $\bar{C}_\rho^{[0,n(q)]}$ with slope x_j with probability $1-e^{-\mu_j(q)}$.*
- *Conditionally given there is a face of slope x_j, the number of blocks of $\Xi^K_{[0,n]}$ with associated slope x_j has the Poisson distribution with parameter $\mu_j(q)$, conditional on the value being at least one.*
- *Conditionally given there are $k_{q,j}$ blocks of $\Xi^K_{[0,n]}$ with associated slope x_j, the number of excursion blocks in each of the $k_{q,j}$ segment blocks has the log-series distribution with parameter $1-e^{-\mu_j(q)}$, independently for each segment.*
- *The length of each excursion of slope x_j is independent of all other excursions and has distribution with generating function*

$$G_{L^H_{q,j}}(z) = (1-e^{-\mu_j(zq)})/(1-e^{-\mu_j(q)}).$$

Any face block with associated slope x such that $x\neq x_j$ for any $j\geqslant 1$ will be composed of exactly one segment block, which will also be composed of exactly one excursion block. The lengths and increments of faces with slope x such that $x\neq x_j$ for any $j\geqslant 1$ form a Poisson point process on

$\{1,2,\dots\}\times\mathbb{R}$ *with intensity* $i^{-1}\mathbb{P}(S_i\in ds)$ *for* $i\geqslant 1, s\in\mathbb{R}$, *but restricted to the region*

$$\{(i,s)\in\{1,2,\dots\}\times\mathbb{R}\ :\ s\neq ix_j\ \text{for any } j\geqslant 1\}.$$

Three nested compositions with the joint law of $\Xi^H_{[0,n(q)]}$, $\Xi^K_{[0,n(q)]}$ *and* $\Xi^F_{[0,n(q)]}$ *are created by uniformly randomly ordering the excursions within each segment, uniformly randomly ordering the segments within each face, arranging the faces in order of decreasing slope, and then looking at the induced compositions of excursion blocks, segment blocks and face blocks.*

Theorem 6.8 implies that the compositions $\Xi^H_{[0,n(q)]}\preceq\Xi^K_{[0,n(q)]}\preceq\Xi^F_{[0,n(q)]}$ can be generated by nested renewal processes on $\mathbb{N}$ that terminate at some geometric time. There would be three types of renewal epochs. The first would be when a new face block started, which implies a new segment block and excursion block would also start. The second would be when only a new segment block and excursion block started, and the third would be when only a new excursion block started. Unlike in previous investigations into nested renewal sequences [1, 4], the distributions of the length until the next renewal may change with time, and after a renewal has occurred, the number of future renewals may depend on how many have already occurred.

Theorem 6.8 allows us to readily compute the probability of many fluctuation events for $S^{[0,n(q)]}$. Some examples are as follows.

- For each $j\geqslant 1$, the probability that $\bar{C}^{[0,n(q)]}$ consists of only one face of slope x_j is $(1-q)^{-1}e^{-\mu_j(q)}$.
- The probability that $S^{[0,n(q)]}$ has a unique minimum, *i.e.*, the probability that $\bar{C}^{[0,n(q)]}$ has no face of slope zero, is $\exp[-\sum_{k=1}^{\infty}k^{-1}q^k\mathbb{P}(S_k=0)]$.
- For each $j\geqslant 1$, the expected length of the face of $\bar{C}^{[0,n(q)]}$ of slope x_j is

$$\sum_{k=1}^{\infty}q^k\mathbb{P}(S_k=kx_j).$$

7. $S^{[0,n]}$ conditional on its concave majorant

To complete the rearrangement problem stated in the Introduction, we now give a description of the law of $S^{[0,n]}$ conditional on $\bar{C}^{[0,n]}=\bar{c}^{[0,n]}$. It is a generalization of the well-known Vervaat transform for turning a bridge of a random walk into an excursion [21, Theorem 5]. It relies on first choosing a segment composition $\Xi^K_{[0,n]}$ conditional on $\bar{C}^{[0,n]}_\rho=\bar{c}^{[0,n]}$ and then choosing a walk conditional on $\Xi^K_{[0,n]}$.

Let $\mathrm{Supp}(\bar{C}^{[0,n]})$ be the support of the measure on concave functions on $[0,n]$ that represents the law of $\bar{C}^{[0,n]}$. For any composition $(n_1,\dots,n_k)$ of n, we say that $\sigma\in\Sigma_n$ is an $(n_1,\dots,n_k)$-cyclic permutation of $[n]$ if its only action is to cyclically permute the first n_1 elements of $[n]$, cyclically permute the next n_2 elements of $[n]$ and so on. For example, 234175689 is a $(4,3,2)$-cyclic permutation of [9]. Recall that in Section 6 we defined $\mathcal{N}_n$ to be the set of compositions of n, and $\mathcal{N}(\bar{c}^{[0,n]})\subseteq\mathcal{N}_n$ to be the set of possible values of $\Xi^K_{[0,n]}$ conditionally given $\bar{C}^{[0,n]}_\rho=\bar{c}^{[0,n]}$.

Theorem 7.1. *Let* $S_0=0$ *and* $S_j=\sum_{\ell=1}^{j}X_\ell$ *for* $1\leqslant j\leqslant n$, *where* $X_1,\dots,X_n$ *are exchangeable random variables. Let* $S^{[0,n]}=\{(j,S_j):0\leqslant j\leqslant n\}$ *and let* $\bar{C}^{[0,n]}$ *be the concave majorant*

https://doi.org/10.1017/S0963548311000307 Published online by Cambridge University Press

of $S^{[0,n]}$*. Suppose* $\bar{c}^{[0,n]} \in \mathrm{Supp}(\bar{C}^{[0,n]})$*. Let* $q(\cdot)$ *be the probability density function on* $\mathcal{N}_n$*, that is, the regular conditional distribution of* $\Xi_{[0,n]}$ *conditionally given* $\bar{C}^{[0,n]}_\rho = \bar{c}^{[0,n]}$*. Let* $(N_{n,1}, N_{n,2}, \ldots, N_{n,K_n})$ *be a composition of n chosen according to the density function* $q(\cdot)$*, independently of* $\{X_j : 1 \leqslant j \leqslant n\}$*.*

Conditionally given $\{K_n = k\}$ *and* $\{N_{n,i} = n_i : 1 \leqslant i \leqslant k\}$*, let* $Y_1, \ldots, Y_n$ *be random variables, independent of all previously introduced random variables, whose joint law is the regular conditional joint distribution of* $X_1, \ldots, X_n$*, conditionally given*

$$\left\{ S_j \in d\bar{c}^{[0,n]}(j), j = \sum_{i=1}^{m} n_i, 1 \leqslant m \leqslant k \right\}.$$

Conditionally given $Y_1, \ldots, Y_n$*, let B be the random set of* $(n_1, \ldots, n_k)$*-cyclic permutations of* $[n]$ *such that*

$$Y_{\sigma(j)} \geqslant \bar{c}^{[0,n]}(j) \quad \text{for } 1 \leqslant j \leqslant n$$

if and only if $\sigma \in B$*. Let* $\hat{\rho}$ *be an independently chosen uniform random element of B, and let* $S^{\hat{\rho}}_j = \sum_{\ell=1}^{j} Y_{\hat{\rho}(\ell)}$ *for* $1 \leqslant j \leqslant n$*. Then* $S^{[0,n]}_{\hat{\rho}} := \{(j, S^{\hat{\rho}}_j) : 1 \leqslant j \leqslant n\}$ *has the regular conditional distribution of* $S^{[0,n]}$ *conditionally given* $\bar{C}^{[0,n]} = \bar{c}^{[0,n]}$*.*

The theorem is direct result of Bayes' rule and Theorem 1.1. Note that when $X_1, \ldots, X_n$ satisfy assumption **A**, $\mathcal{N}(\bar{c}^{[0,n]})$ has only one element, the composition induced by the lengths of the faces of $\bar{c}^{[0,n]}$, and A also only contains one element by Lemma 2.1, so the theorem simplifies significantly. It remains to describe $q(\cdot)$.

Lemma 7.2. *Suppose* $\bar{c}^{[0,n]} \in \mathrm{Supp}(\bar{C}^{[0,n]})$ *and that* $X_1, \ldots, X_n$ *are exchangeable. The regular conditional distribution of* $\Xi_{[0,n]}$ *conditionally given* $\bar{C}^{[0,n]}_\rho = \bar{c}^{[0,n]}$ *is given by*

$$\mathbb{P}(\bar{C}^{[0,n]}(j) \in d\bar{c}^{[0,n]}(j), 1 \leqslant j \leqslant n)\mathbb{P}(\Xi^K_{[0,n]} = (n_1, \ldots, n_k) | \bar{C}^{[0,n]}_\rho = \bar{c}^{[0,n]})$$
$$= 1_{(n_1,\ldots,n_k) \in \mathcal{N}(\bar{c}^{[0,n]})} \frac{\prod_{i=1}^{k} n_i}{\prod_{j=1}^{f(\bar{c}^{[0,n]})} k_j(n_1, \ldots, n_k)!} \mathbb{P}(S_j \in d\bar{c}^{[0,n]}(j), j = \textstyle\sum_{i=1}^{l} n_i, 1 \leqslant l \leqslant k),$$

where S_j*,* $1 \leqslant j \leqslant n$ *is as in Theorem 7.1.*

Proof. Let $(n_1, \ldots, n_k) \in \mathcal{N}(\bar{c}^{[0,n]})$. Following the construction in Theorem 1.1, by the Ewens sampling formula the probability that $\{L_{n,1}, \ldots, L_{n,K_n}\}$ is a list of the elements of $(n_1, \ldots, n_k)$ in non-increasing order is

$$\left(\prod_{j=1}^{n} (a_j!)^{-1} \right) \left(\prod_{i=1}^{k} n_i^{-1} \right),$$

where $a_j = \#\{i : 1 \leqslant i \leqslant k, n_i = j\}$ for $1 \leqslant j \leqslant n$. Conditionally given $\{L_{n,1}, \ldots, L_{n,K_n}\}$ is a list of the elements of $(n_1, \ldots, n_k)$ in non-increasing order, the probability of the event $\{\Xi^K = (n_1, \ldots, n_k), \bar{C}^{[0,n]} = \bar{c}^{[0,n]}\}$ is

$$\left(\frac{\prod_{j=1}^{n} a_j!}{\prod_{j=1}^{f(\bar{c}^{[0,n]})} k_j(n_1, \ldots, n_k)!} \right) \mathbb{P}\left(S_j \in d\bar{c}^{[0,n]}(j), j = \sum_{i=1}^{l} n_i, 1 \leqslant l \leqslant k \right),$$

where the denominator in the multiplicative factor in the parentheses is due to the restrictions on the orderings of partition segments within each face, and the numerator is due to repeated segment lengths. □

We say that the concave majorant of a walk is *trivial* if it has only one face. A particularly useful form of Theorem 7.1 arises from the special case when the increments $X_1,\ldots,X_n$ are independent, the probability that the concave majorant of $S^{[0,n]}$ is trivial with slope zero is positive, and we want the conditional distribution of the walk $S^{[0,n]}$ given that it has trivial concave majorant of slope zero. By subtraction of a line of constant slope, this gives us the conditional distribution of the walk $S^{[0,n]}$ given that it has trivial concave majorant of any slope, as long as the probability that the concave majorant of $S^{[0,n]}$ is trivial with that slope is positive. In the case where we want the regular conditional distribution for $S^{[0,n]}$ conditional on having trivial concave majorant of a slope that has zero probability, then the only possible value for $\Xi_{[0,n]}$ is the trivial composition (n).

Corollary 7.3. *Let $S_0 = 0$ and $S_j = \sum_{i=1}^{j} X_i$ for $1 \leqslant j \leqslant n$, where $X_1,\ldots,X_n$ are independent identically distributed random variables, and let $S^{[0,n]} = \{(j,S_j) : 0 \leqslant j \leqslant n\}$. Suppose that*

$$p_{\text{triv}} := \mathbb{P}(\textit{concave majorant of } S^{[0,n]} \textit{ is trivial with slope zero}) > 0.$$

Define a probability density function $q(\cdot)$ on $\mathcal{N}_n$ by

$$q\big((n_1,\ldots,n_k)\big) = \frac{1}{p_{\text{triv}}k!}\prod_{i=1}^{k} n_i u_{n_i},$$

where $u_j = \mathbb{P}(S_j = 0)$ for $1 \leqslant j \leqslant n$. Let $(N_{n,1},N_{n,2},\ldots,N_{n,K_n})$ be a composition of n chosen according to the density function $q(\cdot)$, independently of $\{X_j : 1 \leqslant j \leqslant n\}$.

Conditionally given $\{K_n = k\}$ and $\{N_{n,i} = n_i : 1 \leqslant i \leqslant k\}$, independently for each $1 \leqslant i \leqslant k$, let $Y_{n_1+\cdots+n_{i-1}+1},\ldots,Y_{n_1+\cdots+n_i}$ be random variables, independent of all previously introduced random variables, whose joint law is the regular conditional joint distribution of $X_1,\ldots,X_{n_i}$ conditionally given $\sum_{\ell=1}^{n_i} X_\ell = 0$.

Conditionally given $Y_1,\ldots,Y_n$, let B be the random set of $(n_1,\ldots,n_k)$-cyclic permutations of $[n]$ such that

$$Y_{\sigma(j)} \leqslant 0 \quad \textit{for } 1 \leqslant j \leqslant n$$

if and only if $\sigma \in B$. Let $\hat\rho$ be an independently chosen uniform random element of B, and let $S_j^{\hat\rho} = \sum_{\ell=1}^{j} Y_{\hat\rho(\ell)}$ for $1 \leqslant j \leqslant n$. Then $S_{\hat\rho}^{[0,n]} := \{(j,S_j^{\hat\rho}) : 1 \leqslant j \leqslant n\}$ has the regular conditional distribution of $S^{[0,n]}$ conditionally given that $S^{[0,n]}$ has trivial concave majorant with slope zero.

8. A path transformation

This section provides an important path transformation which, by taking scaling limits, is used by Pitman and Uribe Bravo to completely describe the concave majorant (or as in that paper, convex minorant) of a Lévy process and the excursions of that process beneath its concave majorant [13]. Essentially, the idea is that a uniformly sampled face of the concave majorant should have

https://doi.org/10.1017/S0963548311000307 Published online by Cambridge University Press

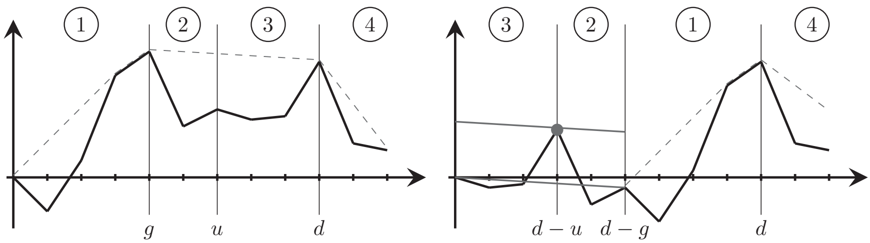

Figure 3. An example of the '3214' path transformation of Theorem 8.1. The walk on the right is the transformed version of the walk on the left. Note how, given $d-g$, the transform is easily inverted: the index at which the first $d-g$ increments should start after cyclic permutation is marked, and can be found by lowering a line with the slope the mean of the first $d-g$ increments.

uniform length and the walk over it should be a Vervaat-like transform of some walk of the same length.

Let $S_0 = 0$ and $S_j = \sum_{i=1}^{n} X_i$ for $1 \leqslant j \leqslant n$, where $X_i, i = 1, \ldots, n$ are exchangeable random variables satisfying assumption **A**. We introduce the following path transformation for the random walk $S^{[0,n]} = \{(j, S_j),\ 1 \leqslant j \leqslant n\}$. Let U be distributed uniformly on $[n]$. Let g and d be the left and right end points, respectively, of the face of the concave majorant of $S^{[0,n]}$ containing the Uth increment X_U. Define S_j^U for $1 \leqslant j \leqslant n$ by

$$S_j^U = \begin{cases} S_{U+j} - S_U & \text{for } 0 \leqslant j < d - U, \\ S_{g+j-(d-U)} + S_d - S_g - S_U & \text{for } d - U \leqslant j < d - g, \\ S_{j-(d-g)} + S_d - S_g & \text{for } d - g \leqslant j < d, \\ S_j & \text{for } d \leqslant j \leqslant n, \end{cases} \tag{8.1}$$

and let $S_U^{[0,n]} = \{(j, S_j^U),\ 1 \leqslant j \leqslant n\}$.

Theorem 8.1.

$$(U, S^{[0,n]}) \stackrel{d}{=} (d - g, S_U^{[0,n]}).$$

In fact, Theorem 8.1 provides an alternative method of proving Theorem 1.1 under assumption **A**, since by applying the transformation again to the $S_U^{[0,n]}$ restricted to the interval $[d-g, n]$, and then doing this repeatedly until there is nothing left to transform, we are actually performing the inverse of the transformation given in Theorem 1.1. However, this method does not extend to cover the general case as considered in Section 6, so we will not expand on it.

Proof. As in the proof of Theorem 1.1 under assumption **A** in Section 2, it is enough to show that the equality in distribution holds when $X_1, \ldots, X_n$ are samples without replacement from $x_1, \ldots, x_n$ satisfying assumption **A**. $S^{[0,n]}$ and $S_U^{[0,n]}$ may thus be thought of as permutations of n, so we may think of the mapping $(U, S^{[0,n]}) \mapsto (d-g, S_U^{[0,n]})$ as a mapping from $[n] \times \Sigma_n$ to itself. Since U is uniform on $[n]$, and the ordering of $X_1, \ldots, X_n$ is a uniform random permutation of $x_1, \ldots, x_n$, it is enough to show that this mapping is a bijection. To do this, it suffices to show that the mapping is surjective. This can be seen visually in Figure 3 since it is clear from the

https://doi.org/10.1017/S0963548311000307 Published online by Cambridge University Press

figure and its description that the map is easily inverted. More formally, to show that the map is surjective, it is sufficient to show that for $k \in [n]$ there exists $u \in [n]$ and $\sigma \in \Sigma_n$ such that

$$\left(u, \left\{(0,0),(1,x_{\sigma(1)}),(2,x_{\sigma(1)}+x_{\sigma_2}),\ldots,\left(n,\sum_{i=1}^{n} x_{\sigma(i)}\right)\right\}\right)$$
$$\mapsto \left(k, \left\{(0,0),(1,x_1),(2,x_1+x_2),\ldots,\left(n,\sum_{i=1}^{n} x_i\right)\right\}\right).$$

Let f be the number of faces of the concave majorant of the walk of length $n-k$ with increments $x_{k+1},\ldots,x_n$, and let the lengths and increments of these faces in order of appearance be $(\ell_1,s_1),\ldots,(\ell_f,s_f)$. Let r be the unique $r \in [k]$ such that the walk with increments

$$(x_{r+1}, x_{(r+1) \bmod k+1}, x_{(r+2) \bmod k+1}, \ldots, x_{(r+k-2) \bmod k+1}, x_r)$$

remains below its concave majorant. Let $s^* = \sum_{i=1}^{k} x_i$, and let m be the unique $m \in \{0,\ldots,f\}$ such that

$$\frac{s_m}{\ell_m} > \frac{s^*}{k} > \frac{s_{m+1}}{\ell_{m+1}},$$

where we say that $s_0/\ell_0 = +\infty$ and $s_{f+1}/\ell_{f+1} = \infty$. The appropriate $(\sigma(1),\ldots,\sigma(n))$ is given by

$$(k+1, k+2, \ldots, k+\textstyle\sum_{i=1}^{m} \ell_i,$$
$$r+1, (r+1) \bmod k+1, (r+2) \bmod k+1, \ldots, r, k+\textstyle\sum_{i=1}^{m} \ell_i + 1, \ldots, n). \qquad \square$$

References

[1] Bertoin, J. (1999) Renewal theory for embedded regenerative sets. *Ann. Probab.* **27** 1523–1535.

[2] Bertoin, J. (2000) The convex minorant of the Cauchy process. *Electron. Comm. Probab.* **5** 51–55.

[3] Brunk, H. D. (1964) A generalization of Spitzer's combinatorial lemma. *Z. Wahrsch. Verw. Gebiete* **2** 395–405.

[4] Diaconis, P., Holmes, S., Janson, S., Lalley, S. P. and Pemantle, R. (1996) Metrics on compositions and coincidences among renewal sequences. In *Random Discrete Structures*, Vol. 76 of *IMA Volumes in Mathematics and its Applications*, Springer, pp. 81–101.

[5] Feller, W. (1971) *An Introduction to Probability Theory and its Applications*, Vol. II, second edition, Wiley.

[6] Gnedin, A. and Pitman, J. (2005) Regenerative composition structures. *Ann. Probab.* **33** 445–479.

[7] Goldie, C. M. (1989) Records, permutations and greatest convex minorants. *Math. Proc. Cambridge Philos. Soc.* **106** 169–177.

[8] Greenwood, P. and Pitman, J. (1980) Fluctuation identities for Lévy processes and splitting at the maximum. *Adv. Appl. Probab.* **12** 893–902.

[9] Greenwood, P. and Pitman, J. (1980) Fluctuation identities for random walk by path decomposition at the maximum. *Adv. Appl. Probab.* **12** 291–293.

[10] Groeneboom, P. (1983) The concave majorant of Brownian motion. *Ann. Probab.* **11** 1016–1027.

[11] Kac, M. (1954) Toeplitz matrices, translation kernels and a related problem in probability theory. *Duke Math. J.* **21** 501–509.

[12] Pitman, J. (2006) *Combinatorial Stochastic Processes*, Vol. 1875 of *Lecture Notes in Mathematics*, Springer.

[13] Pitman, J. and Uribe Bravo, G. (2011) The convex minorant of a Lévy process. *Ann. Probab.*, to appear.

[14] Qiao, Z. and Steele, J. M. (2005) Random walks whose concave majorants often have few faces. *Statist. Probab. Lett.* **75** 97–102.
[15] Shepp, L. A. and Lloyd, S. P. (1966) Ordered cycle lengths in a random permutation. *Trans. Amer. Math. Soc.* **121** 340–357.
[16] Sherman, S. (1964) Fluctuation and periodicity. *J. Math. Anal. Appl.* **9** 468–476.
[17] Sparre Andersen, E. (1954) On the fluctuations of sums of random variables II. *Math. Scand.* **2** 195–223.
[18] Sparre Andersen, E. (1959) On the distribution of the random variable H_n. Tech. Sci. Note no. 1, Contract no. AF 61(052)-42.
[19] Spitzer, F. (1956) A combinatorial lemma and its application to probability theory. *Trans. Amer. Math. Soc.* **82** 323–339.
[20] Steele, J. M. (2002) The Bohnenblust–Spitzer algorithm and its applications. *J. Comput. Appl. Math.* **142** 235–249.
[21] Vervaat, W. (1979) A relation between Brownian bridge and Brownian excursion. *Ann. Probab.* **7** 143–149.