The Annals of Probability
2012, Vol. 40, No. 4, 1535–1576
DOI: 10.1214/11-AOP656

TIGHT MARKOV CHAINS AND RANDOM COMPOSITIONS[1]

BY BORIS PITTEL

Ohio State University

Dedicated to the memory of Philippe Flajolet

For an ergodic Markov chain $\{X(t)\}$ on $\mathbb{N}$, with a stationary distribution π, let $T_n > 0$ denote a hitting time for $[n]^c$, and let $X_n = X(T_n)$. Around 2005 Guy Louchard popularized a conjecture that, for $n \to \infty$, T_n is almost Geometric(p), $p = \pi([n]^c)$, X_n is almost stationarily distributed on $[n]^c$ and that X_n and T_n are almost independent, if $p(n) := \sup_i p(i, [n]^c) \to 0$ exponentially fast. For the chains with $p(n) \to 0$, however slowly, and with $\sup_{i,j} \|p(i, \cdot) - p(j, \cdot)\|_{\mathrm{TV}} < 1$, we show that Louchard's conjecture is indeed true, even for the hits of an arbitrary $S_n \subset \mathbb{N}$ with $\pi(S_n) \to 0$. More precisely, a sequence of k consecutive hit locations paired with the time elapsed since a previous hit (for the first hit, since the starting moment) is approximated, within a total variation distance of order $k \sup_i p(i, S_n)$, by a k-long sequence of independent copies of (ℓ_n, t_n), where $t_n = \text{Geometric}(\pi(S_n))$, ℓ_n is distributed stationarily on S_n and ℓ_n is independent of t_n. The two conditions are easily met by the Markov chains that arose in Louchard's studies as likely sharp approximations of two random compositions of a large integer ν, a column-convex animal (cca) composition and a Carlitz (C) composition. We show that this approximation is indeed very sharp for each of the random compositions, read from left to right, for as long as the sum of the remaining parts stays above $\ln^2 \nu$. Combining the two approximations, a composition—by its chain, and, for $S_n = [n]^c$, the sequence of hit locations paired each with a time elapsed from the previous hit—by the independent copies of (ℓ_n, t_n), enables us to determine the limiting distributions of $\mu = o(\ln \nu)$ and $\mu = o(\nu^{1/2})$ largest parts of the random cca-composition and the random C-composition, respectively. (Submitted to Annals of Probability in June 2009.)

1. Introduction. Consider a Markov chain $X(t)$ on $\mathbb{N}$. Given $S \subset \mathbb{N}$, let $T(S)$ be the hitting time, that is, $T(S) = \min\{t > 0 : X(t) \in S\}$. Keilson [14] proved that if a state i is positive-recurrent, and a nested sequence $S_1 \supseteq S_2 \supseteq \cdots$ is such that $i \notin S_1$ and $E_i[T(S_n)] \to \infty$, then

$$P_i\left\{\frac{T(S_n)}{E_i[T(S_n)]} \ge t\right\} \to e^{-t} \qquad \forall t \ge 0. \tag{1.1}$$

Received February 2011.

[1]Supported in part by NSF Grant DMS-08-05996.

MSC2010 subject classifications. 05A15, 05A17, 11P99, 60C05, 60F05, 60J05.

Key words and phrases. Random compositions, Markov chains, rarity, exponentiality, extreme values, limit theorems.

The basic idea of the proof was that the probability of hitting S_n between two consecutive returns to i is small, of order $1/E_i[T(S_n)]$, and so $T(S_n)$ is roughly the sum of the geometrically distributed number of i.i.d. times between those returns to i.

If a chain is ergodic, with a stationary distribution π, the condition $E_i[T(S_n)] \to \infty$ is met if (and only if) $\pi(S_n) := \sum_{i \in S_n} \pi(i) \to 0$. Indeed, by Derman's theorem [9] (see Durrett [10], Chapter 5), the expected number of visits to S_n between two returns to i is $\pi(S_n)/\pi(i)$. So the probability of hitting S_n between two returns to i is $\pi(S_n)/\pi(i)$ at most, whence $E_i[T(S_n)] \geq \pi(i)/\pi(S_n)$.

Aldous [1] estimated accuracy of the exponential approximation of the hitting time for a finite-state ergodic Markov chain, when an initial state is chosen at random, in accordance with the stationary distribution π. Roughly, the discrepancy is small if the expected hitting time far exceeds a relaxation time $\tau = \max_i \min\{t : \|p^t(i, \cdot) - \pi(\cdot)\|_{\mathrm{TV}} \leq \rho\}$, $\rho < 1/2$. τ "measures the time taken for the chain to approach stationarity" in a sense that $\max_i \|p^t(i, \cdot) - \pi(\cdot)\|_{\mathrm{TV}} \leq (2\rho)^{\lfloor t/\tau \rfloor}$.

Precisely because these results are so strikingly general, more subtle questions remain open. Is there a geometrically distributed random variable close to $T(S)$ in terms of the total variation distance? What is, asymptotically, the joint distribution of the hitting time $T(S)$ and the hit location $X(T(S))$? Is there an explicit convergence rate in terms of the total variation distance? Are $X(T(S))$ and $T(S)$ almost independent? How does one describe asymptotic behavior of the first k visits to the rare set S, if $k = k(S)$ is not too large?

For an ergodic Markov chain $\{X(t)\}$ on $\mathbb{N}$, with a stationary distribution π, let $T_n > 0$ denote a hitting time for $[n]^c = \mathbb{N} \setminus [n]$, and let $X_n = X(T_n)$. Around 2005 Guy Louchard [18] popularized the following conjecture. If $p(n) := \sup_i p(i, [n]^c) = O(q^n)$, $q < 1$, then T_n is almost Geometric(p), ($p = \pi([n]^c)$), X_n is almost stationarily distributed on $[n]^c$, and X_n and T_n are almost independent. The Markov chains with $p(n) = O(q^n)$ arose in the studies of two random compositions, Louchard [19, 20] and Louchard and Prodinger [21], as possibly sharp approximations of those random compositions. Louchard's thought-provoking idea was that if the conjecture and approximability of each random compositions by a chain would be proved, potentially one could obtain the limiting distributions, marginal and joint, of extreme-valued parts and, possibly, of other related characteristics of the random compositions.

In this paper we introduce a class of Markov chains that contains the chains from [19–21] for which we can give full answers to the questions posed above and, in particular, fully confirm Guy Louchard's conjecture. We also prove that the chains in [19–21] indeed provide a good approximation of the random compositions. The two approximations made in tandem lead to the asymptotic distributions of the extreme-valued parts of the compositions, together with the convergence rates.

Let us give a more specific description of our results.

DEFINITION 1.1. An ergodic Markov chain on $\mathbb{N}$, with a transition probability matrix $P = \{p(i,k)\}_{i,k\in\mathbb{N}}$ and a stationary distribution π, is called tight if the family of row probability measures $\{p(i,\cdot)\}_{i\in\mathbb{N}}$ is tight, that is,

$$\lim_{n\to\infty} \sup_i \sum_{k>n} p(i,k) = 0. \tag{1.2}$$

For a tight P, we will prove that if $\varnothing \neq S_n \subset \mathbb{N}$ is such that $\pi(S_n) \to 0$, then uniformly for all initial states i,

$$E_i[T^k(S_n)] \sim \frac{k!}{\pi^k(S_n)}, \qquad k \geq 1, \tag{1.3}$$

so $E_i[T(S_n)] \sim \pi^{-1}(S_n)$ in particular. Thus all the moments of $T(S_n)/E_i[T(S_n)]$ converge, uniformly over i, to the moments of the exponential random variable, which implies convergence in distribution as well. As for the hit location $X(T(S_n))$, given $U_n \subseteq S_n$,

$$\lim_{n\to\infty} \left| P_i\{X(T(S_n)) \in U_n\} - \frac{\pi(U_n)}{\pi(S_n)} \right| = 0, \tag{1.4}$$

uniformly for $i \in \mathbb{N}$. Thus, *marginally*, $T(S_n)$ and $X(T(S_n))$ behave in the limit as if $X(t)$ is a Bernoulli sequence with each trial outcome having distribution π.

Now suppose that, besides being tight, the chain meets a condition

$$\delta_0 := \inf_{i,j\in\mathbb{N}} \sum_{k\in\mathbb{N}} p(i,k)p(j,k) > 0. \tag{1.5}$$

For a tight chain, this condition is equivalent to

$$\rho_0 := \sup_{i,j\in\mathbb{N}} \|p(i,\cdot) - p(j,\cdot)\|_{\mathrm{TV}} < 1,$$

which implies that

$$\|p^n(i,\cdot) - \pi\|_{\mathrm{TV}} \leq \rho_0^n.$$

(So, for the relaxation time τ in [1], we have $\tau = \lceil \ln 2/\ln(1/\rho_0) \rceil$.)

Given a random vector $\mathbf{Y}$ with integer components, we denote its probability distribution by $d(\mathbf{Y})$. Under conditions (1.2) and (1.5), we show that, uniformly for the initial state $i \in \mathbb{N}$,

$$\begin{aligned} \|d((X(T(S_n)), T(S_n))) - d((\ell_n, t_n))\|_{\mathrm{TV}} &= O(p(S_n)), \\ p(S) &:= \sup_{k\in\mathbb{N}} p(k, S), \end{aligned} \tag{1.6}$$

where ℓ_n and t_n are *independent*,

$$\begin{aligned} P\{t_n = \tau\} &= \pi(S_n)\bigl(1 - \pi(S_n)\bigr)^{\tau-1}, \qquad \tau \geq 1, \\ P\{\ell_n = k\} &= \frac{\pi(k)}{\pi(S_n)}, \qquad k \in S_n. \end{aligned}$$

More generally, the k-long sequence of chronologically ordered locations of first k hits of S_n, each paired with the time elapsed since a preceding hit [paired with $T(S_n)$ in the case of the first hit] is approximated by the k-long sequence of *independent* copies of (ℓ_n, t_n), within the total variation distance of order $O(kp(S_n))$. (Aldous and Brown [2, 3] had used Stein's method to show that, for a stationary, continuous-time, reversible Markov process, the hitting times for a subset A of states after prolonged excursions outside of A form an approximately Poisson process.)

Equation (1.6) yields, rather directly, the limiting distributions of the extreme values for $\{X(t)\}_{1\le t\le N}$. Given μ, let $X^{(\mu)}$ denote the μth largest among $X(1), \dots, X(N)$. Then

$$P_i\{X^{(\mu)} \le n\} = P\{\mathrm{Poisson}(N\pi(S_n)) < \mu\} + O(\mu^2/N + Np^2(S_n)), \tag{1.7}$$

and we have an extended version of (1.7) for the joint distribution of $X^{(1)}, \dots, X^{(\mu)}$.

Turn now to the application of these results to the random compositions studied in [19–21].

A composition of a positive integer ν is $\mathbf{y} = (y_1, \dots, y_\mu)$, $\mu \le \nu$, such that $y_1, \dots, y_\mu$ are positive integers satisfying

$$\sum_{i=1}^{\mu} y_i = \nu. \tag{1.8}$$

Since, for each μ, there are $\binom{\nu-1}{\mu-1}$ compositions, we have $2^{\nu-1}$ compositions overall. Assuming that a solution of (1.8) is chosen uniformly at random (*uar*), we have a *random* composition $\mathbf{Y}$ of ν, its dimension M being random as well. It is known (Andrews [4]) that

$$\mathbf{Y} \overset{\mathcal{D}}{\to\equiv} (Z_1, \dots, Z_{\mathcal{M}-1}, \hat{Z}_{\mathcal{M}}), \tag{1.9}$$

($\overset{\mathcal{D}}{\to\equiv}$ meaning equality of distributions), where $Z_1, Z_2, \dots$ are independent geometrics with success probability $1/2$,

$$\mathcal{M} = \min\{m : Z_1 + \cdots + Z_m \ge \nu\} \tag{1.10}$$

and

$$\hat{Z}_{\mathcal{M}} := \nu - \sum_{j=1}^{\mathcal{M}-1} Z_j. \tag{1.11}$$

Hitczenko and Savage [12] used this connection to the well-studied success runs in a fair coin-tossing process as an efficient tool for asymptotic analysis of various characteristics of the random composition.

If a random composition $\mathbf{Y}$ is not uniformly distributed on the set (1.8), one can only hope for *asymptotic* independence of most of the parts. Lowering expectations then, one may search for a Markov chain that approximates the behavior of $\mathbf{Y}$ in question; ergodicity of such a chain would mean near independence of parts Y_{t_1} and Y_{t_2} with $|t_1 - t_2|$ sufficiently large.

Here are two examples of such random compositions. A column-convex-animal (cca) composition of ν is a collection of lengths of an ordered sequence of contiguous columns on $\mathbb{Z}^2$, whose total sum is ν, such that every two successive columns have a common boundary consisting of at least one vertical edge of $\mathbb{Z}^2$; Klarner [15], Privman and Forgacs [22], Privman and Svrakic [23], Louchard [19, 20]. A Carlitz (C) composition meets a condition that no two adjacent parts coincide; Carlitz [8], Knopfmacher and Prodinger [16], Louchard and Prodinger [21], Hitczenko and Louchard [11].

One obtains a certain, *nonuniform*, distribution on the set of solutions of (1.8), if a column-convex animal is chosen uar from among all such creatures. One obtains another nonuniform distribution, if a composition of ν is chosen uar from among all C-compositions. We call these objects a random cca-composition $\mathbf{Y}$, and a random C-composition $\mathbf{Y}$, and denote the random number of components of $\mathbf{Y}$ by M. For both schemes, Louchard [19, 20] and Louchard and Prodinger [21], determined a limiting joint distribution of two successive parts, Y_t and Y_{t+1}, in the case when t and $M - t$ are of order ν, and also the limiting distribution, π_1, of the first (last) part Y_1 (Y_M). These results strongly suggest, though do not actually prove that, in both cases,

$$p(i,k) := \lim_{\nu,t\to\infty} \mathrm{P}\{Y_{t+1} = k | Y_t = i\} \qquad (i,k \in \mathbb{N}), \tag{1.12}$$

might well be the transition probabilities of a Markov chain, with an initial distribution π_1, that closely approximates the *whole* random composition. We fully confirm this conjecture, proving an approximational counterpart of (1.9)–(1.11). The chains turn out to be tight, exponentially mixing and this enables us to use our results for asymptotic analysis of extreme-valued parts of both random compositions. Let $Y^{(\mu)}$ denote the μth largest part of the random composition in question. For the random cca-composition, we show that, for $\mu = o(\ln \nu)$,

$$Y^{(\mu)} = \frac{\ln(\mu^{-1}\nu \ln^2 \nu)}{\ln(1/z^*)} + O_p(1), \tag{1.13}$$

where $z_* = 0.31\ldots$ is the smallest-modulus root of

$$4z^3 - 7z^2 + 5z - 1 = 0.$$

For the random C-composition, if $\mu = o(\nu^{1/2})$, then

$$Y^{(\mu)} = \frac{\ln(\mu^{-1}\nu)}{\ln(1/z^*)} + O_p(1), \tag{1.14}$$

where $z_* = 0.57\dots$ is the smallest-modulus root of

$$\sum_{j\ge 1} \frac{z^j}{1+z^j} - 1 = 0.$$

[$O_p(1)$ stands for a random variable bounded in probability.] It follows from (1.13) and (1.14) that the number of distinct values among $X^{(1)}, \dots, X^{(\mu)}$ is likely to be at most $(1+o(1))\ln\mu/\ln(1/z_*)$, for the corresponding z_*, for $\mu = o(\ln\nu)$ and $\mu = o(\nu^{1/2})$, respectively. It can be shown that, in fact, the range is asymptotic to $\ln\mu/\ln(1/z_*)$, in probability. (See Hitczenko and Louchard [11] regarding a limiting distribution of a "distinctness" (range size) of the random C-composition.)

We plan to extend this approach to other constrained compositions, such as quite general Carlitz-type compositions studied by Bender and Canfield [6].

The rest of the paper is organized as follows. In Section 2 we show that, for the tight Markov chains $\{X(t)\}$, the hitting time of a rare set S_n, that is, with $\pi(S_n) \to 0$, scaled by $\pi^{-1}(S_n)$ converges, with all its moments, to the exponentially distributed random variable of unit mean, while the hit location has, in the limit, a stationary distribution restricted to S_n. And convergence is uniform over all initial states. In Section 3 we add a second condition that guarantees exponential mixing, calling such chains tight, exponentially mixing (t.e.m.) chains. Significantly sharpening the results of Section 2, we demonstrate that the hitting time and the hit location are asymptotic, with respect to the total variation distance, to a pair of *independent* random variables, one being geometrically distributed with success probability $\pi(S_n)$, and another having the restricted stationary distribution. The error term is $O(p(S_n))$; see (1.6) for definition of $p(\cdot)$. We extend this result to the first k hits of S_n, and then state and prove the claims about the limiting distribution of the μ largest values among $X(1), \dots, X(N)$, useful for $\mu = o(N^{1/2})$. In Section 4 we apply these claims to the extreme-valued parts of two random compositions of a large ν, the cca-composition and the C-composition. Specifically, in Section 4.1 we briefly survey the basic known facts about the compositions. In Section 4.2 we show that each composition is sharply approximated, in terms of total variation distance, by a related Markov chain, for as long as the current sum of parts does not exceed $\nu - \ln^2\nu$. In Section 4.3, for each composition, we derive the limiting distributions of the μ largest values of a random composition parts, assuming that $\mu = o(\ln\nu)$ for the cca-composition and $\mu = o(\nu^{1/2})$ for the C-composition. In the Appendix we prove an auxiliary result on large deviations of the number of parts in each of the random compositions.

2. Tight Markov chains. Consider an ergodic Markov chain $X(t)$ on $\mathbb{N}$ with the stationary distribution $\boldsymbol{\pi} = \{\pi(j)\}_{j\in\mathbb{N}}$. Given $S \subset \mathbb{N}$, we denote $\pi(S) = \sum_{j\in S}\pi(j)$. Introduce $T(S)$ the *positive* hitting time of S, that is, $T(S) = \min\{t > 0 | X(t) \in S\}$, and the hit location $X(T(S))$. Our focus is on a *rare* S, that is, with a small $\pi(S)$.

Assuming that the chain satisfies a tightness condition (1), namely

$$\lim_{n\to\infty} \sup_i \sum_{k>n} p(i,k) = 0, \tag{2.1}$$

we will show that, uniformly for an initial state in $\mathbb{N}$, (1) $T(S)$ is asymptotically exponential, with mean $\pi^{-1}(S)$, and (2) the distribution of $X(T(S))$ is asymptotic to $\{\pi(s)/\pi(S)\}_{s\in S}$.

As a first step we prove the following.

LEMMA 2.1. *Let a possibly infinite $S_n \neq \varnothing$ be such that* $\lim_{n\to\infty}\pi(S_n)=0$. *Under condition* (2.1),

$$E_i[T(S_n)] \sim \frac{1}{\pi(S_n)}, \qquad n\to\infty, \tag{2.2}$$

uniformly for $i \in \mathbb{N}$.

NOTE. Consider a simple asymmetric random walk on $\mathbb{N}$, that is, the Markov chain with $p(1,1)=q,\ p(1,2)=p$ and $p(i,i-1)=q,\ p(i,i+1)=p$ for $i\ge 2$. For $p<q$ this chain is ergodic, with the stationary distribution $\pi(j) = (1-p/q)(p/q)^{j-1}$, but it is clearly not tight. For $i=1$, $T(\{n+1\}) = T(\{n+1, n+2,\dots\})$, but $\pi(\{n+1\}) \not\sim \pi(\{n+1,n+2,\dots\})$. So (2.2) cannot hold for all S_n with $\pi(S_n)\to 0$. In fact, the expected common hitting time for these two sets is not asymptotic to the reciprocal of either of these stationary probabilities.

PROOF OF LEMMA 2.1. By tightness condition (2.1), there exists K such that

$$\sum_{j\le K} p(i,j) \ge 1/2 \qquad \forall i\ge 1.$$

Then, for $t\ge 1$,

$$P_i\{T([K])>t\} \le \frac{1}{2^t} \Longrightarrow E_i[T([K])]\le 2.$$

Now, one (possibly not the shortest) way of hitting S_n, starting at i, is to hit the set $[K]$ and from there to hit S_n. By the strong Markov property, conditionally on $X(T([K])) = j\,(j\in[K])$, the residual travel time $\hat{T}(S_n)$ till hitting S_n is distributed as $T(S_n)$ under P_j. So

$$E[\hat{T}(S_n)|X(T([K]))=j] = E_j[T(S_n)], \qquad j\in[K].$$

Then, introducing $\ell\in[K]$ such that

$$E_\ell[T(S_n)] = \max_{j\in[K]} E_j[T(S_n)],$$

we have

$$
\begin{aligned}
E_i[T(S_n)] &\le E_i[T[K]] + \sum_{j\in[K]} P_i\{X(T([K])) = j\} E_j[T(S_n)] \\
&\le 2 + E_\ell[T(S_n)];
\end{aligned}
\tag{2.3}
$$

in particular, $\sup_i E_i[T(S_n)] < \infty$.

By the Markov property,

$$
E_j[T(S_n)] = 1 + \sum_{k\in S_n^c} p(j,k) E_k[T(S_n)], \qquad j \in \mathbb{N}. \tag{2.4}
$$

Multiplying both sides of (2.4) by $\pi(j)$ and summing for $j \in \mathbb{N}$, we get

$$
\begin{aligned}
\sum_{j\in\mathbb{N}} \pi(j) E_j[T(S_n)] &= 1 + \sum_{k\in S_n^c} E_k[T(S_n)] \sum_{j\in\mathbb{N}} \pi(j) p(j,k) \\
&= 1 + \sum_{k\in S_n^c} \pi(k) E_k[T(S_n)],
\end{aligned}
$$

as $\pi(\cdot)$ is stationary. So, as both series converge,

$$
\sum_{k\in S_n} \pi(k) E_k[T(S_n)] = 1. \tag{2.5}
$$

(We note that (2.5) is a special case of a well-known result, due to Kac [13], with inevitably harder proof, for a general discrete-time stationary process; see also Breiman [7], Section 6.9.) Then, by (2.3),

$$
E_\ell[T(S_n)] + 1 \ge \frac{1}{\pi(S_n)} \Longrightarrow E_\ell[T(S_n)] \gtrsim \frac{1}{\pi(S_n)}. \tag{2.6}
$$

Now, given a state k, we have

$$
E_\ell[T(S_n)] \le E_\ell[T(\{k\})] + E_k[T(S_n)], \tag{2.7}
$$

$T(\{k\})$ being the hitting time for the singleton $\{k\}$. Combining (2.6) and (2.7), we obtain that for every *fixed* k,

$$
E_k[T(S_n)] \gtrsim \frac{1}{\pi(S_n)}. \tag{2.8}
$$

Picking arbitrary L, by (2.4), we have that for $n \ge n(L)$,

$$
E_j[T(S_n)] \ge 1 + \sum_{k\le L} p(j,k) E_k[T(S_n)], \qquad j \in \mathbb{N}.
$$

Therefore, by (2.8),

$$
\liminf_{n\to\infty} \Bigl(\inf_{j\in\mathbb{N}} E_j[T(S_n)]\Bigr) \pi(S_n) \ge \liminf_{n\to\infty} \inf_{j\in\mathbb{N}} \sum_{k\le L} p(j,k),
$$

where, by (2.1), the RHS approaches 1 as $L \uparrow \infty$. So

$$E_j[T(S_n)] \gtrsim \frac{1}{\pi(S_n)}, \tag{2.9}$$

uniformly for $j \in \mathbb{N}$.

It remains to show that

$$E_j[T(S_n)] \lesssim \frac{1}{\pi(S_n)},$$

uniformly for $j \in \mathbb{N}$. Using (2.4)–(2.5), we obtain then

$$\sum_{j \in S_n} \frac{\pi(j)}{\sum_{i \in S_n} \pi(i)} \biggl(1 + \sum_{k \in S_n^c} p(j,k) E_k[T(S_n)]\biggr) = \frac{1}{\pi(S_n)}. \tag{2.10}$$

Suppose that there exists a subsequence $n_m \to \infty$ and $\delta > 0$, such that

$$\lim_{n \in \{n_m\}} E_\ell[T(S_n)]\pi(S_n) \geq 1 + \delta.$$

Then, by (2.7),

$$\lim_{n \in \{n_m\}} E_k[T(S_n)]\pi(S_n) \geq 1 + \delta,$$

for every fixed k. Picking $M > 0$ and dropping the summands for $k > M$ in (2.10), we get then: for $n = n_m$ large enough,

$$(1 + \delta/2) \sum_{j \in S_n} \frac{\pi(j)}{\sum_{i \in S_n} \pi(i)} \biggl(\sum_{k \leq M} p(j,k)\biggr) \leq 1.$$

This is impossible if M is chosen so large that

$$\inf_j \sum_{k \leq M} p(j,k) \geq \frac{1}{1 + \delta/3}.$$

Therefore

$$E_\ell[T(S_n)] \lesssim \frac{1}{\pi(S_n)},$$

and so, invoking (2.3),

$$E_k[T(S_n)] \lesssim \frac{1}{\pi(S_n)}, \tag{2.11}$$

uniformly for $k \in \mathbb{N}$.

Combining (2.9) and (2.11), we complete the proof of Lemma 2.1. □

The fact that $E_i[T(S_n)] \to \infty$ already implies, via Keilson's theorem [14], that, for each *fixed* initial state i, $T(S_n)/E_i[T(S_n)]$ is, in the limit, exponentially distributed, with parameter 1. The tightness condition allowed us to estimate the scaling parameters $E_i[T(S_n)]$ asymptotically, uniformly for $i \in \mathbb{N}$. Interestingly, this uniformity can be used for a simple alternative proof of asymptotic exponentiality of $T(S_n)/E_i[T(S_n)]$.

LEMMA 2.2. *Under condition* (2.1), *for each fixed* $k \geq 1$,

$$E_i[T^k(S_n)] \sim k!/\pi^k(S_n), \tag{2.12}$$

uniformly for $i \in \mathbb{N}$. *Consequently, uniformly for* $i \in \mathbb{N}$,

$$P_i\{T(S_n)\pi(S_n) > x\} \to e^{-x} \qquad \forall x \geq 0. \tag{2.13}$$

PROOF. Introduce the moment generating functions

$$\phi_i(u) = \sum_{r \geq 0} \frac{u^r}{r!} E_i[T^r(S_n)], \qquad i \in \mathbb{N}.$$

As formal power series, these functions satisfy

$$\phi_i(u) = e^u \left(\sum_{j \in S_n} p(i, j) + \sum_{j \in S_n^c} p(i, j)\phi_j(u) \right), \qquad i \in \mathbb{N}. \tag{2.14}$$

Differentiating both sides of (2.14) k times at $u = 0$, we get

$$\begin{aligned} E_i[T^k(S_n)] &= b(i,k) + \sum_{j \in S_n^c} p(i, j)E_j[T^k(S_n)], \qquad i \in \mathbb{N}, \\ b(i,k) &:= 1 + \sum_{r=1}^{k-1} \binom{k}{r} \sum_{j \in S_n^c} p(i, j)E_j[T^r(S_n)]. \end{aligned} \tag{2.15}$$

For $k = 1$ we get (2.4). Let $k \geq 2$, and suppose that, for $r < k$,

$$E_i[T^r(S_n)] = \bigl(1 + o(1)\bigr) \frac{r!}{\pi^{r-1}(S_n)} E_i[T(S_n)], \tag{2.16}$$

uniformly for $i \in \mathbb{N}$. (This is obviously true for $k = 2$.) Then

$$\begin{aligned} b(i,k) &= 1 + \bigl(1 + o(1)\bigr) \sum_{r=1}^{k-1} \frac{(k)_r}{\pi^{r-1}(S_n)} \sum_{j \in S_n^c} p(i, j)E_j[T(S_n)] \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad [\text{using } (2.4)] \\ &= 1 + \bigl(1 + o(1)\bigr) \sum_{r=1}^{k-1} \frac{(k)_r}{\pi^{r-1}(S_n)} \bigl(E_i[T(S_n)] - 1\bigr) \\ &= \bigl(1 + o(1)\bigr) \frac{(k)_{k-1}}{\pi^{k-1}(S_n)} = \bigl(1 + o(1)\bigr) \frac{k!}{\pi^{k-1}(S_n)}, \end{aligned} \tag{2.17}$$

uniformly for $i \in \mathbb{N}$. Using (2.17), we rewrite (2.15) as

$$E_i[T^k(S_n)] = \bigl(1 + o(1)\bigr) \frac{k!}{\pi^{k-1}(S_n)} + \sum_{j \in S_n^c} p(i, j)E_j[T^k(S_n)],$$

uniformly for $i \in \mathbb{N}$. Now if we define $x^0(i,k) = b(i,k)$ and, for $t \geq 0$,

$$x^{t+1}(i,k) = b(i,k) + \sum_{j \in S_n^c} p(i,j)x^t(j,k), \qquad i \in \mathbb{N},$$

then $x^t(i,k) \uparrow E_i[T^k(S_n)]$, $i \in \mathbb{N}$. In particular, for $k = 1$, we have $b(i,1) = 1$ and $x^t(i,1) \uparrow E_i[T(S_n)]$. Using this observation and (2.17), and $E_i[T(S_n)] \sim 1/\pi(S_n)$ uniformly for $i \in \mathbb{N}$, we conclude

$$E_i[T^k(S_n)] = \bigl(1 + o(1)\bigr)\frac{k!}{\pi^{k-1}(S_n)} E_i[T(S_n)],$$

uniformly for $i \in \mathbb{N}$. Thus (2.16) holds for all $r \geq 1$, and so

$$E_i[T^k(S_n)] = \bigl(1 + o(1)\bigr)\frac{k!}{\pi^k(S_n)}, \qquad k \geq 1,$$

uniformly for $i \in \mathbb{N}$.

Since $\limsup k^{-1}(k!)^{1/k}/k < \infty$, the exponential distribution is the only one with the moments $k!$ (Durrett [10]). The proof of Lemma 2.2 is complete. $\square$

Turn now to $H_n := X(T(S_n))$, H reminding us that $X(T(S_n))$ is the *hit* location.

LEMMA 2.3. *Let $U_n \subseteq S_n$. Uniformly for $i \in \mathbb{N}$,*

$$\lim_{n\to\infty} \left| P_i\{H_n \in U_n\} - \frac{\pi(U_n)}{\pi(S_n)} \right| = 0. \tag{2.18}$$

PROOF. By the Markov property,

$$P_i\{H_n \in U_n\} = p(i,U_n) + \sum_{j \in S_n^c} p(i,j)P_j\{H_n \in U_n\}, \qquad i \in \mathbb{N}, \tag{2.19}$$

where we use the notation $p(i,A) = \sum_{k \in A} p(i,k)$, $A \subseteq \mathbb{N}$.

(a) Assuming only that $\{p(i,k)\}$ is ergodic, let us show that, for all *fixed* $i, j \in \mathbb{N}$,

$$\lim_{n\to\infty} |P_i\{H_n \in U_n\} - P_j\{H_n \in U_n\}| = 0. \tag{2.20}$$

By Cantor's diagonalization device, any subsequence $\{n_m\}$ of $1, 2, \ldots$ contains a further subsequence $\{n_{m_\ell}\}$ such that for $n \to \infty$ along this subsequence, there exists

$$f_i = \lim_{n\to\infty} P_i\{H_n \in U_n\} \in [0,1], \qquad i \in \mathbb{N}.$$

The limits f_i may well depend on $\{n_m\}$, of course. Letting $n = n_{m_\ell} \to \infty$ in (2.19), we obtain

$$f_i = \sum_{j \in \mathbb{N}} p(i,j) f_j, \qquad i \in \mathbb{N}. \tag{2.21}$$

Since the matrix $\{p(i,j)\}$ is ergodic, f_i does not depend on i; Durrett [10], Exercise 3.9. So (2.20) follows.

(b) By tightness, given $\varepsilon \in (0,1)$, there exists $J = J(\varepsilon)$ such that

$$\sum_{j \le J} p(i,j) \ge 1 - \varepsilon \qquad \forall i \in \mathbb{N}.$$

For $n \ge n(J)$, $[J] \subseteq S_n^c$. So, by (2.19),

$$\inf_{i \in \mathbb{N}} P_i\{H_n \in U_n\} \ge (1-\varepsilon) \min_{i \le J} P_i\{H_n \in U_n\}$$

and

$$\sup_{i \in \mathbb{N}} P_i\{H_n \in U_n\} \le \varepsilon + \max_{i \le J} P_i\{H_n \in U_n\}.$$

So

$$\begin{aligned} &\limsup_n \Big[\sup_{i \in \mathbb{N}} P_i\{H_n \in U_n\} - \inf_{i \in \mathbb{N}} P_i\{H_n \in U_n\}\Big] \\ &\qquad \le 2\varepsilon + \lim_{n \to \infty}\Big[\max_{i \le J} P_i\{H_n \in U_n\} - \min_{i \le J} P_i\{H_n \in U_n\}\Big] = 2\varepsilon. \end{aligned}$$

Thus

$$\lim_{n \to \infty}\Big[\sup_{i \in \mathbb{N}} P_i\{H_n \in U_n\} - \inf_{i \in \mathbb{N}} P_i\{H_n \in U_n\}\Big] = 0. \tag{2.22}$$

(c) Multiplying both sides of (2.19) by $\pi(i)$, summing for $i \in \mathbb{N}$ and using stationarity of $\pi(\cdot)$, we obtain

$$\begin{aligned} \sum_{i \in \mathbb{N}} \pi(i) P_i\{H_n \in U_n\} &= \sum_{i \in \mathbb{N}} \pi(i) p(i, U_n) + \sum_{j \in S_n^c} P_j\{H_n \in U_n\} \sum_{i \in \mathbb{N}} \pi(i) p(i,j) \\ &= \pi(U_n) + \sum_{j \in S_n^c} \pi(j) P_j\{H_n \in U_n\}, \end{aligned}$$

so that

$$\sum_{i \in S_n} \pi(i) P_i\{H_n \in U_n\} = \pi(U_n). \tag{2.23}$$

Consequently

$$\inf_{i \in S_n} P_i\{H_n \in U_n\} \le q(n, U_n) \le \sup_{i \in S_n} P_i\{H_n \in U_n\},$$

where

$$q(n, U_n) := \frac{\pi(U_n)}{\pi(S_n)}.$$

Combining (2.22) and the double inequality we conclude that

$$\lim_{n\to\infty}\left|P_i\{H_n \in U_n\} - \frac{\pi(U_n)}{\pi(S_n)}\right| = 0,$$

uniformly for $i \in \mathbb{N}$. The proof of Lemma 2.3 is complete. $\square$

Thus, considered separately, $T(S_n)$ and $X(T(S_n))$ asymptotically behave as if $X(t)$ is a Bernoulli sequence with each trial outcome having distribution π. Of course, the Bernoulli sequence possesses finer properties; in particular, $T(S_n)$ and $X(T(S_n))$ are independent of each other. We are about to impose an additional condition on $\{p(i,k)\}$. It will be used to to establish a limit distribution of the vector $(T(S_n), X(T(S_n)))$, together with a convergence rate in terms of the $\|\cdot\|_{\mathrm{TV}}$ distance. In particular, under the two conditions, $T(S_n)$ and $X(T(S_n))$ turn out to be *asymptotically* independent.

3. Tight, exponentially mixing Markov chains. The extra condition (2) is

$$\rho := \sup_{i,j\in\mathbb{N}} \sum_{k\in\mathbb{N}} |p(i,k) - p(j,k)| < 2. \tag{3.1}$$

(Of course, $\rho \le 2$ always.) Then (Durrett [10], Exercise 5.11),

$$\sum_{k\in\mathbb{N}} |p^n(i,k) - p^n(j,k)| \le 2(\rho/2)^n,$$

where $p^n(\cdot,\cdot)$ are the n-step transition probabilities. Consequently, multiplying by $\pi(j)$ and summing over $j \in \mathbb{N}$,

$$\sum_{k\in\mathbb{N}} |p^n(i,k) - \pi(k)| \le 2(\rho/2)^n.$$

Equivalently, denoting $\mathbf{e} = (\{1\}_{i\in\mathbb{N}})^T$,

$$\|(P^n - \mathbf{e}\boldsymbol{\pi})^T\|_{L_1(\mathbb{N})} = \|P^n - \mathbf{e}\boldsymbol{\pi}\|_{L_\infty(\mathbb{N})} \le 2(\rho/2)^n. \tag{3.2}$$

We call the chains meeting (3.2) exponentially mixing, and use the abbreviation “t.e.m.” chains for tight, exponentially mixing Markov chains. Now

$$\frac{1}{2}\sum_{k\in\mathbb{N}} |p(i,k) - p(j,k)| = \min_{(X,Y)} P\{X \ne Y\},$$

where minimum is over all random vectors (X,Y) such that $P\{X = k\} = p(i,k)$, $P\{Y = k\} = p(j,k)$, $k \in \mathbb{N}$; see Durrett [10]. Therefore, selecting independent X and Y,

$$\frac{1}{2}\sum_{k\in\mathbb{N}} |p(i,k) - p(j,k)| \le 1 - P\{X = Y\} = 1 - \sum_{k\in\mathbb{N}} p(i,k)p(j,k).$$

Hence condition (3.1) is met if

$$\delta_0 := \inf_{i,j\in\mathbb{N}} \sum_{k\in\mathbb{N}} p(i,k)p(j,k) > 0, \tag{3.3}$$

in which case $\rho/2 \le 1-\delta_0$. In fact, for the tight chains the converse is true: (3.1) implies (3.3). Suppose not. Then there exists $\{(i_r, j_r)\}_{r\ge 1}$ such that

$$\lim_{r\to\infty} \sum_{k\in\mathbb{N}} p(i_r,k)p(j_r,k) = 0. \tag{3.4}$$

By the tightness condition, we may assume that $p(i_r,\cdot)$ and $p(j_r,\cdot)$ converge, weakly, to some probability distributions, p_1 and p_2, respectively, that is,

$$p(i_r,k) \to p_1(k), \qquad p(j_r,k) \to p_2(k), \qquad k\in\mathbb{N}.$$

Combining this with (3.1) and (3.4), we obtain

$$\sum_{k\in\mathbb{N}} |p_1(k)-p_2(k)| < 2, \qquad \sum_{k\in\mathbb{N}} p_1(k)p_2(k) = 0.$$

This is impossible, since the second condition implies that

$$|p_1(k)-p_2(k)| = p_1(k)+p_2(k).$$

THEOREM 3.1. *Let $\partial_{n,i}$ denote the joint distribution of $X(T(S_n))$ and $T(S_n)$ for an initial state $i\in\mathbb{N}$. Let ∂_n denote the product probability measure on $S_n\times\mathbb{N}$, such that*

$$\partial_n(A\times B) = \frac{\pi(A)}{\pi(S_n)} \cdot \sum_{\tau\in B} \pi(S_n)\bigl(1-\pi(S_n)\bigr)^{\tau-1}, \qquad A\subseteq S_n, B\subseteq\mathbb{N}.$$

Under conditions (1) *and* (2), *uniformly for $i\in\mathbb{N}$,*

$$\|\partial_{n,i} - \partial_n\|_{\mathrm{TV}} = O(p(S_n)), \tag{3.5}$$

where

$$p(S_n) := \sup_{i\in\mathbb{N}} p(i,S_n). \tag{3.6}$$

PROOF. Introduce

$$\varepsilon_n = \sup_{i\in S_n^c} p(i,S_n), \qquad p(i,A) := \sum_{k\in A} p(i,k);$$

by the tightness and $\lim \pi(S_n) = 0$, we have $\lim \varepsilon_n = 0$. Let $P_n = \{p(i,k)\}_{i,k\in S_n^c}$.

As a first step let us prove the following claim.

LEMMA 3.2. *For n large enough, P_n has an eigenvalue $\lambda_n\in[1-\varepsilon_n, 1)$ and a corresponding eigenvector $\mathbf{f}_n = (\{f_n(i)\}_{i\in S_n^c})^T$, such that*

$$1 \le f_n(i) \le \frac{1}{1-(6/\delta_0)\varepsilon_n}, \tag{3.7}$$

with δ_0 coming from condition (3.3).

PROOF. Given $m > 0$, introduce $S_{n,m} = S_n \cup \{m+1, m+2, \dots\}$, so that $S_{n,m}^c = S_n^c \cap [m]$, which is a finite set. Denote $P_{n,m} = \{p(i,k)\}_{i,k \in S_{n,m}^c}$. By conditions (2.1) and (3.3), there exist n_0 and m_0 such that

$$\inf_{\substack{n \ge n_0 \\ m \ge m_0}} \min_{i \in S_{n,m}^c} p(i, S_{n,m}^c) > 1 - \delta_0/3, \tag{3.8}$$

$$\inf_{\substack{n \ge n_0 \\ m \ge m_0}} \min_{i,j \in S_{n,m}^c} \sum_{k \in S_{n,m}^c} p(i,k)p(j,k) \ge \delta_0/2. \tag{3.9}$$

Let $n \ge n_0$, $m \ge m_0$. Call $\varnothing \ne A \subseteq S_{n,m}^c$ closed (in $S_{n,m}^c$), if $p(i, S_{n,m}^c \setminus A) = 0$ for each $i \in A$. Call a closed set minimal if it does not contain a closed subset. Condition (3.9) clearly ensures that there exists exactly one minimal closed subset A, which may be the whole set $S_{n,m}^c$. A submatrix $P_A := \{p(i,k)\}_{i,k \in A}$ is irreducible; so it has a positive eigenvalue $\lambda(A)$ with a positive eigenvector $\mathbf{f}_A$, and the absolute values of the remaining eigenvalues of P_A do not exceed $\lambda(A)$. In fact, those absolute values are strictly less than $\lambda(A)$. Otherwise, by Frobenius's theorem, there exists a partition $A = \biguplus_{r=1}^h A_r$, $h > 1$, such that, for $r = 1, \dots, h$, $\{k \in A | \exists i \in A_r, p(i,k) > 0\} = A_{r+1}$ $(h+1 := 1)$. So, for $i \in A_1$, $j \in A_2$,

$$\sum_{k \in L} p(i,k)p(j,k) = \sum_{k \in A} p(i,k)p(j,k) = \sum_{k \in A_2 \cap A_3} p(i,k)p(j,k) = 0.$$

And this contradicts (3.9). Furthermore, by (3.8),

$$\lambda(A) \ge \min_{i \in A} p(i, A) = \min_{i \in A} p(i, S_{n,m}^c) > 1 - \delta_0/3,$$

while, denoting $A' = S_{n,m}^c \setminus A$ and using (3.9),

$$\begin{aligned} \max_{i \in A'} p(i, A') &\le 1 - \min_{i \in A'} p(i, A) \\ &\le 1 - \min_{i \in A', j \in A} \sum_{k \in A} p(i,k)p(j,k) \\ &\le 1 - \min_{i \in A', j \in S_{n,m}^c} \sum_{k \in S_{n,m}^c} p(i,k)p(j,k) \le 1 - \delta_0/2. \end{aligned}$$

Therefore $\lambda(A)$ is strictly larger than $\lambda(A')$, the largest eigenvalue of $P_{A'}$. Denoting $P_{A',A} = \{p(i,k)\}_{i \in A', k \in A}$, let $\mathbf{f}_{A'}$ be a solution of

$$P_{A',A}\mathbf{f}_A + P_{A'}\mathbf{f}_{A'} = \lambda(A)\mathbf{f}_{A'}.$$

Since $P_{A',A}\mathbf{f}^A > \mathbf{0}$ and $\lambda(A) > \lambda(A')$, $\mathbf{f}_{A'}$ exists uniquely and is positive. The combined vector $\mathbf{f}_{n,m} = (\mathbf{f}_A, \mathbf{f}_{A'})$ is a unique, positive, eigenvector of $P_{n,m}$ for its largest eigenvalue

$$\lambda_{n,m} = \lambda(A) \ge 1 - \delta_0/3.$$

Let us bound $\max_i \mathbf{f}_{n,m}(i)/\min_i \mathbf{f}_{n,m}(i)$. Introduce $p_i(\tau)$ ($i \in S^c_{n,m}, \tau \ge 0$), the probability that, starting at state i, the Markov process $X(t)$ stays in $S^c_{n,m}$ for all $t \le \tau$. The sequence $\mathbf{p}(\tau) = \{p_i(\tau)\}_{i \in S^c_{n,m}}$, satisfies a recurrence

$$\mathbf{p}(\tau+1) = P_{S^c_{n,m}}\mathbf{p}(\tau), \qquad \mathbf{p}(0) = (1, \dots, 1)^T.$$

Moreover, there exists $C_{n,m} > 0$ such that

$$\mathbf{p}(\tau) \sim C_{n,m}\lambda^{\tau}_{n,m}\mathbf{f}_{n,m}, \qquad \tau \to \infty. \tag{3.10}$$

To exploit this connection, let us first use a coupling device to derive a recurrence for the differences $p_i(\tau) - p_j(\tau)$, $i \ne j$. Consider two independent processes, $X(t)$ and $Y(t)$, starting at i and j in $S^c_{n,m}$. Introduce the events $U(t) = \{X(t) \in S^c_{n,m}\}$ and $V(t) = \{Y(t) \in S^c_{n,m}\}$, and let $\mathbf{1}(W)$ denote the indicator of an event W. By the Markov property,

$$\begin{aligned}
&p_i(\tau+1) - p_j(\tau+1)\\
&\quad = E_{(i,j)}\Bigl[\prod_{t\le\tau+1}\mathbf{1}(U(t)) - \prod_{t\le\tau+1}\mathbf{1}(V(t))\Bigr]\\
&\quad = \sum_{\substack{k_1,k_2\in S^c_{n,m}\\ k_1\ne k_2}} p(i,k_1)p(i,k_2)E_{(k_1,k_2)}\Bigl[\prod_{t\le\tau}\mathbf{1}(U(t)) - \prod_{t\le\tau}\mathbf{1}(V(t))\Bigr]\\
&\qquad + \sum_{k_1\in S^c_{n,m},k_2\in S_{n,m}} p(i,k_1)p(i,k_2)E_{k_1}\Bigl[\prod_{t\le\tau}\mathbf{1}(U(t))\Bigr]\\
&\qquad - \sum_{k_1\in S_{n,m},k_2\in S^c_{n,m}} p(i,k_1)p(i,k_2)E_{k_2}\Bigl[\prod_{t\le\tau}\mathbf{1}(V(t))\Bigr]\\
&\quad = \sum_{\substack{k_1,k_2\in S^c_{n,m}\\ k_1\ne k_2}} p(i,k_1)p(i,k_2)[p_{k_1}(\tau) - p_{k_2}(\tau)]\\
&\qquad + p(j,S_{n,m})\sum_{k_1\in S^c_{n,m}} p(i,k_1)p_{k_1}(\tau) - p(i,S_{n,m})\sum_{k_2\in S^c_{n,m}} p(j,k_2)p_{k_2}(\tau).
\end{aligned}$$

Letting $\tau \uparrow \infty$ and using (3.10) we obtain

$$\begin{aligned}
\lambda_{n,m}\bigl(f_{n,m}(i) - f_{n,m}(j)\bigr) = {}& \sum_{\substack{k_1,k_2\in S^c_{n,m}\\ k_1\ne k_2}} p(i,k_1)p(j,k_2)\bigl(f_{n,m}(k_1) - f_{n,m}(k_2)\bigr)\\
& + p(j,S_{n,m})\sum_{k_1\in S^c_{n,m}} p(i,k_1)f_{n,m}(k_1)\\
& - p(i,S_{n,m})\sum_{k_2\in S^c_{n,m}} p(j,k_2)f_{n,m}(k_2).
\end{aligned}\tag{3.11}$$

Let $f_{n,m}(i_1) = \max_{i \in S^c_{n,m}} f_{n,m}(i)$, $f_{n,m}(i_2) = \min_{i \in S^c_{n,m}} f_{n,m}(i)$. Then it follows from (3.11) that

$$\bigl(f_{n,m}(i_1) - f_{n,m}(i_2)\bigr)\Biggl(\lambda_{n,m} - \sum_{\substack{k_1,k_2 \in S^c_{n,m}\\ k_1 \neq k_2}} p(i_1,k_1)p(i_2,k_2)\Biggr) \leq p(i_2, S_{n,m}) f_{n,m}(i_1).$$

Here [see (3.3)]

$$\sum_{\substack{k_1,k_2 \in S^c_{n,m}\\ k_1 \neq k_2}} p(i_1,k_1)p(i_2,k_2) \leq 1 - \sum_{k \in S^c_{n,m}} p(i_1,k)p(j,k) \leq 1 - \delta_0/2$$

and

$$p(i_2, S_{n,m}) \leq \varepsilon_{n,m} := \max_{i \in S^c_n} p(i, S_{n,m}).$$

As $\lambda_{n,m} \geq 1 - \delta_0/3$, we obtain

$$f_{n,m}(i_1)\biggl(1 - \frac{6}{\delta_0}\varepsilon_{n,m}\biggr) \leq f_{n,m}(i_2) \qquad \forall n \geq n_0, \forall m \geq m_0. \tag{3.12}$$

Now

$$0 \leq \varepsilon_{n,m} - \varepsilon_n \leq \sup_i \sum_{k>m} p(i,k),$$

so that $\lim_{m\to\infty} \varepsilon_{n,m} = \varepsilon_n$, uniformly for n, and $\lim_{n\to\infty} \varepsilon_n = 0$. So there exist $n_1 > n_0$, $m_1 \geq m_0$ such that $\varepsilon_{n,m} \leq \delta_0/7$ for $n \geq n_1$, $m \geq m_1$. For those n, m, relation (3.12), with $f_{n,m}(i_2) = \min_i f_{n,m}(i) = 1$, yields

$$1 \leq f_{n,m}(i) \leq \frac{1}{1 - (6/\delta_0)\varepsilon_{n,m}}, \qquad i \in S^c_{n,m}. \tag{3.13}$$

A standard argument shows then existence of a subsequence $m_s \uparrow \infty$ such that (1) for each $i \in S^c_n$, there exists $f_n(i) = \lim_{m_s\to\infty} f_{n,m_s}(i)$, which necessarily satisfies

$$1 \leq f_n(i) \leq \frac{1}{1 - (6/\delta_0)\varepsilon_n}, \qquad i \in S^c_n,$$

and (2) there exists $\lambda_n = \lim_{m_s\to\infty} \lambda_{n,m_s} \in [1 - \varepsilon_n, 1)$. Clearly then $\mathbf{f}_n := (\{f_n(i)\}_{i \in S^c_n})^T \in L_\infty(S^c_n)$ is an eigenvector of P_n, λ_n being a corresponding eigenvalue. The proof of Lemma 3.2 is complete. □

Let F_n be a diagonal $S^c_n \times S^c_n$ matrix with $F_n(i,i) = f_n(i)$, $i \in S^c_n$. Define a $S^c_n \times S^c_n$ matrix

$$Q_n = \lambda_n^{-1} F_n^{-1} P_n F_n = \lambda_n^{-1}\{(f_n(i))^{-1} p(i,k) f_n(k)\}_{i,k \in S^c_n}.$$

Let $\mathbf{e}_n = (\{1\}_{i \in S^c_n})^T$. Since $F_n \mathbf{e}_n = \mathbf{f}_n$, we have

$$Q_n \mathbf{e}_n = \lambda_n^{-1} F_n^{-1} P_n \mathbf{f}_n = F_n^{-1} \mathbf{f}_n = \mathbf{e}_n,$$

so that Q_n is stochastic. From tightness of $P=\{p(i,k)\}_{i,k\in\mathbb{N}}$ and (3.9) it follows that, for each fixed n, and even uniformly over n, Q_n is tight as well, that is,

$$\delta(K):=\sup_{n,i\in S_n^c}\sum_{k\in S_n^c:\,k>K}Q_n(i,k)\to 0,\qquad K\uparrow\infty.$$

Now from

$$(Q_n)^{\nu}(i,k)=\sum_{j\in S_n^c}Q_n(i,j)(Q_n)^{\nu-1}(j,k),$$

by induction on ν it follows that

$$\sup_{\nu,i\in S_n^c}\sum_{k\in S_n^c:\,k>K}(Q_n)^{\nu}(i,k)\le\delta(K).$$

Hence, given n, the rows of all matrices $(Q_n)^{\nu}$ form a tight set of probability distributions. Therefore there exists $\nu_s\to\infty$ and a family of probability distributions $\pi_n(i,\cdot)$ on S_n^c, $(i\in S_n^c)$, such that, for $i,k\in S_n^c$,

$$(Q_n)^{\nu}(i,k)\to\pi_n(i,k),\qquad \nu\to\infty.$$

In addition, by (3.7) and (3.9), for $i,j\in S_n^c$,

$$\begin{aligned}\sum_{k\in S_n^c}Q_n(i,k)Q_n(j,k)&=\lambda_n^{-2}\sum_{k\in S_n^c}\frac{(f_k^n)^2}{f_i^nf_j^n}p(i,k)p(j,k)\\&\ge\bigl(1-(6/\delta_0)\varepsilon_n\bigr)^2\delta_0/2\ge\delta_0/3,\end{aligned}$$

for n large enough. Therefore [cf. (3.2)]

$$\sum_{k\in S_n^c}|(Q_n)^{\nu}(i,k)-(Q_n)^{\nu}(j,k)|\le 2(1-\delta_0/3)^{\nu}\to 0,\qquad \nu\to\infty. \tag{3.14}$$

Letting $\nu\to\infty$ along $\{\nu_s\}$ in (3.14) we obtain that the family $\{\pi_n(i,\cdot)\}_{i\in S_n^c}$ consists of a single probability distribution $\pi_n(\cdot)$ on S_n^c. Thus, for any distribution $q(\cdot)$ on S_n^c, $q(Q_n)^{\nu_s}\to\pi_n$. Applying this to $q=\pi_n$, and then to $q=\pi_nQ_n$,

$$\pi_nQ_n=\lim_{\nu_s\to\infty}\pi_n(Q_n)^{\nu_s}Q_n=\lim_{\nu_s\to\infty}(\pi_nQ_n)(Q_n)^{\nu_s}=\pi_n,$$

that is, π_n is a stationary distribution of Q_n. Using stationarity of π_n and (3.14) we obtain

$$\begin{aligned}\|[(Q_n)^{\nu}-\mathbf{e}_n\boldsymbol{\pi}_n]^T\|_{L_1(S_n^c)}&=\|(Q_n)^{\nu}-\mathbf{e}_n\boldsymbol{\pi}_n\|_{L_\infty(S_n^c)}\\&\le 2(1-\delta_0/3)^{\nu};\end{aligned}\tag{3.15}$$

cf. (3.2). Since $(Q_n)^{\nu}=\lambda_n^{-\nu}F_n^{-1}P_n^{\nu}F_n$, the combination of (3.3) and (3.2) implies that

$$(P_n)^{\nu}=\lambda_n^{\nu}\mathbf{f}_n\boldsymbol{\sigma}_n+R_{n,\nu},\qquad \boldsymbol{\sigma}_n:=\{\pi_n(i)/f_n(i)\}_{i\in S_n^c},\tag{3.16}$$

where

$$\|R^T_{n,\nu}\|_{L_1(S_n^c)} = \|R_{n,\nu}\|_{L_\infty(S_n^c)} \le 2(1-\delta_0/3)^\nu. \tag{3.17}$$

Estimates (3.16)–(3.17) enable us to determine the limiting joint distribution of $T(S_n)$ and $X(T(S_n))$. Given $A \subseteq S_n$, and z with $|z| \le 1$, define

$$\psi_i(z) = E_i\big[z^{T(S_n)}\mathbf{1}\big(X(T(S_n)) \in A\big)\big], \qquad i \in S_n^c,$$

and $\boldsymbol{\psi}(z) = [\{\psi_i(z)\}_{i\in S_n^c}]^T$. Using the Markov property, we have

$$\psi_i(z) = zp(i,A) + z\sum_{k\in S_n^c} p(i,k)\psi_k(z), \qquad i \in S_n^c,$$

or

$$\boldsymbol{\psi}(z) = z\mathbf{p}_n + zP_n\boldsymbol{\psi}(z), \qquad \mathbf{p}_n := [\{p(i,A)\}_{i\in S_n^c}]^T.$$

Therefore, introducing the $S_n^c \times S_n^c$ identity matrix I_n and using (3.16)–(3.17),

$$\begin{aligned}\boldsymbol{\psi}(z) &= z(I_n - zP_n)^{-1}\mathbf{p}_n = z\sum_{\nu\ge 0} z^\nu (P_n)^\nu \mathbf{p}_n \\ &= \frac{z}{1-z\lambda_n}\mathbf{f}_n(\boldsymbol{\sigma}_n\mathbf{p}_n) + \mathbf{R}_n(z);\end{aligned} \tag{3.18}$$

$$\mathbf{R}_n(z) := z\sum_{\nu\ge 0} z^\nu R_{n,\nu}\mathbf{p}_n.$$

By (3.17) and

$$\|\mathbf{p}_n\|_{L_\infty(S_n^c)} \le p(A) := \sup_{i\in\mathbb{N}} p(i,A),$$

we have that each component of $\mathbf{R}_n(z)$ is analytic for $|z| < (1-\delta_0/3)^{-1}$, and

$$\|\mathbf{R}_n(z)\|_{L_\infty(S_n^c)} \le \frac{2p(A)}{1-|z|(1-\delta_0/3)}.$$

Therefore each $\psi_i(z)$ initially defined in the unit disk admits a *meromorphic* extension to the open disk of radius $(1-\delta_0/3)^{-1} > 1$, with a single, simple pole $z = 1/\lambda_n$ in that disk.

As for the explicit term in (3.18),

$$[\mathbf{f}_n(\boldsymbol{\sigma}_n\mathbf{p}_n)]_i = C(A)f_n(i), \qquad C(A) := \sum_{j\in S_n^c} \sigma_n(j)p(j,A).$$

In particular, setting $z = 1$,

$$P_i\{X(T(S_n)) \in A\} = \psi_i(1) = \frac{C(A)f_n(i)}{1-\lambda_n} + O(p(A)).$$

Since $P_i\{X(T(S_n)) \in S_n\} = 1$, we then obtain

$$1 - \lambda_n = \frac{C(S_n) f_n(i)}{1 + \varepsilon_i(n)}, \qquad \varepsilon_i(n) = O(p(S_n)). \tag{3.19}$$

Therefore

$$P_i\{X(T(S_n)) \in A\} = \frac{C(A)}{C(S_n)} + O(p(S_n)), \qquad i \in S_n^c.$$

This uniform estimate and (2.23), with $U_n = A_n$, easily imply that

$$\frac{C(A)}{C(S_n)} = \frac{\pi(A)}{\pi(S_n)} + O(p(S_n)). \tag{3.20}$$

Furthermore, given a positive integer τ,

$$P_i\{T(S_n) = \tau, X(T(S_n)) \in A\} = [z^\tau]\psi_i(z) = \frac{1}{2\pi i} \oint_L \frac{\psi_i(z)}{z^{\tau+1}}\, dz,$$

where L is a circular contour $|z| = 1$. By (3.18), the *extended* $\psi_i(z)$ has a unique singularity, a simple pole, in a ring between L and L_1, which is the circular contour of radius $(1 - \delta_0/4)^{-1}$. Using (3.18) and the residue theorem, we obtain

$$\begin{aligned}
\frac{1}{2\pi i} \oint_L \frac{\psi_i(z)}{z^{\tau+1}}\, dz &= -\frac{C(A_n) f_n(i)}{2\pi i} \oint_{L_1} \frac{1}{(1 - z\lambda_n) z^\tau}\, dz + O\big((1 - \delta_0/4)^\tau p(A)\big) \\
&= C(A_n) f_n(i) \lambda_n^{\tau-1} + O\big((1 - \delta_0/4)^\tau p(A)\big).
\end{aligned}$$

Thus, by (3.19) and (3.20),

$$\begin{aligned}
&P_i\{T(S_n) = \tau, X(T(S_n)) \in A\} \\
&\quad = (1 - \lambda_n)\lambda_n^{\tau-1} \frac{\pi(A)}{\pi(S_n)} \\
&\qquad + O[(1 - \lambda_n)\lambda_n^{\tau-1} p(S_n)] + O[(1 - \delta_0/4)^\tau p(S_n)].
\end{aligned} \tag{3.21}$$

In particular, for $A = S_n$,

$$\begin{aligned}
P_i\{T(S_n) = \tau\} &= (1 - \lambda_n)\lambda_n^{\tau-1} \\
&\quad + O[(1 - \lambda_n)\lambda_n^{\tau-1} p(S_n)] + O[(1 - \delta_0/4)^\tau p(S_n)].
\end{aligned} \tag{3.22}$$

Now we had proved already that, under the tightness only,

$$E_i[T^k(S_n)] \sim k! E_i^k[T(S_n)] \sim k! \pi^{-k}(S_n),$$

so that

$$E_i[(T(S_n))_k] \sim k! \pi^{-k}(S_n).$$

According to (3.22), we also have

$$E_i[(T(S_n))_k] = \frac{k!}{(1 - \lambda_n)^k} + O\big(p(S_n)(1 - \lambda_n)^{-k}\big).$$

Comparing the two formulas we see that $1-\lambda_n \sim \pi(S_n)$. In fact, we can say more. From (3.22) it follows that, uniformly for $i \in S_n^c$,

$$E_i[T(S_n)] = \frac{1+O(p(S_n))}{1-\lambda_n}.$$

Combining this with (2.10), we get

$$1-\lambda_n = \pi(S_n)\bigl(1+O(p(S_n))\bigr). \tag{3.23}$$

The rest is short. Let $C \subseteq S_n \times \mathbb{N}$, and $C_\tau = \{k \in S_n : (k,\tau) \in C\}$. From (3.21) it follows that, uniformly for $i \in S_n^c$ *and* C,

$$P_i\{(X(T(S_n)), T(S_n)) \in C\} = \sum_{\tau\in\mathbb{N}} (1-\lambda_n)\lambda_n^{\tau-1}\frac{\pi(C_\tau)}{\pi(S_n)} + O(p(S_n)). \tag{3.24}$$

And, by (3.23),

$$\begin{aligned}
&\left|\sum_{\tau\in\mathbb{N}}(1-\lambda_n)\lambda_n^{\tau-1}\frac{\pi(C_\tau)}{\pi(S_n)} - \sum_{\tau\in\mathbb{N}}\pi(S_n)\bigl(1-\pi(S_n)\bigr)^{\tau-1}\frac{\pi(C_\tau)}{\pi(S_n)}\right| \\
&\quad\le |(1-\lambda_n)-\pi(S_n)| \sum_{\tau\in\mathbb{N}} x^{\tau-2}[1+\tau(1-x)]\frac{\pi(C_\tau)}{\pi(S_n)}, \\
&\qquad\qquad [x \text{ between } \lambda_n \text{ and } 1-\pi(S_n)] \\
&\quad\le |(1-\lambda_n)-\pi(S_n)|\left[2(1-x)\sum_{\tau\ge 1}\tau x^{\tau-1} + 2\sum_{\tau\ge1} x^{\tau-1}\right] \\
&\quad= |(1-\lambda_n)-\pi(S_n)| \cdot 4(1-x)^{-1} = O(p(S_n)).
\end{aligned}$$

So (3.24) becomes

$$\|\partial_{n,i}-\partial_n\|_{\mathrm{TV}} = O(p(S_n)), \qquad i \in S_n^c.$$

Suppose that $i \in S_n$. Then

$$\begin{aligned}
&P_i\{(X(T(S_n)), T(S_n)) \in C\} \\
&\quad= \sum_{k\in S_n^c} p(i,k)P_k\bigl\{\bigl(X(T(S_n)), T(S_n)+1\bigr) \in C\bigr\} + O(p(S_n)),
\end{aligned} \tag{3.25}$$

where, by (3.24),

$$\begin{aligned}
P_k\bigl\{\bigl(X(T(S_n)), T(S_n)+1\bigr)\in C\bigr\} &= \sum_{\tau\in\mathbb{N}}(1-\lambda_n)\lambda_n^{\tau-1}\frac{\pi(C_{\tau+1})}{\pi(S_n)} + O(p(S_n)) \\
&= \sum_{\tau\ge 2}(1-\lambda_n)\lambda_n^{\tau-2}\frac{\pi(C_\tau)}{\pi(S_n)} + O(p(S_n))
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{\tau \geq 2} (1-\lambda_n)\lambda_n^{\tau-1} \frac{\pi(C_\tau)}{\pi(S_n)} + O\big(1 - \lambda_n + p(S_n)\big) \\
&= \sum_{\tau \in \mathbb{N}} (1-\lambda_n)\lambda_n^{\tau-1} \frac{\pi(C_\tau)}{\pi(S_n)} + O\big(1 - \lambda_n + p(S_n)\big) \\
&= \sum_{\tau \in \mathbb{N}} (1-\lambda_n)\lambda_n^{\tau-1} \frac{\pi(C_\tau)}{\pi(S_n)} + O(p(S_n)).
\end{aligned}
$$

Therefore, by (3.25), (3.24) holds for $i \in S_n$ as well. This completes the proof of Theorem 3.1. □

Let $T_{n,r}$ be the time intervals between consecutive visits to S_n. So $T_{n,1} = T(S_n)$, and, for $r > 1$,

$$
\begin{aligned}
T_{n,r} &= \min\{t > \mathcal{T}_{n,r-1} : X(t) \in S_n\} - \mathcal{T}_{n,r-1}, \\
\mathcal{T}_{n,r-1} &:= \sum_{k<r} T_{n,k},
\end{aligned}
$$

that is, $\mathcal{T}_{n,r}$ is the time of rth visit to S_n. Let $X_{n,r} = X(\mathcal{T}_{n,r})$, that is, $X_{n,r}$ is a state in S_n visited at time $\mathcal{T}_{n,r}$. Introduce a random sequence $\{\ell_r; t_r\}_{r\geq 1}$, where all $\ell_1, t_1, \ell_2, t_2, \dots$ are independent and, for each r,

$$
P\{\ell_r \in A\} = \frac{\pi(A)}{\pi(S_n)}, \qquad A \subseteq S_n,
$$

while t_r is distributed geometrically, with success probability $\pi(S_n)$. Also, for two random vectors, $\mathbf{Y}$ and $\mathbf{Z}$, of a common dimension ν, let $d_{\mathrm{TV}}(\mathbf{Y}, \mathbf{Z})$ denote the total variation distance between the distributions of $\mathbf{Y}$ and $\mathbf{Z}$, that is,

$$
d_{\mathrm{TV}}(\mathbf{Y}; \mathbf{Z}) = \sup_{B \in \mathcal{B}^\nu} |P\{\mathbf{Y} \in B\} - P(\mathbf{Z} \in B)|.
$$

Since $|x|$ is convex,

$$
\begin{aligned}
0.5 &\sup_{f : \|f\|_{L_\infty(\mathbb{N}^\nu)} \leq 1} |E[f(\mathbf{Y})] - E[f(\mathbf{Z})]| \\
&\leq d_{\mathrm{TV}}(\mathbf{Y}; \mathbf{Z}) \\
&\leq \sup_{f : \|f\|_{L_\infty(\mathbb{N}^\nu)} \leq 1} |E[f(\mathbf{Y})] - E[f(\mathbf{Z})]|.
\end{aligned}
$$

Theorem 3.1 implies the following.

THEOREM 3.3. *Uniformly for an initial state $i \in \mathbb{N}$,*

$$
d_{\mathrm{TV}}(\{X_{n,r}; T_{n,r}\}_{1\leq r\leq k}; \{\ell_r; t_r\}_{1\leq r\leq k}) = O(kp(S_n)). \tag{3.26}
$$

Thus, if $k = k(n)$ is such that $kp(S_n) \to 0$, the random sequence $\{X_{n,r}; T_{n,r}\}_{1\leq r\leq k}$ is asymptotic, with respect to the total variation distance, to the Bernoulli sequence $\{\ell_r, t_r\}_{1\leq r\leq k}$.

PROOF. We prove (3.26) by induction on k. For $k=1$, it is the statement of Theorem 3.1. Assume (3.26) holds for some $k \ge 1$. Let $f : \mathbb{N}^{k+1} \times \mathbb{N}^{k+1} \to \mathbb{R}$ have $\|f\|_{L_\infty(\mathbb{N}^{k+1}\times\mathbb{N}^{k+1})} \le 1$. Denote

$$\begin{aligned} &\mathbf{X} = \{X_{n,r}\}_{1\le r\le k+1}, \qquad \mathbf{X}^{(k)} = \{X_{n,r}\}_{1\le r\le k}, \\ &\mathbf{Y} = \{T_{n,r} - \delta(i,r)\}_{1\le r\le k+1}, \qquad \mathbf{Y}^{(k)} = \{T_{n,r} - \delta(i,r)\}_{1\le r\le k}, \\ &\mathbf{x}^{(k)} = \{x_r\}_{1\le r\le k}, \qquad \mathbf{y}^{(k)} = \{y_r\}_{1\le r\le k}. \end{aligned} \tag{3.27}$$

We write first

$$\begin{aligned} &E_i[f(\mathbf{X};\mathbf{Y})] \\ &\quad = E_i\big[E_i\big[f\big((\mathbf{X}^{(k)}, X_{n,k+1}); (\mathbf{Y}^{(k)}, T_{n,k+1}-1)\big)|(\mathbf{X}^{(k)}, \mathbf{Y}^{(k)})\big]\big]. \end{aligned} \tag{3.28}$$

By the strong Markov property,

$$\begin{aligned} &E_i\big[f\big((\mathbf{X}^{(k)}, X_{n,k+1}); (\mathbf{Y}^{(k)}, T_{n,k+1}-1)\big)|(\mathbf{X}^{(k)}, \mathbf{Y}^{(k)})\big] \\ &\quad = E_{x_k}\big[f\big((\mathbf{x}^{(k)}, X), (\mathbf{y}^{(k)}, T-1)\big)\big]|_{\mathbf{x}^{(k)}=\mathbf{X}^{(k)}, \mathbf{y}^{(k)}=\mathbf{Y}^{(k)}}; \end{aligned} \tag{3.29}$$

here X, T are the location and the time of the first hit of S_n for the chain starting at $x_k \in S_n$. Using (3.26) for $k=1$, we have

$$\begin{aligned} &\big|E_{x_k}\big[f\big((\mathbf{x}^{(k)}, X), (\mathbf{y}^{(k)}, T-1)\big)\big] - E\big[f\big((\mathbf{x}^{(k)}, \ell), (\mathbf{y}^{(k)}, t)\big)\big]\big| \\ &\quad = O(p(S_n)), \end{aligned} \tag{3.30}$$

uniformly for $j \in S_n$. [Here $(\ell, t) \to \overset{\mathcal{D}}{\equiv} (\ell_r, t_r)$.] So, introducing $\tilde f : \mathbb{N}^k \times \mathbb{N}^k \to \mathbb{R}$ by

$$\tilde f(\mathbf{x}^{(k)}, \mathbf{y}^{(k)}) = E\big[f\big((\mathbf{x}^{(k)}, \ell), (\mathbf{y}^{(k)}, t)\big)\big],$$

and using (3.28)–(3.30), we have

$$\big|E_i[f(\mathbf{X}, \mathbf{Y})] - E_i\big[\tilde f(\mathbf{X}^{(k)}, \mathbf{Y}^{(k)})\big]\big| = O(p(S_n)). \tag{3.31}$$

Besides, applying the inductive hypothesis to $\tilde f$, we also have

$$E_i\big[\tilde f(\mathbf{X}^{(k)}, \mathbf{Y}^{(k)})\big] - E[\tilde f(\{\ell_r, t_r\}_{1\le r\le k})] = O(kp(S_n)). \tag{3.32}$$

It follows from Fubini's theorem and (3.31)–(3.32), that

$$\begin{aligned} E_i[f(\mathbf{X}, \mathbf{Y})] - E[f(\{\ell_r, t_r\}_{1\le r\le k+1})] &= E_i[f(\mathbf{X}, \mathbf{Y})] - E[\tilde f(\{\ell_r, t_r\}_{1\le r\le k})] \\ &= O\big((k+1)p(S_n)\big), \end{aligned}$$

which proves the inductive step. So (3.26) holds for all k. $\square$

Let us apply Theorem 3.3 to the extreme values for the t.e.m. chains. Given a large N, let $X^{(j)} = X^{(N,j)}$ denote the jth largest among $X(1), \dots, X(N)$; in particular, $X^{(1)} = \max_{1\le t\le N} X(t)$. From now on we will use a notation $S_n = \{n+1, n+2, \dots\}$.

COROLLARY 3.4. *Uniformly for $i = X(0)$,*

$$P_i\{X^{(\mu)} \le n\} = P\{\text{Poisson}(N\pi(S_n)) < \mu\} + O(\mu^2/N + Np^2(S_n)). \tag{3.33}$$

PROOF. $X^{(\mu)} \le n$ iff during $[1, N]$ the chain visited S_n at most $\mu - 1$ times. So, by Theorem 3.3,

$$P_i\{X^{(\mu)} \le n\} = \sum_{j<\mu} \binom{N}{j} \pi^j(S_n)\big(1 - \pi(S_n)\big)^{N-j} + O(\mu p(S_n)). \tag{3.34}$$

Here

$$(N)_j\big(1 - \pi(S_n)\big)^{-j} = N^j\big(1 + O(\mu^2/N) + O(\mu p(S_n))\big)$$

and

$$\begin{aligned}\big(1 - \pi(S_n)\big)^N - e^{-N\pi(S_n)} &\le 2Ne^{-N\pi(S_n)}\big(e^{-\pi(S_n)} - \big(1 - \pi(S_n)\big)\big)\\ &\le N\pi^2(S_n)e^{-N\pi(S_n)}.\end{aligned}$$

So, as $\pi(S_n) \le p(S_n)$, (3.34) becomes (3.33). □

Corollary (3.4) is a special case of the following result. Given $a < b \le \infty$, denote $S_{a,b} = (a, b]$, that is, $S_{a,b} = S_a \setminus S_b$. Let $V_{a,b} = V_{N,a,b}$ denote the number of visits to $S_{a,b}$ during $[1, N]$, and $\lambda_{a,b} = \lambda_{N,a,b} = N\pi(S(a, b))$.

THEOREM 3.5. *Let $(a_1, b_1], \dots, (a_k, b_k]$ be disjoint. Uniformly for $i = X(0)$,*

$$\begin{aligned}P_i\Big\{\bigcap_{1\le\ell\le k}\{V_{a_\ell,b_\ell} \le \mu_\ell\}\Big\} &= \prod_{1\le\ell\le k} P\{\text{Poisson}(\lambda_{a_\ell,b_\ell}) \le \mu_\ell\}\\ &\quad + O\big(\mu^2/N + Np^2(S_a)\big),\end{aligned} \tag{3.35}$$

where $\mu = \mu_1 + \cdots + \mu_k$, $a = \min_\ell a_\ell$. Thus, if $\mu = o(N^{1/2})$ and $Np^2(S_a) = o(1)$, the numbers of visits to nonoverlapping intervals $(a_\ell, b_\ell]$ are asymptotically independent Poissons with parameters $N\pi(S_{a_\ell,b_\ell})$.

PROOF. Applying Theorem 3.3 to $S := \bigcup_{1\le\ell\le k} S_{a_\ell,b_\ell}$,

$$\begin{aligned}P_i\Big\{\bigcap_{1\le\ell\le k}\{V_{a_\ell,b_\ell} \le \mu_\ell\}\Big\} &= \sum_{j_1\le\mu_1;\dots;j_k\le\mu_k} \binom{N}{j_1, \dots, j_k} \prod_{1\le\ell\le k} \pi^{j_\ell}(S_{a_\ell,b_\ell})\\ &\quad\times\Big(1 - \sum_{1\le\ell\le k}\pi(S_{a_\ell,b_\ell})\Big)^{N-j} + O(\mu p(S_a)),\end{aligned}$$

where $j = j_1 + \cdots + j_\ell$, and

$$\binom{N}{j_1, \dots, j_k} = \frac{N!}{j_1! \cdots j_k!(N-j)!}.$$

The rest runs parallel with the proof of Corollary 3.4. $\square$

Analogously we obtain a relatively simple asymptotic formula for the joint distribution of $X^{(1)}, \dots, X^{(\mu)}$.

THEOREM 3.6. *Let $\infty = n_0 \geq n_1 \geq n_2 \geq \cdots \geq n_\mu$. Uniformly for $i = X(0)$,*

$$\begin{aligned} &P_i\Biggl\{\bigcap_{1\le\ell\le\mu}\{X^{(\ell)} \le n_\ell\}\Biggr\} \\ &\qquad = \sum_{\substack{\nu_1,\dots,\nu_\mu \\ \forall r\le\mu:\sum_{j=1}^{r}\nu_j\le r-1}} \prod_{1\le r\le\mu} P\{\text{Poisson}(\lambda_{n_r,n_{r-1}}) = \nu_r\} \\ &\qquad\quad + O\bigl(\mu^2/N + Np^2(S_{n_\mu})\bigr). \end{aligned} \tag{3.36}$$

More generally, let

$$B \subseteq \{\mathbf{x} = (x_1, \dots, x_\mu) \in \mathbb{N} : x_1 \geq \cdots \geq x_\mu\}.$$

Given $\mathbf{x}$, let $y_1(\mathbf{x}) > \cdots > y_m(\mathbf{x})$ denote all the distinct values (range) of the sequence $x_1, \dots, x_\mu$, and let $a_j = a_j(\mathbf{x}) > 0$ be the multiplicity of $y_j = y_j(\mathbf{x})$. So $m = m(\mathbf{x}) \leq \mu$, and $a_1 + \cdots + a_m = \mu$. Then, denoting $n(B) = \inf_{\mathbf{x}\in B} x_\mu$ and setting $y_0 = \infty$,

$$\begin{aligned} &P_i\{(X^{(1)}, \dots, X^{(\mu)}) \in B\} \\ &\quad = \sum_{\mathbf{x}\in B} \prod_{1\le r\le m} e^{-N\pi([y_r, y_{r-1}))} \frac{(N\pi(y_r))^{a_r}}{a_r!} + O\bigl(\mu^2/N + Np^2(S_{n(B)})\bigr) \\ &\quad = \sum_{\mathbf{x}\in B} e^{-N\pi([y_m,\infty))}(N\pi([y_m,\infty)))^{\mu} \prod_{1\le r\le m} \frac{\sigma^{a_r}(y_r)}{a_r!} \\ &\qquad + O\bigl(\mu^2/N + Np^2(S_{n(B)})\bigr), \end{aligned} \tag{3.37}$$

where $\sigma(y) = \pi(y)/\pi([y_m, \infty))$, $y \in [y_m, \infty)$.

In the next section we will describe two models of a random constrained composition, and show that each random composition is sharply approximated by a t.e.m. chain. It will enable us to use Corollary 3.4 and Theorems 3.5, 3.6 for analysis of the limiting distribution of the larger parts.

4. Two random constrained compositions and Markov chain approximations. We focus on two interesting cases of such compositions, the column-convex-animals (cca) compositions and the Carlitz (C) compositions.

4.1. *Definitions and some basic facts*. (a) A column-convex animal (cca) is a sequence of contiguous vertical segments of unit squares in $\mathbb{Z}^2$, ordered from left to right, such that every two successive columns have a common boundary consisting of at least one vertical edge of $\mathbb{Z}^2$. If the total number of unit squares involved is ν, then the lengths of the vertical segments form a composition of ν; we call it a cca-composition. Let $T(\nu, \mu)$ denote the total number of the cca-compositions of ν with μ parts; then $T(\nu) := \sum_{\mu \ge 1} T(\nu, \mu)$ is the total number of the cca-compositions of ν. Introduce $f(w, z)$, the bivariate generating function (BGF) of $T(\nu, \mu)$,

$$f(w, z) = \sum_{\mu, \nu \ge 1} T(\nu, \mu) w^{\mu} z^{\nu}.$$

Louchard [19] found that

$$\begin{aligned} f(w, z) &= \frac{wz(z-1)^3}{h(w, z)}, \\ h(w, z) &:= z^4(w-1) + z^3(w^2 - w + 4) - z^2(w+6) + z(w+4) - 1. \end{aligned} \tag{4.1}$$

Therefore $f(z)$, the GF of $T(\nu)$, is

$$f(z) = f(1, z) = \frac{z(z-1)^3}{h(1, z)} = \frac{(z-1)^3}{4z^3 - 7z^2 + 5z - 1}, \tag{4.2}$$

a formula discovered earlier by Klarner [15]. Privman and Forgacs [22] used (4.2), and Darboux theorem, to show that

$$T(\nu) = \frac{C}{z_*^{\nu}}\bigl(1 + O(\gamma^{\nu})\bigr), \tag{4.3}$$

where $\gamma < 1$, $C = 0.18\ldots$, and $z_* = 0.31\ldots$ is the smallest-modulus solution of $h(1, z) = 0$.

We get a uniformly random cca-composition of ν, if we assume that each composition has the same probability, $1/T(\nu)$. It was discovered in [19, 20] that the distribution of the last (first) part is asymptotic to

$$\begin{aligned} \pi_1(k) &= z_*^k (k + a), \\ a &:= \frac{1 - z_*}{z_*} - \frac{1}{1 - z_*} = 0.75\ldots, \end{aligned} \tag{4.4}$$

which is directly seen as a probability distribution. Besides, the joint distribution of *two* consecutive parts Y_t and Y_{t+1}, with both t and $\mathcal{M} - t$ of order $\Theta(\nu)$, was shown to be asymptotic to that of two consecutive states of an ergodic Markov chain on $\mathbb{N}$, in a stationary regime, with transition probabilities

$$p(i, k) = z_*^k (i + k - 1)\frac{k + a}{i + a}, \tag{4.5}$$

and a stationary distribution

$$\begin{aligned} \pi(k) &= A^{-1} z_*^k (k+a)^2, \\ A &:= \sum_{k\geq 1} z_*^k (k+a)^2 = \frac{z_*^2}{(1-z_*)^3} + \frac{1-z_*}{z_*}. \end{aligned} \tag{4.6}$$

That $\sum_{k\geq 1} p(i,k) = 1$ follows from another formula for a,

$$a = \frac{2z_*^2}{(1-2z_*)(1-z_*)}.$$

(The given formulation is slightly different from, but equivalent to that in [19, 20].) One way to derive (4.4) is to use (4.3) and a formula for $f_k(z)$, the generating function of the cca-compositions with the first (last) part equal k,

$$f_k(z) = z^k + z^k f(z)\left[k + \frac{z^3 - z^2 + z}{(1-z)^3}\right], \tag{4.7}$$

which can be read out of [19]. Comparing the first line in (4.4) and (4.7) we must also have yet another formula for a, namely,

$$a = \frac{z_*^3 - z_*^2 + z_*}{(1-z_*)^3}, \tag{4.8}$$

which is indeed the case.

(b) A Carlitz (C) composition of ν is defined as a composition such that every two consecutive parts are distinct from each other. The counterparts of the cited results for the cca-compositions are as follows. Carlitz [8] proved that

$$\begin{aligned} f(w,z) &= -1 + \frac{1}{h(w,z)}, \\ h(w,z) &:= 1 - \sum_{j\geq 1} (-1)^{j+1} \frac{w^j z^j}{1-z^j}; \end{aligned} \tag{4.9}$$

for $|w| \leq 1$, h, as a function of z is analytic for $|z| < 1$, and for $|w| \geq 1$, h is analytic for $|z| < 1/|w|$. Louchard and Prodinger [21] found a rather more tractable expression for h, namely,

$$h(w,z) = 1 - \sum_{j\geq 1} \frac{wz^j}{1+wz^j}. \tag{4.10}$$

(4.9) and (4.10) were used in [20] to show that

$$T(\nu) = \frac{C}{z_*^\nu}\left(1 + O(\gamma^\nu)\right), \tag{4.11}$$

where $\gamma < 1$, $C = 0.456\ldots$, and $z_* = 0.57\ldots$ is the smallest-modulus solution of $h(1,z) = 0$.

We get a uniformly random C-composition of ν, if we assume that each C-composition has the same probability, $1/T(\nu)$. In a striking analogy with the random cca-composition, the two consecutive parts Y_t and Y_{t+1}, deep inside the composition, are also jointly asymptotic to the two consecutive states of an ergodic Markov chain, with transition probabilities

$$p(i,k)=\begin{cases} z_*^k \dfrac{1+z_*^i}{1+z_*^k}, & i\neq k,\\ 0, & i=k\end{cases} \tag{4.12}$$

and a stationary distribution

$$\pi(k)=A^{-1}\frac{z_*^k}{(1+z_*^k)^2}, \qquad A:=\sum_{k\ge 1}\frac{z_*^k}{(1+z_*^k)^2}. \tag{4.13}$$

And the limiting distribution of Y_1 is

$$\pi_1(k)=\frac{z_*^k}{1+z_*^k}, \tag{4.14}$$

which follows from (4.11) and a counterpart of (4.7),

$$f_k(z)=\frac{z^{k+1}}{1+z^{k+1}}+f(z)\frac{z^k}{1+z^k}. \tag{4.15}$$

[That (4.12) and (4.14) are indeed probability distributions follows from the definition of z_* as a root of $h(1,z)=0$ and (4.10).]

For each of the compositions, an equation $h(w,z)=0$ [for the attendant function $h(w,z)$] determines a root $z(w)$, well defined for w sufficiently close to 1, such that $z(1)=z_*$, $z(w)$ is infinitely differentiable, and $z'(1)<0$. The number of parts $\mathcal{M}$ for each of the random compositions was shown, in [19] and [21], respectively, to be Gaussian in the limit $\nu\to\infty$, with mean $\alpha\nu$ and variance $\beta\nu$, where

$$\alpha=-\frac{z'(1)}{z(1)}=-\frac{z'(1)}{z_*}, \qquad \beta=\alpha^2+\alpha-\frac{z''(1)}{z_*}. \tag{4.16}$$

In particular,

$$\alpha=\begin{cases} -\dfrac{12z_*^2-14z_*+5}{z_*^4+z_*^3-z_*^2+z_*}=0.45\dots, & \text{(for cca)},\\ -\dfrac{\sum_{j\ge 1} j z_*^{j-1}/(1+z_*^j)^2}{\sum_{j\ge 1} z_*^j/(1+z_*^j)^2}=0.35\dots, & \text{(for C)}; \end{cases} \tag{4.17}$$

needless to say, in each case z_* is the root of the corresponding equation $h(1,z)=0$.

In the Appendix we will prove the following large deviation result.

LEMMA 4.1. *For each of the compositions, there exists an absolute constant* $c > 0$ *such that*

$$P\{|\mathcal{M} - \alpha\nu| \ge s\} \le c\nu \exp(-s^2/3\beta\nu),$$

provided that $s = o(\nu)$. *Thus*

$$P\{|\mathcal{M} - \alpha\nu| \le \nu^{1/2} \ln \nu\} \ge 1 - \nu^{-K} \qquad \forall K > 0. \tag{4.18}$$

NOTE. Borrowing a term from Knuth, Motwani and Pittel [17], the event on the left of (4.18) happens *quite surely* (q.s.).

4.2. *Approximating the random compositions by the Markov chains*. The results cited above strongly suggest, though not actually prove, that the random cca-composition and the random C-composition considered as random processes are each asymptotic to its own Markov chain, defined in (4.4)–(4.5) and (4.11)–(4.13), respectively.

The following theorem confirms this natural conjecture.

THEOREM 4.2. *Let* $\mathbf{Y} = \{Y_t\}_{t\ge1}$ *be either the random cca-composition, or the random C-composition of* ν. *Let* $\mathbf{Z} = \{Z(t)\}_{t\ge1}$ *be the corresponding Markov chain with the transition probabilities* $p(i,k)$, *and* $Z(1)$ *having the distribution* $\{\pi_1(i)\}_{i\ge1}$. *Introduce*

$$\begin{aligned} \hat{\mathcal{M}} &= \max\{1 \le m < \mathcal{M} : Y_1 + \cdots + Y_m \le \nu - \ln^2 \nu\}, \\ \hat{M} &= \max\{m \ge 1 : Z(1) + \cdots + Z(m) \le \nu - \ln^2 \nu\}; \end{aligned} \tag{4.19}$$

in particular, $\hat{\mathcal{M}} \in (\mathcal{M} - \ln^2 \nu - 1, \mathcal{M})$. *Let* $\hat{\partial}$ *and* $\hat{d}$ *denote the probability distribution of* $(\hat{\mathcal{M}}, (Y_1, \dots, Y_{\hat{\mathcal{M}}}))$ *and* $(\hat{M}, (Z(1), \dots, Z(\hat{M}))$, *respectively. For each chain*,

$$\|\hat{\partial} - \hat{d}\|_{\mathrm{TV}} = O(\nu^{-K}) \qquad \forall K > 0. \tag{4.20}$$

So, the random composition of ν, read from left to right, is closely approximated by the corresponding Markov chain, as long as the accumulated sum of parts stays below $\nu - \ln^2 \nu$. (A restriction of this sort is unavoidable: like the first part, the last part of the random composition has the distribution π_1, which differs from the stationary distribution π.) Now, we will see that, with high probability, the extreme-valued parts are in this "bulk" of the composition, implying that they are well approximated by the extreme-valued states of the $\hat{M}$-long segment of the corresponding Markov chain. It is easy to verify that

$$\sup_i \sum_{k>n} p(i,k) = \begin{cases} O(z_*^n n^2), & \text{cca-chain}, \\ O(z_*^n), & \text{C-chain}, \end{cases} \tag{4.21}$$

where $z_* = 0.31\dots$ for the cca-chain and $z_* = 0.57\dots$ for the C-chain. That is, the chains meet the tightness condition (2.1). And the exponential mixing property in the form of (3.3) is easily verified as well. So we are able to use Corollary 3.4 and Theorem 3.6, say, for derivation of the limiting distribution of those extreme values, and then the last theorem for a quick proof of the corresponding results regarding extreme-valued parts of each of the random compositions.

Turning the tables, we can also use Theorem 4.2 and Lemma 4.1 to determine the very likely bounds of $\hat{M}$ with sufficient accuracy. Since $\hat{\mathcal{M}} \in (\mathcal{M} - \ln^2 \nu - 1, \mathcal{M})$, Lemma 4.1 implies that q.s.

$$|\hat{\mathcal{M}} - \alpha\nu| \le 2\nu^{1/2} \ln \nu.$$

So, applying Theorem 4.2, we immediately see that

$$|\hat{M} - \alpha\nu| \le 2\nu^{1/2} \ln \nu \tag{4.22}$$

q.s. as well. (!)

PROOF. The key element is the following claim.

LEMMA 4.3. *Let* $\mathbf{Y}$ *be either the random cca-composition or the random C-composition of* ν. *Let* $k \ge 1$, $\mathbf{i} = (i_1, \dots, i_k) \in \mathbb{N}^k$, *where* $i_1 + \cdots + i_k < \nu$. *Denote* $P_\nu(\mathbf{i}) = P\{Y_1 = i_1, \dots, Y_k = i_k\}$ *and* $P(\mathbf{i}) = P\{Z(1) = i_1, \dots, Z(k) = i_k\}$. *Then, uniformly for* k *and* $\mathbf{i}$,

$$P_\nu(\mathbf{i}) = P(\mathbf{i}) \exp\bigl(O\bigl(k\gamma^{\nu - |\mathbf{i}|}\bigr)\bigr), \qquad |\mathbf{i}| = i_1 + \cdots + i_k, \tag{4.23}$$

where γ *comes from either* (4.3) *or* (4.10).

PROOF. Let $\mathbf{Y}$ be the random cca-composition of ν. We will prove (4.23) by induction on k.

For $k = 1$,

$$P_\nu(i_1) = \frac{[z^\nu] f_{i_1}(z)}{[z^\nu] f(z)}, \tag{4.24}$$

where $f(z)$ and $f_{i_1}(z)$ are given by (4.2) and (4.7), respectively. Here, by (4.3),

$$[z^\nu] f(z) = T(\nu) = \frac{C}{z_*^\nu} \exp(O(\gamma^\nu)). \tag{4.25}$$

Further, by (4.7),

$$\begin{aligned}
[z^\nu] f_{i_1}(z) &= \delta_{\nu, i_1} + [z^{\nu - i_1}] f(z) \biggl[i_1 + \frac{z^3 - z^2 + z}{(1-z)^3} \biggr] \\
&= i_1 T(\nu - i_1) + [z^{\nu - i_1}] f(z) \frac{z^3 - z^2 + z}{(1-z)^3} \qquad (4.26)\\
&= i_1 \frac{C}{z_*^{\nu - i_1}} \exp(O(\gamma^{\nu - i_1})) + \frac{C}{z_*^{\nu - i_1}} \frac{z_*^3 - z_*^2 + z_*}{(1 - z^*)^3} \exp(O(\gamma^{\nu - i_1})).
\end{aligned}$$

[z_* is the smallest modulus pole of $f(z)(z^3-z^2+z)(1-z)^{-3}$, as well.] It follows from (4.24)–(4.26) and (4.8) that

$$\begin{aligned}P_\nu(i_1) &= z_*^{i_1}\left[i_1+\frac{z_*^3-z_*^2+z_*}{(1-z_*)^3}\right]\exp(O(\gamma^{\nu-i_1}))\\ &= z_*^{i_1}(i_1+a)\exp(O(\gamma^{\nu-i_1})) = P(i_1)\exp(O(\gamma^{\nu-i_1})),\end{aligned}$$

which is (4.23) for $k=1$.

Suppose that (4.23) holds for some $k \ge 1$. Let $\mathbf{i}=(i_1,\dots,i_{k+1})$ be such that $|\mathbf{i}|<\nu$. Let $\mathbf{i}'=(i_2,\dots,i_{k+1})$; then $|\mathbf{i}'|<\nu-i_1$. Let $T(\mathbf{i},\nu)$ and $T(\mathbf{i}',\nu-i_1)$ denote the total number of the cca of area ν ($\nu-i_1$, resp.) with the first $k+1$ parts $i_1,\dots,i_{k+1}$ (the first k parts $i_2,\dots,i_{k+1}$, resp.). By the definition of the cca-composition,

$$T(\mathbf{i},\nu) = (i_1+i_2-1)T(\mathbf{i}',\nu-i_1).$$

Therefore,

$$\begin{aligned}P_\nu(\mathbf{i}) &= \frac{T(\mathbf{i},\nu)}{T(\nu)}\\ &= \frac{[z^\nu]f_{i_1}(z)}{T(\nu)}\cdot\frac{T(\nu-i_1)}{[z^\nu]f_{i_1}(z)}\cdot\frac{(i_1+i_2-1)T(\mathbf{i}',\nu-i_1)}{T(\nu-i_1)}\\ &= P_\nu(i_1)\frac{\exp(O(\gamma^{\nu-i_1}))}{i_1+a}\cdot(i_1+i_2-1)P_{\nu-i_1}(\mathbf{i}')\\ &= P(i_1)\frac{\exp(O(\gamma^{\nu-i_1}))}{i_1+a}\cdot(i_1+i_2-1)P(\mathbf{i}')\exp\bigl(O\bigl(k\gamma^{\nu-i_1-|\mathbf{i}'|}\bigr)\bigr)\\ &= P(i_1)\frac{i_1+i_2-1}{i_1+a}P(\mathbf{i}')\exp\bigl(O\bigl((k+1)\gamma^{\nu-|\mathbf{i}|}\bigr)\bigr),\end{aligned}$$

and we observe that

$$\begin{aligned}\frac{i_1+i_2-1}{i_1+a}P(\mathbf{i}') &= \frac{i_1+i_2-1}{i_1+a}P(i_2)\prod_{r=2}^{k}p(i_r,i_{r+1})\\ &= p(i_1,i_2)\prod_{r=2}^{k}p(i_r,i_{r+1}) = \prod_{r=1}^{k}p(i_r,i_{r+1}).\end{aligned}$$

Hence

$$\begin{aligned}P_\nu(\mathbf{i}) &= P(i_1)\prod_{r=1}^{k}p(i_r,i_{r+1})\exp\bigl(O\bigl((k+1)\gamma^{\nu-|\mathbf{i}|}\bigr)\bigr)\\ &= P(\mathbf{i})\exp\bigl(O\bigl((k+1)\gamma^{\nu-|\mathbf{i}|}\bigr)\bigr),\end{aligned}$$

which completes the inductive proof of (4.23) for the random cca-composition. The proof for the random C-composition is similar, and we omit it. $\square$

Lemma 4.3 implies the bound (4.20) of Theorem 4.2 without much difficulty. Consider, for instance, the random cca-composition of ν. Let m, $\mathbf{i} = (i_1, \dots, i_m)$ be given. Clearly

$$P\{\hat{\mathcal{M}} \ge m, Y_1 = i_1, \dots, Y_m = i_m\}$$
$$= P\{\hat{M} \ge m, Z(1) = i_1, \dots, Z(m) = i_m\} = 0,$$

unless $|\mathbf{i}| \le \nu - \ln^2 \nu$. In the latter case $m \le \nu - \ln^2 \nu$, and, by Lemma 4.3,

$$P\{\hat{\mathcal{M}} \ge m, Y_1 = i_1, \dots, Y_m = i_m\}$$
$$= P(\mathbf{i}) \exp\bigl(O\bigl(m\gamma^{\nu - |\mathbf{i}|}\bigr)\bigr)$$
$$= P\{\hat{M} \ge m, Z(1) = i_1, \dots, Z(m) = i_m\} \exp(O(\nu\gamma^{\ln^2 \nu})),$$

uniformly for m and $\mathbf{i}$ in question. Consequently, uniformly for all m and $B \subseteq \mathbb{N}^m$,

$$P\{\hat{\mathcal{M}} \ge m, (Y_1, \dots, Y_m) \in B\}$$
$$= P\{\hat{M} \ge m, (Z(1), \dots, Z(m)) \in B\} \exp(O(\gamma^{0.5 \ln^2 \nu})),$$

whence

$$\begin{aligned} &P\{\hat{\mathcal{M}} = m, (Y_1, \dots, Y_m) \in B\} \\ &\qquad = P\{\hat{M} = m, (Z(1), \dots, Z(m)) \in B\} + O(\gamma^{0.5 \ln^2 \nu}). \end{aligned} \tag{4.27}$$

Let $D \subseteq \mathbb{N}^{\nu+1}$ be given. For $\mathbf{z} \in \mathbb{N}^k$, $k \le \nu + 1$, we write $\mathbf{z} \in D$ if $\mathbf{z}$ is a projection of a point in D on the first k coordinates. Noticing that $\hat{\mathcal{M}} \le \nu$ and $\hat{M} \le \nu$, we obtain from (4.26): uniformly for all $D \in \mathbb{N}^{\nu+1}$,

$$P\{(\hat{\mathcal{M}}, (Y_1, \dots, Y_{\hat{\mathcal{M}}})) \in D\}$$
$$= P\{(\hat{M}, (Z(1), \dots, Z(\hat{M}))) \in D\} + O(\gamma^{0.5 \ln^2 \nu}).$$

This completes the proof of Theorem 4.2. $\square$

4.3. *Limiting distributions of the extreme parts of the random compositions.* By (4.21), for each of the two chains, q.s.

$$N_1 + 1 \le \hat{M} \le 1 + N_2, \qquad N_{1,2} = \lfloor \alpha\nu \pm 2\nu^{1/2} \ln \nu \rfloor. \tag{4.28}$$

So q.s. the extreme values of $\{Z(t)\}_{0<t\le\hat{M}}$ are sandwiched between those of $\{Z(t)\}_{0<t\le N_1+1}$ and $\{Z(t)\}_{0<t\le N_2+1}$. Picking a generic $N \in [N_1, N_2]$, introduce $\{X(t)\}_{0\le t\le N} = \{Z(t)\}_{0<t\le N+1}$. Here $X(0)$ has distribution $\pi_1(\cdot)$.

Let $X^{(\mu)}$ be the μth largest among $X(t)$, $t \in [1, N]$, for $X(0) = i$, $i \in \mathbb{N}$. By Corollary 3.4,

$$P_i\{X^{(\mu)} \le n\} = P\{\text{Poisson}(N\pi(S_n)) < \mu\} + O\bigl(\mu^2/N + Np^2(S_n)\bigr), \tag{4.29}$$

where

$$\pi(S_n) = \sum_{k>n} \pi(k), \qquad p(S_n) = \sup_i \sum_{k>n} p(i,k).$$

Here $p(S_n) = O(n^2 z_*^n)$ for the cca-chain and $p(S_n) = O(z_*^n)$ for the C-chain, see (4.21). (Again, $z_* = 0.31\dots$ for the cca-chain and $z_* = 0.57\dots$ for the C-chain.) Turn to $\pi(S_n)$. For the cca chain by (4.6),

$$\pi(S_n) = n^2 z_*^n B\bigl(1 + O(n^{-1})\bigr), \qquad B := \frac{z_*^2(1-z_*)^2}{z_*^3 + (1-z_*)^4}. \tag{4.30}$$

For the C-chain, by (4.13),

$$\pi(S_n) = B z_*^n\bigl(1 + O(z_*^n)\bigr), \qquad B := A^{-1}\frac{z_*}{1-z_*}. \tag{4.31}$$

LEMMA 4.4 (cca-chain). *Suppose that*

$$n = \frac{\ln[\lambda^{-1} B N (\ln N/\ln z_*)^2]}{\ln(1/z_*)} \in \mathbb{N}, \tag{4.32}$$

where $\lambda = o(\ln N)$. *If* $\mu = o(\ln N)$, *then*, *uniformly for* $i \in \mathbb{N}$,

$$P_i\{X^{(\mu)} \le n\} = P\{\text{Poisson}(\lambda) < \mu\} + O[(\lambda + \mu)/\ln N]. \tag{4.33}$$

Equivalently, *define* $W_{N,\mu}$ *by*

$$X^{(\mu)} = \frac{\ln[W_{N,\mu}^{-1} B N (\ln N/\ln z_*)^2]}{\ln(1/z_*)}; \tag{4.34}$$

then, *for* $s = o(\ln N)$ *such that*

$$\frac{\ln[s^{-1} B N (\ln N/\ln z_*)^2]}{\ln(1/z_*)} \in \mathbb{N}, \tag{4.35}$$

we have

$$P_i\{W_{N,\mu} \ge s\} = P\{W_\mu \ge s\} + O[(s + \mu)/\ln N]; \tag{4.36}$$

here $W_\mu = V_1 + \cdots + V_\mu$, *and* $V_1, \dots, V_\mu$ *are independent exponentials with unit mean.*

This lemma implies the following cruder result. [We use a symbol $O_p(1)$ to denote a random variable bounded in probability as $N \to \infty$.]

COROLLARY 4.5. *If* $\mu = o(\ln N)$, *then*, *uniformly for* $i \in \mathbb{N}$,

$$X^{(\mu)} = \frac{\ln(\mu^{-1} N \ln^2 N)}{\ln(1/z^*)} + O_p(1). \tag{4.37}$$

Here are the counterparts for the chain associated with the random C-composition.

LEMMA 4.6 (C-chain). *Suppose that*

$$n = \frac{\ln(\lambda^{-1} BN)}{\ln(1/z_*)} \in \mathbb{N}, \tag{4.38}$$

where $\lambda = o(N^{1/2})$. *If* $\mu = o(N^{1/2})$, *then, uniformly for* $i \in \mathbb{N}$,

$$P_i\{X^{(\mu)} \le n\} = P\{\mathrm{Poisson}(\lambda) < \mu\} + O[(\lambda^2 + \mu^2)/N]. \tag{4.39}$$

Equivalently, define $W_{N,\mu}$ *by*

$$X^{(\mu)} = \frac{\ln(W_{N,\mu}^{-1} BN)}{\ln(1/z_*)}; \tag{4.40}$$

then, for $s = o(N^{1/2})$ *such that*

$$\frac{\ln(s^{-1} BN)}{\ln(1/z_*)} \in \mathbb{N}, \tag{4.41}$$

we have

$$P_i\{W_{N,\mu} \ge s\} = P\{W_\mu \ge s\} + O[(s^2 + \mu^2)/N].$$

COROLLARY 4.7 (C-chain). *If* $\mu = o(N^{1/2})$, *then, uniformly for* $i \in \mathbb{N}$,

$$X^{(\mu)} = \frac{\ln(\mu^{-1} N)}{\ln(1/z^*)} + O_p(1). \tag{4.42}$$

PROOF OF LEMMA 4.4 AND COROLLARY 4.5. (a) By (4.30), (4.32) and (4.21),

$$N\pi(S_n) = \lambda + O(\lambda/\ln N), \qquad Np^2(S_n) = O[N^{-1}(N\pi(S_n))^2] = O(\lambda^2/N).$$

Then, for $j \le \mu$,

$$(N\pi(S_n))^j = \lambda^j\bigl(1 + O(\mu/\ln N)\bigr).$$

So, by Corollary 3.4, (3.33) and (4.21), (4.30),

$$\begin{aligned} P_i\{X^{(\mu)} \le n\} &= \sum_{j<\mu} e^{-N\pi(S_n)} \frac{(N\pi(S_n))^j}{j!} + O\bigl((\mu^2 + \lambda^2)/N\bigr) \\ &= \sum_{j<\mu} e^{-\lambda} \frac{\lambda^j}{j!} + O\bigl((\lambda + \mu)/\ln N\bigr). \end{aligned}$$

(b) Given $s > 0$,

$$\left\lfloor \frac{\ln[s^{-1}BN(\ln n/\ln z_*)^2]}{\ln(1/z_*)} \right\rfloor = \frac{\ln[s_1^{-1}BN(\ln n/\ln z_*)^2]}{\ln(1/z_*)},$$

where $s_1 \in [s, sz_*^{-1})$. Using the definition of $W_{N,\mu}$ in (4.34) and the asymptotic formula (4.36) we obtain then, for $s = o(\ln N)$,

$$\begin{aligned} &P\{z_* W_\mu \geq s\} + O[(s+\mu)/\ln N] \\ &\quad \leq P_i\{W_{N,\mu} \geq s\} \\ &\quad \leq P\{W_\mu \geq s\} + O[(s+\mu)/\ln N]. \end{aligned} \tag{4.43}$$

For μ fixed, (4.43) implies that

$$\lim_{A\to\infty} \liminf_{N\to\infty} P_i\{W_{N,\mu} \in [A^{-1}, A]\} = 1,$$

that is, in probability, $W_{N,\mu}$ is bounded away from zero and infinity, whence $\ln W_{N,\mu} = O_p(1)$. Suppose $\mu \to \infty$. Then $(W_\mu - \mu)/\mu^{1/2}$ is asymptotically normal, with zero mean and unit variance. Consequently,

$$\ln W_\mu = \ln \mu + O_p(1). \tag{4.44}$$

Let $y = y(N) \to \infty$ so slow that $s = \mu e^y = o(\ln N)$ as well. Using the right-hand side of (4.43), we obtain

$$\begin{aligned} &P_i\{\ln W_{N,\mu} \geq \ln \mu + y\} \\ &\quad = P_i\{W_{N,\mu} \geq \mu e^y\} \\ &\quad = P\{W_\mu \geq \mu e^y\} + O\bigl((s+\mu)/\ln N\bigr) \\ &\quad = P\{\ln W_\mu \geq \ln \mu + y\} + O\bigl((s+\mu)/\ln N\bigr) = o(1). \end{aligned} \tag{4.45}$$

Analogously, the left-hand side of (4.43) delivers

$$\lim_{N\to\infty} P_i\{\ln W_{N,\mu} \geq \ln \mu - y\} = 0. \tag{4.46}$$

The relations (4.44)–(4.46), together with (4.34) prove (4.37). $\square$

The proof of Lemma 4.6 and Corollary 4.7 is similar and we omit it.

Recall that $N \in [N_1, N_2]$, $N_{1,2} = \lfloor \alpha\nu \pm 2\nu^{1/2} \ln \nu \rfloor$. Introduce $N_0 = \lfloor \alpha\nu \rfloor$. It is easy to check that the proof of Lemma 4.4 and Corollary 4.5 goes through with very minor changes if, instead of (4.32), we define an integer n by

$$n = \frac{\ln[\lambda^{-1}BN_0(\ln N_0/\ln z_*)^2]}{\ln(1/z_*)}.$$

[The key is that

$$N(\ln N)^2 = \bigl(1 + O(N_0^{-1/2} \ln N_0)\bigr) N_0 (\ln N_0)^2,$$

uniformly for N in question.] The same change can be made in the formulation of Lemma 4.6 and Corollary 4.7 for the C-chain. This observation coupled with the fact that $\hat{X}_+^{(\mu)}$, the μth largest value of $\{X(t)\}_{0<t<\hat{M}}$, is sandwiched between those for $\{X(t)\}_{0<t\le N_1}$ and $\{X(t)\}_{0<t\le N_2}$, show that in Lemma 4.4, Corollary 4.5, Lemma 4.6 and Corollary 4.7 we can put $\hat{X}_+^{(\mu)}$ instead of $X^{(\mu)}$. Below, by relations (4.33) and (4.39) we will mean their modifications, that is, with $\hat{X}_+^{(\mu)}$ on their LHS.

Turn to $\hat{X}^{(\mu)}$, the μth largest value among $X(0), X(1), \dots, X(\hat{M})$. $X(0)$ has the distribution π_1 given by either by (4.4) or by (4.14). Hence

$$P\{X(0)\ge n\}=\begin{cases} O(nz_*^n), & \text{for cca,}\\ O(z_*^n), & \text{for C.}\end{cases} \tag{4.47}$$

Now

$$\hat{X}_+^{(\mu)}\le \hat{X}^{(\mu)}\le X(0)+\hat{X}_+^{(\mu)};$$

so, for the cca case, we use n defined by (4.32) and add an extra error term coming from (4.47), that is,

$$nz_*^n=O\bigl(z_*^{\ln(\nu\ln^2\nu)/(\lambda\ln(1/z_*))}\ln\nu\bigr)=O(\nu^{-1}),$$

to the RHS of (4.33), to obtain the corresponding claim for $\hat{X}^{(\mu)}$. Likewise, in the C-case we need to add an error term $O(\nu^{-1/2})$ to the RHS of (4.39). Again, we will refer to these new relations as (4.33) and (4.39).

But then, according to Theorem 4.2, the μth largest among the *parts* $Y_1,\dots,Y_{\hat{\mathcal{M}}}$ of the corresponding random composition can replace $\hat{X}^{(\mu)}$ on the LHS of (4.33) and (4.39), respectively. These are our newest, (4.33) and (4.39).

Finally, if we include the rightmost parts $Y_{\hat{\mathcal{M}}+1}, Y_{\hat{\mathcal{M}}+2},\dots,$ it will not substantially affect the the limiting behavior of the μth largest overall part either. Here is why. The number of these parts is $m:=\lceil \ln^2\nu\rceil$, at most. The total number of parts is q.s. of order $\nu\gg m$, which means the last m parts are q.s. well defined. Those parts, read from right to left, and the first m parts, read from left to right, are equidistributed. By Theorem 4.2, these m first parts are within the total variation distance $O(\nu^{-K})$ $(\forall K>0)$, from $Z(1),\dots,Z(m)$. We know that $Z(1)$ has the distribution π_1. Since

$$\sup_{i\in\mathbb{N},t\ge 1}\sum_{k\ge n}p^t(i,k)\le \sup_{i\in\mathbb{N}}\sum_{k\ge n}p(i,k),\qquad t\ge 1,$$

we see that

$$P\{Z(t)\ge n\}\le \sup_{i\in\mathbb{N}}\sum_{k\ge n}p(i,k),\qquad t\ge 2.$$

In view of (4.4) and (4.21), we obtain then, for the cca-chain,

$$\begin{aligned}P\left\{\max_{1\le t\le m} Z(t)\ge \frac{\ln[\lambda^{-1}BN_0(\ln N_0/\ln z_*)^2]}{\ln(1/z_*)}\right\}&=O\bigl((\ln^4\nu)z_*^{\ln(\nu\ln^2\nu)/(\lambda\ln(1/z_*))}\bigr)\\ &=O(\nu^{-1}\ln^3\nu).\end{aligned}$$

For the C-chain, the analogous probability is of order $\nu^{-1/2}\ln^2\nu$. Therefore, by adding the error terms $O(\nu^{-1}\ln^3\nu)$ and $O(\nu^{-1/2}\ln^2\nu)$ to the RHS of (4.33) and (4.39) (where $N = N_0 = \lfloor\alpha\nu\rfloor$, of course), we obtain the limiting distributions of the μth largest part of both random compositions, together with explicit error terms. [For the cca-composition, the order of the total error term remains unchanged, that is, $O((\lambda+\mu)/\ln\nu)$.]

In summary, we have proved the following.

THEOREM 4.8. *For a random composition* $\mathbf{Y}$ *of* ν, *let* $Y^{(\mu)}$ *denote the* μ*th largest part. Let* $N_0 = \lfloor\alpha\nu\rfloor$, α *being defined in* (4.17). *Let* W_μ *be the sum of* μ *independent exponentials with unit mean.*

(i) *For the random cca-composition, define* $W_{\nu,\mu}$ *by*

$$Y^{(\mu)} = \frac{\ln[W_{\nu,\mu}^{-1} BN_0(\ln N_0/\ln z_*)^2]}{\ln(1/z_*)},$$

B *being defined in* (4.30). *Then, for* $s = o(\ln\nu)$ *such that*

$$\frac{\ln[s^{-1}BN_0(\ln N_0/\ln z_*)^2]}{\ln(1/z_*)} \in \mathbb{N},$$

we have

$$P\{W_{\nu,\mu} \ge s\} = P\{W_\mu \ge s\} + O[(s+\mu)/\ln\nu].$$

(ii) *For the random C-composition, define* $W_{\nu,\mu}$ *by*

$$Y^{(\mu)} = \frac{\ln[W_{\nu,\mu}^{-1} BN_0]}{\ln(1/z_*)},$$

B *being defined in* (4.31). *Then, for* $s = o(\nu^{1/2})$ *such that*

$$\frac{\ln(s^{-1}BN_0)}{\ln(1/z_*)} \in \mathbb{N},$$

we have

$$P\{W_{\nu,\mu} \ge s\} = P\{W_\mu \ge s\} + O\big(\nu^{-1/2}\ln^2\nu + (s^2+\mu^2)/\nu\big).$$

Here is a cruder estimate implied by Theorem 4.8.

COROLLARY 4.9. (i) *For the random cca-composition,*

$$Y^{(\mu)} = \frac{\ln(\mu^{-1}\nu\ln^2\nu)}{\ln(1/z_*)} + O_p(1) \qquad [\mu = o(\ln\nu)].$$

(ii) *For the random C-composition,*

$$Y^{(\mu)} = \frac{\ln(\mu^{-1}\nu)}{\ln(1/z_*)} + O_p(1) \qquad [\mu = o(\nu^{1/2})].$$

(iii) *So, for both cases*,

$$Y^{(1)} - Y^{(\mu)} = \frac{\ln \mu}{\ln(1/z_*)} + O_p(1),$$

if $\mu = o(\ln \nu)$ *and* $\mu = o(\nu^{1/2})$, *respectively.*

APPENDIX

PROOF OF LEMMA 4.1. Consider the case of the random C-composition. The BGF of $T(\mu, \nu)$, the number of C-compositions of ν with μ parts and ν and μ marked by z and w, respectively, is given by (4.9)–(4.10)

$$f(w, z) = -1 + \frac{1}{h(w, z)}, \qquad h(w, z) = 1 - \sum_{j \ge 1} \frac{w z^j}{1 + w z^j}.$$

This bivariate series converges for $|z| < 1$ and $|w| < 1/|z|$. So, choosing $r_1 < 1$, and $r_2 < 1/r_1$, we have

$$\begin{aligned} P\{\mathcal{M} = \mu\} &= \frac{[z^\nu w^\mu] f(w, z)}{T(\nu)} \\ &= \frac{1}{T(\nu)} \frac{1}{(2\pi i)^2} \oint_{z \in \mathcal{C}_1} \oint_{w \in \mathcal{C}_2} \frac{f(w, z)}{z^{\nu+1} w^{\mu+1}} \, dw \, dz, \end{aligned}$$

where $\mathcal{C}_1$, $\mathcal{C}_2$ are circles of radius r_1 and r_2, respectively. In essence, it is this formula that, via Bender's method [5], enabled Louchard [19] and Louchard and Prodinger [21] to establish a sharp local limit theorem for $\mathcal{M}$ for the cca-composition and the C-composition. Since our goal is to bound the probability of large deviations, we use a considerably less analytical argument, which is a bivariate extension of Chernoff's method.

As a preparation, we need to define a differentiable extension of $z_* = 0.57\ldots$, the smallest-module root of $h(1, z) = 0$. To this end, we compute

$$h_z(1, z) = -\sum_{j \ge 1} \frac{j z^{j-1}}{(1 + z^j)^2}, \qquad h_w(1, z) = -\sum_{j \ge 1} \frac{z^j}{(1 + z^j)^2}. \tag{A.1}$$

So $h_z(1, z) < 0$, $h_w(1, z) < 0$ for $z \in (0, 1)$. By continuity of $h_z(z, w)$, $h_w(z, w)$, we obtain: there exists $\varepsilon \in (0, 1 - z_*)$ such that (1) $(z_* + \varepsilon)(1 + \varepsilon) < 1$, and (2)

$$h_z(z, w) < 0, \qquad h_w(z, w) < 0 \qquad \forall (z, w) \in \mathbb{R}_+^2 : z \le z + \varepsilon, |w - 1| \le \varepsilon. \tag{A.2}$$

Consequently, for $|w - 1| \le \varepsilon$, the equation $h(z, w) = 0$ has a unique root $z = z(w)$, of multiplicity 1, in $[0, z_* + \varepsilon]$, which is infinitely differentiable as a function of w, and $z(1) = z_*$. In particular,

$$z'(w) = -\frac{h_w(z(w), w)}{h_z(z(w), w)} < 0,$$

that is, $z(w)$ is strictly decreasing. So $z(w) > z_*$ for $w < 1$ and $z(w) < z_*$ for $w > 1$.

Now, the series for the bivariate generating function $f(w,z)$ converges for $|w-1| \le \varepsilon$ and $|z| < z(w)$. Since all the coefficients in the series are nonnegative,

$$\sum_{\ell \ge m} [z^\nu w^\ell] f(w,z) \le \frac{f(w,z)}{z^\nu w^m}, \qquad w \in [1, 1+\varepsilon_0], z \in (0, z(w)).$$

Likewise,

$$\sum_{\ell \le m} [z^\nu w^\ell] f(w,z) \le \frac{f(w,z)}{z^\nu w^m}, \qquad w \in [1-\varepsilon_0, 1], z \in (0, z(w)).$$

Here, by the definition of $f(w,z)$ and $z(w)$,

$$f(w,z) \le \frac{c}{z(w)-z}, \qquad z < z(w).$$

Therefore, for each m,

$$P(\mathcal{M} \ge m) \le c\frac{z^{-n}w^{-m}(z(w)-z)^{-1}}{T(\nu)}, \qquad w \in [1, 1+\varepsilon_0], z \in (0, z(w)), \tag{A.3}$$

and

$$P(\mathcal{M} \le m) \le c\frac{z^{-n}w^{-m}(z(w)-z)^{-1}}{T(\nu)}, \qquad w \in [1-\varepsilon_0, 1], z \in (0, z(w)). \tag{A.4}$$

Consider (A.3). To get the most out of this upper bound we need to determine z and w that minimize the RHS, that is,

$$H^{(m)}(w,z) := -\nu \ln z - m \ln w - \ln\bigl(z(w) - z\bigr).$$

Let us find a stationary point $(\bar{w}, \bar{z})$ of $H^{(m)}(w,z)$ in the region $w \in [1, 1+\varepsilon_0]$, $z \in (0, z(w))$. From the equations

$$H_z^{(m)} = -\frac{\nu}{z} + \frac{1}{z(w)-z} = 0,$$

$$H_w^{(m)} = -\frac{m}{w} - \frac{z'(w)}{z(w)-z} = 0,$$

we obtain that

$$\bar{z} = \frac{\nu}{\nu+1} z(\bar{w}),$$

where $\bar{w} = \bar{w}(m)$ must be a root of

$$\frac{wz'(w)}{z(w)} = -\frac{m}{\nu+1}. \tag{A.5}$$

Equation (A.5) has a solution $w = 1$, if

$$m = \bar{m} := (\nu + 1)\mu, \qquad \mu := -\frac{z'(1)}{z(1)}.$$

Furthermore, in [19] it was shown that

$$\frac{d}{dw}\frac{wz'(w)}{z(w)}\bigg|_{w=1} = \frac{z''(1)}{z(1)} - \mu - \mu^2$$

is negative; this is, $-\beta$, β defined in (4.16). Since

$$\frac{d}{dm}\left(-\frac{m}{\nu+1}\right) = -\frac{1}{\nu+1} < 0$$

as well, for

$$0 \le m - \bar{m} = o(\nu),$$

equation (A.5) defines a strictly increasing $\bar{w}(m)$; so $\bar{w}(m) > 1$ for $m > \bar{m}$. More precisely,

$$\begin{aligned}\bar{w}(m) &= 1 + \frac{\beta}{\nu+1}(m - \bar{m}) + O\big((m-\bar{m})^2/\nu^2\big)\\ &= 1 + \frac{\beta}{\nu}(m - \mu\nu) + O\big((m-\mu\nu)^2/\nu^2\big).\end{aligned}$$

Now

$$\begin{aligned}H^{(\bar{m})}(\bar{w}(\bar{m}), \bar{z}(\bar{w}(\bar{m}))) &= -\nu\ln\left(\frac{\nu}{\nu+1}z_*\right) - \ln\left(\frac{z_*}{\nu+1}\right)\\ &= -\nu\ln z_* + \ln\nu + O(1).\end{aligned}$$

Also

$$\begin{aligned}&\frac{d}{dm}H^{(m)}(\bar{w}(m), \bar{z}(m))\\ &\quad = H_m^{(m)}(\bar{w}(m), \bar{z}(m)) + H_w^{(m)}(\bar{w}(m), \bar{z}(m)) + H_z^{(m)}(\bar{w}(m), \bar{z}(m))\\ &\quad = H_m^{(m)}(\bar{w}(m), \bar{z}(m)) = -\ln\bar{w}(m),\end{aligned}$$

which implies that

$$\frac{d}{dm}H^{(m)}(\bar{w}(m), \bar{z}(m))\bigg|_{m=\bar{m}} = -\ln\bar{w}(\bar{m}) = 0,$$

and also that

$$\begin{aligned}\frac{d^2}{dm^2}H^{(m)}(\bar{w}(m), \bar{z}(m))\bigg|_{m=\bar{m}} &= -\frac{\bar{w}'(m)}{\bar{w}(m)}\bigg|_{m=\bar{m}}\\ &= -\bar{w}'(\bar{m})\\ &= -\frac{\beta}{\nu+1}.\end{aligned}$$

Therefore, for $0 \le m - \bar{m} = o(\nu)$,

$$H^{(m)}(\bar{w}(m), \bar{z}(m)) = -\nu \ln z_* + \ln \nu - \big(1 + o(1)\big)\frac{\beta}{2(\nu+1)}(m - \bar{m})^2 + O(1)$$
$$\le -\nu \ln z_* + \ln \nu - \frac{\beta}{3\nu}(m - \mu\nu)^2 + O(1).$$

Using this bound in (A.3) for $w = \bar{w}(m)$, $z = \bar{z}(\bar{w}(m))$, and recalling that $T(\nu)$ is of order $z_*^{-\nu}$, we obtain

$$P(\mathcal{M} \ge m) \le c\nu \exp\biggl(-\frac{\beta}{3\nu}(m - \mu\nu)^2\biggr), \qquad 0 < m - \mu\nu = o(\nu).$$

Likewise,

$$P(\mathcal{M} \le m) \le c\nu \exp\biggl(-\frac{\beta}{3\nu}(m - \mu\nu)^2\biggr), \qquad 0 < \mu\nu - m = o(\nu).$$

The case of the random cca-composition is quite analogous, so we omit the proof. $\square$

Acknowledgments. It is my genuine pleasure to thank Guy Louchard for introducing me to the random compositions and to his conjecture on hitting times for Markov chains with uniformly exponential tails of the row distributions. His strong belief in the conjecture, kind encouragement and insightful feedback helped to sustain my effort during months of work on this study. I thank a referee for a painstaking effort to evaluate the paper and for a series of penetrating critical comments. I am grateful to an Associate Editor and to the Editor for their thoughtful consideration.

REFERENCES

[1] Aldous, D. J. (1982). Markov chains with almost exponential hitting times. *Stochastic Process. Appl.* **13** 305–310. MR0671039

[2] Aldous, D. J. and Brown, M. (1992). Inequalities for rare events in time-reversible Markov chains. I. In *Stochastic Inequalities* (*Seattle*, *WA*, 1991). *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **22** 1–16. IMS, Hayward, CA. MR1228050

[3] Aldous, D. J. and Brown, M. (1993). Inequalities for rare events in time-reversible Markov chains. II. *Stochastic Process. Appl.* **44** 15–25. MR1198660

[4] Andrews, G. E. (1976). *The Theory of Partitions*. *Encyclopedia of Mathematics and Its Applications* **2**. Addison-Wesley, Reading, MA. MR0557013

[5] Bender, E. A. (1973). Central and local limit theorems applied to asymptotic enumeration. *J. Combinatorial Theory Ser. A* **15** 91–111. MR0375433

[6] Bender, E. A. and Canfield, E. R. (2005). Locally restricted compositions. I. Restricted adjacent differences. *Electron. J. Combin.* **12** 1–27 (electronic). MR2180794

[7] Breiman, L. (1968). *Probability*. Addison-Wesley, Reading, MA. MR0229267

[8] Carlitz, L. (1976). Restricted compositions. *Fibonacci Quart.* **14** 254–264. MR0414479

[9] Derman, C. (1954). A solution to a set of fundamental equations in Markov chains. *Proc. Amer. Math. Soc.* **5** 332–334. MR0060757

[10] DURRETT, R. (2005). *Probability*: *Theory and Examples*, 3rd ed. Brooks/Cole, Thomson Learning, Belmont, CA.

[11] HITCZENKO, P. and LOUCHARD, G. (2001). Distinctness of compositions of an integer: A probabilistic analysis. *Random Structures Algorithms* **19** 407–437. MR1871561

[12] HITCZENKO, P. and SAVAGE, C. D. (2004). On the multiplicity of parts in a random composition of a large integer. *SIAM J. Discrete Math*. **18** 418–435 (electronic). MR2112515

[13] KAC, M. (1947). On the notion of recurrence in discrete stochastic processes. *Bull. Amer. Math. Soc*. **53** 1002–1010. MR0022323

[14] KEILSON, J. (1979). *Markov Chain Models—Rarity and Exponentiality*. *Applied Mathematical Sciences* **28**. Springer, New York. MR0528293

[15] KLARNER, D. A. (1965). Some results concerning polyominoes. *Fibonacci Quart*. **3** 9–20. MR0186569

[16] KNOPFMACHER, A. and PRODINGER, H. (1998). On Carlitz compositions. *European J. Combin*. **19** 579–589. MR1637748

[17] KNUTH, D. E., MOTWANI, R. and PITTEL, B. (1990). Stable husbands. *Random Structures Algorithms* **1** 1–14. MR1068488

[18] LOUCHARD, G. (2005). Private communication.

[19] LOUCHARD, G. (1997). Probabilistic analysis of column-convex and directed diagonally-convex animals. *Random Structures Algorithms* **11** 151–178. MR1610265

[20] LOUCHARD, G. (1999). Probabilistic analysis of column-convex and directed diagonally-convex animals. II. Trajectories and shapes. *Random Structures Algorithms* **15** 1–23. MR1698406

[21] LOUCHARD, G. and PRODINGER, H. (2002). Probabilistic analysis of Carlitz compositions. *Discrete Math. Theor. Comput. Sci*. **5** 71–95 (electronic). MR1902415

[22] PRIVMAN, V. and FORGÁCS, G. (1987). Exact solution of the partially directed compact lattice animal model. *J. Phys. A* **20** L543–L547. MR0893296

[23] PRIVMAN, V. and ŠVRAKIĆ, N. M. (1988). Exact generating function for fully directed compact lattice animals. *Phys. Rev. Lett*. **60** 1107–1109. MR0932171

DEPARTMENT OF MATHEMATICS
OHIO STATE UNIVERSITY
COLUMBUS, OHIO 43210
USA
E-MAIL: bgp@math.ohio-state.edu