SORTING USING COMPLETE SUBINTERVALS AND THE MAXIMUM NUMBER OF RUNS IN A RANDOMLY EVOLVING SEQUENCE.

SVANTE JANSON

Abstract. We study the space requirements of a sorting algorithm where only items that at the end will be adjacent are kept together. This is equivalent to the following combinatorial problem: Consider a string of fixed length n that starts as a string of 0's, and then evolves by changing each 0 to 1, with the n changes done in random order. What is the maximal number of runs of 1's?

We give asymptotic results for the distribution and mean. It turns out that, as in many problems involving a maximum, the maximum is asymptotically normal, with fluctuations of order $n^{1/2}$, and to the first order well approximated by the number of runs at the instance when the expectation is maximized, in this case when half the elements have changed to 1; there is also a second order term of order $n^{1/3}$.

We also treat some variations, including priority queues and sock-sorting.

The proofs use methods originally developed for random graphs.

1. Introduction

Gunnar af Hällström [1] considered, as indicated at the end of his paper, the following algorithm for sorting an unordered pile of student exams in alphabetic order. (It is said that he used this procedure himself.)

The exams are taken one by one from the input. The first exam is put in a new pile. For each following exam (x, say), if the name on it is immediately preceding the name on an exam y at the top of one of the piles, the new exam x is put on top of y. (The professor knows the names of all the students, and can thus see that there are no names between x and y.) Similarly, if the name on x is immediately succeeding the name on an exam z at the bottom of a pile, x is put under z. If both cases apply, with y on top of one pile and z at the bottom of another, the two piles are merged with x inserted between z and y. Finally, if there is no pile matching x in one of these ways, x is put in a new pile.

The algorithm thus maintains a list of sorted piles, each being an interval without gaps of the set of exams. At the end, there is a single sorted pile.

Date: January 10, 2007; revised February 8, 2007.

2000 *Mathematics Subject Classification.* 60C05; 68W40.

Key words and phrases. sorting algorithm, runs, priority queues, sock-sorting, evolution of random strings, Brownian motion.

The problem is the space requirement of this algorithm; more precisely, the maximum number of sorted piles during the execution. The input is assumed to be in random order, so this is a random variable, and we are interested in its mean and distribution.

Remark 1.1. As a sorting method, this algorithm has drawbacks. First, it requires that all names are known from the beginning; mathematically it can be seen as sorting the numbers $1, \dots, n$. Secondly, the space requirement turns out to be quite high, see below. This also implies that the number of comparisons necessary for each insertion is high, of the order of n. The algorithm might be useful when blocks of sorted items can be manipulated as easily as individual items, and we do not want to make insertions inside the blocks, for example when sorting physical objects that are to be glued together in order.

af Hällström [1] gave the following mathematical reformulation, where we also introduce some notation. Consider a deck of n cards numbered $1, \dots, n$ in random order, and a sequence of n places with the same numbers in order. Take the cards one by one and put them at their respective places. When we have placed m cards, $0 \le m \le n$, we see $X_{n,m}$ "islands", i.e. uninterrupted blocks of cards. What is $X_n^* := \max_m X_{n,m}$?

Alternatively, we can use the language of parking cars, which is popular for some related problems in computer science: n cars park, one by one, on n available places along a street; each car parks at a random free place. What is the maximum number of uninterrupted blocks of cars during the process?

Let, for $n \ge 1$, $0 \le m \le n$ and $1 \le k \le n$, the indicator $I_{n,m}(k)$ be 1 if the item (exam or card) with number k is one of the m first in the input, and 0 otherwise. Thus, $X_{n,m}$ is the number of runs of 1's in the random sequence $I_{n,m}(1), \dots, I_{n,m}(n)$ of $n - m$ 0's and m 1's. We can express $X_{n,m}$ algebraically as

$$
\begin{aligned}
X_{n,m} &= I_{n,m}(1) + \sum_{k=1}^{n-1} (1 - I_{n,m}(k)) I_{n,m}(k+1) \\
&= m - \sum_{k=1}^{n-1} I_{n,m}(k) I_{n,m}(k+1). \qquad (1.1)
\end{aligned}
$$

If the input is given by the permutation σ of $\{1, \dots, n\}$, so that item k has position $\sigma^{-1}(k)$,

$$I_{n,m}(k) = \mathbf{1}[\sigma^{-1}(k) \le m],$$

where $\mathbf{1}[\dots]$ denotes the indicator of the indicated event. We assume that σ is a (uniformly chosen) random permutation; thus so is σ^{-1}. Hence, each random sequence $(I_{n,m}(k))_{k=1}^n$ is uniformly distributed over all $\binom{n}{m}$ possibilities; moreover, for each $m < n$ we obtain $(I_{n,m+1}(k))_{k=1}^n$ from $(I_{n,m}(k))_{k=1}^n$ by changing a single randomly chosen 0 to 1, this random choice being uniform among the $n - m$ 0's, and independent of the previous history.

It is easy to see that $\mathbb{E}\, X_{n,m} = m(n-m+1)/n$, see (3.1); it follows that the maximum of $\mathbb{E}\, X_{n,m}$ for a given n is attained for $m = \lceil n/2 \rceil$, and that $\mathbb{E}\, X_{n,\lceil n/2\rceil} > n/4$. Since obviously

$$\mathbb{E}\, X_n^* = \mathbb{E} \max_m X_{n,m} \geq \max_m \mathbb{E}\, X_{n,m}, \tag{1.2}$$

this yields $\mathbb{E}\, X_n^* > n/4$ as observed by af Hällström [1]. Moreover, he observed that $\mathbb{E}\, X_n^*$ is subadditive, and thus the limit

$$\gamma := \lim_{n\to\infty} \mathbb{E}\, X_n^*/n$$

exists and equals $\inf_n \mathbb{E}\, X_n^*/n$; he further showed that $1/4 \leq \gamma \leq 1/3$, where the lower bound comes from (1.2). Based on simulations with $n = 13$ and $n = 52$, af Hällström [1] concluded that γ seems to be very close to or equal to $1/4$. We will show that, indeed, $\gamma = 1/4$. We also show that the distribution of X_n^* is asymptotically normal, with a variance of order n.

Theorem 1.2. *As $n \to \infty$,*

$$n^{-1/2}\big(X_n^* - n/4\big) \xrightarrow{\mathrm{d}} N(0, 1/16), \tag{1.3}$$

with convergence of all moments. In particular,

$$\mathbb{E}\, X_n^* = n/4 + o(n^{1/2}),$$
$$\operatorname{Var} X_n^* = n/16 + o(n).$$

This theorem says that to the first order, the maximum number of piles (runs) X_n^* behaves like the number $X_{n,m}$ with $m = \lceil n/2 \rceil$. A more refined analysis shows that the difference $X_n^* - X_{n,\lceil n/2\rceil}$ is of order $n^{1/3}$. Let $B(t)$, $-\infty < t < \infty$, be a standard two-sided Brownian motion; thus $B(0) = 0$ and $B(t)$, $t \geq 0$, and $B(-t)$, $t \geq 0$, are two independent Brownian motions.

Theorem 1.3. *As $n \to \infty$,*

$$n^{-1/3}\big(X_n^* - X_{n,\lceil n/2\rceil}\big) \xrightarrow{\mathrm{d}} \tfrac{1}{2} V, \tag{1.4}$$

where the random variable V is defined by $V := \max_t\big(B(t) - t^2/2\big)$, and

$$\mathbb{E}\, X_n^* = \mathbb{E}\, X_{n,\lceil n/2\rceil} + \tfrac{1}{2}\, \mathbb{E}\, V n^{1/3} + o(n^{1/3}) = \tfrac{1}{4} n + \tfrac{1}{2}\, \mathbb{E}\, V n^{1/3} + o(n^{1/3}).$$

The random variable V is studied by Barbour [2], Daniels and Skyrme [7] and Groeneboom [10]. Note that $0 < V < \infty$ a.s. We have, see [7] (using `Maple` to improve the numerical values in [2, 3, 7, 6]), with Ai the Airy function,

$$\mathbb{E}\, V = -\frac{2^{-1/3}}{2\pi} \int_{-\infty}^{\infty} \frac{\mathrm{i}y \,\mathrm{d}y}{\mathrm{Ai}(\mathrm{i}y)^2} \approx 0.996193.$$

The numerical values $X_{13}^* \approx 4.22$ and $X_{52}^* \approx 14.66$ found experimentally by af Hällström [1] differ from $n/4$ by about 18% and 10% less than the correction term $\frac{1}{2}\, \mathbb{E}\, V n^{1/3}$ in Theorem 1.3, which is a reasonable agreement for such rather small n.

Remark 1.4. af Hällström [1] considered also the cyclic case, when we regard $\{1,\dots,n\}$ as a circle, which sometimes is slightly simpler to study because of the greater symmetry. In this case we define $I_{n,m}(k)$ for all $k \in \mathbb{Z}$ by $I_{n,m}(k+n) := I_{n,m}(k)$, i.e. we interpret k modulo n, and we sum to n in (1.1). Since the number of runs in the linear and cyclic version differ by at most 1, all our asymptotic results remain the same, and we will only consider the linear case. (Moreover, the cyclic case with n items corresponds exactly to the linear with $n-1$ by fixing the last element, see [1].)

We prove these theorem by studying asymptotics of the entire (random) process $(X_{n,m})_{m=0}^{n}$. The natural time here is m/n, so we take $m = \lfloor nt \rfloor$ for $0 \le t \le 1$ and consider the process $X_{n,\lfloor nt \rfloor}$ with a continuous parameter $t \in [0,1]$. The following theorem shows that this process asymptotically is Gaussian. (The space $D[0,1]$ is defined in Section 4, see [4] for a detailed treatment.)

Theorem 1.5. *As $n \to \infty$, in the space $D[0,1]$ of functions on $[0,1]$,*

$$n^{-1/2}\bigl(X_{n,\lfloor nt \rfloor} - nt(1-t)\bigr) \xrightarrow{\mathrm{d}} Z(t), \tag{1.5}$$

where Z is a continuous Gaussian process on $[0,1]$ with mean $\mathbb{E}\, Z(t) = 0$ and covariances

$$\mathbb{E}\bigl(Z(s)Z(t)\bigr) = s^2(1-t)^2, \qquad 0 \le s \le t \le 1. \tag{1.6}$$

The behaviour of X_n^* shown in Theorems 1.2 and 1.3, with an asymptotic normal distribution with a mean of order n and random fluctuations of order $n^{1/2}$, and with a second order term for the mean of order $n^{1/3}$, is common for this type of random variables defined as the maximum of some randomly evolving process. For various examples, both combinatorial and others, and general results see for example Daniels [5, 6], Daniels and Skyrme [7], Barbour [2, 3] and Louchard, Kenyon and Schott [21]. Indeed, paraphrasing the explanations in these papers, in many such problems, the first order asymptotic of a random process $X_n(t)$ (after suitable scaling) is a deterministic function $f(t)$, say, defined on a compact interval I (typically scaled to be $[0,1]$ as here). Hence the first order asymptotic of the maximum of the process is just the maximum of this function f. Moreover, it is often natural to expect that the random fluctuations around this function $f(t)$ asymptotically form a Gaussian process $G(t)$; this is then a second order term of smaller order as in our Theorem 1.5. If we assume that f is continuous on I and has a unique maximum at a point $t_0 \in I$, then the maximum of the process $X_n(t)$ is attained close to t_0, so the first order approximation of the maximum is the constant $f(t_0) = \max_t f(t)$, while the next approximation is just $X_n(t_0)$, giving a normal limit law as in our Theorem 1.2. The Gaussian fluctuations in this limit have mean 0, so in order to find the next term for the mean $\mathbb{E}\, X_n^*$, we study more closely the difference $\max_t X_n(t) - X_n(t_0)$ by studying the difference $X_n(t) - X_n(t_0)$ close to t_0. Assuming that t_0 is an interior point of I and that f is twice differentiable at t_0 with $f''(t_0) \neq 0$, we can locally at t_0 approximate f by a parabola and $G(t) - G(t_0)$ by a two-sided Brownian motion (with some scaling), and thus $\max_t X_n(t) - X_n(t_0)$ is approximated by a scaling constant times the variable V above, see Barbour [2] and, in our case, Corollary 4.5 below. In the typical case where the mean of $X_n(t)$ is of order n and the Gaussian fluctuations are of order $n^{1/2}$, it is easily seen that the correct scaling gives, as in Theorem 1.3 above, a correction to $\mathbb{E}\, X_n^*$ of order $n^{1/3}$, see [2, 5, 6] and Section 6.

The method used in the present paper is a simple adaption of the method used in [13] and [14] to study the number of subgraphs of a given isomorphism type in a random graph. These papers study the random graphs $G(n,p)$ and $G(n,m)$ that can be constructed by random deletion of edges in the complete graph K_n (with the deletions being independent for $G(n,p)$ and such that a fixed number of edges are deleted for $G(n,m)$). The method applies more generally to random graphs constructed by random edge deletions in these ways from any fixed initial graph F_n. The problem treated in this paper can be regarded as an instance of this when the initial graph is the path P_n with n edges. In particular, Theorem 1.2 corresponds to [14, Theorem 24], which gives the asymptotic distribution of the maximum number of induced subgraphs of a given type during the evolution of $G(n,p)$ or $G(n,m)$; see also [14, Theorem 33] (isolated edges) and [14, Theorem 17] (a general result) for related results. Conversely, we expect that these results for random graphs can be complemented by the analogues of Theorem 1.3 above, using the method of proof in the present paper, but we have not verified the details.

Our method applies also to other problems. First, let $X_{n,m}^{(1)}$ be the number of piles with a single exam (runs with a single 1) in the process studied above. Then we obtain similar results for the maximum $X_n^{(1)*} := \max_m X_{n,m}^{(1)}$, see Section 7. The same applies to the number $X_{n,m}^{(d)}$ of piles with any other fixed number d of exams (runs of a fixed length d).

Another example is given by *priority queues*, where Louchard [19] and Louchard, Kenyon and Schott [21] have proved asymptotic results very similar to the Theorems 1.2–1.5 above, see Section 8 for details. In particular, they found the same asymptotic covariance (1.6) except for a normalizing constant. (See also Flajolet, Françon and Vuillemin [8] and Flajolet, Puech and Vuillemin [9] for combinatorial results on generating functions involving Hermite polynomials; these results, however, do not easily yield asymptotics.)

Priority queues can be defined as follows. Suppose that n items are to be temporarily stored (or processed); let item i arrive at time A_i and be deleted at time D_i. We assume that the $2n$ times A_i and D_i are distinct; thus they can be arranged in a sequence of the $2n$ events A_i and D_i, with A_i coming before D_i for each i. We assume further, as our probabilistic model, that all $(2n)!/2^n$ such sequences are equally probable. Ignoring the labels, we can equivalently consider sequences of n A and n D (or n + and n −), where each A is paired with a D coming later; there is a 1–1 correspondence between such sequences and pairings of $1, \dots, 2n$ into n pairs, and there are $(2n-1)!! = (2n)!/(2^n n!)$ such sequences (with pairings), again taken with equal probability.

Let, for $m = 0, \dots, 2n$, $Y_{n,m}$ be the number of items stored after m of these events, i.e. the number of A's minus the number of D's among the m first events, and let $Y_n^* := \max_{0 \le m \le 2n} Y_{n,m}$. The sequence $(Y_{n,m})_0^{2n}$ is a Dyck path, but note that its distribution is not uniform; for a given Dyck path (or a given sequence of A and D without labels), the number of ways to pair a given D with a preceding A, i.e. the number of ways to choose which item to delete, equals the current number of items stored before this deletion. Thus, the weight of the Dyck path equals the product of these numbers $\prod_{m: Y_{n,m+1} < Y_{n,m}} Y_{n,m}$. Alternatively, which better explains the name priority queue, we can keep the stored items in a list showing the order in which they eventually will be deleted; then there is only one choice for each deletion but each new item can be inserted in $Y+1$ ways if there are Y items stored before the insertion, and thus $Y+1$ after it; hence the weight can also be written as $\prod_{m: Y_{n,m} > Y_{n,m-1}} Y_{n,m}$. (It is easily to see directly that the two products are equal.)

An equivalent example is *sock-sorting*, studied by Li and Pritchard [18] and Steinsaltz [23]. Suppose that we have $2n$ socks; the socks form n pairs with the two socks in each pair identical but different from all others. All socks are mixed and we pick them in random order. If the picked sock is from a pair that we have not yet seen, it is put put on a bench; on the other hand, if we already have picked the other sock in the pair, that sock is taken from the bench, paired with its twin, and put away in permanent storage. What is the maximum number of socks on the bench? It is easily seen that this is equivalent to a priority queue.

We will in Section 8 show how our method applies to priority queues and socks, and explain why we obtain the same asymptotic results as for $X_{n,m}$ and X_n^*. (Note that there is no exact correspondence for finite n, since the natural sample spaces have $n!$ elements for $X_{n,m}$ but $(2n-1)!!$ elements for $Y_{n,m}$.) Again, we can regard the problem as an instance of subgraph counts for randomly deleting edges from a given initial graph F_n; in this case taking F_n to be a multigraph consisting of n double edges.

Yet another example is a model suggested by Van Wyk and Vitter [25] as a model for *hashing with lazy deletion*, and further studied by Louchard [20] and Louchard, Kenyon and Schott [21]. In this model, n item arrives and are deleted as above, but now the arrival and deletion times A_i and D_i are random numbers, with the n pairs (A_i, D_i) mutually independent and each pair distributed as $(T_i \wedge \tilde{T}_i, T_i \vee \tilde{T}_i)$, where T_i and $\tilde{T}_i$ are independent random variables uniformly distributed on [0,1]. (We use $\wedge$ and $\vee$ as notations for min and max of two numbers.) We let $Y_n(t)$ be the number of items present at time t, and again we are especially interested in its maximum $\max_t Y_n(t)$. Again, the asymptotic results for the maximum found by Louchard, Kenyon and Schott [21] are the same as in our Theorems 1.2 and 1.3, except for a constant factor, while the asymptotic result for the process $Y_n(t)$ found by Louchard [20] differs somewhat from the one in Theorem 1.5, it corresponds instead to the one in Corollary 4.2 below; see Section 8. Indeed, as explained by Kenyon and Vitter [17], see also Section 8, this model can be seen as a priority queue with randomized times for insertions and deletions, which explains why the results for the maximum are the same as for priority queues.

We assume in the sequel that $n \geq 2$, to avoid some trivialities. All unspecified limits are as $n \to \infty$. We use the standard notations $\overset{\mathrm{p}}{\longrightarrow}$ and $\overset{\mathrm{d}}{\longrightarrow}$ for convergence in probability and distribution, respectively, of random variables, and a.s. for *almost surely*, i.e. with probability 1.

2. Randomizing time

We will use the standard method of randomizing the time. More precisely, we let $T_1, \dots, T_n$ be independent random variables, each uniformly distributed on $(0, 1)$. We interpret T_k as the time item k arrives, and note that a.s. there are no ties. We define

$$I(t; k) = \mathbf{1}[T_k \leq t],$$

i.e., $I(t; k) = 1$ if item k has arrived by time t. We further define $N_n(t)$ as the number of items that have arrived at time t, and $X_n(t)$ as the number of runs of 1's at time t, i.e., cf. (1.1),

$$N_n(t) = \sum_{k=1}^{n} I(t; k), \tag{2.1}$$

$$X_n(t) = I(t; 1) + \sum_{k=1}^{n-1} \bigl(1 - I(t; k)\bigr) I(t; k + 1) \tag{2.2}$$

$$= N_n(t) - \sum_{k=1}^{n-1} I(t; k) I(t; k + 1). \tag{2.3}$$

Clearly, the items arrive in random order, so the process remains the same except that the insertions occur at the random times $T_{(n;1)}, \dots, T_{(n;n)}$, where $T_{(n;j)}$ is the j:th order statistic of $T_1, \dots, T_n$. We thus have $I_{n,m}(k) = I(T_{(n;m)}; k)$ and $X_n(t) = X_{n,m}$ when $T_{(n;m)} \leq t < T_{(n;m+1)}$ (with $T_{(n;0)} := 0$ and $T_{(n;n+1)} := 1$ for convenience). In particular,

$$X_n^* = \max_{0 \leq t \leq 1} X_n(t). \tag{2.4}$$

Note that $X_n(0) = X_{n,0} = 0$ and $X_n(1) = X_{n,n} = 1$.

The importance of this randomization is that the variables $I(t; k)$, $k = 1, \dots, n$, are independent (both for a fixed t and as stochastic processes, i.e.

as random functions of t). For every n, k and $t \in [0,1]$,

$$\mathbb{P}\big(I(t;k)=1\big)=\mathbb{P}(T_k \le t)=t, \tag{2.5}$$

i.e. $I(t;k)$ has the Bernoulli distribution $\mathrm{Be}(t)$. $X_n(t)$ thus is the number of runs of 1 in a sequence of *independent* 0's and 1's, each with the distribution $\mathrm{Be}(t)$. Furthermore, the number of items sorted at time t is $N_n(t) \sim \mathrm{Bi}(n,t)$.

Define further, for $0 \le t \le 1$, the centralized variables

$$I'(t;k) := I(t;k) - \mathbb{E}\, I(t;k) = I(t;k) - t \tag{2.6}$$

and the sums

$$S_{n,1}(t) := \sum_{k=1}^{n} I'(t;k) = N_n(t) - \mathbb{E}\, N_n(t) = N_n(t) - nt, \tag{2.7}$$

$$S_{n,2}(t) := \sum_{k=1}^{n-1} I'(t;k) I'(t;k+1). \tag{2.8}$$

Thus $S_{n,1}(0) = S_{n,2}(0) = S_{n,1}(1) = S_{n,2}(1) = 0$ and $\mathbb{E}\, S_{n,1}(t) = \mathbb{E}\, S_{n,2}(t) = 0$ for all $t \in [0,1]$. We have

$$N_n(t) = \sum_{k=1}^{n} \big(I'(t;k)+t\big) = S_{n,1}(t) + nt, \tag{2.9}$$

$$\begin{aligned}\sum_{k=1}^{n-1} I(t;k)I(t;k+1) &= \sum_{k=1}^{n-1}\big(I'(t;k)+t\big)\big(I'(t;k+1)+t\big)\\ &= S_{n,2}(t) + t\big(2S_{n,1}(t) - I'(t;1) - I'(t;n)\big) + (n-1)t^2,\end{aligned}$$

and thus from (2.3) the representation

$$\begin{aligned}X_n(t) &= n(t-t^2) + t^2 + (1-2t)S_{n,1}(t) - S_{n,2}(t) + tI'(t;1) + tI'(t;n)\\ &= nt(1-t) + (1-2t)S_{n,1}(t) - S_{n,2}(t) + R_n(t),\end{aligned} \tag{2.10}$$

where $R_n(t) := t^2 + tI'(t;1) + tI'(t;n)$ and thus $|R_n(t)| \le 3$.

We will in Section 4 study the asymptotic distribution of the stochastic processes (i.e., random functions) $S_{n,1}(t)$ and $S_{n,2}(t)$; our main results then follow easily from (2.4) and (2.10).

Note that for any fixed t, the variables $I'(t;k)$ are independent and have means 0; hence the terms in the sums in (2.7) and (2.8) have means and all covariances 0. (They are thus orthogonal in L^2.) It follows immediately that

$$\mathrm{Var}\,(S_{n,1}(t)) = n\,\mathbb{E}\big(I'(t;1)\big)^2 = n\,\mathrm{Var}\big(I(t;1)\big) = nt(1-t), \tag{2.11}$$

$$\begin{aligned}\mathrm{Var}\,(S_{n,2}(t)) &= (n-1)\,\mathbb{E}\big(I'(t;1)I'(t;2)\big)^2 = (n-1)\big(\mathrm{Var}\big(I(t;1)\big)\big)^2\\ &= (n-1)t^2(1-t)^2,\end{aligned} \tag{2.12}$$

$$\mathrm{Cov}\big(S_{n,1}(t), S_{n,2}(t)\big) = 0. \tag{2.13}$$

3. Exact results

We first give some exact results for finite n. It is easy to find the exact distribution of $X_{n,m}$ for given n and m, see for example Stevens [24] or Mood [22]. For $m \geq 1$ and $k \geq 1$ we have $X_{n,m} = k$ if there are k runs of 1's separated by $k-1$ runs of 0's and possibly preceded and/or succeeded by additional runs of 0's. Considering the bivariate generating function for such sequences of arbitrary length, we easily find

$$\begin{aligned}\mathbb{P}(X_{n,m} = k) &= [x^m y^{n-m}]\Big(\frac{x}{1-x}\Big)^k\Big(\frac{y}{1-y}\Big)^{k-1}\Big(\frac{1}{1-y}\Big)^2\\ &= [x^{m-k}y^{n-m-k+1}](1-x)^{-k}(1-y)^{-k-1}\\ &= \binom{m-1}{k-1}\binom{n-m+1}{k}.\end{aligned}$$

The mean can be computed from this [22], [1], but simpler from (1.1):

$$\begin{aligned}\mathbb{E}\,X_{n,m} &= m - \sum_{k=1}^{n-1}\mathbb{E}\big(I_{n,m}(k)I_{n,m}(k+1)\big)\\ &= m-(n-1)\frac{m(m-1)}{n(n-1)} = \frac{m(n-m+1)}{n}. \qquad (3.1)\end{aligned}$$

A similar computation of the variance yields, omitting the details,

$$\operatorname{Var} X_{n,m} = \frac{m(m-1)(n-m)(n-m+1)}{n^2(n-1)}.$$

If we instead randomize the insertion times as in Section 2 and consider the process at a fixed time t, we have by (2.2), (2.5) and the independence of $I(t;k)$ for $k = 1, \dots, n$,

$$\mathbb{E}\,X_n(t) = t + \sum_{k=1}^{n-1}(1-t)t = nt(1-t) + t^2. \qquad (3.2)$$

Similarly, using (2.2), again omitting details,

$$\operatorname{Var} X_n(t) = nt(1-t)(1-3t+3t^2) + t^2(1-t)(3-5t). \qquad (3.3)$$

To find the exact distribution of X_n^* seems much more complicated. Exact values of $\mathbb{P}(X_n^* = h)$ are easily calculated for small n, see af Hällström [1], but we do not know any general formula. It would be interesting to find such a formula by combinatorial methods.

4. The asymptotic distribution of $S_{n,1}(t)$ and $S_{n,2}(t)$

To state our results on the asymptotic distribution of the stochastic processes $S_{n,1}(t)$ and $S_{n,2}(t)$, we need a suitable topological space of functions. We use, for an interval $I \subseteq \mathbb{R}$, the standard space $D(I)$ of right-continuous functions on I that have left-hand limits, equipped with the Skorohod topology. For a precise definition of this (metrizable) topology, see e.g. Billingsley

[4] ($I = [0,1]$), Jacod and Shiryaev [12] ($I = [0,\infty)$), Kallenberg [16, Appendix A.2] ($I = [0,\infty)$), or Janson [14]. For our purposes it is sufficient to know that if f is continuous on I, then $f_n \to f$ in $D(I)$ if and only if $f_n \to f$ uniformly on every compact subinterval. In particular, if I is compact, for example $I = [0,1]$, and f is continuous on I, then $f_n \to f$ in $D(I)$ if and only if $f_n \to f$ uniformly.

Our main result on the asymptotic global behaviour of $S_{n,1}(t)$ and $S_{n,2}(t)$ then can be stated as follows.

Theorem 4.1. *As $n \to \infty$, in $D[0,1]$,*

$$n^{-1/2} S_{n,1}(t) \xrightarrow{\text{d}} Z_1(t), \tag{4.1}$$

$$n^{-1/2} S_{n,2}(t) \xrightarrow{\text{d}} Z_2(t), \tag{4.2}$$

jointly, where Z_1 and Z_2 are two independent continuous Gaussian processes on $[0,1]$ with means $\mathbb{E}\, Z_1(t) = \mathbb{E}\, Z_2(t) = 0$ and covariances

$$\mathbb{E}\big(Z_1(s)Z_1(t)\big) = s(1-t), \qquad 0 \le s \le t \le 1, \tag{4.3}$$

$$\mathbb{E}\big(Z_2(s)Z_2(t)\big) = s^2(1-t)^2, \qquad 0 \le s \le t \le 1. \tag{4.4}$$

Thus, Z_1 is a standard Brownian bridge, and the limit (4.1) is just the well-known theorem that the empirical distribution function asymptotically is distributed as a Brownian bridge, see e.g. Billingsley [4, Theorem 16.4].

The proof of Theorem 4.1, and of all other results in this section, are postponed to Section 6.

Using (2.10), Theorem 4.1 yields the asymptotic distribution of the process $X_n(t)$.

Corollary 4.2. *As $n \to \infty$, in $D[0,1]$,*

$$n^{-1/2}\big(X_n(t) - nt(1-t)\big) \xrightarrow{\text{d}} Z(t), \tag{4.5}$$

where Z is a continuous Gaussian process on $[0,1]$ with mean $\mathbb{E}\, Z(t) = 0$ and covariances, for $0 \le s \le t \le 1$,

$$\mathbb{E}\big(Z(s)Z(t)\big) = s(1-2s)(1-t)(1-2t) + s^2(1-t)^2 \tag{4.6}$$

$$= s(1-t)(1-s-2t+3st). \tag{4.7}$$

In particular, this implies the limit (4.5) for each fixed t, with $\operatorname{Var}(Z(t)) = t(1-t)(1-3t+3t^2)$, which also follows more easily from (2.2), (3.2), (3.3) and the Central Limit Theorem for 1-dependent sequences.

These results are stated using the randomized insertions described in Section 2. We can also return to the original deterministic insertion times and obtain asymptotics of the discrete process $(X_{n,m})_{m=0}^n$, which yields Theorem 1.5 stated in the introduction. Note that the limit processes in Theorem 1.5 and Corollary 4.2 are different, due to the additional random variation introduced when randomizing the time. (The variance of the limit in Theorem 1.5 is strictly smaller than in Corollary 4.2 at every $t \notin \{0, \frac{1}{2}, 1\}$.)

We will also need a moment estimate. It is easy to see that $n^{1/2}S_{n,i}(t)$ has moments that are bounded as $n \to \infty$, for every fixed $t \in [0,1]$. We extend that to the supremum over all t.

Theorem 4.3. *Let* $S^*_{n,i} := \sup_{0\le t\le 1} |S_{n,i}(t)|$ *for* $i = 1,2$. *Then, for each fixed* $r > 0$, $\mathbb{E}\bigl(S^*_{n,i}\bigr)^r = O(n^{r/2})$.

We are primarily interested in the maximum X^*_n of $X_n(t)$. It is evident from Corollary 4.2 that the maximum is attained close to the maximum point of $t(1-t)$, i.e., close to $t = 1/2$. We use a magnifying glass and study the processes close to $t = 1/2$ in greater detail. The correct scaling turns out to be $t = \frac{1}{2} + xn^{-1/3}$, and we have the following asymptotic behaviour on that scale.

Theorem 4.4. *As* $n \to \infty$, *in* $D(-\infty,\infty)$,

$$n^{-1/3}\bigl(S_{n,1}(\tfrac{1}{2} + xn^{-1/3}) - S_{n,1}(\tfrac{1}{2})\bigr) \xrightarrow{\mathrm{d}} B_1(x), \tag{4.8}$$

$$n^{-1/3}\bigl(S_{n,2}(\tfrac{1}{2} + xn^{-1/3}) - S_{n,2}(\tfrac{1}{2})\bigr) \xrightarrow{\mathrm{d}} 2^{-1/2}B_2(x), \tag{4.9}$$

jointly, where B_1 *and* B_2 *are two independent Brownian motions on* $(-\infty,\infty)$. *Furthermore, for any fixed* $A < \infty$ *and* $i = 1,2$,

$$\mathbb{E}\max_{|x|\le A}\Bigl(S_{n,i}(\tfrac{1}{2} + xn^{-1/3}) - S_{n,i}(\tfrac{1}{2})\Bigr)^2 = O(n^{2/3}). \tag{4.10}$$

Corollary 4.5. *As* $n \to \infty$, *in* $D(-\infty,\infty)$,

$$n^{-1/3}\bigl(X_n(\tfrac{1}{2} + xn^{-1/3}) - X_n(\tfrac{1}{2})\bigr) \xrightarrow{\mathrm{d}} 2^{-1/2}B(x) - x^2, \tag{4.11}$$

where B *is a Brownian motion on* $(-\infty,\infty)$.

5. TIME-REVERSAL

In the proofs below, we will introduce factors that blow up at the endpoint $t = 1$. To see that there is no real problem at this endpoint, we will use a time reversal trick which enables us to transfer results from the other endpoint.

If we replace each T_k by $1-T_k$, which of course has the same distribution, then $I(t;k)$ becomes $1 - I(1-t;k)$, except at the jump point, and thus, see (2.6)–(2.8), $I'(t;k)$ becomes $-I'(1-t;k)$ and $S_{n,i}(t)$ becomes $(-1)^i S_{n,i}(1-t)$, again excepting the jump points. To be precise, let for a function f on $[0,1]$, $f(t-) := \lim_{s\uparrow t} f(s)$ (when this exists), with $f(0-) := f(0)$. Then $S_{n,i}(t)$ becomes $(-1)^i S_{n,i}((1-t)-)$ under this time-reversal, and thus

$$S_{n,i}(t) \overset{\mathrm{d}}{=} (-1)^i S_{n,i}((1-t)-), \tag{5.1}$$

as functions in $D[0,1]$ and jointly for $i = 1,2$.

6. Proofs

The proofs are based on martingale theory, in particular a continuous time martingale limit theorem by Jacod and Shiryaev [12]. We will use the *quadratic variation* $[X, X]_t$ of a martingale X (in continuous time) and its bilinear extension $[X, Y]_t$ to two martingales X and Y. For a general definition see e.g. [12]; for us it will suffice to know that, if X and Y are martingales of pathwise finite variation, then

$$[X, Y]_t = \sum_{0<s\le t} \Delta X(s)\Delta Y(s), \tag{6.1}$$

where $\Delta X(s) := X(s) - X(s-)$ is the jump of X at s and, similarly, $\Delta Y(s) := Y(s) - Y(s-)$. The sum in (6.1) is formally uncountable, but in reality countable since there is only a countable number of jumps; in the applications below, the sum will be finite.

A real-valued martingale $X(s)$ on $[0, t]$ is an L^2-martingale if and only if $\mathbb{E}[X, X]_t < \infty$ and $\mathbb{E}\,|X(0)|^2 < \infty$, and then

$$\mathbb{E}\,|X(t)|^2 = \mathbb{E}[X, X]_t + \mathbb{E}\,|X(0)|^2. \tag{6.2}$$

We will use the following general result based on [12]; see [15, Proposition 9.1] for a detailed proof (for $I = [0, \infty)$; the general case is the same). (See also [13] and [14] for similar versions).

Proposition 6.1. *Let $I = [a, b]$ or $I = [a, b)$, with $-\infty < a < b \le \infty$. Assume that for each n, $\mathcal{M}_n(t) = (\mathcal{M}_{ni}(t))_{i=1}^q$ is a q-dimensional martingale on I with $\mathcal{M}_n(a) = 0$, and that $\Sigma(t) = (\Sigma_{ij}(t))_{i,j=1}^q$ is a (non-random) continuous matrix-valued function on I such that for every fixed $t \in I$ and $i, j = 1, \dots, q$,*

$$[\mathcal{M}_{ni}, \mathcal{M}_{nj}]_t \xrightarrow{\text{p}} \Sigma_{ij}(t) \quad \text{as } n \to \infty, \tag{6.3}$$

$$\sup_n \mathbb{E}[\mathcal{M}_{ni}, \mathcal{M}_{ni}]_t < \infty. \tag{6.4}$$

Then $\mathcal{M}_n \xrightarrow{\text{d}} \mathcal{M}_\infty$ as $n \to \infty$, in $D(I)$, where $\mathcal{M}_\infty = (\mathcal{M}_{\infty i})_{i=1}^q$ is a continuous q-dimensional Gaussian martingale with $\mathbb{E}\,\mathcal{M}_\infty(t) = 0$ and covariances

$$\mathbb{E}\big(\mathcal{M}_{\infty i}(s)\mathcal{M}_{\infty j}(t)\big) = \Sigma_{ij}(s \wedge t), \qquad s, t \in I.$$

In other words, the components $\mathcal{M}_{ni}(t)$ converge jointly to $\mathcal{M}_{\infty i}(t)$ in $D(I)$.

Remark 6.2. By (6.2), (6.4) is equivalent to $\sup_n \mathbb{E}\,|\mathcal{M}_n(t)|^2 < \infty$, the form used in e.g. [15].

Proof of Theorem 4.1. We first construct martingales from $S_{n,1}(t)$ and $S_{n,2}(t)$. We define, for $0 \le t < 1$,

$$\widehat{I}(t;k) := \frac{I'(t;k)}{1-t} = \begin{cases} 1, & I(t;k) = 1, \\ -t/(1-t), & I(t;k) = 0; \end{cases}$$

$$\widehat{S}_{n,1}(t) := \sum_{k=1}^{n} \widehat{I}(t;k) = (1-t)^{-1} S_{n,1}(t); \tag{6.5}$$

$$\widehat{S}_{n,2}(t) := \sum_{k=1}^{n-1} \widehat{I}(t;k)\widehat{I}(t;k+1) = (1-t)^{-2} S_{n,2}(t). \tag{6.6}$$

We have $\mathbb{E}\,\widehat{I}(t;k) = 0$ and

$$\mathbb{E}\big(\widehat{I}(t;k)^2\big) = \operatorname{Var}\big(\widehat{I}(t;k)\big) = (1-t)^{-2}\operatorname{Var}\big(I(t;k)\big) = \frac{t}{1-t}. \tag{6.7}$$

It is easily checked that each $\widehat{I}(t;k)$ is a martingale on $[0,1)$ [14, Lemma 2.1]; since these martingales for different k are independent, the products $\widehat{I}(t;k)\widehat{I}(t;k+1)$ are martingales too, and thus $\widehat{S}_{n,1}(t)$ and $\widehat{S}_{n,2}(t)$ are martingales on $[0,1)$ with $\widehat{S}_{n,1}(0) = \widehat{S}_{n,2}(0) = 0$. To calculate their quadratic variations and covariation, note that $\Delta\widehat{I}(t;k) = (1-t)^{-1}$ when $t = T_k$ and 0 otherwise. Further, with $\widehat{I}(t;0) := \widehat{I}(t;n+1) := 0$,

$$\Delta\widehat{S}_{n,1}(t) = \sum_{k=1}^{n} \Delta\widehat{I}(t;k),$$

$$\Delta\widehat{S}_{n,2}(t) = \sum_{k=1}^{n} \Delta\widehat{I}(t;k)\big(\widehat{I}(t;k-1) + \widehat{I}(t;k+1)\big),$$

and thus

$$[\widehat{S}_{n,1},\widehat{S}_{n,1}]_t = \sum_{s\le t}\sum_{k=1}^{n} \big|\Delta\widehat{I}(s;k)\big|^2 = \sum_{k=1}^{n} \frac{1}{(1-T_k)^2}\mathbf{1}[T_k \le t], \tag{6.8}$$

$$\begin{aligned}[\widehat{S}_{n,1},\widehat{S}_{n,2}]_t &= \sum_{s\le t}\sum_{k=1}^{n} \big|\Delta\widehat{I}(s;k)\big|^2 \big(\widehat{I}(s;k-1) + \widehat{I}(s;k+1)\big) \\ &= \sum_{k=1}^{n} \frac{1}{(1-T_k)^2}\big(\widehat{I}(T_k;k-1) + \widehat{I}(T_k;k+1)\big)\mathbf{1}[T_k \le t], \end{aligned} \tag{6.9}$$

$$\begin{aligned}[\widehat{S}_{n,2},\widehat{S}_{n,2}]_t &= \sum_{s\le t}\sum_{k=1}^{n} \big|\Delta\widehat{I}(s;k)\big|^2 \big(\widehat{I}(s;k-1) + \widehat{I}(s;k+1)\big)^2 \\ &= \sum_{k=1}^{n} \frac{1}{(1-T_k)^2}\big(\widehat{I}(T_k;k-1) + \widehat{I}(T_k;k+1)\big)^2\mathbf{1}[T_k \le t]. \end{aligned} \tag{6.10}$$

Hence, since the T_k are independent and uniformly distributed on [0,1], and using (6.7),

$$\mathbb{E}[\widehat{S}_{n,1},\widehat{S}_{n,1}]_t = n\int_0^t \frac{\mathrm{d}s}{(1-s)^2} = n\left[\frac{1}{1-s}\right]_0^t = n\frac{t}{1-t}, \tag{6.11}$$

$$\mathbb{E}[\widehat{S}_{n,1},\widehat{S}_{n,2}]_t = \sum_{k=1}^n \int_0^t \frac{\mathrm{d}s}{(1-s)^2}\,\mathbb{E}\bigl(\widehat{I}(s;k-1)+\widehat{I}(s;k+1)\bigr) = 0, \tag{6.12}$$

$$\begin{aligned}\mathbb{E}[\widehat{S}_{n,2},\widehat{S}_{n,2}]_t &= \sum_{k=1}^n \int_0^t \frac{\mathrm{d}s}{(1-s)^2}\,\mathbb{E}\bigl(\widehat{I}(s;k-1)+\widehat{I}(s;k+1)\bigr)^2 \\ &= ((n-2)\cdot 2+2\cdot 1)\int_0^t \frac{\mathrm{d}s}{(1-s)^2}\frac{s}{1-s} = (n-1)\frac{t^2}{(1-t)^2}\end{aligned} \tag{6.13}$$

(Indeed, these formulas also follow directly from (2.11)–(2.13) by (6.5), (6.6) and (6.2) together with its polarized version for two martingales.)

Moreover, the k:th and l:th terms in the sums in (6.8)–(6.10) are independent when $|k-l|>2$, and each term is $O\bigl((1-t)^{-2}\bigr)$. Hence, for $i,j\in\{1,2\}$,

$$\operatorname{Var}\bigl([\widehat{S}_{n,i},\widehat{S}_{n,j}]_t\bigr) = O\bigl(n(1-t)^{-4}\bigr). \tag{6.14}$$

Define now, for $i=1,2$ and $0\le t<1$,

$$\widetilde{S}_{n,i}(t) := n^{-1/2}\widehat{S}_{n,i}(t). \tag{6.15}$$

By (6.11)–(6.14), for every fixed $t\in[0,1)$,

$$\begin{aligned}[\widetilde{S}_{n,1},\widetilde{S}_{n,1}]_t &\overset{\mathrm{p}}{\longrightarrow} \frac{t}{1-t},\\ [\widetilde{S}_{n,1},\widetilde{S}_{n,2}]_t &\overset{\mathrm{p}}{\longrightarrow} 0,\\ [\widetilde{S}_{n,2},\widetilde{S}_{n,2}]_t &\overset{\mathrm{p}}{\longrightarrow} \frac{t^2}{(1-t)^2}.\end{aligned}$$

Proposition 6.1 thus applies with $I=[0,1)$, with (6.4) verified by (6.15), (6.11) and (6.13), which shows that

$$n^{-1/2}\widehat{S}_{n,i}(t) = \widetilde{S}_{n,i}(t) \overset{\mathrm{d}}{\longrightarrow} \widehat{Z}_i(t), \qquad i=1,2, \tag{6.16}$$

jointly in $D[0,1)$, where $\widehat{Z}_1(t)$ and $\widehat{Z}_2(t)$ are continuous Gaussian processes on $[0,1)$ with means 0 and covariances, for $0\le s\le t<1$,

$$\mathbb{E}\bigl(\widehat{Z}_1(s)\widehat{Z}_1(t)\bigr) = \frac{s}{1-s}, \tag{6.17}$$

$$\mathbb{E}\bigl(\widehat{Z}_1(s)\widehat{Z}_2(t)\bigr) = 0, \tag{6.18}$$

$$\mathbb{E}\bigl(\widehat{Z}_2(s)\widehat{Z}_2(t)\bigr) = \frac{s^2}{(1-s)^2}. \tag{6.19}$$

Note that (6.18) implies that $\widehat{Z}_1$ and $\widehat{Z}_2$ are independent.

We define $Z_i(t) := (1-t)^i \widehat{Z}_i(t)$ for $t \in [0,1)$, and $Z_i(1) := 0$. Then (6.16) implies, by (6.5) and (6.6), that (4.1) and (4.2) hold jointly in $D[0,1)$. Furthermore, (6.17)–(6.19) imply that Z_1 and Z_2 have the covariances (4.3) and (4.4).

It remains to extend this from $[0,1)$ to $[0,1]$. We use the time-reversal trick in Section 5 and have by (5.1) and the result just shown

$$n^{-1/2} S_{n,i}((1-t)-) \overset{\mathrm{d}}{=} (-1)^i n^{-1/2} S_{n,i}(t) \overset{\mathrm{d}}{\longrightarrow} (-1)^i Z_i(t) \overset{\mathrm{d}}{=} Z_i(t)$$

in $D[0,1)$, and thus $n^{-1/2} S_{n,i}(t) \overset{\mathrm{d}}{\longrightarrow} Z_i(1-t)$ in $D(0,1]$. Clearly $Z_i(1-t) \overset{\mathrm{d}}{=} Z_i(t)$, as processes on $[0,1]$; since Z_i is continuous on $[0,1)$, this implies continuity at 1 too, and thus Z_i is continuous on [0,1]. We have shown that the limits (4.1) and (4.2) hold (jointly) in both $D[0,1)$ and $D(0,1]$, which easily implies that they hold in $D[0,1]$ too, see e.g. [14, Lemma 2.3]. □

Proof of Corollary 4.2. Immediate by Theorem 4.1 and (2.10), with the limit $Z(t) := (1-2t)Z_1(t) - Z_2(t)$, since $n^{-1/2} R_n(t) \to 0$ uniformly. The covariances (4.6)–(4.7) of $Z(t)$ follow from (4.3), (4.4) and the independence of Z_1 and Z_2. □

Proof of Theorem 4.3. By Hölder's inequality, it suffices to prove the result when r is an even integer. Since $S^*_{n,i} \le \sup_{0 \le t \le 1/2} |S_{n,i}(t)| + \sup_{1/2 \le t \le 1} |S_{n,i}(t)|$ and the time-reversal symmetry (5.1) implies

$$\sup_{1/2 \le t \le 1} |S_{n,i}(t)| \overset{\mathrm{d}}{=} \sup_{0 \le t \le 1/2} |S_{n,i}(t)|, \tag{6.20}$$

it is sufficient to consider $\sup_{0 \le t \le 1/2} |S_{n,i}(t)|$. Moreover, $|S_{n,i}(t)| \le \left|\widehat{S}_{n,i}(t)\right|$, and by Doob's maximal inequality for martingales, see e.g. [16, Proposition 7.16],

$$\mathbb{E}\Bigl(\sup_{0 \le t \le 1/2} |S_{n,i}(t)|\Bigr)^r \le \mathbb{E}\Bigl(\sup_{0 \le t \le 1/2} \left|\widehat{S}_{n,i}(t)\right|\Bigr)^r \le C_r\, \mathbb{E}\Bigl(\left|\widehat{S}_{n,i}(\tfrac{1}{2})\right|\Bigr)^r, \tag{6.21}$$

for some constant C_r $(= (r/(r-1))^r)$.

Finally, $\widehat{S}_{n,1}(\frac{1}{2})$ is the sum of n independent random variables $\widehat{I}(\frac{1}{2};k)$, each with values ± 1 and mean 0, and it is easily verified that, with $r = 2\ell$,

$$\mathbb{E}\Bigl(\left|\widehat{S}_{n,i}(\tfrac{1}{2})\right|\Bigr)^r = O(n^\ell) = O(n^{r/2}). \tag{6.22}$$

Similarly, $\widehat{S}_{n,2}(\frac{1}{2})$ is the sum of the $n-1$ random variables $\widehat{I}(\frac{1}{2};k)\widehat{I}(\frac{1}{2};k+1)$; these variables too have values ± 1 and mean 0; moreover, it is easily verified that they too are independent. Hence $\widehat{S}_{n,2}(\frac{1}{2}) \overset{\mathrm{d}}{=} S_{n-1,1}(\frac{1}{2})$, and (6.22) implies the same estimate for $\widehat{S}_{n,2}(\frac{1}{2})$ too.

The result follows by this, (6.21) and (6.20). □

Proof of Theorem 1.2. We claim that Corollary 4.2 implies that

$$n^{-1/2}\Bigl(\max_{0 \le t \le 1} X_n(t) - \tfrac{1}{4}n\Bigr) \overset{\mathrm{d}}{\longrightarrow} Z(\tfrac{1}{2}) := -Z_2(\tfrac{1}{2}), \tag{6.23}$$

which gives (1.3) by (4.4). (We could use Theorem 1.5 instead.) The argument was sketched in the introduction, and this is an application of [14, Theorem 16], but for completeness we give the details in our case. We may for simplicity use the Skorohod coupling theorem [16, Theorem 4.30], which says that we can assume that (4.5) holds with convergence a.s. and not just in distribution. Thus, for (almost) every point in our probability space, $n^{-1/2}\bigl(X_n(t)-nt(1-t)\bigr)\to Z(t)$ in $D[0,1]$, which since $Z(t)$ is continuous means uniform convergence on [0,1]. In other words, uniformly in $t\in[0,1]$,

$$X_n(t)=nt(1-t)+n^{1/2}Z(t)+o(n^{1/2})=\tfrac14 n-n(\tfrac12-t)^2+n^{1/2}Z(t)+o(n^{1/2}). \tag{6.24}$$

In particular,

$$X_n^*\ge X_n(\tfrac12)=\tfrac14 n+n^{1/2}Z(\tfrac12)+o(n^{1/2}). \tag{6.25}$$

Conversely, (6.24) yields for $|t-\frac12|<n^{-1/8}$, since Z is continuous,

$$X_n(t)\le \tfrac14 n+n^{1/2}Z(t)+o(n^{1/2})=\tfrac14 n+n^{1/2}Z(\tfrac12)+o(n^{1/2}), \tag{6.26}$$

and for $|t-\frac12|\ge n^{-1/8}$, since Z is bounded,

$$X_n(t)\le \tfrac14 n-n^{1-1/4}+O(n^{1/2})\le \tfrac14 n+n^{1/2}Z(\tfrac12) \tag{6.27}$$

for large n. It follows from (6.25), (6.26) and (6.27) that

$$X_n^*=\tfrac14 n+n^{1/2}Z(\tfrac12)+o(n^{1/2}),$$

and (6.23) follows.

To prove moment convergence in (1.3), it is, as is well-known, see e.g. [11, Theorems 5.4.2 and 5.5.9], enough to prove that for each fixed $r>0$, the r:th absolute moment of the left hand side is bounded, as $n\to\infty$. By (2.10),

$$\begin{aligned}\bigl|X_n^*-\tfrac14 n\bigr|&=\bigl|\sup_t X_n(t)-\sup_t nt(1-t)\bigr|\le\sup_t\bigl|X_n(t)-nt(1-t)\bigr|\\&\le S_{n,1}^*+S_{n,2}^*+3,\end{aligned}$$

and the required estimate follows by Theorem 4.3. □

Proof of Theorem 1.5. Recall the order statistics $T_{(n;m)}$ from Section 2. Since $X_{n,m}=X_n(T_{(n;m)})$, we are studying the process $X_{n,\lfloor nt\rfloor}=X_n(T_{(n;\lfloor nt\rfloor)})$. The idea of the proof is to use the functional limit results just shown and replace t by the random time $T_{(n;\lfloor nt\rfloor)}$. Note first that $N_n(T_{(n;m)})=m$ and thus

$$N_n(T_{(n;\lfloor nt\rfloor)})=\lfloor nt\rfloor=nt+O(1). \tag{6.28}$$

By (2.7),

$$\begin{aligned}\sup_{0\le m\le n}\bigl|N_n(T_{(n;m)})/n-T_{(n;m)}\bigr|\le\sup_{0\le t\le1}\bigl|N_n(t)/n-t\bigr|&=\sup_{0\le t\le1}\bigl|n^{-1}S_{n,1}(t)\bigr|\\&=n^{-1}S_{n,1}^*,\end{aligned} \tag{6.29}$$

which by (4.1) (or Theorem 4.3, or the Glivenko–Cantelli theorem [16, Proposition 4.24]) tends to 0 in probability. Thus, by (6.28),

$$\sup_{0\le t\le 1}\bigl|t-T_{(n;\lfloor nt\rfloor)}\bigr| \le \sup_{0\le t\le 1}\bigl|N_n(T_{(n;\lfloor nt\rfloor)})/n-T_{(n;\lfloor nt\rfloor)}\bigr|+n^{-1}\overset{\mathrm{p}}{\longrightarrow}0. \tag{6.30}$$

The proof of Corollary 4.2 shows that (4.5) holds jointly with (4.1) and (4.2), with $Z(t)=(1-2t)Z_1(t)-Z_2(t)$. Furthermore, by (2.7), $N_n(t)/n = t+S_{n,1}(t)/n$, and a Taylor expansion of the function $t\mapsto nt(1-t)$ yields

$$N_n(t)\bigl(1-N_n(t)/n\bigr)=nt(1-t)+(1-2t)S_{n,1}(t)-S_{n,1}(t)^2/n.$$

Consequently, by (4.1), in $D[0,1]$, still jointly with (4.5),

$$n^{-1/2}\Bigl(N_n(t)\bigl(1-N_n(t)/n\bigr)-nt(1-t)\Bigr)\overset{\mathrm{d}}{\longrightarrow}(1-2t)Z_1(t),$$

and subtracting this from (4.5) yields

$$n^{-1/2}\Bigl(X_n(t)-N_n(t)\bigl(1-N_n(t)/n\bigr)\Bigr)\overset{\mathrm{d}}{\longrightarrow}Z(t)-(1-2t)Z_1(t)=-Z_2(t). \tag{6.31}$$

Because (6.30) holds and $Z_2(t)$ is continuous, we may replace t by $T_{(n;\lfloor nt\rfloor)}$ on the left hand side; for a formal verification of this we may again use the Skorohod coupling theorem [16, Theorem 4.30] and thus assume that (6.30) and (6.31) hold a.s., i.e. that the functions in (6.30) and (6.31) converge uniformly on [0,1] to their limits. Consequently,

$$n^{-1/2}\Bigl(X_{n,\lfloor nt\rfloor}-N_n(T_{(n;\lfloor nt\rfloor)})\bigl(1-N_n(T_{(n;\lfloor nt\rfloor)})/n\bigr)\Bigr)\overset{\mathrm{d}}{\longrightarrow}-Z_2(t), \tag{6.32}$$

which by (6.28) yields (1.5) with $Z(t)=-Z_2(t)$. □

The fact that the terms with $S_{n,1}$ cancel in the proof above is no coincidence. $S_{n,1}$ measures by (2.7) the random fluctuations introduced by used random insertion times T_k, and it is very intuitive that this term will appear in the limits for $X_n(t)$ but not for $X_{n,m}$. A theorem verifying that this cancellation happens in general in a situation closely related to the one studied here is given in [14, Theorem 7].

Proof of Theorem 4.4. Fix $A>0$, and define for $n>(2A)^3$ and $x\in[0,2A]$,

$$W_{n,i}(x):=\widehat{S}_{n,i}\bigl(\tfrac12+(x-A)n^{-1/3}\bigr)-\widehat{S}_{n,i}\bigl(\tfrac12-An^{-1/3}\bigr). \tag{6.33}$$

Then $W_{n,i}$ is a martingale on $[0,2A]$ with $W_{n,i}(0)=0$, and its quadratic variation is by (6.1)

$$[W_{n,i},W_{n,i}]_x=[\widehat{S}_{n,i},\widehat{S}_{n,i}]_{\frac12+(x-A)n^{-1/3}}-[\widehat{S}_{n,i},\widehat{S}_{n,i}]_{\frac12-An^{-1/3}}. \tag{6.34}$$

Hence, by (6.11)–(6.13), for $0 \le x \le 2A$,

$$\mathbb{E}[W_{n,1}, W_{n,1}]_x = n \int_{\frac{1}{2}-An^{-1/3}}^{\frac{1}{2}+(x-A)n^{-1/3}} \frac{\mathrm{d}s}{(1-s)^2} = n^{2/3}x(4+O(n^{-1/3})), \quad (6.35)$$

$$\mathbb{E}[W_{n,1}, W_{n,2}]_x = 0, \quad (6.36)$$

$$\mathbb{E}[W_{n,2}, W_{n,2}]_x = (2n-2) \int_{\frac{1}{2}-An^{-1/3}}^{\frac{1}{2}+(x-A)n^{-1/3}} \frac{s\,\mathrm{d}s}{(1-s)^3} = 2n^{2/3}x(4+O(n^{-1/3})). \quad (6.37)$$

Moreover, by (6.34) and (6.14), for $n > (4A)^3$, say,

$$\operatorname{Var}[W_{n,i}, W_{n,j}]_x = O(n).$$

Consequently, Proposition 6.1 applies to $n^{-1/3}W_{n,i}$, and shows that in $D[0, 2A]$ and jointly for $i = 1, 2$,

$$n^{-1/3}W_{n,i}(x) \xrightarrow{\mathrm{d}} W_i(x), \quad (6.38)$$

where W_1 and W_2 are independent Gaussian stochastic processes with means 0 and

$$\mathbb{E}\big(W_1(x)W_1(y)\big) = 4x, \qquad \mathbb{E}\big(W_2(x)W_2(y)\big) = 8x, \qquad 0 \le x \le y \le 2A.$$

In other words, $W_1(x) = 2B_1(x)$ and $W_2(x) = \sqrt{8}\,B_2(x)$, where B_1 and B_2 are independent Brownian motions on $[0, 2A]$. We may assume that B_1 and B_2 actually are independent two-sided Brownian motions defined on the entire real line. Note that $B_i(x+A) - B_i(A) \overset{\mathrm{d}}{=} B_i(x)$ (as processes on $\mathbb{R}$). Hence we can make a translation and obtain from (6.33) and (6.38), in $D[-A, A]$ and jointly for $i = 1, 2$,

$$n^{-1/3}\big(\widehat{S}_{n,i}\big(\tfrac{1}{2} + xn^{-1/3}\big) - \widehat{S}_{n,i}\big(\tfrac{1}{2}\big)\big) = n^{-1/3}\big(W_{n,i}(x+A) - W_{n,i}(A)\big) \xrightarrow{\mathrm{d}} 2^{(i+1)/2}\big(B_i(x+A) - B_i(A)\big) \overset{\mathrm{d}}{=} 2^{(i+1)/2}B_i(x). \quad (6.39)$$

By (6.5) and (6.6) we further have, uniformly for $n > (4A)^3$ and $x \in [-A, A]$,

$$S_{n,i}\big(\tfrac{1}{2} + xn^{-1/3}\big) - S_{n,i}\big(\tfrac{1}{2}\big) = (\tfrac{1}{2} - xn^{-1/3})^i \widehat{S}_{n,i}\big(\tfrac{1}{2} + xn^{-1/3}\big) - (\tfrac{1}{2})^i \widehat{S}_{n,i}\big(\tfrac{1}{2}\big) = 2^{-i}\Big(\widehat{S}_{n,i}\big(\tfrac{1}{2} + xn^{-1/3}\big) - \widehat{S}_{n,i}\big(\tfrac{1}{2}\big)\Big) + O(n^{-1/3}S^*_{n,i}). \quad (6.40)$$

and thus, using (6.39) and Theorem 4.3, in $D[-A, A]$ and jointly for $i = 1, 2$,

$$n^{-1/3}\big(S_{n,i}\big(\tfrac{1}{2} + xn^{-1/3}\big) - S_{n,i}\big(\tfrac{1}{2}\big)\big) \xrightarrow{\mathrm{d}} 2^{-i}2^{(i+1)/2}B_i(x) = 2^{(1-i)/2}B_i(x).$$

Since convergence in $D[-A, A]$ for every $A > 0$ implies convergence in $D(-\infty, \infty)$, this proves (4.8) and (4.9).

For the second moment estimate (4.10), we first note that (6.2), (6.35) and (6.37) show that, for each fixed A,

$$\mathbb{E}\,|W_{n,i}(2A)|^2 = \mathbb{E}[W_{n,i}, W_{n,i}]_{2A} = O(n^{2/3}), \tag{6.41}$$

and thus by Doob's maximal inequality [16, Proposition 7.16],

$$\mathbb{E}\Bigl(\max_{0\le x\le 2A} |W_{n,i}(x)|^2\Bigr) = O(n^{2/3}).$$

Hence, by (6.33) and translation again,

$$\begin{aligned}\mathbb{E}\max_{|x|\le A}\bigl|\widehat{S}_{n,i}\bigl(\tfrac12 + xn^{-1/3}\bigr) - \widehat{S}_{n,i}\bigl(\tfrac12\bigr)\bigr|^2 &= \mathbb{E}\max_{|x|\le A}\bigl|W_{n,i}(x+A) - W_{n,i}(A)\bigr|^2 \\ &\le 4\,\mathbb{E}\max_{0\le x\le 2A}\bigl|W_{n,i}(x)\bigr|^2 = O(n^{2/3}),\end{aligned}$$

and (4.10) follows by (6.40) and Theorem 4.3. □

Remark 6.3. We can extend (4.10) to arbitrary powers $r > 0$, with the estimate $O(n^{r/3})$, by the argument above with the Burkholder–Davis–Gundy inequalities [16, Theorem 26.12] replacing (6.41); we omit the details.

To study $X_n(t)$ close to $t = 1/2$, we rewrite (2.10) as, for $|x| \le 1/2$,

$$X_n\bigl(\tfrac12 + x\bigr) = \tfrac14 n - nx^2 - 2xS_{n,1}\bigl(\tfrac12 + x\bigr) - S_{n,2}\bigl(\tfrac12 + x\bigr) + R_n\bigl(\tfrac12 + x\bigr). \tag{6.42}$$

Hence, still for $|x| \le 1/2$,

$$\begin{aligned}X_n\bigl(\tfrac12 + x\bigr) - X_n\bigl(\tfrac12\bigr) = -nx^2 - 2xS_{n,1}\bigl(\tfrac12 + x\bigr) - \bigl(S_{n,2}\bigl(\tfrac12 + x\bigr) - S_{n,2}\bigl(\tfrac12\bigr)\bigr) \\ + R_n\bigl(\tfrac12 + x\bigr) - R_n\bigl(\tfrac12\bigr)\end{aligned} \tag{6.43}$$

and thus, for $|x| \le n^{1/3}/2$, recalling $|R_n(t)| \le 3$,

$$\begin{aligned}n^{-1/3}\Bigl(X_n\bigl(\tfrac12 + xn^{-1/3}\bigr) - X_n\bigl(\tfrac12\bigr)\Bigr) = -x^2 - 2n^{-2/3}xS_{n,1}\bigl(\tfrac12 + xn^{-1/3}\bigr) \\ - n^{-1/3}\bigl(S_{n,2}\bigl(\tfrac12 + xn^{-1/3}\bigr) - S_{n,2}\bigl(\tfrac12\bigr)\bigr) + O\bigl(n^{-1/3}\bigr).\end{aligned} \tag{6.44}$$

Proof of Corollary 4.5. Fix $A > 0$. Then (4.11) follows in $D[-A, A]$ by (6.44), (4.9) and Theorem 4.3 (which implies $n^{-2/3}S^*_{n,i} \overset{\mathrm{p}}{\longrightarrow} 0$), with $B(x) := -B_2(x)$. Since $A > 0$ is arbitrary, this yields convergence in $D(-\infty, \infty)$. □

Let $x_+ := x \vee 0$.

Lemma 6.4. *Let $x_1 > 0$ and suppose that $\mathcal{M}(x)$ is a martingale on $[0, x_1]$ with $\mathcal{M}(0) = 0$ such that for some constant K and all $x \in [0, x_1]$*

$$\operatorname{Var}\mathcal{M}(x) \le Kx.$$

Then, for every $a > 0$ and $x_0 \in (0, x_1]$,

$$\mathbb{E}\Bigl(\max_{x_0\le x\le x_1}\bigl(\mathcal{M}(x) - ax^2\bigr)_+\Bigr) \le \frac{4K}{ax_0}.$$

Proof. We may for convenience extend $\mathcal{M}$ to a martingale on $[0,\infty)$ by letting $\mathcal{M}(x) := \mathcal{M}(x_1)$ for $x > x_1$. Let $y > 0$ and $t > 0$. Then, by Kolmogorov-Doob's inequality [16, Proposition 7.16], [11, Theorem 10.9.1],

$$\begin{aligned}\mathbb{P}\Bigl(\sup_{x\in[y,2y]}\bigl(\mathcal{M}(x)-ax^2\bigr)>t\Bigr) &\le \mathbb{P}\Bigl(\sup_{x\in[0,2y]}\mathcal{M}(x)>t+ay^2\Bigr)\\ &\le \frac{\mathbb{E}\,\mathcal{M}(2y)^2}{(t+ay^2)^2}\le\frac{2Ky}{(t+ay^2)^2}.\end{aligned}$$

Integrating with respect to t from 0 to ∞ yields

$$\mathbb{E}\bigl(\sup_{x\in[y,2y]}\bigl(\mathcal{M}(x)-ax^2\bigr)_+\bigr)\le\int_0^\infty\frac{2Ky}{(t+ay^2)^2}=\frac{2Ky}{ay^2}=\frac{2K}{ay},$$

and the result follows by summing over $y = 2^k x_0$, $k = 0, 1, \dots$. □

Proof of Theorem 1.3. We begin by showing that we can replace $X_{n,\lceil n/2\rceil}$ by $X_n(\frac12)$ in the statement. By (6.29) and Theorem 4.3,

$$n^{1/3}\bigl|N_n(T_{(n;\lceil n/2\rceil)})/n-T_{(n;\lceil n/2\rceil)}\bigr|\le n^{-2/3}S^*_{n,1}\xrightarrow{\text{p}}0$$

and thus by (6.28)

$$n^{1/3}\bigl|T_{(n;\lceil n/2\rceil)}-\tfrac12\bigr|\xrightarrow{\text{p}}0. \tag{6.45}$$

It now follows from Corollary 4.5, arguing as for (6.32) and using (6.45) and the fact that the limit in (4.11) is continuous, that we can substitute $x = n^{1/3}\bigl(T_{(n;\lceil n/2\rceil)}-\frac12\bigr)$ in (4.11) and obtain

$$n^{-1/3}\bigl(X_{n,\lceil n/2\rceil}-X_n(\tfrac12)\bigr)=n^{-1/3}\bigl(X_n\bigl(T_{(n;\lceil n/2\rceil)}\bigr)-X_n(\tfrac12)\bigr)\xrightarrow{\text{p}}0. \tag{6.46}$$

Furthermore, by (3.1) and (3.2),

$$\mathbb{E}\,X_{n,\lceil n/2\rceil}=\tfrac14 n+O(1)=\mathbb{E}\,X_n(\tfrac12)+O(1). \tag{6.47}$$

Hence, it is enough to prove Theorem 1.3 with $X_{n,\lceil n/2\rceil}$ replaced by $X_n(\frac12)$; we thus study

$$M_n := X^*_n - X_n(\tfrac12)=\max_{t\in[0,1]}\bigl(X_n(t)-X_n(\tfrac12)\bigr). \tag{6.48}$$

We would like to take the supremum over all real x in (4.11), but that is not allowed without further arguments since the supremum is not a continuous functional on $D(-\infty,\infty)$ (the topology is too weak). We therefore fix a large $A > 0$ and study the following five intervals separately (assuming $n > (4A)^3$):

$$\begin{aligned}I_{-2} &:= [0,\tfrac14],\\ I_{-1} &:= [\tfrac14,\tfrac12-An^{-1/3}],\\ I_0 &:= [\tfrac12-An^{-1/3},\tfrac12+An^{-1/3}],\\ I_1 &:= [\tfrac12+An^{-1/3},\tfrac34],\\ I_2 &:= [\tfrac34,1].\end{aligned}$$

We denote further

$$M_{n,j} := \max_{t\in I_j}\bigl(X_n(t) - X_n(\tfrac12)\bigr)_+$$

and have thus, since $M_n \ge 0$,

$$M_n = \max_{-2\le j\le 2} M_{n,j} \le \sum_{j=-2}^{2} M_{n,j}. \tag{6.49}$$

On I_0 we use (4.11). Since the maximum is a continuous functional on $D(I)$ for any compact interval I, we obtain from (4.11) on $D[-A,A]$ immediately

$$n^{-1/3}M_{n,0} \xrightarrow{\mathrm{d}} V_A := \max_{|x|\le A}\bigl(2^{-1/2}B(x) - x^2\bigr). \tag{6.50}$$

Furthermore, it follows from (6.44), Theorem 4.3 and (4.10) that

$$\mathbb{E}\bigl(n^{-1/3}M_{n,0}\bigr)^2 \le C(A),$$

for some constant $C(A)$ depending on A but not on n. Hence the random variables $n^{-1/3}M_{n,0}$ are uniformly integrable, and (6.50) implies, see e.g. [11, Theorems 5.4.2 and 5.5.9],

$$\mathbb{E}\bigl(n^{-1/3}M_{n,0}\bigr) \to \mathbb{E}\,V_A. \tag{6.51}$$

On $I_{\pm2}$ we have by (6.43), with $\frac14 \le |x| \le \frac12$,

$$X_n\bigl(\tfrac12 + x\bigr) - X_n\bigl(\tfrac12\bigr) \le -n\bigl(\tfrac14\bigr)^2 + S^*_{n,1} + 2S^*_{n,2} + 6.$$

We use the elementary inequality, for $a > 0$ and $b \in \mathbb{R}$,

$$-a + b = -\frac{(a-b/2)^2}{a} + \frac{b^2}{4a} \le \frac{b^2}{4a}, \tag{6.52}$$

and obtain

$$M_{n,\pm2} \le \frac{(S^*_{n,1} + 2S^*_{n,2})^2}{4n/16} + 6 \le 8\frac{(S^*_{n,1})^2}{n} + 32\frac{(S^*_{n,2})^2}{n} + 6$$

and thus by Theorem 4.3

$$\mathbb{E}\,M_{n,\pm2} = O(1). \tag{6.53}$$

For I_1 we define

$$U_n(x) := -\tfrac14\bigl(\widehat{S}_{n,2}(\tfrac12 + x) - \widehat{S}_{n,2}(\tfrac12)\bigr); \tag{6.54}$$

this is a martingale on $[0,1/2)$. For $0 \le x \le \frac14$, we have

$$\tfrac14\widehat{S}_{n,2}\bigl(\tfrac12 + x\bigr) = (1-2x)^{-2}S_{n,2}\bigl(\tfrac12 + x\bigr) = \bigl(1 + O(x)\bigr)S_{n,2}\bigl(\tfrac12 + x\bigr),$$

and thus, using (6.43) and (6.52), for some constants $C_1, C_2, \dots,$

$$\begin{aligned} X_n\bigl(\tfrac12+x\bigr)-X_n\bigl(\tfrac12\bigr) &= -nx^2 - 2xS_{n,1}\bigl(\tfrac12+x\bigr) + U_n(x) + O(x)S_{n,2}\bigl(\tfrac12+x\bigr) + O(1) \\ &\le U_n(x) - \tfrac12 nx^2 + S^*_{n,1} + C_1 S^*_{n,2} - \tfrac12 nx^2 + O(1) \\ &\le U_n(x) - \tfrac12 nx^2 + C_2\frac{(S^*_{n,1})^2}{n} + C_3\frac{(S^*_{n,2})^2}{n} + O(1). \end{aligned} \tag{6.55}$$

By (6.1), (6.54) and (6.13) we further have, for $0 \le x \le 1/4$,

$$\begin{aligned} \operatorname{Var}\bigl(U_n(x)\bigr) = \mathbb{E}[U_n,U_n]_x &= \tfrac1{16}\,\mathbb{E}\Bigl([\widehat{S}_{n,2},\widehat{S}_{n,2}]_{\frac12+x} - [\widehat{S}_{n,2},\widehat{S}_{n,2}]_{\frac12}\Bigr) \\ &= \frac{2n-2}{16}\int_{\frac12}^{\frac12+x}\frac{s}{(1-s)^3}\,\mathrm{d}s \le C_4 nx. \end{aligned}$$

Hence, Lemma 6.4 yields, for $0 < x_0 \le \frac14$,

$$\mathbb{E}\Bigl(\max_{x_0\le x\le 1/4}\bigl(U_n(x)-\tfrac12 nx^2\bigr)_+\Bigr) \le \frac{C_5 n}{nx_0} = \frac{C_5}{x_0}.$$

Taking $x_0 = An^{-1/3}$ we thus obtain from (6.55), using Theorem 4.3 again,

$$\mathbb{E}\,M_{n,1} \le \frac{C_5}{An^{-1/3}} + O(1) = \frac{C_5}{A}n^{1/3} + O(1). \tag{6.56}$$

We obtain the same estimate for $M_{n,-1}$ by the time-reversal $t \mapsto 1-t$ and (5.1).

By (6.49) and the estimates (6.53) for $M_{n,\pm2}$ and (6.56) for $M_{n,\pm1}$ we find

$$\mathbb{E}\,|M_n - M_{n,0}| \le \mathbb{E}\,M_{n,-2} + \mathbb{E}\,M_{n,-1} + \mathbb{E}\,M_{n,1} + \mathbb{E}\,M_{n,2} \le C_6 + C_7 n^{1/3}/A,$$

and thus

$$\limsup_{n\to\infty}\mathbb{E}\bigl|n^{-1/3}M_n - n^{-1/3}M_{n,0}\bigr| \le C_7/A. \tag{6.57}$$

Now let $A \to \infty$; then

$$V_A \to V_\infty := \max_{x\in\mathbb{R}}\bigl(2^{-1/2}B(x) - x^2\bigr). \tag{6.58}$$

Note that, letting $x = y/2$, with V as in the statement of the theorem,

$$V_\infty = \max_{y\in\mathbb{R}}\bigl(2^{-1/2}B(y/2) - (y/2)^2\bigr) \overset{\mathrm{d}}{=} \max_{y\in\mathbb{R}}\bigl(2^{-1}B(y) - \tfrac14 y^2\bigr) = \tfrac12 V. \tag{6.59}$$

It follows from (6.57) and (6.58) that we may let $A \to \infty$ in (6.50) and obtain

$$n^{-1/3}M_n \overset{\mathrm{d}}{\longrightarrow} V_\infty, \tag{6.60}$$

see [4, Theorem 4.2] (we may change the notation and denote $M_{n,0}$ by $M_{n;A}$ for $n > (4A)^3$; for smaller n we simply let $M_{n;A} = 0$).

Similarly, by (6.51) and (6.57),

$$\limsup_{n\to\infty}\mathbb{E}\bigl|n^{-1/3}M_n - \mathbb{E}\,V_\infty\bigr| \le C_7 A^{-1} + \bigl|\mathbb{E}\,V_\infty - \mathbb{E}\,V_A\bigr|.$$

As $A \to \infty$, $\mathbb{E}\, V_A \to \mathbb{E}\, V_\infty$ by monotone convergence, and thus we obtain $\limsup_{n\to\infty} \mathbb{E}\bigl|n^{-1/3} M_n - \mathbb{E}\, V_\infty\bigr| = 0$, i.e.,

$$n^{-1/3}\, \mathbb{E}\, M_n \to \mathbb{E}\, V_\infty. \tag{6.61}$$

The theorem follows by (6.60), (6.61), (6.48), (6.46), (6.47) and (6.59). $\square$

7. Further results

Consider $X^{(1)}_{n,m}$, the number of piles with a single exam (runs of length 1) mentioned in Section 1. If we for simplicity consider the cyclic case, see Remark 1.4, to avoid edge effects (these are $O(1)$ only and do not affect the asymptotics), we have

$$X^{(1)}_{n,m} = \sum_{k=1}^{n} (1 - I_{n,m}(k)) I_{n,m}(k+1)(1 - I_{n,m}(k+2)).$$

After randomizing the time as in Section 2, we get (with $I(t; k+n) = I(t; k)$)

$$\begin{aligned} X^{(1)}_{n}(t) &= \sum_{k=1}^{n} \bigl(1 - I(t;k)\bigr) I(t;k+1) \bigl(1 - I(t;k+2)\bigr) \\ &= \sum_{k=1}^{n} \bigl(1 - t - I'(t;k)\bigr)\bigl(t + I'(t;k+1)\bigr)\bigl(1 - t - I'(t;k+2)\bigr) \\ &= nt(1-t)^2 + (1-3t)(1-t) S_{n,1}(t) - 2(1-t) S_{n,2}(t) + t S'_{n,2}(t) + S_{n,3}(t), \end{aligned} \tag{7.1}$$

where we now define $S_{n,2}$ by summing to n in (2.8), and we introduce two new stochastic processes

$$S'_{n,2}(t) := \sum_{k=1}^{n} I'(t;k) I'(t;k+2), \tag{7.2}$$

$$S_{n,3}(t) := \sum_{k=1}^{n} I'(t;k) I'(t;k+1) I'(t;k+2). \tag{7.3}$$

The proof of Theorem 4.1 extends to these and yields, in $D[0,1]$ and jointly with each other and (4.1) and (4.2),

$$n^{-1/2} S_{n,2}(t) \xrightarrow{\mathrm{d}} Z'_2(t), \tag{7.4}$$

$$n^{-1/2} S_{n,3}(t) \xrightarrow{\mathrm{d}} Z_3(t), \tag{7.5}$$

where Z'_2 and Z_3 are two continuous Gaussian processes on $[0,1]$ with means 0 and covariances

$$\mathbb{E}\bigl(Z'_2(s) Z'_2(t)\bigr) = s^2(1-t)^2, \qquad 0 \le s \le t \le 1, \tag{7.6}$$

$$\mathbb{E}\bigl(Z_3(s) Z_3(t)\bigr) = s^3(1-t)^3, \qquad 0 \le s \le t \le 1. \tag{7.7}$$

Furthermore, all four processes Z_1, Z_2, Z'_2 and Z_3 are independent. (Note that Z_2 and Z'_2 have the same distribution but are independent.)

By the arguments in Section 6, which extend without any new difficulties, this yields the following results, corresponding to our results for $X_{n,m}$ and $X_n(t)$ in Sections 1 and 4. We define $X_n^{(1)*} := \max_m X_{n,m}^{(1)} = \max_t X_n^{(1)}(t)$, and note that $\mathbb{E}\, X_n^{(1)}(t) = nt(1-t)^2$ has (on $[0,1]$) a unique maximum at $t = 1/3$.

Theorem 7.1. *As $n \to \infty$, in $D[0,1]$,*

$$n^{-1/2}\big(X_n^{(1)}(t) - nt(1-t)^2\big) \\ \xrightarrow{\mathrm{d}} Z(t) := (1-t)(1-3t)Z_1(t) - 2(1-t)Z_2(t) + tZ_2'(t) + Z_3(t);$$

Z is a continuous Gaussian process on $[0,1]$ with mean $\mathbb{E}\, Z(t) = 0$.

Theorem 7.2. *As $n \to \infty$, in $D[0,1]$,*

$$n^{-1/2}\big(X_{n,\lfloor nt \rfloor}^{(1)} - nt(1-t)^2\big) \xrightarrow{\mathrm{d}} Z(t) := -2(1-t)Z_2(t) + tZ_2'(t) + Z_3(t);$$

where Z is a continuous Gaussian process on $[0,1]$ with mean $\mathbb{E}\, Z(t) = 0$.

We leave the explicit formulas for (co)variances in these theorems to the reader.

Theorem 7.3. *As $n \to \infty$,*

$$n^{-1/2}\big(X_n^{(1)*} - \tfrac{4}{27}n\big) \xrightarrow{\mathrm{d}} N(0, \tfrac{76}{729}),$$

with convergence of all moments. In particular,

$$\mathbb{E}\, X_n^{(1)*} = \frac{4}{27}n + o(n^{1/2}),$$
$$\operatorname{Var} X_n^{(1)*} = \frac{76}{729}n + o(n).$$

Theorem 7.4. *As $n \to \infty$, in $D(-\infty,\infty)$,*

$$n^{-1/3}\big(X_n^{(1)}(\tfrac{1}{3} + xn^{-1/3}) - X_n^{(1)}(\tfrac{1}{3})\big) \xrightarrow{\mathrm{d}} \sqrt{\frac{80}{81}}B(x) - x^2,$$

where B is a Brownian motion on $(-\infty,\infty)$.

Theorem 7.5. *As $n \to \infty$,*

$$n^{-1/3}\big(X_n^{(1)*} - X_{n,\lfloor n/3 \rfloor}^{(1)}\big) \xrightarrow{\mathrm{d}} \beta V,$$

where the random variable V is as in Theorem 1.3 and $\beta := 2^{7/3}3^{-8/3}5^{2/3} = \frac{4}{27}(150)^{1/3}$. Furthermore,

$$\mathbb{E}\, X_n^{(1)*} = \mathbb{E}\, X_{n,\lfloor n/3 \rfloor}^{(1)} + \beta\, \mathbb{E}\, V n^{1/3} + o(n^{1/3}) = \tfrac{4}{27}n + \beta\, \mathbb{E}\, V n^{1/3} + o(n^{1/3}).$$

These results are easily extended to the number $X_{n,m}^{(d)}$ of piles with exactly d items (runs with exactly d 1's) for any fixed d. We may also count occurrences of any other fixed pattern, and more generally any functional of the type

$$\bar{X}_{n,m} := \sum_{k=1}^{n} \psi\bigl(I_{n,m}(k), \dots, I_{n,m}(k+\ell-1)\bigr) \tag{7.8}$$

for some fixed $\ell \geq 1$ and function $\psi : \{0,1\}^{\ell} \to \mathbb{R}$. We will pursue this in some detail, leave some other details to the reader, because the more general version illuminates the arguments above and the structure of our method. First, randomizing the time yields

$$\bar{X}_n(t) := \sum_{k=1}^{n} \Psi_k(t), \tag{7.9}$$

where we define $\Psi_k(t) = \psi\bigl(I(t;k), \dots, I(t;k+\ell-1)\bigr)$. We note that we will need more processes of the type $S_{n,j}$. It turns out that it is natural to use finite sequences of 0's and 1's to index these processes; we thus change the notation and define a stochastic process $S_{n;\alpha}(t)$ for each such sequence $\alpha = \alpha_1 \cdots \alpha_\ell$ by

$$S_{n;\alpha}(t) := \sum_{k=1}^{n} \prod_{\substack{j\in\{1,\dots,\ell\},\\ \alpha_j=1}} I'(t;k+j). \tag{7.10}$$

We thus now denote $S_{n,1}(t)$, $S_{n,2}(t)$, $S'_{n,2}(t)$, $S_{n,3}(t)$ by $S_{n;1}(t)$, $S_{n;11}(t)$, $S_{n;101}(t)$, $S_{n;111}(t)$. Initial and final 0's in α do not affect $S_{n;\alpha}$, so it is enough to consider α that begin and end with 1; let $\mathcal{A}$ be the set of all such strings α.

Let $\ell(\alpha)$ denote the length of α and $\nu(\alpha)$ the number of 1's in α, and consider only $n \geq 2\ell(\alpha)$. Then the terms in the sum in (7.10) are orthogonal and we obtain $\mathbb{E}\bigl(S_{n;\alpha}(t)\bigr)^2 = n\bigl(t(1-t)\bigr)^{\nu(\alpha)}$. Moreover, $\widehat{S}_{n;\alpha}(t) := (1-t)^{-\nu(\alpha)}S_{n;\alpha}(t)$ is a martingale on $[0,1)$, and the proof of Theorem 4.1 extends immediately and shows that, in $D[0,1]$ and jointly for all $\alpha \in \mathcal{A}$,

$$n^{-1/2}S_{n;\alpha}(t) \xrightarrow{\mathrm{d}} Z_\alpha(t), \tag{7.11}$$

where Z_α, $\alpha \in \mathcal{A}$, are independent continuous Gaussian processes with means 0 and covariances

$$\mathbb{E}\bigl(Z_\alpha(s)Z_\alpha(t)\bigr) = s^{\nu(\alpha)}(1-t)^{\nu(\alpha)}, \qquad 0 \leq s \leq t \leq 1. \tag{7.12}$$

Furthermore, the estimate Theorem 4.3 extends to every $S_{n;\alpha}$ (with the implicit constant possibly depending on α).

A functional $\bar{X}_{n,m}$ of the type (7.8) yields after randomizing the time the functional $\bar{X}_n(t)$ in (7.9), which always can be expanded as a finite sum (with orthogonal terms)

$$\bar{X}_n(t) = g_0(t)n + \sum_{\alpha\in\mathcal{A},\, \ell(\alpha)\leq\ell} g_\alpha(t)S_{n;\alpha}(t), \tag{7.13}$$

for some polynomials $g_0(t)$ and $g_\alpha(t)$, $\alpha \in \mathcal{A}$; this is seen by the same argument as in [14, Proposition 4.1]. Note that, for any $n \geq 2\ell$ and any k,

$$g_0(t) = \frac{\mathbb{E}\,\bar{X}_n(t)}{n} = \mathbb{E}\,\Psi_k(t). \tag{7.14}$$

It follows from (7.11) and (7.13) that, cf. Corollary 4.2, in $D[0,1]$,

$$n^{-1/2}\bigl(\bar{X}_n(t) - ng_0(t)\bigr) \xrightarrow{\mathrm{d}} Z(t) := \sum_\alpha g_\alpha(t) Z_\alpha(t), \tag{7.15}$$

which is a continuous Gaussian process with mean 0 and covariance function

$$\mathbb{E}\bigl(Z(s)Z(t)\bigr) = \sigma(s,t) := \sum_\alpha g_\alpha(s) g_\alpha(t) (s \wedge t)^{\nu(\alpha)} (1 - s \vee t)^{\nu(\alpha)}. \tag{7.16}$$

We have moment convergence in (7.15); moreover, the variance of $n^{-1/2}\bar{X}_n(t)$ is independent of $n \geq 2\ell$ and we have, for any k and $n \geq 2\ell$,

$$\sigma(s,t) = n^{-1} \operatorname{Cov}\bigl(\bar{X}_n(s), \bar{X}_n(t)\bigr) = \sum_{j=-(\ell-1)}^{\ell-1} \operatorname{Cov}\bigl(\Psi_k(s), \Psi_{k+j}(t)\bigr). \tag{7.17}$$

Similarly, arguing as in the proof of Theorem 1.5 and observing that the $S_{n;1}$ terms cancel because $g_1(t) = g_0'(t)$, we see that, in $D[0,1]$,

$$n^{-1/2}\bigl(\bar{X}_{n,\lfloor nt \rfloor} - ng_0(t)\bigr) \xrightarrow{\mathrm{d}} Z'(t) := \sum_{\alpha \neq 1} g_\alpha(t) Z_\alpha(t), \tag{7.18}$$

another continuous Gaussian process with mean 0.

Now suppose that $g_0(t)$ has a unique maximum on $[0,1]$ at an interior point t_0, with $g_0''(t_0) < 0$. Then all remaining proofs in Section 6 extend too without difficulties. In particular, if we define $\bar{X}_n^* := \max_m \bar{X}_{n,m}$, we have the following generalization of Theorem 1.2.

Theorem 7.6. *As $n \to \infty$,*

$$n^{-1/2}\bigl(\bar{X}_n^* - g(t_0)n\bigr) \xrightarrow{\mathrm{d}} N(0,\sigma^2),$$

with convergence of all moments, with, see (7.16),

$$\sigma^2 := \sigma(t_0,t_0) = \sum_\alpha g_\alpha(t_0)^2 \bigl(t_0(1-t_0)\bigr)^{\nu(\alpha)}.$$

Furthermore, cf. Theorem 4.4, in $D(-\infty,\infty)$ and jointly for all α,

$$n^{-1/3}\bigl(S_{n;\alpha}(t_0 + xn^{-1/3}) - S_{n;\alpha}(t_0)\bigr) \xrightarrow{\mathrm{d}} \sigma_\alpha B_\alpha(x), \tag{7.19}$$

where $\sigma_\alpha^2 := \nu(\alpha)\bigl(t_0(1-t_0)\bigr)^{\nu(\alpha)-1}$ and B_α, $\alpha \in \mathcal{A}$, are independent Brownian motions on $(-\infty,\infty)$. As a consequence, cf. Corollary 4.5, in $D(-\infty,\infty)$,

$$n^{-1/3}\bigl(\bar{X}_n(t_0 + xn^{-1/3}) - \bar{X}_n(t_0)\bigr) \xrightarrow{\mathrm{d}} \sigma_* B(x) - \tfrac{1}{2}|g_0''(t_0)| x^2, \tag{7.20}$$

where B is a Brownian motion on $(-\infty,\infty)$ and

$$\sigma_*^2 := \sum_\alpha g_\alpha(t_0)^2 \sigma_\alpha^2 = \sum_\alpha g_\alpha(t_0)^2 \nu(\alpha) \bigl(t_0(1-t_0)\bigr)^{\nu(\alpha)-1}. \tag{7.21}$$

Finally, substituting $x = (\sigma_*/|g_0''(t_0)|)^{2/3}y$, we see that $\sup_{x\in\mathbb{R}}\bigl(\sigma_* B(x) - \frac{1}{2}|g_0''(t_0)|x^2\bigr) \overset{\mathrm{d}}{=} \beta V$, with

$$\beta := (\sigma_*^4/|g_0''(t_0)|)^{1/3} = (\sigma_*^2)^{2/3}|g_0''(t_0)|^{-1/3}, \tag{7.22}$$

and we obtain the following, where $\bar{X}_n(t_0)$ may be replaced by $\bar{X}_{n,m_0}$, where either $m_0 := \lfloor t_0 n\rfloor$ or m_0 is chosen in $\{0,\dots,n\}$ to maximize $\mathbb{E}\,\bar{X}_{n,m_0}$.

Theorem 7.7. *As $n \to \infty$,*

$$n^{-1/3}\bigl(\bar{X}_n^* - \bar{X}_n(t_0)\bigr) \overset{\mathrm{d}}{\longrightarrow} \beta V, \tag{7.23}$$

where the random variable V is as in Theorem 1.3, and β is given by (7.22) *and* (7.21). *Furthermore,*

$$\mathbb{E}\,\bar{X}_n^* = \mathbb{E}\,\bar{X}(t_0) + \beta\,\mathbb{E}\,V n^{1/3} + o(n^{1/3}) = g_0(t_0)n + \beta\,\mathbb{E}\,V n^{1/3} + o(n^{1/3}).$$

For calculation of the asymptotic variances σ^2 and σ_*^2 above, the given formulas using the coefficients $g_\alpha(t_0)$ in the decomposition (7.13) are often not very convenient. For σ^2, it is usually simpler to use (7.17) with $s = t = t_0$.

For σ_*^2 we first observe that if we take the difference of the left derivative of $\sigma(s,t)$ with respect to s and the right derivative with respect to t at (t_0,t_0) (thus considering $s \le t$ only), we obtain from (7.16) and (7.21) easily

$$\sigma_*^2 = \frac{\partial}{\partial s}\sigma(s,t_0)\Big|_{s=t_0-} - \frac{\partial}{\partial t}\sigma(t_0,t)\Big|_{t=t_0+}\ , \tag{7.24}$$

a formula given by Daniels [6] (in a slightly different setting). It follows by the mean value theorem and (7.17) that, for any fixed $n \ge 2\ell$,

$$\sigma_*^2 = \lim_{h\downarrow 0}\frac{1}{h}\bigl(\sigma(t_0+h,t_0+h) - 2\sigma(t_0,t_0+h) + \sigma(t_0,t_0)\bigr) \tag{7.25}$$

$$= \lim_{h\downarrow 0}\frac{1}{hn}\operatorname{Var}\bigl(\bar{X}_n(t_0+h) - \bar{X}_n(t_0)\bigr). \tag{7.26}$$

For fixed n, the probability that exactly one $I(t;k)$ changes from 0 to 1 in the interval $[t_0, t_0+h]$ is $nh + O(h^2)$ and the probability that more than one will change is $O(h^2)$. Hence, if $\Delta_k\bar{X}_n(t)$ is the function of $\{I(t;k+j)\}_{1\le|j|<\ell}$ that gives the jump in $\bar{X}_n(t)$ (for $n \ge 2\ell$) if $I(t;k)$ is changed from 0 to 1, keeping all other indicators fixed, then (7.26) implies that, for any k and $n \ge 2\ell$,

$$\sigma_*^2 = \operatorname{Var}\bigl(\Delta_k\bar{X}_n(t_0)\bigr). \tag{7.27}$$

For example, for the number of runs,

$$\Delta_k X_n(t) = \bigl(1 - I(t;k-1)\bigr) - I(t;k+1)$$

and $\sigma_*^2 = \operatorname{Var}\bigl(I(t_0;k-1) + I(t_0;k+1)\bigr) = 2t_0(1-t_0) = \frac{1}{2}$, in accordance with Corollary 4.5. For runs of length 1 we similarly get, from (7.1),

$$\begin{aligned}\Delta_k X_n^{(1)}(t) = &-\bigl(1 - I(t;k-2)\bigr)I(t;k-1)\\ &+ \bigl(1 - I(t;k-1)\bigr)\bigl(1 - I(t;k)\bigr) - I(t;k+1)\bigl(1 - I(t;k+2)\bigr)\end{aligned}$$

and (7.27) yields, in accordance with Theorem 7.4,

$$\sigma_*^2 = \operatorname{Var}\bigl(\Delta_k X_n^{(1)}(t_0)\bigr) = \frac{80}{81}.$$

More generally, for $X_{n,m}^{(d)}$ (runs of length exactly d), we find from (7.14), (7.17) and (7.27), after straightforward calculations, $g_0(t) = t^d(1-t)^2$, $t_0 = d/(d+2)$, $g_0''(t_0) = -2d^{d-1}/(d+2)^{d-1}$, and

$$\begin{aligned}
\sigma^2 &= \frac{4d^d}{(d+2)^{d+2}}\Bigl(1-(d+1)\frac{4d^d}{(d+2)^{d+2}}\Bigr),\\
\sigma_*^2 &= 8\frac{d^d}{(d+2)^{d+1}}\Bigl(1+\frac{d^d}{(d+2)^{d+1}}\Bigr),\\
\beta &= \Bigl(32\frac{d^{d+1}}{(d+2)^{d+3}}\Bigl(1+\frac{d^d}{(d+2)^{d+1}}\Bigr)^2\Bigr)^{1/3}.
\end{aligned}$$

8. Priority queues, sock-sorting and lazy hashing

As said above, priority queues are equivalent to sock-sorting. For priority queues, the $2n$ events A_i and D_i come in random order, with the restriction that A_i comes before D_i for each i. Since only the order of the events matters, we may randomize the times as in Section 2 and assume that the times A_i and D_i, $i = 1, \dots, n$, are independent random variables uniformly distributed on $(0, 1)$, conditioned on $A_i < D_i$ for all i. For two independent random variables $T, \tilde{T} \sim U(0, 1)$, the distribution of $(T, \tilde{T})$ conditioned on $T < \tilde{T}$ equals the distribution of $(T \wedge \tilde{T}, T \vee \tilde{T})$, and this randomization of the times in a priority queue thus gives exactly the model for lazy hashing defined in Section 1, as found by Kenyon and Vitter [17]. In particular, $\max_t Y_n(t) \overset{\mathrm{d}}{=} Y_n^*$.

In analogy with the definitions in Section 2, we now let

$$\begin{aligned}
I(t;k) &:= \mathbf{1}[T_k \le t], & \tilde{I}(t;k) &:= \mathbf{1}[\tilde{T}_k \le t],\\
I'(t;k) &:= I(t;k) - t, & \tilde{I}'(t;k) &:= \tilde{I}(t;k) - t,\\
S_{n,1}(t) &:= \sum_{k=1}^n I'(t;k) + \sum_{k=1}^n \tilde{I}'(t;k),\\
S_{n,2}(t) &:= \sum_{k=1}^n I'(t;k)\tilde{I}'(t;k).
\end{aligned}$$

We further let $N_n(t)$ be the number of events (A_k or D_k) up to t. Then, cf. (2.10),

$$N_n(t) = \sum_{k=1}^{n} I(t;k) + \sum_{k=1}^{n} \tilde{I}(t;k) = S_{n,1}(t) + 2nt, \tag{8.1}$$

$$\begin{aligned} Y_n(t) &= \sum_{k=1}^{n} \mathbf{1}[A_k \le t < D_k] = \sum_{k=1}^{n} \bigl(\mathbf{1}[T_k \le t < \tilde{T}_k] + \mathbf{1}[\tilde{T}_k \le t < T_k]\bigr) \\ &= \sum_{k=1}^{n} \Bigl(I(t;k)\bigl(1-\tilde{I}(t;k)\bigr) + \tilde{I}(t;k)\bigl(1-I(t;k)\bigr)\Bigr) \\ &= 2nt(1-t) + (1-2t)S_{n,1}(t) - 2S_{n,2}(t). \end{aligned} \tag{8.2}$$

We introduce martingales $\widehat{S}_{n,1}$ and $\widehat{S}_{n,2}$ as above by (6.5) and (6.6).

All proofs in Section 6 now go through with no or minor changes; the main differences are that (8.1) and (8.2) contain some factors 2 not appearing in (2.9) and (2.10) and that there will be a factor 2 on the right hand side of (6.11); thus Theorem 4.1 holds with the difference that (4.3) is replaced by

$$\mathbb{E}\bigl(Z_1(s)Z_1(t)\bigr) = 2s(1-t), \qquad 0 \le s \le t \le 1;$$

similarly, (4.4) holds with $B_1(x)$ replaced by $2^{1/2}B_1(x)$ in (4.8). This yields the following results, corresponding to our results for $X_{n,m}$ and $X_n(t)$ in Sections 1 and 4.

Theorem 8.1. *As $n \to \infty$, in $D[0,1]$,*

$$n^{-1/2}\bigl(Y_n(t) - 2nt(1-t)\bigr) \xrightarrow{\text{d}} Z(t) := (1-2t)Z_1(t) - 2Z_2(t),$$

where Z is a continuous Gaussian process on $[0,1]$ with mean $\mathbb{E}\,Z(t) = 0$ and covariances, for $0 \le s \le t \le 1$,

$$\begin{aligned} \mathbb{E}\bigl(Z(s)Z(t)\bigr) &= 2s(1-2s)(1-t)(1-2t) + 4s^2(1-t)^2 \\ &= 2s(1-t) - 4s(1-s)t(1-t). \end{aligned}$$

Theorem 8.2. *As $n \to \infty$, in $D[0,1]$,*

$$n^{-1/2}\bigl(Y_{n,\lfloor 2nt \rfloor} - 2nt(1-t)\bigr) \xrightarrow{\text{d}} Z(t) := -2Z_2(t),$$

where Z is a continuous Gaussian process on $[0,1]$ with mean $\mathbb{E}\,Z(t) = 0$ and covariances

$$\mathbb{E}\bigl(Z(s)Z(t)\bigr) = 4s^2(1-t)^2, \qquad\qquad 0 \le s \le t \le 1.$$

Theorem 8.3. *As $n \to \infty$, in $D(-\infty,\infty)$,*

$$n^{-1/3}\bigl(Y_n(\tfrac{1}{2} + xn^{-1/3}) - Y_n(\tfrac{1}{2})\bigr) \xrightarrow{\text{d}} 2^{1/2}B(x) - 2x^2,$$

where B is a Brownian motion on $(-\infty,\infty)$.

Theorem 8.4. *As $n \to \infty$,*

$$n^{-1/2}\big(Y_n^* - n/2\big) \xrightarrow{\mathrm{d}} N(0, 1/4),$$

with convergence of all moments. In particular,

$$\mathbb{E}\, Y_n^* = n/2 + o(n^{1/2}),$$
$$\operatorname{Var} Y_n^* = n/4 + o(n).$$

Theorem 8.5. *As $n \to \infty$,*

$$n^{-1/3}\big(Y_n^* - Y_{n,n}\big) \xrightarrow{\mathrm{d}} V,$$

where the random variable V is as in Theorem 1.3, and

$$\mathbb{E}\, Y_n^* = \mathbb{E}\, Y_{n,n} + \mathbb{E}\, V n^{1/3} + o(n^{1/3}) = \tfrac{1}{2}n + \mathbb{E}\, V n^{1/3} + o(n^{1/3}).$$

Theorem 8.1 is given by Louchard [20], Theorem 8.2 by Louchard [19] (with a deterministic change of time, making the problem equivalent to a queueing problem), and Theorems 8.4 and 8.5 by Louchard, Kenyon and Schott [21] (with different proofs).

Note that in the proof of Theorem 8.2 the terms with $S_{n,1}$ cancel, as discussed for Theorem 1.5 above. In both theorems the limit is thus given by Z_2, which explains why we obtain exactly the same covariances in the two theorems, except for a normalization factor. (Unlike Corollary 4.2 and Theorem 8.1, where the variances of the limits are $t(1-t)(1-3t+3t^2)$ and $2t(1-t)(1-2t+2t^2)$.)

Acknowledgement. I thank Göran Högnäs for telling me about this problem and providing me with the reference [1]. I also thank Persi Diaconis, Anders Martin-Löf and Guy Louchard for interesting discussions.

References

[1] G. af Hällström, Ein lineares Inselproblem der kombinatorischen Wahrscheinlichkeitsrechnung. *Ann. Acad. Sci. Fennicae. Ser. A. I. Math.-Phys.* **123** (1952), 9 pp.

[2] A.D. Barbour, A note on the maximum size of a closed epidemic. *J. Roy. Statist. Soc. Ser. B* **37** (1975), no. 3, 459–460.

[3] A.D. Barbour, Brownian motion and a sharply curved boundary. *Adv. Appl. Probab.* **13** (1981), no. 4, 736–750.

[4] P. Billingsley, *Convergence of Probability Measures.* Wiley, New York, 1968.

[5] H. E. Daniels, The maximum size of a closed epidemic. *Adv. Appl. Probab.* **6** (1974), 607–621.

[6] H. E. Daniels, The maximum of a Gaussian process whose mean path has a maximum, with an application to the strength of bundles of fibres. *Adv. Appl. Probab.* **21** (1989), no. 2, 315–333.

[7] H. E. Daniels and T. H. R. Skyrme, The maximum of a random walk whose mean path has a maximum. *Adv. Appl. Probab.* **17** (1985), no. 1, 85–99.

[8] P. Flajolet, J. Françon & J. Vuillemin, Sequence of operations analysis for dynamic data structures. *J. Algorithms* **1** (1980), no. 2, 111–141.

[9] P. Flajolet, C. Puech & J. Vuillemin, The analysis of simple list structures. *Inform. Sci.* **38** (1986), no. 2, 121–146.

[10] P. Groeneboom, Brownian motion with a parabolic drift and Airy functions. *Probab. Theory Related Fields* **81** (1989), no. 1, 79–109.

[11] A. Gut, *Probability: A Graduate Course.* Springer, New York, 2005.

[12] J. Jacod & A. N. Shiryaev, *Limit Theorems for Stochastic Processes.* Springer-Verlag, Berlin, 1987.

[13] S. Janson, A functional limit theorem for random graphs with applications to subgraph count statistics, *Random Struct. Alg.* **1** (1990), 15–37.

[14] S. Janson, *Orthogonal Decompositions and Functional Limit Theorems for Random Graph Statistics.* Mem. Amer. Math. Soc., vol. 111, no. 534, American Mathematical Society, Providence, R.I., 1994.

[15] S. Janson, Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Processes Appl.* **110** (2004), no. 2, 177–245.

[16] O. Kallenberg, *Foundations of Modern Probability.* 2nd ed., Springer-Verlag, New York, 2002.

[17] C. M. Kenyon & J. S. Vitter, Maximum queue size and hashing with lazy deletion. *Algorithmica* **6** (1991), no. 4, 597–619.

[18] W. V. Li & G. Pritchard, A central limit theorem for the sock-sorting problem. *High dimensional probability (Oberwolfach, 1996)*, Progr. Probab., 43, Birkhäuser, Basel, 1998, 245–248.

[19] G. Louchard, Random walks, Gaussian processes and list structures. *Theoret. Comput. Sci.* **53** (1987), no. 1, 99–124.

[20] G. Louchard, Large finite population queueing systems. I. The infinite server model. *Comm. Statist. Stochastic Models* **4** (1988), no. 3, 473–505.

[21] G. Louchard, C. Kenyon & R. Schott, Data structures' maxima. *SIAM J. Comput.* **26** (1997), no. 4, 1006–1042.

[22] A. M. Mood, The distribution theory of runs. *Ann. Math. Statistics* **11** (1940), 367–392.

[23] D. Steinsaltz, Random time changes for sock-sorting and other stochastic process limit theorems. *Electron. J. Probab.* **4** (1999), no. 14, 25 pp.

[24] W. L. Stevens, Distribution of groups in a sequence of alternatives. *Ann. Eugenics* **IX** (1939), 10–17.

[25] C. J. Van Wyk & J. S. Vitter, The complexity of hashing with lazy deletion. *Algorithmica* **1** (1986), no. 1, 17–29.

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden

E-mail address: svante.janson@math.uu.se

URL: http://www.math.uu.se/~svante/