Adv. Appl. Probab. 1–35 (2024)
*doi:*10.1017/apr.2024.37

GAUSSIAN FLUCTUATIONS FOR THE TWO-URN MODEL

KONRAD KOLESKO,* *University of Wrocław*
ECATERINA SAVA-HUSS,** *Universität Innsbruck*

Abstract

We introduce a modification of the generalized Pólya urn model containing two urns, and we study the number of balls $B_j(n)$ of a given color $j \in \{1, \ldots, J\}$ added to the urns after n draws, where $J \in \mathbb{N}$. We provide sufficient conditions under which the random variables $(B_j(n))_{n\in\mathbb{N}}$, properly normalized and centered, converge weakly to a limiting random variable. The result reveals a similar trichotomy as in the classical case with one urn, one of the main differences being that in the scaling we encounter 1-periodic continuous functions. Another difference in our results compared to the classical urn models is that the phase transition of the second-order behavior occurs at $\sqrt{\rho}$ and not at $\rho/2$, where ρ is the dominant eigenvalue of the mean replacement matrix.

Keywords: Multitype branching processes; Crump–Mode–Jagers processes; Pólya urn; functional limit theorems; martingale CLT

2010 Mathematics Subject Classification: Primary 60J80; 60J85
Secondary 60B20; 60F05; 60F17

1. Introduction

A brief overview of classical Pólya urn models. For $J \in \mathbb{N}$, a J-dimensional Pólya urn process $(B(n))_{n\in\mathbb{N}}$ is an $\mathbb{N}^J$-valued stochastic process which represents the evolution of an urn containing balls of J different colors denoted by $1, 2, \ldots, J$. The initial composition of the urn can be specified by a J-dimensional vector $B(0)$ given by $B(0) = \big(B_1(0), \ldots, B_J(0)\big)$, the jth coordinate $B_j(0)$ of $B(0)$ representing the number of balls of color j present in the urn at the beginning of the process, i.e. at time 0. At each subsequent time step $n \geq 1$, we pick a ball uniformly at random, inspect its color, and put it back into the urn together with a random collection of additional balls, whose colors are given by $L^{(j)} = (L^{(j,1)}, \ldots, L^{(j,J)})$ if the selected ball has color j (which happens with probability proportional to the number of balls of color j already present in the urn). This rule for adding balls can be summed up by the so-called replacement matrix L, which in our case is a random matrix L defined as follows. For $J \in \mathbb{N}$, we write $[J] := \{1, \ldots, J\}$, and we consider a sequence $(L^{(j)})_{j\in[J]}$ of J independent $\mathbb{N}^J$-valued random (column) vectors. We denote by L the $J \times J$ random matrix with column vectors $L^{(j)}$, so $L = \big(L^{(1)}, L^{(2)}, \ldots, L^{(J)}\big)$, and by $a_{ij} = \mathbb{E}\left[L^{(j,i)}\right]$ the expectation of $L^{(j,i)}$, for all $i, j \in [J]$;

Received 20 January 2023; accepted 14 August 2024.
* Postal Address: Department of Mathematics, University of Wrocław, 50-383 Wrocław, Poland. Email address: Konrad.Kolesko@math.uni.wroc.pl
** Postal Address: Institut für Mathematik, Universität Innsbruck, 6020 Innsbruck, Austria. Email address: Ecaterina.Sava-Huss@uibk.ac.at

finally, we let $A=(a_{ij})_{i,j\in[J]}$, so that $\mathbb{E}L=A$. We then continue the process, each time taking an independent (of everything else) copy of the replacement matrix L. Note that the model described involves replacement, meaning that the selected ball is placed back into the urn after each draw. However, it is also possible to study a model without replacement by considering the replacement matrix $L-I$, where I is the $J\times J$ identity matrix. In this case, it might happen that some diagonal entries of $L-I$ are equal to -1, which means a ball is removed from the urn. The *urn process* is the sequence $(B(n))_{n\geq 1}$ of J-dimensional random vectors with nonnegative integer coordinates, and the jth coordinate $B_j(n)$ of $B(n)$ represents the number of balls of color j in the urn after the nth draw, for $j\in[J]$. We also define $B^\circ(n)$ to be the number of balls drawn up to time n; that is, $B^\circ_j(n)$ represents the number of balls of color j drawn up to the nth draw (in particular, $B^\circ_1(n)+\cdots+B^\circ_J(n)=n$). As one expects, the limit behavior of $(B(n))_{n\geq 1}$ and $(B^\circ(n))_{n\geq 1}$ depends on the distribution of the replacement matrix L, and in particular on the spectral properties of its mean value matrix A.

The literature on limit theorems for Pólya urn models is enormous and any attempt to give a complete survey here is hopeless, but we mention some relevant references and results in this direction. For additional results, the reader is referred to the cited articles and the references therein. In 1930, in his original article [10], Pólya investigates a two-color urn process with replacement matrix L being the identity. If L is a non-random, irreducible matrix with exclusively nonnegative entries, then it is well established that the sequence $B^\circ(n)/n$ converges almost surely to u as n goes to infinity, where u is the left eigenvector associated with ρ, the spectral radius of $A=\mathbb{E}L$. The coordinates of u are all nonnegative and normalized in such a way that they sum up to one; see [2, 5, 9, 12] for more details. The second-order behavior of the sequence $(B^\circ(n))_{n\in\mathbb{N}}$ depends on the second eigenvalue λ_2 (ordered by real parts) of A. If $\mathrm{Re}(\lambda_2)$, the real part of the second-largest eigenvalue, is less than or equal to $\rho/2$, then the fluctuations around the limit u are Gaussian (with a random variance). The magnitude of the fluctuations is of order $\sqrt{n}$ when $\mathrm{Re}(\lambda_2)<\rho/2$ and of order $\log(n)\sqrt{n}$ in the critical case $\mathrm{Re}(\lambda_2)=\rho/2$. Conversely, if $\mathrm{Re}(\lambda_2)>\rho/2$, then the fluctuations are non-Gaussian and of higher order. See Janson [5] for this trichotomy and [11] for an approach based on the spectral decomposition of a suitable finite-difference transition operator on polynomial functions.

Apart from these seminal results, the model of Pólya urns has been extended and more precise asymptotics are known. Several generalizations are considered in [5]. Another possible extension is to consider measure-valued Pólya processes; see the recent work [6] for second-order asymptotics of such processes for infinitely many colors and the literature cited there for additional results.

The model and our contribution. In the current paper we consider a modification of the Pólya urn model containing two urns marked U_1 and U_2. For a fixed $J\in\mathbb{N}$ representing the number of colors, we consider the random $J\times J$ matrix L as above, with independent column vectors $L^{(j)}$ and expectation matrix $A=\mathbb{E}L$. With these initial conditions, we define the $\mathbb{N}^J$-valued stochastic process $(B(n))_{n\in\mathbb{N}}$ as follows, with $B(n)=(B_1(n),\ldots,B_J(n))$. Suppose that at time 0 we have one ball of type (color) j_0 in urn U_1. We draw this ball from urn U_1, and we put into urn U_2 a collection of balls $L^{(j_0)}=(L^{(j_0,1)},\ldots,L^{(j_0,J)})$ (this notation means that for each $i\in[J]$, we put $L^{(j_0,i)}$ balls of type i into urn U_2). Thus we now have $\sum_{i=1}^J L^{(j_0,i)}$ balls in urn U_2. At the next step, we draw balls from urn U_2 uniformly at random, one after another, and for any ball of type j drawn we add an independent collection of balls with distribution $L^{(j)}$ into urn U_1. We continue until urn U_2 is empty, and then we exchange the roles of urns U_1 and U_2. We emphasize that, unlike in the Pólya urn model, in the two-urn model it is more convenient to consider drawing without replacement; that is, the ball drawn in each step is not returned to either urn, but only determines the future addition of balls. In particular, all

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

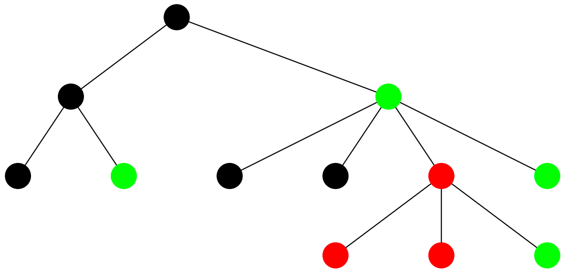

FIGURE 1. The model with two alternating urns and deterministic replacement matrix after $n = 4$ draws.

coefficients of L are nonnegative. For $j \in [J]$, by $B_j(n)$ we denote the total number of balls of type j that have been added to one of these two urns up to (and including) the nth draw.

Graphically, we can draw a random tree in order to visualize the step-by-step evacuation of one urn and the refilling of the other one, as follows: we color the nodes of the tree in colors $\{1, \ldots, J\}$; the content of urn U_1 represents the root of the tree colored with some fixed $j_0 \in \{1, \ldots, J\}$, i.e. the zero generation; after one draw of the node of type j_0, the content $L^{(j_0)}$ of U_2 represents the first generation. Then, after choosing balls (i.e. nodes of the tree at level one) uniformly at random without replacement and putting their offspring in the other urn U_1, we create step by step the second generation of the tree, and step by step we fill up U_1 again. Thus what we propose here is a more refined branching process where the transition from generation k to generation $k+1$ is considered after each member of generation k reproduces. If we visualize the process as a random tree that grows after each node is chosen, the quantity $B_j(n)$ represents the number of nodes of type j and $\sum_{j=1}^{J} B_j(n)$ represents the total number of nodes in the tree after n steps of the process, that is, after n balls have been drawn from U_1 and U_2. For better understanding, we illustrate this process in an example in Appendix A and Figure 1.

The main focus of the current work is to investigate first- and second-order asymptotics of $B_j(n)$ as $n \to \infty$, $j \in [J]$. It may happen that with positive probability $L^{(j,i)}$ vanishes for all $i \in [J]$. In such a case we do not add any new balls to the urn; we just remove the selected ball of type j. In particular, it can happen that after a finite number n_0 of steps, both urns are empty; in such a case we define $B(n) = B(n_0)$, for $n \geq n_0$. Since we are interested in the long-term behavior of the urn process, we restrict the analysis to a set where this does not happen, i.e. to the survival event $\mathcal{S} = \{|B(n)| \to \infty\}$.

We can also define the corresponding sequence $(B^\circ(n))_{n\geq 0}$ that represents the types of the balls drawn up to time n. While it is possible to ask about limit theorems for $B^\circ(n)$, the method developed in this paper for studying $B(n)$ is directly applicable to $B^\circ(n)$. Therefore, we focus our attention exclusively on the sequence $(B(n))_{n\in\mathbb{N}}$.

The approach we use to investigate $(B_j(n))_{n\geq 0}$, $j \in [J]$, is to embed it into a multi-type discrete-time branching process $(Z_n)_{n\in\mathbb{N}}$ with offspring distribution matrix L. A similar approach using the Athreya–Karlin embedding allowed Janson [5] to study $(B_j^\circ(n))_{n\geq 0}$ for the Pólya urn model. One difference between our model and the one in [5] is that, in the latter, the process is embedded into a multitype continuous-time Markov branching process, and an individual reproduces after an exponential clock rings. In our model, in the embedded branching process, an individual reproduces after deterministic time 1. The lattice nature of the model manifests itself in the second-order behavior of $(B_j(n))_{n\geq 0}$.

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

Assumptions. In this work we use the following assumptions:

(GW1) The matrix A has spectral radius $\rho > 1$.

(GW2) The matrix A is positively regular.

(GW3) We have $\mathbf{0} \neq \sum_{j=1}^{J} \mathrm{Cov}\left[L^{(j)}\right]$ and $\mathrm{Var}[L^{(i,j)}] < \infty$ for all $i, j \in [J]$.

(GW4) For every $i, j \in [J]$, the expectation $\mathbb{E}[L^{(i,j)} \log L^{(i,j)}]$ is finite.

In the third condition above, $\mathbf{0}$ is the $J \times J$ zero matrix, and for $j \in [J]$, $\mathrm{Cov}\left[L^{(j)}\right]$ represents the $J \times J$ covariance matrix of the vector $L^{(j)}$. If the matrix A is irreducible, the Perron–Frobenius theorem ensures that the dominant eigenvalue ρ of A is real, positive, and simple. If $\rho > 1$, this means that the multitype branching process with offspring distribution matrix L is supercritical, i.e. $\mathbb{P}(\mathcal{S}) > 0$. If u is the corresponding eigenvector for ρ, then clearly all the entries u_j are strictly greater than zero for any $j \in [J]$, and we assume that u is normalized in such a way that $\sum_{j\in[J]} \mathrm{u}_j = 1$. First-order asymptotics of $B_j(n)$, $j \in [J]$, are determined by ρ and the vector u.

Theorem 1. *Assume (GW1), (GW2), and (GW4) hold. Then for any $j \in [J]$ we have the following strong law of large numbers for the total number of balls of type j after n draws:*

$$\lim_{n\to\infty} \frac{B_j(n)}{n} = \rho \mathrm{u}_j, \qquad \mathbb{P}^{\mathcal{S}}\text{-almost surely.}$$

Thus, the first-order behavior of $B_j(n)$, $j \in [J]$ resembles the first-order behavior of $B_j^{\circ}(n)$ from the model with one urn [5]. In order to understand the second-order asymptotics of $B_j(n)$, we need full information on the spectral decomposition of the mean replacement matrix A. We denote by σ_A the spectrum of the matrix A and define $\sigma_A^1 = \{\lambda \in \sigma_A : |\lambda| > \sqrt{\rho}\}$, $\sigma_A^2 = \{\lambda \in \sigma_A : |\lambda| = \sqrt{\rho}\}$, and finally $\sigma_A^3 = \{\lambda \in \sigma_A : |\lambda| < \sqrt{\rho}\}$. Then we can write

$$\sigma_A = \sigma_A^1 \cup \sigma_A^2 \cup \sigma_A^3.$$

For a simple eigenvalue $\lambda \in \sigma_A$, we denote by u^λ and v^λ the corresponding left and right eigenvectors. We set

$$\gamma = \max\{|\lambda| : \lambda \in \sigma_A \setminus \{\rho\}\} \quad \text{and} \quad \Gamma = \{\lambda \in \sigma_A : |\lambda| = \gamma\},$$

so $\rho - \gamma$ is the *spectral gap* of A, and Γ is the set of eigenvalues of A where the spectral gap is achieved. For a complex number z we set $\log_\rho z = \frac{\log z}{\log \rho}$, where $z \mapsto \log z$ is the principal branch of the natural logarithm. Denote by $\{\mathrm{e}_j\}_{j\in[J]}$ the standard basis vectors in $\mathbb{R}^J$. For $j \in [J]$, consider the following row vector in $\mathbb{R}^{1\times J}$:

$$\mathrm{w}_j = \mathrm{e}_j^\top A - \rho \mathrm{u}_j \mathbb{1}, \tag{1}$$

where the ith entry is $\mathrm{w}_{ji} = a_{ji} - \rho \mathrm{u}_j = \mathbb{E}[L^{(i,j)}] - \rho \mathrm{u}_j$, for $1 \leq i \leq J$. In the above equation, $\mathbb{1}$ denotes the vector in $\mathbb{R}^{1\times J}$ with all entries equal to one. Our main result provides the second-order behavior of $(B_j(n))_{n\in\mathbb{N}}$.

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

Theorem 2. *Assume that (GW1)–(GW3) hold, all $\lambda \in \Gamma$ are simple, and there exists $\delta > 0$ such that $\mathbb{E}\big[\big(L^{(i,j)}\big)^{2+\delta}\big] < \infty$ for all $i, j \in [J]$. Then for any $j \in [J]$ the following trichotomy holds:*

(i) *If $\gamma > \sqrt{\rho}$, then for any $\lambda \in \Gamma$ there exist a 1-periodic, continuous function $f_\lambda : \mathbb{R} \to \mathbb{C}$ and random variables X, X_λ such that the following holds:*

$$B_j(n) = \rho \mathrm{u}_j \cdot n + \sum_{\lambda \in \Gamma} n^{\log_\rho \lambda} f_\lambda(\log_\rho n - X) X_\lambda + o_{\mathbb{P}}\big(n^{\log_\rho \gamma}\big).$$

(ii) *If $\gamma = \sqrt{\rho}$ and for some $\lambda \in \sigma_A^2$ and $i \in [J]$ we have $\mathrm{w}_j \mathrm{u}^\lambda \neq 0$ and $\mathrm{Var}[\mathrm{v}^\lambda L e_i] > 0$ for w_j defined as in (1), then there exist a 1-periodic, continuous function $\Psi : \mathbb{R} \to (0, \infty)$ and a random variable X such that, conditionally on $\mathcal{S}$, the following convergence in distribution holds:*

$$\frac{B_j(n) - \rho \mathrm{u}_j \cdot n}{\sqrt{n \log_\rho n}\ \Psi(\log_\rho n - X)} \xrightarrow{\mathrm{d}} \mathcal{N}(0, 1), \qquad \text{as } n \to \infty.$$

(iii) *If $\gamma < \sqrt{\rho}$ and for some $\lambda \in \sigma_A^3$ and some $i \in [J]$ we have $\mathrm{w}_j \mathrm{u}^\lambda \neq 0$ and $\mathrm{Var}[\mathrm{v}^\lambda L e_i] > 0$ with w_j defined as in (1), then there exist a 1-periodic, continuous function $\Psi : \mathbb{R} \to (0, \infty)$ and a random variable X such that, conditionally on $\mathcal{S}$, the following convergence in distribution holds:*

$$\frac{B_j(n) - \rho \mathrm{u}_j \cdot n}{\sqrt{n}\ \Psi(\log_\rho n - X)} \xrightarrow{\mathrm{d}} \mathcal{N}(0, 1), \qquad \text{as } n \to \infty.$$

A more general and quantified result where the periodic functions are explicitly defined is provided in Theorem 4.

Note that the result above slightly differs from that of the one-urn model, where the functions Ψ and f_λ are actually constants. What might be even more surprising is that, in our model, the phase transition occurs at $\sqrt{\rho}$ rather than at $\rho/2$ as observed in the Pólya urn model. The heuristic explanation is as follows: the growth in mean of the corresponding continuous-time branching process is driven by the semigroup e^{tA}. In particular, the leading asymptotic is $e^{t\rho}$ and the next order is $|e^{t\lambda_2}| = e^{t\mathrm{Re}\lambda_2}$. We anticipate observing Gaussian fluctuations in the branching process at the scale $\sqrt{e^{t\rho}}$ (with possible polynomial corrections), providing a natural threshold for $\mathrm{Re}\lambda_2$ in relation to $\rho/2$. On the other hand, the two-urn model is embedded into a discrete-time branching process whose growth in the mean is driven by the semigroup A^n (or $(I + A)^n$ in the model with no replacement). Thus, the leading term is at scale ρ^n and the subleading term is at scale $|\lambda_2|^n$. As before, we expect to observe Gaussian fluctuations at scale $\sqrt{\rho^n}$, which induce natural distinctions depending on the relative locations of $|\lambda_2|$ and $\sqrt{\rho}$.

Structure of the paper. In Section 2 we introduce multitype branching processes $(Z_n)_{n\in\mathbb{N}}$ and Crump–Mode–Jagers processes $(Z_n^\Phi)_{n\in\mathbb{N}}$ counted with a characteristic $\Phi : \mathbb{Z} \to \mathbb{R}^{1\times J}$. Then in Section 3 we show how to relate our model $(B_j(n))$, with two interacting urns, to a branching process $(Z_n^{\Phi^j})_{n\in\mathbb{N}}$ counted with a characteristic Φ^j. By applying [8, Proposition 4.1] to $(Z_n^{\Phi^j})_{n\in\mathbb{N}}$, we then obtain first-order asymptotics of $(B_j(n))_{n\in\mathbb{N}}$ for any $j \in [J]$ (Theorem 1). The main result is proved in Section 4. We conclude with Appendix A, where the model with two interacting urns is illustrated by an example with deterministic replacement matrix, and Appendix B, where higher-moment estimates for (Z_n^Φ) are given for general characteristics Φ.

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

2. Preliminaries

For the rest of the paper, $(\Omega, \mathcal{A}, \mathbb{P})$ is a probability space on which all the random variables and processes we consider are defined. We write $\xrightarrow{a.s.}$ for almost sure convergence, $\xrightarrow{\mathbb{P}}$ for convergence in probability, $\xrightarrow{\mathrm{d}}$ for convergence in distribution, and $\xrightarrow{\mathrm{st}}$ for stable convergence (cf. [1] for the definition and properties). We also use $\mathbb{P}^{\mathcal{S}}$ to denote the conditional probability $\mathbb{P}[\,\cdot\,|\mathcal{S}]$, and the corresponding convergences are denoted by $\xrightarrow{\mathbb{P}^{\mathcal{S}}}$, $\xrightarrow{\mathrm{d},\mathcal{S}}$, $\xrightarrow{\mathrm{st},\mathcal{S}}$. We use the notation $\mathbb{N} = \{1, 2, \ldots, \}$ and $\mathbb{N}_0 = \{0, 1, 2, \ldots, \}$.

Stochastic processes. Our convergence results for stochastic processes use the usual space $\mathcal{D}$ of right-continuous functions with left-hand limits, always equipped with the Skorokhod $\mathbf{J}_1$ topology. For a finite-dimensional vector space E and any interval $J \subseteq [-\infty, \infty]$, we denote by $\mathcal{D}(J) = \mathcal{D}(J, E)$ the space of all right-continuous functions from J to E with left-hand limits.

For an $n \times m$ matrix $A = (a_{ij})_{i,j}$ with $m, n \in \mathbb{N}$, the Hilbert–Schmidt norm of A, also called the Frobenius norm, is defined as

$$\|A\|_{HS} = \Big(\sum_{i=1}^{n}\sum_{j=1}^{m} |a_{ij}|^2\Big)^{1/2}.$$

Since for any vector its Hilbert–Schmidt norm coincides with its Euclidean norm, for the rest of the paper we write only $\|\cdot\|$ instead of $\|\cdot\|_{HS}$.

Ulam–Harris tree $\mathcal{U}_\infty$. An Ulam–Harris tree $\mathcal{U}_\infty$ is an infinite rooted tree with vertex set $V_\infty = \bigcup_{n\in\mathbb{N}_0} \mathbb{N}^n$, the set of all finite strings or words $i_1 \cdots i_n$ of positive integers over n letters, including the empty word $\varnothing$ (which we take to be the root), and with an edge joining $i_1 \cdots i_n$ and $i_1 \cdots i_{n+1}$ for any $n \in \mathbb{N}_0$ and any $i_1, \cdots, i_{n+1} \in \mathbb{N}$. Thus every vertex $v = i_1 \cdots i_n$ has outdegree ∞, and the children of v are the words $v1, v2, \ldots$. We let them have this order so that $\mathcal{U}_\infty$ becomes an infinite ordered rooted tree. We will identify $\mathcal{U}_\infty$ with its vertex set V_∞, when there is no risk of confusion in doing so. For vertices $v = i_1 \cdots i_n$ we also write $v = (i_1, \ldots, i_n)$, and if $u = (j_1, \ldots, j_m)$ we write uv for the concatenation of the words u and v; that is, $uv = (j_1, \ldots, j_m, i_1, \ldots, i_n)$. The parent of $i_1 \cdots i_n$ is $i_1 \cdots i_{n-1}$. Finally, for $u \in \mathcal{U}_\infty$, we use the notation $|u| = n$ to mean $u \in \mathbb{N}^n$ (i.e. u is a word of length (or height) n; in other words, it is at distance n from the root $\varnothing$). For any $u \in V_\infty$, by $\mathbb{T}_u$ we mean the subtree of $\mathcal{U}_\infty$ rooted at u, that is, u together with all infinite paths going away from u, and for $u, v \in \mathcal{U}_\infty$ we denote by $d(u, v)$ their graph distance. For trees rooted at $\varnothing$, we omit the root and we write only $\mathbb{T}$. For $J \in \mathbb{N}$, a J-type tree is a pair $(\mathbb{T}, \mathrm{t})$ where $\mathbb{T}$ is a rooted subtree of $\mathcal{U}_\infty$ and $\mathrm{t} : \mathbb{T} \to \{1, \ldots, J\}$ is a function defined on the vertices of $\mathbb{T}$ that returns for each vertex v its type $\mathrm{t}(v)$.

Multitype branching processes. Consider the random $J \times J$ matrix L with independent column vectors $L^{(j)}$, for $1 \le j \le J$, as in the introduction. Multitype Galton–Watson trees are random variables taking values in the set of J-type trees $(\mathbb{T}, \mathrm{t})$, where the type function t is random and defined in terms of the matrix L. Let $(L(u))_{u\in\mathcal{U}_\infty}$ be a family of independent and identically distributed (i.i.d.) copies of L indexed over the vertices of $\mathcal{U}_\infty$. For any $i \in [J]$, we define recursively the random labeled tree $\mathbb{T}^i$ rooted at $\varnothing$, with the associated type function $\mathrm{t} = \mathrm{t}^i : \mathbb{T}^i \to \{1, \ldots, J\}$, as follows: $\varnothing \in \mathbb{T}^i$ and $\mathrm{t}(\varnothing) = i$. Now suppose that $u = j_1 \ldots j_m \in \mathbb{T}^i$ with $\mathrm{t}(u) = j$, for some $j \in [J]$. Then

$$j_1 \ldots j_m k \in \mathbb{T}^i \quad \text{if and only if} \quad k \le L^{(j,1)}(u) + \cdots + L^{(j,J)}(u),$$

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

and we set $\mathrm{t}(uk)=\ell$ whenever

$$L^{(j,1)}(u)+\cdots+L^{(j,\ell-1)}(u)<k\le L^{(j,1)}(u)+\cdots+L^{(j,\ell)}(u).$$

The multitype branching process $Z_n=(Z_n^1,\ldots,Z_n^J)$ associated with the pair $(\mathbb{T}^{i_0},\mathrm{t})$, and starting from a single particle (or individual) of type $i_0\in[J]$ at the root $\varnothing$ (i.e., $\mathrm{t}(\varnothing)=i_0$) is defined as follows: $Z_0=\mathrm{e}_{i_0}$, and for $n\ge 1$,

$$Z_n^i=\#\{u\in\mathbb{T}^{i_0}:|u|=n \text{ and } \mathrm{t}(u)=i\},\qquad \text{for } i\in[J],$$

so Z_n^i represents the number of individuals of type i in the nth generation, or the number of vertices $u\in\mathbb{T}^{i_0}$ with $|u|=n$ and $\mathrm{t}(u)=i$. The main results of [8] that we use in the current paper hold under the assumptions (GW1)–(GW3) on $(Z_n)_{n\in\mathbb{N}}$, which we suppose to hold here as well. In particular, $(Z_n)_{n\in\mathbb{N}}$ is a supercritical branching process.

Spectral decomposition of A. Recall the decomposition of the spectrum σ_A of the matrix A as $\sigma_A=\sigma_A^1\cup\sigma_A^2\cup\sigma_A^3$. From the Jordan–Chevalley decomposition of A (which is unique up to the order of the Jordan blocks) we infer the existence of projections $(\pi_\lambda)_{\lambda\in\sigma_A}$ that commute with A and satisfy $\sum_{\lambda\in\sigma_A}\pi_\lambda=I$ and

$$A\pi_\lambda=\pi_\lambda A=\lambda\pi_\lambda+N_\lambda,$$

where $N_\lambda=\pi_\lambda N_\lambda=N_\lambda\pi_\lambda$ is a nilpotent matrix. Moreover, for any $\lambda_1,\lambda_2\in\sigma_A$ it holds that $\pi_{\lambda_1}\pi_{\lambda_2}=\pi_{\lambda_1}\mathbf{1}_{\{\lambda_1=\lambda_2\}}$. If $\lambda\in\sigma_A$ is a simple eigenvalue of A and u^λ, v^λ are the corresponding left and right eigenvectors, normalized in such a way that $\mathrm{v}^\lambda\mathrm{u}^\lambda=1$, then $\pi_\lambda=\mathrm{u}^\lambda\mathrm{v}^\lambda$. If we write $N=\sum_{\lambda\in\sigma_A}N_\lambda$, then N is also a nilpotent matrix and we have $N\pi_\lambda=N_\lambda$. Thus A can be decomposed into its semisimple part $D=\sum_{\lambda\in\sigma_A}\lambda\pi_\lambda$ and a nilpotent part N, as $A=D+N$.

For any $\lambda\in\sigma_A$, we denote by $d_\lambda\ge 0$ the integer such that $N_\lambda^{d_\lambda}\ne 0$ but $N_\lambda^{d_\lambda+1}=0$ (hence $d_\lambda+1$ is at most the multiplicity of λ). So $d_\lambda=0$ if and only if $N_\lambda=0$, and this happens for all λ if and only if A is diagonalizable (that is, A has a complete set of J independent eigenvectors). Since ρ is a simple eigenvalue, we have $N_\rho=0$ and $d_\rho=0$, and $\pi_\rho=\mathrm{uv}$. We set

$$\pi^{(1)}=\sum_{\lambda\in\sigma_A^1}\pi_\lambda,\qquad \pi^{(2)}=\sum_{\lambda\in\sigma_A^2}\pi_\lambda,\qquad \pi^{(3)}=\sum_{\lambda\in\sigma_A^3}\pi_\lambda,$$

and for $i=1,2$, we define

$$A_i=A\pi^{(i)}+\big(I-\pi^{(i)}\big).$$

The process $\big(W_n^{(i)}\big)_{n\in\mathbb{N}}$ defined by

$$W_n^{(i)}=A_i^{-n}\pi^{(i)}Z_n$$

is a $\mathcal{A}_n$-martingale, where $(\mathcal{A}_n)_{n\ge 0}$ is the filtration $\mathcal{A}_n=\sigma(\{L(u):|u|\le n\})$. According to [8, Lemma 2.2], $W_n^{(1)}$ converges in $\mathcal{L}^2(\Omega,\mathcal{A},\mathbb{P})$ to a random variable $W^{(1)}$ whose expectation is $\mathbb{E}W^{(1)}=\pi^{(1)}\mathrm{e}_{i_0}$. In particular, we have the convergence

$$\rho^{-n}Z_n\to\pi_\rho W^{(1)}=W\mathrm{u},\qquad \text{as } n\to\infty,$$

in $\mathcal{L}^2$, and the random variable W is given by $W=\mathrm{v}\cdot W^{(1)}$ with $\mathbb{E}W=\mathrm{v}^{i_0}>0$. The classical Kesten–Stigum theorem [7] states that under (GW1), (GW2), and (GW4) the convergence above holds almost surely. For the rest of the paper, when we use the random variable W, we always mean the limit random variable from the Kesten–Stigum theorem.

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

2.1. Branching processes counted with a characteristic

We recall that a characteristic of dimension one is a function $\Phi : \mathbb{Z} \to \mathbb{R}^{1\times J}$, so that for each $k \in \mathbb{Z}$, $\Phi(k)$ is an $\mathbb{R}^{1\times J}$-valued random variable defined on the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$ where the Galton–Watson process $(Z_n)_{n\in\mathbb{N}}$ and its genealogical tree $\mathbb{T}$ are defined. A *deterministic characteristic* is just a fixed function $\Phi : \mathbb{Z} \to \mathbb{R}^{1\times J}$. For a random function $\Phi : \mathbb{Z} \to \mathbb{R}^{1\times J}$ and the multitype Galton–Watson tree $\mathbb{T}$, the process $(Z_n^{\Phi})_{n\in\mathbb{N}}$ which for any $n \in \mathbb{N}$ is defined as

$$Z_n^{\Phi} = \sum_{u\in\mathbb{T}} \Phi_u(n - |u|)\mathrm{e}_{\mathrm{t}(u)}$$

is called the *multitype Crump–Mode–Jagers (CMJ) process counted with characteristic* Φ, or simply the *branching process counted with characteristic* Φ, where $(\Phi_u)_{u\in\mathcal{U}_\infty}$ is an i.i.d. copy of Φ. First- and second-order asymptotics for $(Z_n^{\Phi})_{n\in\mathbb{N}}$, under mild assumptions on Φ, are considered in [8, Proposition 4.1] and in [8, Theorem 3.5], respectively. We use these two results below for a particular choice of the characteristic Φ, and show how the branching process with this particular choice of characteristic can be related to the two-urn model with alternating urns U_1 and U_2.

The choice of the characteristic. Let $U \sim \mathsf{Unif}[0, 1]$ be a uniform random variable taking values in [0,1] and defined on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$. For every threshold $x \in [0, 1)$ we define the characteristic $\Phi_x^{\mathrm{t}} : \mathbb{N}_0 \to \mathbb{R}^{1\times J}$ by

$$\Phi_x^{\mathrm{t}}(k)\mathrm{e}_i = \mathbf{1}_{\{k\geq 1\}} + \mathbf{1}_{\{k=0\}}\mathbf{1}_{\{U\leq x\}}, \quad \text{for } 1 \leq i \leq J, \tag{2}$$

or, in a simplified way, for $1 \leq i \leq J$,

$$\Phi_x^{\mathrm{t}}(k)\mathrm{e}_i = \begin{cases} 1, & \text{for } k \geq 1, \\ 1 \text{ with probability } \mathbb{P}(U \leq x) = x, & \text{for } k = 0, \\ 0 \text{ with probability } \mathbb{P}(U > x) = 1 - x, & \text{for } k = 0. \end{cases}$$

Similarly, for $j \in [J]$, we define $\Phi_x^j : \mathbb{N}_0 \to \mathbb{R}^{1\times J}$ by

$$\Phi_x^j(k)\mathrm{e}_i = \mathbf{1}_{\{k\geq 1\}} \underbrace{\langle \mathrm{e}_j, L\mathrm{e}_i\rangle}_{=L^{(i,j)}} + \mathbf{1}_{\{k=0\}}\mathbf{1}_{\{U\leq x\}}\langle \mathrm{e}_j, L\mathrm{e}_i\rangle, \quad \text{for } 1 \leq i \leq J, \tag{3}$$

so $\Phi_x^j(k)\mathrm{e}_i = \Phi_x^{\mathrm{t}}(k)\mathrm{e}_i L^{(i,j)}$. For the uniform random variable U on [0,1], let $(U_u)_{u\in\mathcal{U}_\infty}$ be an i.i.d. copy of U. For $u \in \mathcal{U}_\infty$, let $\Phi_{x,u}^{\mathrm{t}}$ (respectively $\Phi_{x,u}^j$) be defined through (2) (respectively (3)) with U, L being replaced by $U_u,\ L(u)$. Then the family $\big((L(u), \Phi_{x,u}^{\mathrm{t}}, \Phi_{x,u}^j)\big)_{u\in\mathcal{U}_\infty}$ is an i.i.d. collection of copies of $(L, \Phi_x^{\mathrm{t}}, \Phi_x^j)$, for $j \in [J]$.

For the characteristic Φ_x^{t} from (2), threshold $x \in [0, 1)$, and multitype Galton–Watson tree $\mathbb{T}$, the process $\big(Z_n^{\Phi_x^{\mathrm{t}}}\big)_{n\geq 0}$ counts the total number of individuals $u \in \mathcal{U}_\infty$ born before time n

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

(including the root), and those born at time n with $U_u \le x$, since we have

$$Z_n^{\Phi_x^{\mathrm{t}}} = \sum_{u\in\mathbb{T}} \Phi_{x,u}^{\mathrm{t}}(n-|u|)\mathrm{e}_{\mathrm{t}(u)} = \sum_{k=0}^{n-1} \sum_{u\in\mathbb{T};|u|=k} 1 + \sum_{u\in\mathbb{T};\ |u|=n} \mathbf{1}_{\{U_u\le x\}}$$

$$= \sum_{k=0}^{n-1} |Z_k| + \sum_{u\in\mathbb{T};\ |u|=n} \mathbf{1}_{\{U_u\le x\}},$$

where $|Z_n| = \sum_{j=1}^{J} Z_n^j$ represents the size of the nth generation of the Galton–Watson process. On the other hand, $L\mathrm{e}_i = L^{(i)}$ represents the random collection of individuals born from an individual of type i, while $L^{(i,j)} = \langle \mathrm{e}_j, L\mathrm{e}_i\rangle$ represents the random number of offspring of type j of an individual of type i for $i, j \in [J]$. Therefore $Z_n^{\Phi_x^j}$ counts the number of individuals of type j born up to time n, and those of type j born at time $n+1$ but with threshold $\le x$, so $Z_n^{\Phi_x^j}$ can be written as

$$Z_n^{\Phi_x^j} = \sum_{k=0}^{n-1} \sum_{u\in\mathbb{T};|u|=k} \langle \mathrm{e}_j, L(u)\mathrm{e}_{\mathrm{t}(u)}\rangle + \sum_{u\in\mathbb{T};|u|=n+1,\mathrm{t}(u)=j} \mathbf{1}_{\{U_u\le x\}},$$

since $Z_{k+1}^j = \sum_{u\in\mathbb{T};|u|=k}\langle \mathrm{e}_j, L(u)\mathrm{e}_{\mathrm{t}(u)}\rangle$ represents the number of offspring of type j in the $(k+1)$th generation and $|\{u \in \mathbb{T};\ |u| = n+1,\ \mathrm{t}(u) = j\}| = Z_{n+1}^j$. Summing up over all $j \in [J]$ gives

$$Z_{n+1}^{\Phi_x^{\mathrm{t}}} - 1 = \sum_{j=1}^{J} Z_n^{\Phi_x^j},$$

and an application of [8, Proposition 4.1] to $Z_n^{\Phi_x^{\mathrm{t}}}$ and $Z_n^{\Phi_x^j}$, for $j \in [J]$, yields the law of large numbers.

Proposition 1. *Under the assumptions (GW1), (GW2), and (GW4), for any threshold $x \in [0, 1)$ and characteristic Φ_x^{t} (respectively Φ_x^j, $j \in [J]$) defined in (2) (respectively (3)) we have the following:*

$$\lim_{n\to\infty} \frac{Z_n^{\Phi_x^{\mathrm{t}}}}{\rho^n} = \left(\frac{1}{\rho-1} + x\right)W, \quad \mathbb{P}^{\mathcal{S}}\text{-almost surely},$$

$$\lim_{n\to\infty} \frac{Z_n^{\Phi_x^j}}{\rho^n} = \left(\frac{1}{\rho-1} + x\right)W\rho\mathrm{u}_j, \quad \mathbb{P}^{\mathcal{S}}\text{-almost surely}.$$

Proof. Since for any fixed $x \in [0, 1)$ the random variables $(\mathbf{1}_{\{U_u\le x\}})_{u\in\mathcal{U}_\infty}$ are i.i.d. and Bernoulli-distributed as $\mathsf{Bern}(x)$, in view of the strong law of large numbers we obtain

$$\lim_{n\to\infty} \frac{1}{|Z_n|} \sum_{u\in\mathbb{T};\ |u|=n} \mathbf{1}_{\{U_u\le x\}} = x, \quad \mathbb{P}^{\mathcal{S}}\text{-almost surely},$$

and similarly

$$\lim_{n\to\infty} \frac{1}{Z_n^j} \sum_{|u|=n,\mathrm{t}(u)=j} \mathbf{1}_{\{U_u\le x\}} = x, \quad \mathbb{P}^{\mathcal{S}}\text{-almost surely}.$$

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

For the deterministic characteristic $\mathbf{1}_{\{k\geq 1\}}$ and the corresponding branching process counted with this characteristic, by applying [8, Proposition 4.1(i)] we get

$$\frac{Z_n^{\Phi_x^{\mathrm{t}}}}{\rho^n} = \frac{1}{\rho}\cdot \overbrace{\frac{1}{\rho^{n-1}}\sum_{k\geq 0}^{n-1}\sum_{|u|=k} 1}^{\xrightarrow{a.s.} W\sum_{k\geq 0}\rho^{-k}} + \overbrace{\frac{1}{\rho^n}\sum_{|u|=n}\mathbf{1}_{\{U_u\leq x\}}}^{\xrightarrow{a.s.} xW}$$
$$\xrightarrow{a.s.} \Big(\frac{1}{\rho}\sum_{k\geq 0}\frac{1}{\rho^k}+x\Big)W = \Big(\frac{1}{\rho-1}+x\Big)W, \qquad \text{as } n\to\infty,$$

and this shows the first part of the claim. For the second one, from the Kesten–Stigum theorem we know that $\frac{Z_n^j}{\rho^n}\xrightarrow{a.s.} W\mathrm{u}_j$ as $n\to\infty$ for any $j\in[J]$, and for the characteristic Φ_x^j, since we have

$$Z_n^{\Phi_x^j} = \sum_{k=0}^{n} Z_k^j + \sum_{u\in\mathbb{T};\ |u|=n+1,\mathrm{t}(u)=j}\mathbf{1}_{\{U_u\leq x\}},$$

we obtain

$$\frac{Z_n^{\Phi_x^j}}{\rho^n} = \overbrace{\frac{1}{\rho^n}\sum_{k=1}^{n} Z_k^j}^{\xrightarrow{a.s.} W\sum_{k\geq 0}\rho^{-k}\mathrm{u}_j} + \overbrace{\rho\frac{1}{\rho^{n+1}}\sum_{|u|=n+1,\mathrm{t}(u)=j}\mathbf{1}_{\{U_u\leq x\}}}^{\xrightarrow{a.s.} \rho x W\mathrm{u}_j}$$
$$\xrightarrow{a.s.} \Big(\frac{1}{\rho}\sum_{k\geq 0}\frac{1}{\rho^k}+x\Big)W\rho\mathrm{u}_j = \Big(\frac{1}{\rho-1}+x\Big)W\rho\mathrm{u}_j, \qquad as\ n\to\infty,$$

and the proof is completed. □

Following the notation from [8], for every characteristic $\Phi:\mathbb{Z}\to\mathbb{R}^{1\times J}$ we define the two vectors $\mathrm{x}_i(\Phi)=\sum_{k\in\mathbb{N}}\mathbb{E}[\Phi(k)]\pi^{(i)}A_i^{-k}$, for $i=1,2$. In particular, since $\mathbb{E}[\Phi_x^{\mathrm{t}}(0)]=x1$, we have

$$\mathrm{x}_i\left(\Phi_x^{\mathrm{t}}\right) = x1\pi^{(i)} + 1\pi^{(i)}\sum_{k=1}^{\infty}A_i^{-k} = 1\pi^{(i)}\left(xI+\sum_{k=1}^{\infty}A_i^{-k}\right)$$
$$= 1\pi^{(i)}\left(x+(A_i-I)^{-1}\right), \qquad \text{for } i=1,2. \tag{4}$$

For any random characteristic $\Phi:\mathbb{Z}\to\mathbb{R}^{1\times J}$ that satisfies the assumptions of [8, Theorem 3.5]—i.e. for which (GW1)–(GW3) hold, $\sum_{k\in\mathbb{Z}}\big\|\mathbb{E}[\Phi(k)]\big\|(\rho^{-k}+\vartheta^{-k})<\infty$ for some $\vartheta<\sqrt{\rho}$, and finally $\sum_{k\in\mathbb{Z}}\|\mathrm{Var}[\Phi(k)]\|\rho^{-k}<\infty$—we set

$$F_n^{\Phi} = \mathrm{x}_1(\Phi)A_1^nW^{(1)} + \mathrm{x}_2(\Phi)A_2^nZ_0. \tag{5}$$

Recall that the constants σ_ℓ^2, for $\ell=0,\ldots,J-1$, are defined as

$$\sigma_\ell^2=\sigma_\ell^2(\Phi)=\frac{\rho^{-\ell}}{(2\ell+1)(\ell!)^2}\sum_{\lambda\in\sigma_A^2}\mathrm{Var}\Big[\mathrm{x}_2(\Phi)\pi_\lambda(A-\lambda I)^\ell L\Big]\mathrm{u}. \tag{6}$$

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

Let ℓ be the maximal integer such that $\sigma_\ell(\Phi) > 0$, and set $\ell = -\frac{1}{2}$ if there is no such integer. Then Theorem 3.5 of [8] states that, for a standard normal variable $\mathcal{N}(0, 1)$ independent of W, the following stable convergence holds:

$$\frac{Z_n^\Phi - F_n^\Phi}{n^{\ell+\frac{1}{2}}\rho^{n/2}\sqrt{W}} \xrightarrow{\text{st},\mathcal{S}} G^\Phi, \quad \text{as } n \to \infty, \tag{7}$$

where $G^\Phi = \sigma_\ell(\Phi)\mathcal{N}(0, 1)$ if not all σ_ℓ are zero, and $G^\Phi = \sigma(\Phi)\mathcal{N}(0, 1)$ otherwise, while $\sigma(\Phi)$ is defined by

$$\sigma^2(\Phi) = \sum_{k\in\mathbb{Z}} \rho^{-k}\mathrm{Var}\left[\Phi(k) + \Psi^\Phi(k)\right]\mathrm{u}, \tag{8}$$

and Ψ^Φ is the centered characteristic given by

$$\Psi^\Phi(k) = \sum_{\ell\in\mathbb{Z}} \mathbb{E}\Phi(k-\ell-1)A^\ell\mathsf{P}(k, \ell)(L-A).$$

Above, the matrices $\mathsf{P}(k, \ell)$, for $k, \ell \in \mathbb{Z}$ are defined as

$$\mathsf{P}(k, \ell) = \begin{cases} -\pi^{(1)}\mathbf{1}_{\{\ell<0\}} + \pi^{(2)}\mathbf{1}_{\{\ell\geq 0\}} + \pi^{(3)}\mathbf{1}_{\{\ell\geq 0\}}, & \text{if } k \leq 0, \\ -\pi^{(1)}\mathbf{1}_{\{\ell<0\}} - \pi^{(2)}\mathbf{1}_{\{\ell<0\}} + \pi^{(3)}\mathbf{1}_{\{\ell\geq 0\}}, & \text{if } k > 0. \end{cases}$$

3. The embedding of the urn model into the branching process

Notation. We slightly abuse notation and write Φ^{t} (respectively Φ^j, $j \in [J]$) for the whole family $\{\Phi_x^{\mathsf{t}}\}_{x\in[0,1)}$ (respectively $\{\Phi_x^j\}_{x\in[0,1)}$) of characteristics indexed over the threshold $x \in [0, 1)$. We denote by $\mathcal{C}$ the set of characteristics Φ which are linear combinations of Φ^{t} and Φ^j, for $j \in [J]$. Again by abuse of notation, by $\Phi \in \mathcal{C}$ we actually refer to the whole family of characteristics $\{\Phi_x\}_{x\in[0,1)}$; that is,

$$\Phi = \left\{\Phi_x \in \mathcal{C} : x \in [0, 1)\right\} = \left\{a\Phi_x^{\mathsf{t}} + b\Phi_x^j : a, b \in \mathbb{R},\ j \in [J],\ x \in [0, 1)\right\}.$$

Extension of $x \in [0, 1)$ to $x \in \mathbb{R}$. Instead of working with thresholds $x \in [0, 1)$ and corresponding characteristics Φ_x, we can extend the domain of x to the whole of $\mathbb{R}$ as follows. For any $\Phi \in \mathcal{C}$ and any $x \in \mathbb{R}$ we define

$$\Phi_x(k) = \Phi_{\{x\}}(k + \lfloor x \rfloor).$$

The corresponding CMJ process $(Z_n^{\Phi_x})_{n\in\mathbb{N}}$ satisfies $Z_n^{\Phi_x} = Z_0^{\Phi_{x+n}}$ for every $n \in \mathbb{N}$ and $x \in \mathbb{R}$, and finally we define

$$\mathcal{Z}^\Phi(x) = Z_0^{\Phi_x},$$

and similarly $\mathcal{F}^\Phi(x) = F_0^{\Phi_x}$. For any $x \in \mathbb{R}$, (7) yields the existence of a Gaussian process $\{\mathcal{G}^\Phi(x); x \in \mathbb{R}\}$ such that the following convergence holds:

$$\frac{\mathcal{Z}^\Phi(x+n) - \mathcal{F}^\Phi(n+x)}{n^{\ell+\frac{1}{2}}\rho^{n/2}\sqrt{W}} \xrightarrow{\text{st},\mathcal{S}} \mathcal{G}^\Phi(x), \quad \text{as } n \to \infty. \tag{9}$$

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

The Cramér–Wold device implies that, in fact, the convergence holds for finite-dimensional distributions and the limiting process $\mathcal{G}^{\Phi}$ is jointly Gaussian with $\mathcal{G}^{\Phi}(x) \stackrel{d}{=} \rho^{\lfloor x \rfloor/2}\sigma_{\ell}(\Phi_{\{x\}})\mathcal{N}(0, 1)$ or $\mathcal{G}^{\Phi}(x) \stackrel{d}{=} \rho^{\lfloor x \rfloor/2}\sigma(\Phi_{\{x\}})\mathcal{N}(0, 1)$ depending on the value of the constants σ_{ℓ} or σ, respectively. Furthermore, we write $\mathcal{Z}^{\mathrm{t}}, \mathcal{F}^{\mathrm{t}}, \mathcal{G}^{\mathrm{t}}$ (respectively $\mathcal{Z}^{j}, \mathcal{F}^{j}, \mathcal{G}^{j}$) for $\mathcal{Z}^{\Phi}, \mathcal{F}^{\Phi}, \mathcal{G}^{\Phi}$ if $\Phi = \Phi^{\mathrm{t}}$ (respectively $\Phi = \Phi^{j}$). Since, with probability one, all the random variables $(U_u)_{u\in\mathcal{U}_\infty}$ are different, the process $(\mathcal{Z}^{\mathrm{t}}(x))_{x\geq 0}$ at its jump point increases by 1. Therefore, the following stopping times are well defined: for $k \in \mathbb{N}$, define τ_k as

$$\tau_k = \inf\left\{x \geq 0 : \mathcal{Z}^{\mathrm{t}}(x) = k\right\}. \tag{10}$$

Remark 1. With the stopping times $(\tau_k)_{k\in\mathbb{N}}$ just introduced, we have

$$B_j(k) = \mathcal{Z}^{j}(\tau_k),$$

and this is exactly the number of balls of type j added to the two urns after k draws, for which we seek first- and second-order asymptotics. Our strategy is as follows: first we prove functional limit theorems for the processes $\{\mathcal{Z}^{\mathrm{t}}(x); x \in \mathbb{R}\}$ and $\{\mathcal{Z}^{j}(x); x \in \mathbb{R}\}$, and then we conclude the corresponding limit theorems for $B_j(k)$, for $j \in [J]$. We start with the description of the leading term in the asymptotics of $\mathcal{Z}^{\Phi}(x)$, for any characteristic $\Phi \in \mathcal{C}$.

Periodic functions. For any $\lambda \in \mathbb{C}$, we introduce the function

$$l_\lambda : [0, \infty) \to \mathbb{R} \quad \text{defined as} \quad l_\lambda(x) = (1 + (\lambda - 1)\{x\})\lambda^{-\{x\}}, \tag{11}$$

where $\lambda^t = e^{t \log \lambda}$ and $z \mapsto \log z$ is the principal branch of the logarithm. The function l_λ is continuous and 1-periodic, and it satisfies

$$\lambda^x l_\lambda(x) = \lambda^{\lfloor x \rfloor}(1 + (\lambda - 1)\{x\}).$$

Moreover, the mapping $x \mapsto \lambda^x l_\lambda(x)$ equals λ^x for integer x and is linear in between.

Proposition 2. *Assume (GW1), (GW2), and (GW4) hold, and for any $j \in [J]$ let $\Phi = a\Phi^{\mathrm{t}} + b\Phi^{j}$, with $a, b \in \mathbb{R}$. Then it holds that*

$$\lim_{x\to\infty} \frac{\mathcal{Z}^{\Phi}(x)}{\rho^x l_\rho(x)} = \lim_{x\to\infty} \frac{\mathcal{Z}^{\Phi}(x)}{\rho^{\lfloor x \rfloor}(1 + (\rho - 1)\{x\})} = (a + b\rho \mathrm{u}_j)\frac{1}{\rho - 1}W, \quad \mathbb{P}^{\mathcal{S}}\text{-almost surely}. \tag{12}$$

Proof. Because of the linearity of CMJ processes, for $x \in \mathbb{R}$ it holds that

$$\mathcal{Z}^{\Phi}(x) = a\mathcal{Z}^{\mathrm{t}}(x) + b\mathcal{Z}^{j}(x) \quad \text{for any } a, b \in \mathbb{R} \text{ and } j \in [J],$$

so it suffices to prove the $\mathbb{P}^{\mathcal{S}}$-almost sure convergence for $\mathcal{Z}^{\mathrm{t}}(x)$ and $\mathcal{Z}^{j}(x)$ separately, as $x \to \infty$.

Case 1: $x \in [0, 1)$. For $n \in \mathbb{N}$ and $x \in [0, 1)$, since $\mathcal{Z}^{\mathrm{t}}(x + n) = Z_n^{\Phi_x^{\mathrm{t}}}$, in view of Proposition 2.1 we get

$$\frac{\mathcal{Z}^{\mathrm{t}}(x + n)}{\rho^n(1 + (\rho - 1)x)} \to \frac{1}{\rho - 1}W, \quad \mathbb{P}^{\mathcal{S}}\text{-almost surely as } n \to \infty. \tag{13}$$

Case 2: $x \in [0, \infty)$. This case can be reduced to the previous one, where $x \in [0, 1)$, as follows. In view of Equation (13), for any $x \in [0, \infty)$, with $m = n + \lfloor x \rfloor$ we obtain

$$\frac{\mathcal{Z}^{\mathrm{t}}(x + n)}{\rho^{x+n}l_\rho(x)} = \frac{\mathcal{Z}^{\mathrm{t}}(\lfloor x \rfloor + n + \{x\})}{\rho^{\lfloor x \rfloor + n}(1 + (\rho - 1)\{x\})} = \frac{\mathcal{Z}^{\mathrm{t}}(\{x\} + m)}{\rho^m(1 + (\rho - 1)\{x\})} \xrightarrow{a.s.} \frac{1}{\rho - 1}W, \text{ as } m \to \infty.$$

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

In the last equation above, we used the fact that

$$\rho^{x+n} l_\rho(x) = \rho^{\lfloor x \rfloor + n}(1 + (\rho - 1)\{x\}) = \rho^{x+n} l_\rho(x+n).$$

We still have to prove that the above $\mathbb{P}^{\mathcal{S}}$-almost sure convergence holds for $x \to \infty$, that is, that

$$\lim_{x\to\infty} \frac{\mathcal{Z}^{\mathrm{t}}(x)}{\rho^x l_\rho(x)} = \frac{1}{\rho - 1} W, \quad \mathbb{P}^{\mathcal{S}}\text{-almost surely.}$$

Indeed, from any sequence tending to infinity we may choose a subsequence (x_n) such that $\{x_n\}$ converges to some $x_0 \in \mathbb{R}$. Then for any $\delta > 0$ and large n, in view of

$$\mathcal{Z}^{\mathrm{t}}(\lfloor x_n \rfloor + x_0 - \delta) = Z^{\Phi^{\mathrm{t}}_{\{\lfloor x_n \rfloor + x_0 - \delta\}}}_{\lfloor \lfloor x_n \rfloor + x_0 - \delta \rfloor} = Z^{\Phi^{\mathrm{t}}_{\{x_0 - \delta\}}}_{\lfloor \lfloor x_n \rfloor + x_0 - \delta \rfloor},$$

we have

$$\frac{\mathcal{Z}^{\mathrm{t}}(\lfloor x_n \rfloor + x_0 - \delta)}{\rho^{\lfloor x_n \rfloor + x_0 - \delta} l_\rho(x_0 - \delta)} \cdot \rho^{x_0 - \{x_n\} - \delta} \cdot \frac{l_\rho(x_0 - \delta)}{l_\rho(x_n)} \le \frac{\mathcal{Z}^{\mathrm{t}}(x_n)}{\rho^{x_n} l_\rho(x_n)}$$
$$\le \frac{\mathcal{Z}^{\mathrm{t}}(\lfloor x_n \rfloor + x_0 + \delta)}{\rho^{\lfloor x_n \rfloor + x_0 + \delta} l_\rho(x_0 + \delta)} \rho^{x_0 - \{x_n\} + \delta} \cdot \frac{l_\rho(x_0 + \delta)}{l_\rho(x_n)}.$$

Taking the limit first as n goes to infinity and then as δ goes to 0, we get the desired convergence, since $x \mapsto l_\rho(x)$ is uniformly continuous. The same argument can be used to show that for any $j \in [J]$ we have

$$\lim_{x\to\infty} \frac{\mathcal{Z}^j(x)}{\rho^x l_\rho(x)} = \frac{1}{\rho - 1} W \rho \mathrm{u}_j, \quad \mathbb{P}^{\mathcal{S}}\text{-almost surely,}$$

and this proves the claim. □

An immediate consequence of Proposition 2 is the following corollary.

Corollary 1. *Under the assumptions of Proposition 2, for $\Phi = \rho \mathrm{u}_j \Phi^{\mathrm{t}} - \Phi^j$, we have*

$$\lim_{x\to\infty} \frac{\mathcal{Z}^\Phi(x)}{\rho^x l_\rho(x)} = \lim_{x\to\infty} \frac{\mathcal{Z}^\Phi(x)}{\rho^{\lfloor x \rfloor}(1 + (\rho - 1)\{x\})} = 0, \quad \mathbb{P}^{\mathcal{S}}\text{-almost surely.}$$

Also, the strong law of large numbers for $(B_j(k))_{k\in\mathbb{N}}$ follows immediately from Proposition 2.

Proof of Theorem 1. Since τ_k goes to infinity as k does, we have

$$\lim_{k\to\infty} \frac{B_j(k)}{k} = \lim_{k\to\infty} \frac{\mathcal{Z}^j(\tau_k)}{\mathcal{Z}^{\mathrm{t}}(\tau_k)} = \lim_{k\to\infty} \frac{\mathcal{Z}^j(\tau_k)}{\rho^{\lfloor \tau_k \rfloor}(1 + (\rho - 1)\{\tau_k\})} \cdot \frac{\rho^{\lfloor \tau_k \rfloor}(1 + (\rho - 1)\{\tau_k\})}{\mathcal{Z}^{\mathrm{t}}(\tau_k)}$$
$$= \frac{1}{\rho - 1} W \rho \mathrm{u}_j \cdot \frac{1}{\frac{1}{\rho - 1} W} = \rho \mathrm{u}_j, \quad \mathbb{P}^{\mathcal{S}}\text{-almost surely,}$$

and this finishes the proof. □

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

4. Proof of the main result

The proof is completed in several steps:

- In Lemma 1 we investigate compositions of the fluctuations $\mathcal{F}^{\mathsf{t}}$ and $\mathcal{F}^{j}$.
- In Theorem 3 we prove weak convergence of the processes $\mathcal{X}^{\mathsf{t}} = \mathcal{Z}^{\mathsf{t}} - \mathcal{F}^{\mathsf{t}}$ and $\mathcal{X}^{j} = \mathcal{Z}^{j} - \mathcal{F}^{j}$ (rescaled appropriately) to Gaussian processes $\mathcal{G}^{\mathsf{t}}$ and $\mathcal{G}^{j}$ respectively, for any $j \in [J]$.
- Continuity and strict positivity of the variances of the limiting processes $\mathcal{G}^{\mathsf{t}}$ and $\mathcal{G}^{j}$ are analyzed in Proposition 4 and in Lemma 2.
- Localization of the stopping times τ_n is done in Proposition 5.
- Finally, the limit theorems for $B_j(n)$ are given in Proposition 6 and in Theorem 4.

4.1. Leading asymptotic terms

We start by describing the leading terms in the asymptotics of $\mathcal{Z}^{\mathsf{t}}$ and of $\mathcal{Z}^{j}$ for any $j \in [J]$. We recall first that for a characteristic $\Phi \in \mathcal{C}$, the leading term in the asymptotics of $\mathcal{Z}^{\Phi}$ is given by $\mathcal{F}^{\Phi}$; for simplicity of notation, for $x \in \mathbb{R}$ we write

$$\mathcal{X}^{\Phi}(x) = \mathcal{Z}^{\Phi}(x) - \mathcal{F}^{\Phi}(x).$$

In particular, for $\Phi = \Phi^{\mathsf{t}}$ and $\Phi = \Phi^{j}$ respectively, we write

$$\mathcal{X}^{\mathsf{t}}(x) = \mathcal{Z}^{\mathsf{t}}(x) - \mathcal{F}^{\mathsf{t}}(x) \quad \text{and} \quad \mathcal{X}^{j}(x) = \mathcal{Z}^{j}(x) - \mathcal{F}^{j}(x).$$

Lemma 1. *Assume (GW1)–(GW3) hold. Then for sufficiently large arguments the inverse function $\mathcal{F}^{\mathsf{inv}} = (\mathcal{F}^{\mathsf{t}})^{-1}$ is well defined, and for every $j \in [J]$ we have*

$$\lim_{t\to\infty} \sup_{s\geq 1} s^{-1} \left| \mathcal{F}^{j}\big(\mathcal{F}^{\mathsf{inv}}(t+s)\big) - \mathcal{F}^{j}\big(\mathcal{F}^{\mathsf{inv}}(t)\big) - \rho \mathrm{u}_j s \right| = 0, \quad \mathbb{P}^{\mathcal{S}}\text{-almost surely.} \tag{14}$$

Proof. For any $k \in \mathbb{N}_0$, $x \in [0, 1)$, the equality

$$\mathbb{E}[\Phi^{\mathsf{t}}_x(k)] = (1-x)\mathbb{E}[\Phi^{\mathsf{t}}_0(k)] + x\mathbb{E}[\Phi^{\mathsf{t}}_0(k+1)]$$

together with Equation (5) gives us

$$F_n^{\Phi^{\mathsf{t}}_x} = (1-x)F_n^{\Phi^{\mathsf{t}}_0} + xF_{n+1}^{\Phi^{\mathsf{t}}_0};$$

that is, for any $x \geq 0$, $\mathcal{F}^{\mathsf{t}}(x)$ is a linear interpolation between $\mathcal{F}^{\mathsf{t}}(\lfloor x \rfloor)$ and $\mathcal{F}^{\mathsf{t}}(\lfloor x \rfloor + 1)$. On the other hand, as, in view of (4) and (5), for $n \in \mathbb{N}$ we have

$$\mathcal{F}^{\mathsf{t}}(n) = \rho^n \mathrm{x}_1(\Phi^{\mathsf{t}}_0) A_1^n \pi_\rho W^{(1)} = \tfrac{\rho^n}{\rho-1} W + o(\rho^n),$$

we conclude that the following holds:

$$\mathcal{F}^{\mathsf{t}}(x) = \tfrac{\rho^x}{\rho-1} l_\rho(x) W + o(\rho^x) \text{ and}$$

$$\big(\mathcal{F}^{\mathsf{t}}\big)'(x) = \rho^{\lfloor x \rfloor} W + o(\rho^x) \text{ for } x \notin \mathbb{Z}.$$

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

By the same argument, for $j \in [J]$ we obtain

$$\mathcal{F}^j(x) = \tfrac{\rho^x}{\rho-1} l_\rho(x)\rho \mathrm{u}_j W + o(\rho^x) \text{ and}$$

$$\big(\mathcal{F}^j\big)'(x) = \rho^{\lfloor x \rfloor}\rho \mathrm{u}_j W + o(\rho^x) \text{ for } x \notin \mathbb{Z}.$$

In particular, $\mathcal{F}^{\mathsf{t}}$ is eventually increasing $\mathbb{P}^{\mathcal{S}}$-almost surely and thus the inverse function $(\mathcal{F}^{\mathsf{t}})^{-1}$ is well defined for large arguments. Furthermore, if $\mathcal{F}^{\mathsf{inv}}(t) \notin \mathbb{N}$ and t is large enough then we have

$$\big(\mathcal{F}^j \circ \mathcal{F}^{\mathsf{inv}}\big)'(t) = (\mathcal{F}^j)'\big(\mathcal{F}^{\mathsf{inv}}(t)\big) \cdot (\mathcal{F}^{\mathsf{inv}}(t))' = \frac{(\mathcal{F}^j)'\big(\mathcal{F}^{\mathsf{inv}}(t)\big)}{(\mathcal{F}^{\mathsf{t}})'\big(\mathcal{F}^{\mathsf{inv}}(t)\big)} \to \rho \mathrm{u}_j, \quad \text{as } t \to \infty,$$

since $\mathcal{F}^{\mathsf{inv}}(t)$ diverges to infinity. Finally, since for large t it holds that

$$\mathcal{F}^j\big(\mathcal{F}^{\mathsf{inv}}(t+s)\big) - \mathcal{F}^j\big(\mathcal{F}^{\mathsf{inv}}(t)\big) = \int_t^{t+s} \big(\mathcal{F}^j \circ \mathcal{F}^{\mathsf{inv}}\big)'(u)du,$$

we obtain (14). □

4.2. Limit theorems for $\mathcal{X}^{\mathsf{t}}$ and $\mathcal{X}^j$

To prove weak convergence of the processes $\left\{\frac{1}{n^{\ell+\frac{1}{2}}\rho^{n/2}\sqrt{W}}\mathcal{X}^j(n+x); x \in \mathbb{R}\right\}$, we follow the well-known technique: we first prove weak convergence of the finite-dimensional distributions, then prove tightness. According to [8, Theorem 3.5], the finite-dimensional distributions of the aforementioned processes converge jointly; see also Equation (9) and the discussion thereafter.

Theorem 3. *Suppose that (GW1)–(GW3) hold and L satisfies* $\mathbb{E}\left[\|L\|^{2+\delta}\right] < \infty$ *for some* $\delta \in (0, 1)$. *Then for every* $j \in [J]$, *the family of distributions*

$$\left\{\frac{1}{n^{\ell+\frac{1}{2}}\rho^{n/2}}\mathcal{X}^j(n+x); x \in \mathbb{R}\right\}$$

with respect to $\mathbb{P}$ *is tight in the Skorokhod space* $\mathcal{D}(\mathbb{R})$ *endowed with the standard* $\mathbf{J}_1$ *topology.*

Proof. First let us observe that for any $k \in \mathbb{Z},\ m \in \mathbb{N}$, the concatenation is a continuous mapping from $\mathcal{D}([k, k+1)) \times \cdots \times \mathcal{D}([k+m-1, k+m))$ to $\mathcal{D}([k, k+m))$. Consequently, it suffices to prove tightness in the space $\mathcal{D}([0, 1))$. For $x \in [0, 1)$, $j \in [J]$, and $n \in \mathbb{N}$ we set

$$\mathcal{Y}^j(n+x) = \frac{1}{n^{\ell+\frac{1}{2}}\rho^{n/2}}\mathcal{X}^j(n+x),$$

where ℓ may be $-\frac{1}{2}$ in Case (i) of [8, Theorem 3.5], and the characteristic $\Phi^j = (\Phi^j_x)_{x\in[0,1)}$ is defined as in (3). In view of [3, Theorem 13.5], it suffices to show that for any $0 \le x \le y \le z < 1$, $\lambda > 0$, and n large enough, it holds that

$$\mathbb{P}\Big(\big|\mathcal{Y}^j(n+y) - \mathcal{Y}^j(n+x)\big| \wedge \big|\mathcal{Y}^j(n+z) - \mathcal{Y}^j(n+y)\big| \ge \lambda\Big) \le C\lambda^{-2p}|z-x|^p,$$

where $p := (2+\delta)/2 \in (1, 3/2)$ and $C > 0$ is some constant. A more restrictive condition is the following inequality:

$$\mathbb{E}\Big[\ \big|\mathcal{Y}^j(n+y) - \mathcal{Y}^j(n+x)\big|^p \cdot \big|\mathcal{Y}^j(n+z) - \mathcal{Y}^j(n+y))\big|^p\ \Big] \le C|z-x|^p.$$

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

By Hölder's inequality, we have

$$\mathbb{E}\Big[|\mathcal{Y}^j(n+y)-\mathcal{Y}^j(n+x)|^p\cdot|\mathcal{Y}^j(n+z)-\mathcal{Y}^j(n+y)|^p\Big]$$
$$\leq \mathbb{E}\left[\mathbb{E}\Big[|\mathcal{Y}^j(n+y)-\mathcal{Y}^j(n+x)|^2\cdot|\mathcal{Y}^j(n+z)-\mathcal{Y}^j(n+y)|^2\Big|\mathcal{A}_n\Big]^{p/2}\right],$$

so it remains to show the estimate

$$\mathbb{E}\Big[|\mathcal{Y}^j(n+y)-\mathcal{Y}^j(n+x)|^2\cdot|\mathcal{Y}^j(n+z)-\mathcal{Y}^j(n+y)|^2\Big|\mathcal{A}_n\Big]\leq H_n|z-x|^2,$$

for some random variables H_n with bounded $p/2$ moment. Recalling that for $0\leq x<1$ we have

$$\mathcal{Z}^j(n+x)=\sum_{k=0}^{n-1}Z_k^j+\sum_{u\in\mathbb{T};|u|=n+1;\mathrm{t}(u)=j}\mathbf{1}_{\{U_u\leq x\}},$$

we deduce that for $0\leq x\leq y<1$,

$$\mathcal{Z}^j(n+y)-\mathcal{Z}^j(n+x)=\sum_{u\in\mathbb{T};|u|=n+1,\mathrm{t}(u)=j}\mathbf{1}_{\{x<U_u\leq y\}}.$$

We also have

$$\begin{aligned}\mathcal{F}^j(n+y)-\mathcal{F}^j(n+x)&=\left(\mathrm{x}_1(\Phi_y^j)-\mathrm{x}_1(\Phi_x^j)\right)A_1^nW^{(1)}+\left(\mathrm{x}_2(\Phi_y^j)-\mathrm{x}_2(\Phi_x^j)\right)A_2^nZ_0\\&=(y-x)A\mathrm{e}_j^\top\pi^{(1)}A_1^nW^{(1)}+(y-x)A\mathrm{e}_j^\top\pi^{(2)}A_2^nZ_0\\&=(y-x)A\mathrm{e}_j^\top\left(\pi^{(1)}A_1^nW^{(1)}+\pi^{(2)}A_2^nZ_0\right)\\&=(y-x)F_n^\Psi=(y-x)(Z_{n+1}^j-(Z_n^\Psi-F_n^\Psi)),\end{aligned}$$

where $\Psi:\mathbb{N}_0\to\mathbb{R}^{1\times J}$ is the characteristic given by $\Psi(k)=\mathbf{1}_{\{k=0\}}\mathrm{e}_j^\top L=\mathbf{1}_{\{k=0\}}\langle\mathrm{e}_j,L\rangle$. Hence Z_n^Ψ counts the number of individuals of type j in the $(n+1)$th generation. We then obtain

$$\begin{aligned}\mathcal{Y}^j(n+y)-\mathcal{Y}^j(n+x)&=\frac{1}{n^{\ell+\frac12}\rho^{n/2}}\left(\mathcal{X}^j(n+y)-\mathcal{X}^j(n+x)\right)\\&=\frac{1}{n^{\ell+\frac12}\rho^{n/2}}\left[(\mathcal{Z}^j(n+y)-\mathcal{Z}^j(n+x))-(\mathcal{F}^j(n+y)-\mathcal{F}^j(n+x))\right]\\&=\frac{1}{n^{\ell+\frac12}\rho^{n/2}}\Bigg[\sum_{u\in\mathbb{T};|u|=n+1;\mathrm{t}(u)=j}\mathbf{1}_{\{x<U_u\leq y\}}-(y-x)Z_{n+1}^j+(y-x)(Z_n^\Psi-F_n^\Psi)\Bigg]\\&=\frac{1}{n^{\ell+\frac12}\rho^{n/2}}\Bigg[\sum_{u\in\mathbb{T};|u|=n+1;\mathrm{t}(u)=j}\left(\mathbf{1}_{\{x<U_u\leq y\}}-(y-x)\right)\cdot1+(y-x)(Z_n^\Psi-F_n^\Psi)\Bigg].\end{aligned}$$

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

Since Z_{n+1}^j and $Z_n^\Psi - F_n^\Psi$ are $\mathcal{A}_n$-measurable, by applying Lemma 5 to the intervals $I = (x, y)$, $J = (y, z)$, with $a_i = 1$, $A = (y - x)(Z_n^\Psi - F_n^\Psi)$, and $B = (z - y)(Z_n^\Psi - F_n^\Psi)$, we finally obtain

$$\mathbb{E}\Big[|\mathcal{Y}^j(n+y) - \mathcal{Y}^j(n+x)|^2 \cdot |\mathcal{Y}^j(n+z) - \mathcal{Y}^j(n+y)|^2 | \mathcal{A}_n\Big]$$
$$\le C\Big(\frac{1}{n^{2\ell+1}\rho^n}\Big)^2 \Big((y-x)(z-y)\big(Z_{n+1}^j\big)^2 + A^4 + B^4\Big)$$
$$\le C(z-x)^2 \Big(\frac{1}{n^{\ell+\frac{1}{2}}\rho^{n/2}}\Big)^4 \Big[(Z_n^\Psi - F_n^\Psi)^4 + (Z_{n+1}^j)^2\Big] =: C|z-x|^2 H_n,$$

for some absolute constant C. We have $Z_{n+1}^j = Z_n^\Psi$, so Theorem 5(i) implies that $\mathbb{E}[(Z_{n+1}^j)^p] = O(\rho^{pn})$. On the other hand, Corollary 4 yields that $\mathbb{E}\big[(Z_n^\Psi - F_n^\Psi)^{2p}\big] = O(n^{(2\ell+1)p}\rho^{pn})$. As a consequence, the random variables H_n have bounded $p/2$ moments, and in turn the process $\Big\{\frac{1}{n^{\ell+\frac{1}{2}}\rho^{n/2}}\mathcal{X}^j(n+x); x \in \mathbb{R}\Big\}$ is tight in $\mathcal{D}(\mathbb{R})$. □

As a consequence of the previous result we obtain the following.

Corollary 2. *Suppose that the assumptions (GW1)–(GW3) hold and that the matrix L satisfies* $\mathbb{E}\left[\|L\|^{2+\delta}\right] < \infty$ *for some* $\delta \in (0, 1)$. *Then the family of distributions*

$$\Big\{\frac{1}{n^{\ell+\frac{1}{2}}\rho^{n/2}}\mathcal{X}^{\mathsf{t}}(n+x); x \in \mathbb{R}\Big\}$$

is tight in the Skorokhod space $\mathcal{D}(\mathbb{R})$ *endowed with the standard* $\mathbf{J}_1$ *topology.*

Proof. In view of the two equalities $\mathcal{Z}^{\mathsf{t}}(n+x) - 1 = \sum_{j=1}^J \mathcal{Z}^j(n-1+x)$ and $\mathcal{F}^{\mathsf{t}}(n+x) = \sum_{j=1}^J \mathcal{F}^j(n-1+x)$, together with

$$\mathcal{Y}^{\mathsf{t}}(n+x) = \frac{1}{n^{\ell+\frac{1}{2}}\rho^{n/2}}\mathcal{X}^{\mathsf{t}}(n+x) = \sum_{j=1}^J \frac{1}{n^{\ell+\frac{1}{2}}\rho^{n/2}}\mathcal{X}^j(n-1+x),$$

we see that $\mathcal{Y}^{\mathsf{t}}(n+x)$ can be written as a finite sum of tight processes, so it is tight as well. □

The convergence of the finite-dimensional distributions together with the tightness gives the weak convergence.

Proposition 3. *Suppose that the assumptions (GW1)–(GW3) hold and that the matrix L satisfies* $\mathbb{E}\left[\|L\|^{2+\delta}\right] < \infty$ *for some* $\delta \in (0, 1)$. *Then we have the following weak convergence of sequences of processes in the Skorokhod space* $\mathcal{D}(\mathbb{R})$ *endowed with the standard* $\mathbf{J}_1$ *topology: for every* $j \in [J]$ *it holds that*

$$\Big\{\frac{1}{n^{\ell+\frac{1}{2}}\rho^{n/2}\sqrt{W}}(\mathcal{X}^{\mathsf{t}}(n+x), \mathcal{X}^j(n+x)); x \in \mathbb{R}\Big\} \xrightarrow{\text{st},\mathcal{S}} \{(\mathcal{G}^{\mathsf{t}}(x), \mathcal{G}^j(x)); x \in \mathbb{R}\}.$$

4.3. Properties of the limiting processes $\mathcal{G}^{\mathsf{t}}$ and $\mathcal{G}^j$

Notice that for $\Phi_0 \in \mathcal{C}$, we have

$$\Phi_0(1) = a\Phi_0^{\mathsf{t}}(1) + b\Phi_0^j(1) = a1 + be_j^\top L \quad \text{and} \quad \mathbb{E}[\Phi_0(1)] = a1 + be_j^\top A,$$

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

where $a, b \in \mathbb{R}$ and $j \in [J]$. On the other hand, taking $a = -\rho \mathrm{u}_j$ and $b = 1$, we recover

$$\mathrm{w}_j = \mathrm{e}_j^\top A - \rho \mathrm{u}_j 1 = \mathbb{E}[\Phi_0^j(1) - \rho \mathrm{u}_j \Phi_0^{\mathrm{t}}(1)],$$

as defined in (1), whose ith entry is given by $\mathrm{w}_{ji} = \mathbb{E}[L^{(i,j)} - \rho \mathrm{u}_j] = a_{ji} - \rho \mathrm{u}_j$.

Proposition 4. *For any $\Phi \in \mathcal{C}$ and $j \in [J]$, assume that $\mathrm{w}_j \neq 0$ and that*

$$\sum_{\lambda \in \sigma_A} \sum_{\ell=0}^{J-1} \|\mathrm{Var}(\mathrm{w}_j N^\ell \pi_\lambda L)\| > 0. \tag{15}$$

(i) *If $\sum_{\lambda \in \sigma_A^2} \sum_{\ell=0}^{J-1} \|\mathrm{Var}(\mathrm{w}_j N^\ell \pi_\lambda L)\| > 0$, then for any $x \in [0, 1)$ it holds that*

$$\max\{0 \le \ell \le J-1 : \sigma_\ell^2(\Phi_x) > 0\} = \max\left\{\ell \ge 0 : \sum_{\lambda \in \sigma_A^2} \|\mathrm{Var}(\mathrm{w}_j N^\ell \pi_\lambda L)\| > 0\right\}. \tag{16}$$

In particular, the largest ℓ such that $\sigma_\ell^2(\Phi_x) > 0$ does not depend on x.

(ii) *Otherwise, for any $x \in [0, 1)$ and $0 \le \ell \le J-1$, it holds that*

$$\sigma_\ell^2(\Phi_x) = 0 \quad \text{and} \quad \sigma^2(\Phi_x) > 0.$$

Proof. (i) For any $x \in [0, 1)$, the vector $\mathrm{x}_2(\Phi_x)$ is given by

$$\mathrm{x}_2(\Phi_x) = \sum_{k=0}^{\infty} \mathbb{E}[\Phi_x(k)] \pi^{(2)} A_2^{-k} = x \mathrm{w}_j \pi^{(2)} + \mathrm{w}_j \pi^{(2)} \sum_{k=1}^{\infty} A_2^{-k}.$$

If k is at least the right-hand side of Equation (16), then for any $\lambda \in \sigma_A^2$ we have $\mathrm{w}_j N_\lambda^k (L - A) = 0$ almost surely. Since A_2 is invertible, we have

$$\sum_{j\ge 1} A_2^{-j} = A_2^{-1} \sum_{j \ge 0} A_2^{-j} = \sum_{j \ge 0} A_2^{-j} - I,$$

so from the last two matrix equations we get $\sum_{j\ge 0} A_2^{-j} = A_2 \sum_{j\ge 0} A_2^{-j} - A_2$, which in turn implies $(A_2 - I) \sum_{j \ge 0} A_2^{-j} = A_2$. Multiplying this equation by $(A_2 - I)^{-1}$ from the right and by A_2^{-1} from the left, we obtain $\sum_{j\ge 1} A_2^{-j} = (A_2 - I)^{-1}$; hence it holds that $(A_2 - I) \sum_{j \ge 1} A_2^{-j} \pi_\lambda = \pi_\lambda$, and finally it follows that $\sum_{j\ge 1} A_2^{-j} \pi_\lambda = ((\lambda - 1)I + N_\lambda)^{-1}$. In view of the definition (6) of $\sigma_\ell(\Phi)$, it is enough to understand the term $\mathrm{x}_2(\Phi)\pi_\lambda (A - \lambda I)^\ell (L - A)$. It holds that

$$\begin{aligned}
\mathrm{x}_2(\Phi_x)\pi_\lambda(A - \lambda I)^k (L - A) &= \Big(x \mathrm{w}_j \pi_\lambda + \mathrm{w}_j \sum_{j=1}^{\infty} A_2^{-j} \pi_\lambda\Big)(A - \lambda I)^k \pi_\lambda (L - A) \\
&= \big(x \mathrm{w}_j \pi_\lambda + \mathrm{w}_j ((\lambda - 1)I + N_\lambda)^{-1}\big) N_\lambda^k (L - A) \\
&= \left(x \mathrm{w}_j \pi_\lambda + \mathrm{w}_j \sum_{j=0}^{d_\lambda - 1} \binom{-1}{j} (\lambda - 1)^{-1-j} N_\lambda^j\right) N_\lambda^k \pi_\lambda (L - A) = 0,
\end{aligned}$$

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

almost surely, and hence $\sigma_\ell^2(\Phi_x)=0$. On the other hand, if $0\le\ell\le J-1$ is the maximal number such that $\mathbb{P}(\mathrm{w}_j N_\lambda^\ell(L-A)\neq 0)>0$ for some $\lambda\in\sigma_A^2$, then for such ℓ and λ, similar calculations give

$$\mathrm{x}_2(\Phi_x)\pi_\lambda(A-\lambda I)^\ell(L-A)=\left(x\mathrm{w}_j\pi_\lambda+\mathrm{w}_j\sum_{j=0}^{d_\lambda-1}\binom{-1}{j}(\lambda-1)^{-1-j}N_\lambda^j\right)N_\lambda^\ell\pi_\lambda(L-A)$$
$$=\mathrm{w}_j\big(x+(\lambda-1)^{-1}\big)N_\lambda^\ell\pi_\lambda(L-A)\neq 0$$

with positive probability, since $x+(\lambda-1)^{-1}\neq 0$ because $\lambda\notin\mathbb{R}$. This implies that $\sigma_\ell^2(\Phi_x)>0$, and this proves (i).

(ii) Assume that $\mathrm{w}_jN_\lambda^\ell(L-A)=0$ almost surely for any $\lambda\in\sigma_A^2$ and any $\ell\ge 0$. Let $\mathrm{t}_j=\Phi_0(1)=a1+b\mathrm{e}_j^\top L\in\mathbb{R}^{1\times J}$ be the random row vector whose expectation is $\mathbb{E}[\mathrm{t}_j]=\mathrm{w}_j\neq 0$ by assumption, and whose ith entry is denoted by t_{ji}. Note that for $x\in(0,1)$, in view of (8), we have

$$\sigma^2(\Phi_x)=\sum_{k=0}^{\infty}\rho^{-k}\mathrm{Var}\left[\Phi_x(k)+\Psi^{\Phi_x}(k)\right]\mathrm{u}\ge\mathrm{Var}\left[\Phi_x(0)+\Psi^{\Phi_x}(0)\right]\mathrm{u}$$
$$\ge\mathbb{E}\Big[\mathrm{Var}\big[\mathrm{t}_j\mathbf{1}_{\{U\le x\}}+\Psi^{\Phi_x}(0)\big|L\big]\mathrm{u}\Big]+\mathrm{Var}\Big[\mathbb{E}\big[\mathrm{t}_j\mathbf{1}_{\{U\le x\}}+\Psi^{\Phi_x}(0)\big|L\big]\mathrm{u}\Big]$$
$$\ge\mathbb{E}\Big[\mathrm{Var}\big[\mathrm{t}_j\mathbf{1}_{\{U\le x\}}\big|L\big]\mathrm{u}\Big]=x(1-x)\mathbb{E}\Big[\sum_{i=1}^{J}\mathrm{t}_{ji}^2\mathrm{u}_i\Big]\ge x(1-x)\sum_{i=1}^{J}\mathrm{w}_{ji}^2\mathrm{u}_i>0,$$

since by assumption $\mathrm{w}_j\neq 0$. Now we prove that $\sigma^2(\Phi_0)$ does not vanish. We have $\Phi_0(k)=\mathrm{w}_j\cdot\mathbf{1}_{\{k\ge1\}}$, so $\Phi_0:\mathbb{N}_0\mapsto\mathbb{R}^{1\times J}$ is completely deterministic. Assume that $\sigma^2(\Phi_0)=0$. Then for any $k\in\mathbb{N}_0$ it holds that $\mathrm{Var}\left[\Psi^{\Phi_0}(k)\right]\mathrm{u}=0$, which in turn implies, as $\mathrm{u}_j>0$ for $j\in[J]$, that for any $k\in\mathbb{N}_0$ and $j\in[J]$ we have $\mathrm{Var}\left[\Psi^{\Phi_0}(k)\mathrm{e}_j\right]=0$. The latter is equivalent to

$$\sum_{\ell\in\mathbb{Z}}\mathbf{1}_{\{k-\ell-1\ge1\}}\mathrm{w}_jA^\ell\mathsf{P}(k,\ell)(L-A)\mathrm{e}_j=0\quad\text{almost surely.}\tag{17}$$

We set $A_\lambda=\pi_\lambda A+\lambda(I-\pi_\lambda)$ and observe that for any $n\in\mathbb{Z}$ and $m\in\mathbb{N}$ we have

$$A_\lambda^n\pi_\lambda=(\lambda I+N_\lambda)^n\pi_\lambda=\lambda^n\sum_{i=0}^{d_\lambda}\lambda^{-i}\binom{n}{i}N_\lambda^i\pi_\lambda$$

and

$$\sum_{0\le\ell\le m}A_1^\ell\pi_1=\sum_{0\le\ell\le m}(I+N_1)^\ell\pi_1=\sum_{0\le\ell\le m}\sum_{i=0}^{\ell}\binom{\ell}{i}N_1^i\pi_1=\sum_{i=0}^{d_1}N_1^i\sum_{l=i}^{m}\binom{l}{i}\pi_1,$$

where in the second equation we have used $\lambda=1$ (if λ would be in the spectrum of A). Suppose now that for some vector $\mathrm{z}\in\mathbb{R}^J$ we have

$$\mathrm{w}_jN_\lambda^i\pi_\lambda\mathrm{z}=0\quad\text{for any }\lambda\in\sigma_A^2\text{ and }\ i=0,\ldots,d_\lambda-1,$$

and for any $k \in \mathbb{N}_0$ it holds that

$$\sum_{\ell\in\mathbb{Z}} \mathbf{1}_{\{k-\ell-1\geq 1\}} \mathrm{w}_j A^\ell \mathsf{P}(k,\ell)\mathrm{z} = 0. \tag{18}$$

Note that the left-hand side of the previous equation can be rewritten as

$$\sum_{\ell\in\mathbb{Z}} \mathbf{1}_{\{k-\ell-1\geq 1\}} \mathrm{w}_j A^\ell \mathsf{P}(k,\ell)\mathrm{z} = \sum_{\lambda\in\sigma_A}\sum_{\ell\in\mathbb{Z}} \mathbf{1}_{\{k-\ell-1\geq 1\}} \mathrm{w}_j A_\lambda^\ell \pi_\lambda \mathsf{P}(k,\ell)\mathrm{z}$$
$$= \sum_{\lambda\in\sigma_A^1\cup\sigma_A^3}\sum_{\ell\in\mathbb{Z}} \mathbf{1}_{\{k-\ell-1\geq 1\}} \mathrm{w}_j A_\lambda^\ell \pi_\lambda \mathsf{P}(k,\ell)\mathrm{z}$$
$$= -\sum_{\lambda\in\sigma_A^1}\sum_{\ell\in\mathbb{Z}} \mathbf{1}_{\{k-\ell-1\geq 1\}} \mathrm{w}_j A_\lambda^\ell \pi_\lambda \mathbf{1}_{\{\ell<0\}}\mathrm{z} + \sum_{\lambda\in\sigma_A^3}\sum_{\ell\in\mathbb{Z}} \mathbf{1}_{\{k-\ell-1\geq 1\}} \mathrm{w}_j A_\lambda^\ell \pi_\lambda \mathbf{1}_{\{\ell\geq 0\}}\mathrm{z}.$$

Setting $-n = k-2 \geq -2$, we get

$$0 = \sum_{\lambda\in\sigma_A^1}\sum_{\ell\leq -n} \mathrm{w}_j A_\lambda^\ell \mathbf{1}_{\{\ell<0\}} \pi_\lambda \mathrm{z} = \sum_{\lambda\in\sigma_A^1} \mathrm{w}_j (I - A_\lambda^{-1})^{-1} A_\lambda^{-n} \pi_\lambda \mathrm{z}$$
$$= \sum_{\lambda\in\sigma_A^1} \mathrm{w}_j (I - A_\lambda^{-1})^{-1} (\lambda\pi_\lambda + N_\lambda)^{-n} \pi_\lambda \mathrm{z}$$
$$= \sum_{\lambda\in\sigma_A^1} \lambda^{-n} \sum_{i=0}^{d_\lambda} \binom{-n}{i} \lambda^{-i} \Big(\mathrm{w}_j (I - A_\lambda^{-1})^{-1} N_\lambda^i \pi_\lambda \mathrm{z}\Big),$$

which in view of Lemma 6 implies that

$$\mathrm{w}_j (I - A_\lambda^{-1})^{-1} N_\lambda^i \pi_\lambda \mathrm{z} = 0 \quad \text{for } \lambda \in \sigma_A^1 \text{ and } i < d_\lambda. \tag{19}$$

Now we can use the decomposition

$$(I - A_\lambda^{-1})^{-1} \pi_\lambda = \sum_{i=0}^{d_\lambda - 1} c_i N_\lambda^i \pi_\lambda,$$

for some $c_0, \ldots, c_{d_\lambda - 1}$, and since $(I - A_\lambda^{-1})^{-1}$ is invertible, we conclude that the matrix $(I - A_\lambda^{-1})^{-1}\pi_\lambda$ is not nilpotent and hence $c_0 \neq 0$. Now, taking $i = d_{\lambda-1}$ in (19), we infer that $c_0 \mathrm{w}_j N_\lambda^{d_\lambda - 1}\mathrm{z} = 0$; that is, we have $\mathrm{w}_j N_\lambda^{d_\lambda - 1}\mathrm{z} = 0$. Recursively, we see that for $i = d_\lambda - 1, d_\lambda - 2, \ldots, 0$ Equation (19) implies $\mathrm{w}_j N_\lambda^i \mathrm{z} = 0$, for any $\lambda \in \sigma_A^1$. Taking this into account and setting

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

$n = k - 2 > 0$, from the condition (18) we get

$$0 = \sum_{\lambda\in\sigma_A^3}\sum_{0\le\ell\le n} \mathrm{w}_j A_\lambda^\ell \pi_\lambda \mathrm{z} = \sum_{\lambda\in\sigma_A^3\setminus\{1\}}\sum_{0\le\ell\le n} \mathrm{w}_j A_\lambda^\ell \pi_\lambda \mathrm{z} + \mathbf{1}_{\{1\in\sigma_A^3\}}\sum_{0\le\ell\le n} \mathrm{w}_j A_1^\ell \pi_1 \mathrm{z}$$

$$= \sum_{\lambda\in\sigma_A^3\setminus\{1\}} \mathrm{w}_j (A_\lambda - I)^{-1}(A_\lambda^{n+1} - I)\pi_\lambda \mathrm{z} + \mathbf{1}_{\{1\in\sigma_A^3\}}\sum_{0\le\ell\le n} \mathrm{w}_j (I + N_1)^\ell \pi_1 \mathrm{z}$$

$$= \sum_{\lambda\in\sigma_A^3\setminus\{1\}}\sum_{i=0}^{d_\lambda-1} \mathrm{w}_j (A_\lambda - I)^{-1} N_\lambda^i \pi_\lambda \mathrm{z} \lambda^n p_{\lambda,i}(n) + \mathbf{1}_{\{1\in\sigma_A^3\}}\sum_{i=0}^{d_1-1} \mathrm{w}_j N_1^i \pi_1 \mathrm{z} p_{1,i+1}(n) + c,$$

where $p_{\gamma,i}$ is some polynomial of degree i and c does not depend on n. Lemma 6 implies that

$$\mathrm{w}_j (A_\lambda - I)^{-1} N_\lambda^i \pi_\lambda \mathrm{z} = 0 \quad \text{for } \lambda \in \sigma_A^3 \setminus \{1\} \text{ and } i < d_\lambda,$$

and also

$$\mathrm{w}_j N_1^i \pi_1 \mathrm{z} = 0 \quad \text{if } 1 \in \sigma_A^3 \text{ and } i < d_1.$$

The same argument as before gives that $\mathrm{w}_j N_\lambda^i \mathrm{z} = 0$, for any $\lambda \in \sigma_A^3$ and $i < d_\lambda$. Suppose now that $\sigma^2(\Phi_0) = 0$ and also $\sigma_\ell^2(\Phi_0) = 0$ for all $0 \le \ell \le J$. Then by setting $\mathrm{z} = (L - A)\mathrm{e}_j$ with $j \in [J]$ we conclude that for any $\lambda \in \sigma_A$ and $\ell \ge 0$,

$$\mathrm{w}_j N_\lambda^\ell \pi_\lambda L = \mathrm{w}_j N_\lambda^\ell \pi_\lambda A \quad \text{almost surely.}$$

This contradicts the assumption. □

Continuity of the limit processes $\mathcal{G}^{\mathrm{t}}$ and $\mathcal{G}^j$

We recall once again the notation $\mathcal{G}^\Phi(x) = \rho^{\lfloor x\rfloor/2} G^{\Phi_{\{x\}}}$, where $G^{\Phi_{\{x\}}} \overset{\mathrm{d}}{=} \sigma_\ell(\Phi_{\{x\}})\mathcal{N}$ or $G^{\Phi_{\{x\}}} \overset{\mathrm{d}}{=} \sigma(\Phi_{\{x\}})\mathcal{N}$, with $\mathcal{N} := \mathcal{N}(0, 1)$ denoting a standard normal variable independent of W, and $\mathcal{G}^{\mathrm{t}}$ (respectively $\mathcal{G}^\Phi$) denoting $\mathcal{G}^\Phi$ if $\Phi = \Phi^{\mathrm{t}}$ (respectively $\Phi = \Phi^i$).

Lemma 2. *For $j \in [J]$, let $\mathcal{H}(x) = \rho \mathrm{u}_j \mathcal{G}^{\mathrm{t}}(x) - \mathcal{G}^j(x)$. Under the assumptions of Theorem 3, $\mathcal{H}(x)$ is continuous for any $x \in [0, \infty)$.*

Proof. Since $\mathcal{H}$ is a linear combination of the Gaussian processes $\mathcal{G}^{\mathrm{t}}$ and $\mathcal{G}^j$, it is enough to show continuity for the two terms separately. We start with the continuity of $\mathcal{G}^j$. For any $0 \le x \le y \le 1$, since either

$$\mathcal{G}^j(y) - \mathcal{G}^j(x) \overset{\mathrm{d}}{=} (\sigma(\Phi_y^j) - \sigma(\Phi_x^j))\mathcal{N} = \sigma(\Phi_y^j - \Phi_x^j)\mathcal{N}$$

or

$$\mathcal{G}^j(y) - \mathcal{G}^j(x) \overset{\mathrm{d}}{=} (\sigma_\ell(\Phi_y^j) - \sigma_\ell(\Phi_x^j))\mathcal{N} = \sigma_\ell(\Phi_y^j - \Phi_x^j)\mathcal{N},$$

depending on whether we are in Case (i) or Case (ii) of Theorem 3.5 in [8], we have to upper-bound $\sigma^2(\Phi_y^j - \Phi_x^j)$ and $\sigma_\ell^2(\Phi_y^j - \Phi_x^j)$ by some power of $|y - x|$. The definition of $\sigma^2(\Phi)$ applied to the characteristic $\Phi_y^j - \Phi_x^j$ yields

$$\sigma^2(\Phi_y^j - \Phi_x^j) = \sum_{k<0} \rho^{-k} \mathrm{Var}[\mathrm{e}_j \pi^{(1)}(y - x)A^{k-1})(L - A)]\mathrm{u}$$

$$+ \mathrm{Var}[\mathrm{e}_j \mathbf{1}_{\{x<U\le y\}} - \mathrm{e}_j \pi^{(1)}(y - x)A^{-1}(L - A)]\mathrm{u} + \rho \mathrm{Var}[(y - x)e_j \pi^{(3)}(L - A)]\mathrm{u} \le C(y - x).$$

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

On the other hand, as for $\Phi_y^j - \Phi_x^j$ it holds that $\mathrm{x}_2 = \mathrm{x}_2(\Phi_y^j - \Phi_x^j) = (y-x)\mathrm{e}_j\pi^{(2)}$, we conclude that

$$\sigma_\ell^2(\Phi_y^j - \Phi_x^j) = \frac{\rho^{-\ell}}{(2\ell+1)(\ell!)^2} \sum_{\lambda\in\sigma_A^2} \mathrm{Var}\left[\mathrm{x}_2\pi_\lambda(A-\lambda I)^\ell L\right]\mathrm{u} \leq C(y-x)^2.$$

In particular, in both of the cases (i) and (ii) of [8, Theorem 3.5] we have

$$\mathbb{E}[|\mathcal{G}^j(y) - \mathcal{G}^j(x)|^2] \leq C(y-x),$$

and therefore, since $\mathcal{G}^j$ is Gaussian, we obtain

$$\mathbb{E}[|\mathcal{G}^j(y) - \mathcal{G}^j(x)|^4] = 6\mathbb{E}[|\mathcal{G}^j(y) - \mathcal{G}^j(x)|^2]^2 \leq C|y-x|^2,$$

which by the Kolmogorov continuity theorem implies that $\mathcal{G}^j$ is continuous. The same calculations as for $\mathcal{G}^j$ can be carried out to prove that $\mathcal{G}^\mathrm{t}$ is continuous, so also $\mathcal{H}(x) = \rho\mathrm{u}_j\mathcal{G}^\mathrm{t}(x) - \mathcal{G}^j(x)$ is continuous. □

Localization of the stopping times

This section addresses the localization of the stopping times $(\tau_k)_{k\in\mathbb{N}}$. On the non-extinction event $\mathcal{S}$, for any $n \in \mathbb{N}$, we define the random variable

$$T_n = \log_\rho \frac{n(\rho-1)}{W},$$

and we define the function $h : \mathbb{R} \to \mathbb{R}$ by

$$h(x) = \lfloor x \rfloor + \frac{\rho^{\{x\}}-1}{\rho-1} = x + \frac{\rho^{\{x\}}-1}{\rho-1} - \{x\}. \tag{20}$$

Note that h^{-1} is uniformly continuous and given by $h^{-1}(x) = \lfloor x \rfloor + \log_\rho \big(1 + (\rho-1)\{x\}\big)$.

Proposition 5. *Under the assumptions (GW1)–(GW3), and if for $k \in \mathbb{N}$ we set $t_k = h(T_k)$, then for (τ_k) defined as in (10), we have*

$$\lim_{k\to\infty}(t_k - \tau_k) = \lim_{k\to\infty}(h(T_k) - \tau_k) = 0, \quad \mathbb{P}^{\mathcal{S}}\text{-almost surely.}$$

Proof. By Proposition 2, the following $\mathbb{P}^{\mathcal{S}}$-almost sure convergence holds:

$$\lim_{x\to\infty} \frac{\mathcal{Z}^\mathrm{t}(x)}{\rho^{\lfloor x\rfloor}(1+(\rho-1)\{x\})} = \lim_{x\to\infty}\frac{\mathcal{Z}^\mathrm{t}(x)}{\rho^x l_\rho(x)} = \frac{1}{\rho-1}W.$$

We recall that $l_\rho : [0,\infty) \to \mathbb{R}$ is defined as $l_\rho(x) = (1+(\rho-1)\{x\})\rho^{-\{x\}}$. Since $\mathcal{Z}^\mathrm{t}(\tau_k) = k$, we infer that for any $\delta > 0$ and large enough k we have

$$\frac{k}{\rho^{\tau_k} l_\rho(\tau_k)} \leq \frac{1}{\rho-1}We^\delta \quad \text{and} \quad \frac{k}{\rho^{\tau_k-\delta} l_\rho(\tau_k-\delta)} \geq \frac{1}{\rho-1}We^{-\delta}.$$

This can be rewritten as

$$\tau_k + \log_\rho l_\rho(\tau_k-\delta) - \log_\rho e^{-\delta} \leq T_k \leq \tau_k + \log_\rho l_\rho(\tau_k) + \log_\rho e^\delta.$$

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

Notice that we have

$$\tau_k + \log_\rho l_\rho(\tau_k) = \lfloor \tau_k \rfloor + \log_\rho (1 + (\rho - 1)\{\tau_k\}) = h^{-1}(\tau_k),$$

and the inverse of the increasing function $h^{-1}(x)$ is given by $\lfloor x \rfloor + \frac{\rho^{\{x\}}-1}{\rho-1} = h(x)$, which then yields the following inequalities:

$$\lfloor T_k - \log_\rho e^{\delta} \rfloor + \frac{\rho^{\{T_k - \log_\rho e^{\delta}\}} - 1}{\rho - 1} \le \tau_k \le \lfloor T_k - \log_\rho e^{-\delta} \rfloor + \frac{\rho^{\{T_k - \log_\rho e^{-\delta}\}} - 1}{\rho - 1} + \delta.$$

From the uniform continuity of $h(x)$, by letting $\delta \to 0$, we obtain the claim. □

The above proposition implies that

$$\tau_n = \log_\rho n + O(1), \qquad \mathbb{P}^{\mathcal{S}}\text{-almost surely.}$$

4.4. Limit theorems for B_j

Proposition 6. *Under the assumptions of Theorem 3, let $\Phi : \mathbb{Z} \to \mathbb{R}^{1\times J}$ be any characteristic such that the following stable convergence holds:*

$$\frac{\mathcal{X}^{\Phi}(n+x)}{n^{\ell+\frac{1}{2}}\rho^{n/2}\sqrt{W}} \xrightarrow{\text{st},\mathcal{S}} \mathcal{G}^{\Phi}(x) \quad \text{in } \mathcal{D}(\mathbb{R}), \tag{21}$$

for some continuous Gaussian process $\mathcal{G}^{\Phi}$ with $\mathrm{Var}\left[\mathcal{G}^{\Phi}(x)\right] > 0$, for any $x \in \mathbb{R}$. Then there exists a continuous, positive, 1-periodic function Ψ^{Φ} such that for τ_n as in (10) and $T_n = \log_\rho \frac{n(\rho-1)}{W}$, it holds that

$$\frac{\mathcal{X}^{\Phi}(\tau_n)}{\sqrt{n}(\log_\rho n)^{l+\frac{1}{2}}\Psi^{\Phi}(T_n)} \xrightarrow{\text{d},\mathcal{S}} \mathcal{N}(0, 1), \qquad \text{as } n \to \infty. \tag{22}$$

Proof. The key idea in the proof is to use the functional limit theorem for $\mathcal{X}^{\Phi}$ to replace $\mathcal{X}^{\Phi}(\tau_n)$ by $\mathcal{X}^{\Phi}(t_n)$. Recall that for h as defined in (20), which is continuous and strictly increasing, we have used the notation $t_n = h(\log_\rho n - \log_\rho \frac{W}{\rho-1}) = h(T_n)$. Furthermore, for any $x \in \mathbb{R}$ and $n \in \mathbb{N}$ it holds that $h(n+x) = n + h(x)$. Consequently, by (21), we have

$$\frac{\mathcal{X}^{\Phi}\big(h(n+x)\big)}{n^{\ell+\frac{1}{2}}\rho^{n/2}\sqrt{W}} = \frac{\mathcal{X}^{\Phi}\big(n+h(x)\big)}{n^{\ell+\frac{1}{2}}\rho^{n/2}\sqrt{W}} \xrightarrow{\text{st},\mathcal{S}} \mathcal{G}^{\Phi}\big(h(x)\big) \quad \text{in } \mathcal{D}(\mathbb{R}),$$

and the latter convergence can be rewritten as

$$\frac{\mathcal{X}^{\Phi}\big(h(n+x)\big)}{\left(n^{2\ell+1}\rho^n W \mathrm{Var}\left[\mathcal{G}^{\Phi}\big(h(x)\big)\right]\right)^{1/2}} \xrightarrow{\text{st},\mathcal{S}} G(x) \quad \text{in } \mathcal{D}(\mathbb{R}),$$

for some stationary and continuous process G with $G(0) \overset{\text{D}}{=} \mathcal{N}(0, 1)$. Note that one consequence of the convergence in (21) is the following property of the limiting process: $\mathcal{G}^{\Phi}(x+1) \overset{\text{d}}{=} \sqrt{\rho}\mathcal{G}^{\Phi}(x)$. Hence, the previous convergence is equivalent to

$$\frac{\mathcal{X}^{\Phi}\big(h(n+x)\big)}{\left(\big((n+x)\vee 1\big)^{2\ell+1}\rho^{n+x}\rho^{-\{x\}} W \mathrm{Var}\left[\mathcal{G}^{\Phi}\big(h(\{x\})\big)\right]\right)^{1/2}} \xrightarrow{\text{st},\mathcal{S}} G(x) \quad \text{in } \mathcal{D}(\mathbb{R}).$$

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

In other words, for the function Ψ^{Φ} defined by

$$\Psi^{\Phi}(x)=\Big((\rho-1)\rho^{-\{x\}}\mathrm{Var}\big[\mathcal{G}^{\Phi}\big(h(\{x\})\big)\big]\Big)^{1/2},$$

which is continuous and positive, and for

$$X(x)=\frac{\mathcal{X}^{\Phi}\big(h(x)\big)}{\Big(\big(x\vee 1\big)^{2\ell+1}\rho^{x}\frac{W}{\rho-1}\Big)^{1/2}\Psi^{\Phi}(x)},$$

it holds that $X(n+x)\xrightarrow{\mathrm{st},\mathcal{S}}G(x)$, and therefore an application of Lemma 3(ii) with $a_n=\log_\rho n$ yields

$$X\big(h^{-1}(t_n)\big)=X\big(\log_\rho n-\log_\rho\tfrac{W}{\rho-1}\big)\xrightarrow{\mathrm{d},\mathcal{S}}G(0). \tag{23}$$

Since $h^{-1}(x)$ is uniformly continuous, by Proposition 5 we get

$$h^{-1}(\tau_n)-h^{-1}(t_n)\to 0,\qquad \mathbb{P}^{\mathcal{S}}\text{-almost surely.} \tag{24}$$

We claim that from Lemma 3(i) with $N_n=\lfloor\log_\rho n\rfloor$ and $\delta_n=2^{-n}$ it follows that

$$X\big(h^{-1}(\tau_n)\big)-X\big(h^{-1}(t_n)\big)\xrightarrow{\mathbb{P}^{\mathcal{S}}}0,\qquad \text{as } n\to\infty. \tag{25}$$

Indeed, for fixed $m\in\mathbb{N}$, on the event $|\log_\rho W|\le k_m-2\delta_m-1-\log_\rho(\rho-1)$ and $|h^{-1}(\tau_n)-h^{-1}(t_n)|\le\delta_m$ we have

$$\big|X\big(h^{-1}(\tau_n)\big)-X\big(h^{-1}(t_n)\big)\big|\le\omega(X_{N_m},k_m,\delta_m).$$

In turn, for any $\varepsilon>0$ we get

$$\mathbb{P}^{\mathcal{S}}\Big(\big|X\big(h^{-1}(\tau_n)\big)-X\big(h^{-1}(t_n)\big)\big|>\varepsilon\Big)\le\mathbb{P}^{\mathcal{S}}\big(\omega(X_{N_m},k_m,\delta_m)>\varepsilon\big)$$
$$+\mathbb{P}^{\mathcal{S}}\big(|\log_\rho W|>k_m-2\delta_m-1-\log_\rho(\rho-1)\big)+\mathbb{P}^{\mathcal{S}}\big(|h^{-1}(\tau_n)-h^{-1}(t_n)|>\delta_m\big).$$

Taking the limit first as $n\to\infty$ and then as $m\to\infty$ and using (24), we get (25). Finally, (25) and (23) imply that

$$\frac{\mathcal{X}^{\Phi}(\tau_n)}{\big(h^{-1}(\tau_n)\vee 1\big)^{\ell+\frac12}\sqrt{n}\rho^{(h^{-1}(\tau_n)-h^{-1}(t_n))/2}\Psi^{\Phi}(h^{-1}(\tau_n))}=X\big(h^{-1}(\tau_n)\big)\xrightarrow{\mathrm{d},\mathcal{S}}G(0),$$

which together with (24) and the $\mathbb{P}^{\mathcal{S}}$-almost sure convergence of $\frac{\Psi^{\Phi}(h^{-1}(\tau_n))}{\Psi^{\Phi}(h^{-1}(t_n))}$ and $\frac{h^{-1}(\tau_n)}{\log_\rho n}$ to 1 yields that

$$\frac{\mathcal{X}^{\Phi}(\tau_n)}{\sqrt{n}(\log_\rho n)^{\ell+\frac12}\Psi^{\Phi}(T_n)}\xrightarrow{\mathrm{d},\mathcal{S}}\mathcal{N}(0,1);$$

that is, (22) holds. This completes the proof. □

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

Using the auxiliary result that we have just proved, we can now state and prove our main result.

Theorem 4. *Suppose that (15) holds and all assumptions from Theorem 3 are satisfied. Then there exists a continuous, positive, and 1-periodic function Ψ such that, for $T_n = \log_\rho \frac{n(\rho-1)}{W}$ and any $j \in [J]$, we have*

$$\frac{B_j(n) - \mathcal{F}^j\big(\mathcal{F}^{\mathsf{inv}}(n)\big)}{\sqrt{n}(\log_\rho n)^{\ell+\frac{1}{2}}\Psi(T_n)} \xrightarrow{\mathrm{d},\mathcal{S}} \mathcal{N}(0,1), \quad \text{as } n \to \infty.$$

Proof. Since for any $j \in [J]$ we have $\mathcal{Z}^j(\tau_n) = B_j(n)$ and $\mathcal{X}^j(n) = \mathcal{Z}^j(n) - \mathcal{F}^j(n)$, we can write

$$\begin{aligned} B_j(n) &= \mathcal{F}^j(\tau_n) + \mathcal{X}^j(\tau_n) \\ &= \mathcal{F}^j\big(\mathcal{F}^{\mathsf{inv}}(n)\big) + \mathcal{X}^j(\tau_n) + (\mathcal{F}^j \circ \mathcal{F}^{\mathsf{inv}} \circ \mathcal{F}^{\mathsf{t}}(\tau_n) - \mathcal{F}^j(\mathcal{F}^{\mathsf{inv}}(n))). \end{aligned}$$

From the fact that $\mathcal{F}^{\mathsf{t}}(\tau_n) = n - \mathcal{X}^{\mathsf{t}}(\tau_n)$ together with Lemma 1, we obtain

$$\mathcal{F}^j \circ \mathcal{F}^{\mathsf{inv}} \circ \mathcal{F}^{\mathsf{t}}(\tau_n) - \mathcal{F}^j \circ \mathcal{F}^{\mathsf{inv}}(n) + \rho \mathrm{u}_j \mathcal{X}^{\mathsf{t}}(\tau_n) = \mathcal{X}^{\mathsf{t}}(\tau_n) o(1), \quad \mathbb{P}^{\mathcal{S}}\text{-almost surely},$$

and therefore

$$B_j(n) - \mathcal{F}^j(\mathcal{F}^{\mathsf{inv}}(n)) = \mathcal{X}^j(\tau_n) - \rho \mathrm{u}_j \mathcal{X}^{\mathsf{t}}(\tau_n) + o(1)\mathcal{X}^{\mathsf{t}}(\tau_n) \quad \mathbb{P}^{\mathcal{S}}\text{-almost surely}.$$

By Proposition 4 (positivity of the variances of the limiting process) and Lemma 2 (continuity of the limit processes), the characteristics $\Phi = \Phi^j - \rho \mathrm{u}_j \Phi^{\mathsf{t}}$ and $\Phi = \Phi^{\mathsf{t}}$ and the corresponding processes $\mathcal{X}^{\Phi}$ with these characteristics satisfy the assumptions of Proposition 6. Thus we can apply Proposition 6 to the processes $\mathcal{X}^{\Phi} = \mathcal{X}^j - \rho \mathrm{u}_j \mathcal{X}^{\mathsf{t}}$ and to $\mathcal{X}^{\mathsf{t}}$ to obtain

$$\frac{\mathcal{X}^j(\tau_n) - \rho \mathrm{u}_j \mathcal{X}^{\mathsf{t}}(\tau_n)}{\sqrt{n}(\log_\rho n)^{\ell+\frac{1}{2}}\Psi(T_n)} \xrightarrow{\mathrm{d},\mathcal{S}} \mathcal{N}(0,1),$$

for some function Ψ which is continuous, positive, and 1-periodic; moreover, $\frac{\mathcal{X}^{\mathsf{t}}(\tau_n)}{\sqrt{n}(\log_\rho n)^{\ell+\frac{1}{2}}} \cdot o(1) \xrightarrow{\mathbb{P}} 0$ as $n \to \infty$, which completes the proof. □

Finally, in view of Theorem 4, it suffices to find an expansion of $\mathcal{F}^j(\mathcal{F}^{\mathsf{inv}}(n))$ up to an error of order $o(n^{\log_\rho \gamma})$ to prove Theorem 2 ; the latter will then follow immediately from the next corollary.

Corollary 3. *Under the assumptions of Theorem 4, suppose that all eigenvalues in Γ are simple. Then the following hold:*

(i) *If $\gamma > \sqrt{\rho}$, then for any $\lambda \in \Gamma$ there exist a 1-periodic, continuous function $f_\lambda : \mathbb{R} \to \mathbb{C}$ and a random variable X_λ such that*

$$B_j(n) = \rho \mathrm{u}_j \cdot n + \sum_{\lambda \in \Gamma} n^{\log_\rho \lambda} f_\lambda(T_n) X_\lambda + o_{\mathbb{P}}\Big(n^{\log_\rho \gamma}\Big).$$

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

(ii) *If* $\gamma = \sqrt{\rho}$, *then there is a* 1*-periodic, continuous function* $\Psi : \mathbb{R} \to (0, \infty)$ *such that the following convergence holds:*

$$\frac{B_j(n) - \rho \mathrm{u}_j \cdot n}{\sqrt{n}(\log_\rho n)^{\frac{1}{2}} \Psi(T_n)} \xrightarrow{\mathrm{d},\mathcal{S}} \mathcal{N}(0, 1), \quad \text{as } n \to \infty.$$

(iii) *If* $\gamma < \sqrt{\rho}$, *then there is a* 1*-periodic, continuous function* $\Psi : \mathbb{R} \to (0, \infty)$ *such that the following convergence holds:*

$$\frac{B_j(n) - \rho \mathrm{u}_j \cdot n}{\sqrt{n}\Psi(T_n)} \xrightarrow{\mathrm{d},\mathcal{S}} \mathcal{N}(0, 1), \quad \text{as } n \to \infty.$$

Proof. We handle the case (i) in detail; the other two cases are identical, so we leave the details to the interested reader. If $\gamma > \sqrt{\rho}$, then by (5) we have

$$\begin{aligned}\mathcal{F}^\Phi(n) &= \mathrm{x}_1(\Phi)A_1^n W^{(1)} + \mathrm{x}_2(\Phi)A_2^n Z_0 = \sum_{\lambda \in \Gamma \cup \{\rho\}} \lambda^n \mathrm{x}_1(\Phi)\pi_\lambda W^{(1)} + o(\gamma^n) \\ &= \rho^n \mathrm{x}_1(\Phi) W \mathrm{u} + \sum_{\lambda \in \Gamma} \lambda^n \mathrm{x}_1(\Phi) W_\lambda \mathrm{u}^\lambda + o(\gamma^n),\end{aligned}$$

where we define $W_\lambda \mathrm{u}^\lambda = \pi_\lambda W^{(\lambda)}$ for some scalar random variable W_λ (this can be done since π_λ is a projection on the space spanned by the eigenvector u^λ). In particular, as $\mathcal{F}^j$ and $\mathcal{F}^{\mathrm{t}}$ are piecewise linear between consecutive integer arguments, we conclude that

$$\mathcal{F}^j(x) = \rho^x l_\rho(x) \mathrm{x}_1(\Phi_0^j) W \mathrm{u} + \sum_{\lambda \in \Gamma} \lambda^x l_\lambda(x) \mathrm{x}_1(\Phi_0^j) W_\lambda \mathrm{u}^\lambda + o(\gamma^x),$$

and taking into account that $\mathrm{x}_1(\Phi_0^j) = \frac{\lambda}{\lambda - 1} \mathrm{e}_j^\top \pi_\lambda$, we finally get

$$\mathcal{F}^j(x) = \frac{\rho^{x+1}}{\rho - 1} l_\rho(x) W \mathrm{u}_j + \sum_{\lambda \in \Gamma} \frac{\lambda^{x+1}}{\lambda - 1} l_\lambda(x) W_\lambda \mathrm{u}_j^\lambda + o(\gamma^x).$$

By the same argument, we also obtain

$$\mathcal{F}^{\mathrm{t}}(x) = \sum_{j=1}^J \mathcal{F}^j(x-1) = \frac{\rho^x}{\rho - 1} l_\rho(x) W + \sum_{\lambda \in \Gamma} \frac{\lambda^x}{\lambda - 1} l_\lambda(x) W_\lambda \Big(\sum_{i=1}^J \mathrm{u}_i^\lambda \Big) + o(\gamma^x).$$

In particular, it holds that

$$\mathcal{F}^j(x) = \rho \mathrm{u}_j \cdot \mathcal{F}^{\mathrm{t}}(x) + \sum_{\lambda \in \Gamma} \frac{\lambda^x}{\lambda - 1} l_\lambda(x) \Big(\lambda \mathrm{u}_j^\lambda - \rho \mathrm{u}_j \Big(\sum_{i=1}^J \mathrm{u}_i^\lambda \Big) \Big) W_\lambda + o(\gamma^x).$$

Since for $\lambda \in \Gamma$ we have

$$\lambda^x = \left(\frac{\rho - 1}{l_\rho(x) W} \mathcal{F}^{\mathrm{t}}(x) \right)^{\log_\rho \lambda} (1 + o(1)) = \left(\frac{\rho - 1}{l_\rho(x) W} \mathcal{F}^{\mathrm{t}}(x) \right)^{\log_\rho \lambda} + o(\gamma^x) \text{ on } \mathcal{S},$$

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

and $\mathcal{F}^{\mathsf{inv}}(n) = t_n + o(1) = h(T_n) + o(1)$, we deduce that

$$\mathcal{F}^j(\mathcal{F}^{\mathsf{inv}}(n)) = \rho \mathrm{u}_j \cdot n + \sum_{\lambda \in \Gamma} n^{\log_\rho \lambda} \frac{l_\lambda(h(T_n))}{l_\rho(h(T_n))^{\log_\rho \lambda}} \cdot \Big(\lambda \mathrm{u}_j^\lambda - \rho \mathrm{u}_j \Big(\sum_{i=1}^{J} \mathrm{u}_i^\lambda\Big)\Big)\Big(\frac{\rho - 1}{W}\Big)^{\log_\rho \lambda} \frac{W_\lambda}{\lambda - 1}$$
$$+ o_{\mathbb{P}}(n^{\log_\rho \gamma}),$$

and thus (i) holds with

$$f_\lambda(x) = \frac{l_\lambda(h(x))}{l_\rho(h(x))^{\log_\rho \lambda}} \quad \text{and} \quad X_\lambda = \big(\lambda \mathrm{u}_j^\lambda - \rho \mathrm{u}_j\big(\sum_{i=1}^{J} \mathrm{u}_i^\lambda\big)\big)\Big(\frac{\rho - 1}{W}\Big)^{\log_\rho \lambda} \frac{W_\lambda}{\lambda - 1},$$

for every $\lambda \in \Gamma$.

In the case $\gamma = \sqrt{\rho}$ we have

$$\mathcal{F}^\Phi(n) = \rho^n \mathrm{x}_1(\Phi) W \mathrm{u} + \sum_{\lambda \in \Gamma} \lambda^n \mathrm{x}_1(\Phi) \pi_\lambda Z_0 + o(\gamma^n) = \rho^n \mathrm{x}_2(\Phi) W \mathrm{u} + o(\rho^{n/2}),$$

and for $\gamma < \sqrt{\rho}$ we have

$$\mathcal{F}^\Phi(n) = \rho^n \mathrm{x}_1(\Phi) W \mathrm{u} + o(\rho^{n/2}).$$

In both cases we can use the same approach as in the proof of (i), after an application of Theorem 4; this proves (ii) and (iii). □

Appendix A. Example

We illustrate here the model with two alternating urns using an example with $J = 3$ colors ($1 =$ black, $2 =$ red, and $3 =$ green in the tree from Figure 1) and deterministic replacement matrix L given by

$$L = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

and $j_0 = 1$; that is, we start with one black ball in U_1 at time 0, so $B(0) = (1, 0, 0)$. The first column $L^{(1)}$ of L tells us that when we draw a black ball from some urn, we add one black and one green ball to the other urn, so U_2 is given by the nodes at level one of the tree and $B(1) = (2, 0, 1)$. Since after one step U_1 has been emptied, we proceed to draw balls step by step from U_2 (from the first level of the tree). After drawing a green ball from U_2, since the third column of the matrix L gives the number of balls of each color added to the other urn, with probability $1/2$ the number of balls added after two steps is $B(2) = (4, 1, 2)$ and with probability $1/2$ it is $B(2) = (3, 0, 2)$. Then $B(3) = (5, 1, 3)$ and U_2 has been emptied, so we proceed again to U_1, which now contains 6 balls, and with probability $1/6$ we have $B(4) = (5, 3, 4)$; now we have started to build the third level of the random tree and to fill U_2 again.

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

Appendix B. Higher-order estimates and additional results

Higher-moment estimates for Z_n^Φ. We provide here, for a general random characteristic Φ, higher-moment estimates for the random variable $Z_n^\Phi - \mathrm{x}_1 A_1^n W^{(1)} - \mathrm{x}_2 A_2^n Z_0$. These estimates are needed in the proof of Theorem 3 for the characteristics Φ^{t} and Φ^j which fulfill the assumptions of the next result. Recall that Φ is called *centered* if $\mathbb{E}[\Phi(k)] = 0 \in \mathbb{R}^{1\times J}$ for any $k \in \mathbb{Z}$.

Theorem 5. *Let $p \in [1, 2]$, and let $\Phi : \mathbb{Z} \to \mathbb{C}^{1\times J}$ be a random characteristic. Moreover, assume that (GW1)–(GW3) hold and the second moment of L is finite. Then the following hold:*

(i) *If $\sum_{k\in\mathbb{Z}} \big(\mathbb{E}\big[\|\Phi(k)\|^p\big]\big)^{1/p} \rho^{-k} < \infty$, then $\mathbb{E}\big[|Z_n^\Phi|^p\big] = O(\rho^{pn})$.*

(ii) *If $\sum_{k\le n} \big(\mathbb{E}\big[\|\Phi(k)\|^p\big]\big)^{1/p} \rho^{-k} = O(n^r)$ for some $r \ge 0$, then $\mathbb{E}\big[|Z_n^\Phi|^p\big] = O(\rho^{pn} n^{pr})$.*

If in addition Φ is centered, then the following hold:

(i) *If $\sum_{k\in\mathbb{Z}} \big(\mathbb{E}\big[\|\Phi(k)\|^{2p}\big]\big)^{1/p} \rho^{-k} < \infty$, then $\mathbb{E}\left[|Z_n^\Phi|^{2p}\right] = O(\rho^{np})$.*

(ii) *If $\sum_{k\le n} \big(\mathbb{E}\big[\|\Phi(k)\|^{2p}\big]\big)^{1/p} \rho^{-k} = O(\rho^n n^r)$, then $\mathbb{E}\left[|Z_n^\Phi|^{2p}\right] = O(\rho^{np} n^{pr})$.*

Proof. From the decomposition

$$Z_n^\Phi = Z_n^{\Phi - \mathbb{E}\Phi} + Z_n^{\mathbb{E}\Phi},$$

it is enough to get the desired bound on each term separately. If v is the right eigenvector of A for the eigenvalue $\rho > 1$, then it holds that

$$\langle \mathbf{1}, Z_n \rangle \le \min_i \mathrm{v}_i^{-1} \langle \mathrm{v}, Z_n \rangle = \rho^n \min_i \mathrm{v}_i^{-1} \langle \mathrm{v}, W_n^{(1)} \rangle,$$

and further we have

$$\begin{aligned}
\big|Z_n^{\mathbb{E}\Phi}\big| &\le \sum_{k\ge 0} \|\mathbb{E}\Phi(n-k)\| \|Z_k\| \le \sum_{k\ge 0} \|\mathbb{E}\Phi(n-k)\| \langle \mathbf{1}, Z_k \rangle \\
&\le \min_i \mathrm{v}_i^{-1} \sum_{k\ge 0} \|\mathbb{E}\Phi(n-k)\| \langle \mathrm{v}, Z_k \rangle \le \min_i \mathrm{v}_i^{-1} \sum_{k\ge 0} \|\mathbb{E}\Phi(n-k)\| \langle \mathrm{v}, \rho^k W_k^{(1)} \rangle \\
&= \rho^n \times \min_i \mathrm{v}_i^{-1} \sum_{k\le n} \|\mathbb{E}\Phi(k)\| \rho^{-k} \langle \mathrm{v}, W_{n-k}^{(1)} \rangle.
\end{aligned}$$

In view of Lemma 2.2 from [8], the random variables $\langle \mathrm{v}, W_k^{(1)} \rangle$ are bounded in $\mathcal{L}^2$, and therefore by Minkowski's inequality the $\mathcal{L}^2$ norm of $Z_n^{\mathbb{E}\Phi}$ is bounded by a multiple of $\rho^n \sum_{k\le n} \|\mathbb{E}\Phi(k)\| \rho^{-k}$. As a result we get

$$\mathbb{E}\Big[\big|Z_n^{\mathbb{E}\Phi}\big|^2\Big] = O(\rho^{2n})$$

in the case (i) and

$$\mathbb{E}\Big[\big|Z_n^{\mathbb{E}\Phi}\big|^2\Big] = O(n^{2r} \rho^{2n})$$

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

in the case (ii). By Jensen's inequality, the pth moment of $Z_n^{\mathbb{E}\Phi}$, for $p \in [1, 2]$, is of order $O(\rho^{pn})$ in the case (i) and of order $O(n^{pr}\rho^{pn})$ in the case (ii).

Now we focus on the case with a centered characteristic $\Psi := \Phi - \mathbb{E}\Phi$. We consider an increasing sequence $(G_n)_{n\in\mathbb{N}}$ of subsets of $\mathcal{U}_\infty$ that satisfies the following: $\cup_{n\geq 1} G_n = \mathcal{U}_\infty$; for any $n \in \mathbb{N}$, $G_n = n$; if $u \in G_n$, then for any $v \leq u$, $v \in G_n$. Such a sequence can be constructed using the diagonal method. If $\mathcal{G}_n = \sigma(\{L(u) : u \in G_n\})$, then one can see that $\sum_{u\in G_k} \Psi_u(n - |u|)\mathrm{e}_{\mathrm{t}(u)}$ is a $\mathcal{G}_k$-martingale. Indeed, for any $u \in G_k$ both $\mathrm{t}(u)$ and Ψ_u are $\mathcal{G}_k$-measurable, and the fact that Ψ is centered gives the martingale property. By the Topchii–Vatutin inequality [13, Theorem 2] applied to $\sum_{u\in G_n} \Psi_u(n - |u|)\mathrm{e}_{\mathrm{t}(u)}$ we get

$$\begin{aligned}
\mathbb{E}\Big[\Big|\sum_{u\in\mathbb{T}}\Psi_u(n-|u|)\mathrm{e}_{\mathrm{t}(u)}\Big|^p\Big] &\leq C_p\mathbb{E}\Big[\sum_{u\in\mathbb{T}}\big|\Psi_u(n-|u|)\mathrm{e}_{\mathrm{t}(u)}\big|^p\Big] \\
&\leq C_p\sum_{k\geq 0}\mathbb{E}\big[\|\Psi(n-k)\|^p\big]\rho^k \leq 2^pC_p\sum_{k\geq 0}\mathbb{E}\big[\|\Phi(n-k)\|^p\big]\rho^k \\
&\leq 2^pC_p\Big(\sum_{k\geq 0}\big(\mathbb{E}\big[\|\Phi(n-k)\|^p\big]\big)^{1/p}\rho^{k/p}\Big)^p \leq 2^pC_p\Big(\sum_{k\geq 0}\big(\mathbb{E}\big[\|\Phi(n-k)\|^p\big]\big)^{1/p}\rho^{k}\Big)^p \\
&\leq 2^pC_p\rho^{pn}\Big(\sum_{k\leq n}\big(\mathbb{E}\big[\|\Phi(k)\|^p\big]\big)^{1/p}\rho^{-k}\Big)^p,
\end{aligned}$$

and the latter expression is of the order $O(\rho^{pn})$ in the case (i) and $O(n^{pr}\rho^{pn})$ in the case (ii). This finishes the proof of the first two statements (i) and (ii).

Now we turn to the proof of (iii) and (iv). Observe that the Burkholder–Davis–Gundy inequality [4, Theorem 1.1] yields

$$\begin{aligned}
\mathbb{E}\Big[\big|Z_n^\Phi\big|^{2p}\Big] &= \mathbb{E}\Big[\Big|\sum_u \Phi_u(n-|u|)\mathrm{e}_{\mathrm{t}(u)}\Big|^{2p}\Big] \\
&\leq C_p\mathbb{E}\Big[\Big|\sum_u\big|\Phi_u(n-|u|)\mathrm{e}_{\mathrm{t}(u)}\big|^2\Big|^p\Big] = C_p\mathbb{E}\Big[\big|Z_n^\Psi\big|^p\Big],
\end{aligned}$$

where Ψ is a new characteristic defined by $\Psi_u(k)\mathrm{e}_i := \big|\Phi_u(k)\mathrm{e}_i\big|^2$, i.e. the components of $\Psi(k)$ are squares of the components of Φ. Clearly $\|\Psi(k)\| \leq \|\Phi(k)\|^2$, and as a consequence (iii) and (iv) follow from (i) and (ii) respectively applied to the characteristic Ψ. □

Corollary 4. *Let $\Phi : \mathbb{Z} \to \mathbb{C}^{1\times J}$ be a random characteristic such that*

$$\sum_{k\in\mathbb{Z}}\big\|\mathbb{E}[\Phi(k)]\big\|(\rho^{-k} + \vartheta^{-k}) < \infty, \tag{26}$$

for some $\vartheta < \sqrt{\rho}$, and

$$\sum_{k\in\mathbb{Z}}\|\mathrm{Var}[\Phi(k)]\|\rho^{-k} < \infty. \tag{27}$$

Suppose that $\mathbb{E}\left[\|L\|^{2p}\right] < \infty$ for some $p \in (1, 2)$. Then, for F_n^Φ defined by (5), it holds that

$$\mathbb{E}\Big[\big|Z_n^\Phi - F_n^\Phi\big|^{2p}\Big] = \begin{cases} O\big(\rho^{np}\big) & \textit{if for all } 0 \leq \ell \leq J-1,\ \sigma_\ell = 0, \\ O\big(n^{(2\ell+1)p}\rho^{np}\big) & \textit{if } 0 \leq \ell \leq J-1 \textit{ is maximal with } \sigma_\ell > 0. \end{cases}$$

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

Proof. The decomposition from Equation (18) in [8] yields

$$Z_n^{\Phi} - F_n^{\Phi} = Z_n^{\Psi_1} + Z_n^{\Psi_2} + o(\rho^{n/2}), \tag{28}$$

where Ψ_1 and Ψ_2 are two random centered characteristics such that

- for any $k \in \mathbb{Z}$ we can decompose Ψ_1 as $\Psi_1(k) = \Psi_1{}'(k)(L - A)$ for some deterministic characteristic $\Psi_1{}'(k)$ (see the paragraph after Equation (19) in [8]) and $\sum_{k\in\mathbb{Z}} \mathbb{E}\left[\|\Psi_1(k)\|^2\right] \rho^{-k} < \infty$, and
- $Z_n^{\Psi_2} = \mathrm{x}_2\pi^{(2)}(Z_n - A_2^n Z_0)$, i.e. $\Psi_2(k) = \mathrm{x}_2\pi^{(2)}A^{k-1}(L-A)\mathbf{1}_{\{k>0\}}$, for some row vector x_2.

Moreover, the last term $o(\rho^{n/2})$ in (28) is deterministic. By Minkowski's inequality, we have

$$\mathbb{E}\Big[\big|Z_n^{\Phi} - F_n^{\Phi}\big|^{2p}\Big]^{1/2p} \le \mathbb{E}\Big[\big(Z_n^{\Psi_1}\big)^{2p}\Big]^{1/2p} + \mathbb{E}\Big[\big(Z_n^{\Psi_2}\big)^{2p}\Big]^{1/2p} + o(\rho^{n/2}). \tag{29}$$

We estimate each of the two terms on the right-hand side separately. In view of Lemma 4, there is a constant $C > 0$ such that

$$\mathbb{E}\left[\|\Psi_1(k)\|^{2p}\right] \le C\Big(\mathbb{E}\left[\|\Psi_1(k)\|^2\right]\Big)^p,$$

and, in particular,

$$\sum_{k\in\mathbb{Z}} \rho^{-k}\Big(\mathbb{E}\left[\|\Psi_1(k)\|^{2p}\right]\Big)^{1/p} \le C^{1/p} \sum_{k\in\mathbb{Z}} \rho^{-k}\mathbb{E}\left[\|\Psi_1(k)\|^2\right] < \infty.$$

Theorem 5(iii) applied to Ψ_1 yields $\mathbb{E}\big[\big(Z_n^{\Psi_1}\big)^{2p}\big] = O(\rho^{np})$. In order to deal with the second term $\mathbb{E}\left[\big(Z_n^{\Psi_2}\big)^{2p}\right]$ on the left-hand side of (29), note that by the definition of ℓ we may write

$$\begin{aligned}\Psi_2(k) &= \mathbf{1}_{\{k>0\}}\mathrm{x}_2\pi^{(2)} \sum_{j=0}^{J-1} \binom{k-1}{j} D^{k-1-j}N^j(L-A)\\ &= \mathbf{1}_{\{k>0\}}\mathrm{x}_2\pi^{(2)} \sum_{j=0}^{(J-1)\wedge\ell} \binom{k-1}{j} D^{k-1-j}N^j(L-A).\end{aligned}$$

In particular, as k goes to infinity we have

$$\Big(\mathbb{E}\left[\|\Psi_2(k)\|^{2p}\right]\Big)^{1/p} = O\big(\rho^k k^{2\ell}\big), \quad \text{so} \quad \sum_{k=0}^{n} \rho^{-k}\Big(\mathbb{E}\left[\|\Psi_2(k)\|^{2p}\right]\Big)^{1/p} = O\big(n^{2\ell+1}\big),$$

and by Theorem 5(iv) applied to Ψ_2, we obtain $\mathbb{E}\big[\big(Z_n^{\Psi_2}\big)^{2p}\big] = O(n^{(2\ell+1)p}\rho^{np})$, which together with (29) proves the desired. □

For any function $f : \mathbb{R} \to \mathbb{R}$, let

$$\omega(f, k, t) := \sup\{|f(x) - f(y)| : x, y \in [-k, k],\ |x-y| \le t\}$$

be the modulus of continuity of f on the interval $[-k, k]$.

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

Lemma 3. *For a stochastic process X taking values in the Skorokhod space $\mathcal{D}(\mathbb{R})$, and for $n \in \mathbb{N}$, let $X_n(t) = X(t+n)$. Furthermore, suppose that $Y = (Y(t))_{t\in\mathbb{R}}$ is a stationary process with almost surely continuous trajectories. Then we have the following:*

(i) *If $X_n \xrightarrow{\mathrm{d}} Y$, N_n is a sequence of natural numbers diverging to infinity, and $\delta_n \searrow 0$, then there is a sequence $k_n \nearrow \infty$ such that*

$$\omega(X_{N_n}, k_n, \delta_n) \xrightarrow{\mathbb{P}} 0, \quad \text{as } n \to \infty. \tag{30}$$

(ii) *If, for some real-valued random variable S independent of Y, it holds that $(S, X_n) \xrightarrow{\mathrm{d}} (S, Y)$ as $n \to \infty$, then for any sequence a_n that diverges to infinity we have*

$$X(a_n + S) \xrightarrow{\mathrm{d}} Y(0), \quad \text{as } n \to \infty. \tag{31}$$

Proof. (i): Fix $\delta > 0$ and $k \in \mathbb{N}$. Then the mapping $\mathcal{D}(\mathbb{R}) \ni f \mapsto \omega(f, k, \delta) \in \mathbb{R}$ is continuous at any $f \in \mathcal{C}(\mathbb{R})$. The continuous mapping theorem yields

$$\lim_{n\to\infty} \mathbb{E}[\omega(X_{N_n}, k, \delta) \wedge 1] = \mathbb{E}[\omega(Y, k, \delta) \wedge 1].$$

Hence, for any $\varepsilon > 0$ we have

$$\limsup_{n\to\infty} \mathbb{E}[\omega(X_{N_n}, k, \delta_n) \wedge 1] \le \mathbb{E}[\omega(Y, k, \varepsilon) \wedge 1].$$

Since ε is arbitrary, we can take the limit as ε goes to 0, and from the continuity of Y we conclude that

$$\mathbb{E}[\omega(X_{N_n}, k, \delta_n) \wedge 1] \to 0.$$

In particular, there is $n(k) > n(k-1)$ such that for all $n \ge n(k)$ it holds that

$$\mathbb{E}[\omega(X_{N_n}, k, \delta_n) \wedge 1] \le 1/k.$$

Now for $n \ge n(1)$ we set $k_n = k$ whenever $n(k) \le n < n(k+1)$. Clearly $k_n \nearrow \infty$ and it holds that

$$\mathbb{E}[\omega(X_{N_n}, k_n, \delta_n) \wedge 1] \le 1/k_n,$$

which implies (30); this proves the first part of the claim.

(ii): The convergence in (31) holds if for any subsequence n_k we can choose a further subsequence n_{k_l} along which the convergence holds. Since we may replace the sequence a_n by a subsequence a_{n_k}, it suffices to show the convergence (31) along some subsequence. Thus, without loss of generality, we may assume that $\{a_n\} \to a$ for some $a \in [0, 1]$. For $N_n = \lfloor a_n \rfloor$ and $\delta_n = 2|\{a_n\} - a|$, we infer from Part (i) the existence of a sequence $k_n \nearrow \infty$ such that (30) holds.

For $Z_n = X(N_n + \{a_n\} + S) - X(N_n + a + S)$, once again by Part (i) of the proof, we have

$$|Z_n| \le |Z_n| \mathbf{1}_{\{|S| \le k_n\}} + |Z_n| \mathbf{1}_{\{|S| > k_n\}} \le \omega(X_{N_n}, k_n, \delta_n) + |Z_n| \mathbf{1}_{\{|S| > k_n\}} \xrightarrow{\mathbb{P}} 0,$$

and thus, by Slutsky's theorem, it suffices to prove

$$X_{N_n}(a + S) = X(N_n + a + S) \xrightarrow{\mathrm{d}} Y(0).$$

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

Next, observe that the mapping $\mathbb{R}\times\mathcal{D}(\mathbb{R})\ni(s,x)\mapsto x(a+s)\in\mathbb{R}$ is continuous at any point $(s,x)\in\mathbb{R}\times\mathcal{C}(\mathbb{R})$. Therefore, by the continuous mapping theorem we have

$$X_{N_n}(a+S)\xrightarrow{\text{d}}Y(a+S)\overset{\text{D}}{=}Y(0),$$

and this completes the proof. $\square$

Lemma 4. *Suppose that X is a random $k\times m$ matrix with $\mathbb{E}[\|X\|^r]<\infty$ for some $r>1$. Then, for any $q<r$, there is a constant $C=C(m,k,q,r,X)>0$ such that for any $m\times k$ deterministic matrix A it holds that*

$$\mathbb{E}[\|AX\|^r]\le C\mathbb{E}[\|AX\|^q]^{r/q}.$$

Proof. Without loss of generality, by the homogeneity of both sides, we may also assume that $\|A\|=1$. Now let $N=\{a\in\mathbb{R}^{m\times k}: aX=0\ \text{ almost surely}\}$ be a linear subspace of $\mathbb{R}^{m\times k}$ and V its orthogonal complement. As both functions

$$a\mapsto\mathbb{E}[\|aX\|^r]\quad\text{and}\quad a\mapsto\mathbb{E}[\|aX\|^q]^{r/q}$$

defined on the compact space $V\cap\{\|x\|=1\}$ are continuous and do not vanish, they achieve their minimum and maximum. We define

$$C=\frac{\max_{a\in V,\|a\|=1}\mathbb{E}[\|aX\|^r]}{\min_{a\in V,\|a\|=1}\mathbb{E}[\|aX\|^q]^{r/q}}<\infty.$$

Finally, by writing $A=A_1+A_2$ with $A_1\in N$ and $A_2\in V$, we obtain

$$\mathbb{E}[\|AX\|^r]\le\mathbb{E}[\|A_2X\|^r]\le C\mathbb{E}[\|A_2X\|^q]^{r/q}\le C\mathbb{E}[\|AX\|^q]^{r/q},$$

and this completes the proof. $\square$

Lemma 5. *Let I, J be two disjoint subintervals of $[0,1]$, let $N\in\mathbb{N}$, and let $U_1,\dots U_N$ be an independent collection of random variables uniformly distributed on $[0,1]$. Then for any sequence $a\in\ell^2$ and any numbers $A,B\in\mathbb{R}$, we have*

$$\begin{aligned}\mathbb{E}\Big[\Big|\sum_{i=1}^{N}(\mathbf{1}_{\{U_i\in I\}}-|I|)a_i+A\Big|^2\cdot\Big|\sum_{i=1}^{N}(\mathbf{1}_{\{U_i\in J\}}-|J|)a_i+B\Big|^2\Big]&\\ \le C|I||J|\|a\|^2(A^2+B^2+\|a\|^2)+A^2B^2&\\ \lesssim|I||J|\|a\|^4+A^4+B^4,&\end{aligned}$$

for some absolute constant $C>0$.

Proof. For ease of notation, for $i=1,\dots,N$ we set

$$q_i=\big(\mathbf{1}_{\{U_i\in I\}}-|I|\big)a_i\quad\text{and } r_i=\big(\mathbf{1}_{\{U_i\in J\}}-|J|\big)a_i,$$

so $\mathbb{E}[q_i]=\mathbb{E}[r_i]=\mathbb{E}[q_iq_j]=\mathbb{E}[r_ir_j]=0$, for $i\ne j$. Simple calculations give

$$\mathbb{E}[q_i^2]=\big(|I|-|I|^2\big)a_i^2\quad\text{and}\quad\mathbb{E}[q_ir_i]=-|I||J|a_i^2,$$

$$\mathbb{E}[q_i^2r_i]=\big(-|I||J|+2|I|^2|J|\big)a_i^3\quad\text{and}\quad\mathbb{E}[q_i^2r_i^2]=\big(|I||J|(|I|-3|I||J|+|J|)\big)a_i^4.$$

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

We first have
$$E\Big[\Big(\sum_{i\le N} q_i + A\Big)^2\Big] = (|I| - |I|^2)N + A^2 \le |I|N + A^2. \tag{32}$$
Expanding the expectation, we get

$$\mathbb{E}\Big[\Big|\sum_{i\le N} q_i + A\Big|^2 \cdot \Big|\sum_{i\le N} r_i + B\Big|^2\Big] = \sum_{i_1,i_2,j_1,j_2} \mathbb{E}[q_{i_1}q_{i_2}r_{j_1}r_{j_2}] + 2A\sum_{i,j_1,j_2}\mathbb{E}[q_i r_{j_1} r_{j_2}]$$
$$+2B\sum_{i_1,i_2,j}\mathbb{E}[q_{i_1}q_{i_2}r_j] + 4AB\sum_{i,j}\mathbb{E}[q_i r_j] + A^2B^2 =: I + II + III + IV + A^2B^2.$$

We show that each of the terms I, II, III, IV is bounded by a multiple of the term $|I||J|N(A^2 + B^2 + N)$. Note that a nontrivial term of the form $\mathbb{E}[q_{i_1}q_{i_2}r_{j_1}r_{j_2}]$ is either of the form $E[q_i^2 r_j^2]$ or of the form $\mathbb{E}[q_{i_1}q_{i_2}r_{i_1}r_{i_2}]$, which in turn implies

$$I = \sum_{i,j}\mathbb{E}[q_i^2 r_j^2] + 4\sum_{i\ne j}\mathbb{E}[q_i r_i]\mathbb{E}[q_j r_j] = |I||J|(|I| - 3|I||J| + |J|)\sum a_i^4$$
$$+ 2(|I| - |I|^2)(|J| - |J|^2)\sum_{i\ne j} a_i^2 a_j^2 + 4|I|^2|J|^2\sum_{i\ne j} a_i^2 a_j^2 \le 8|I||J|\|a\|^4.$$

Next, $\mathbb{E}[q_{i_1}q_{i_2}r_j]$ is nonzero if it is of the form $\mathbb{E}[q_i^2 r_i]$. Hence, we have

$$III = 2B\sum_i \mathbb{E}[q_i^2 r_i] = 2B(-|I||J| + 2|I|^2|J|)\sum_i a_i^3 = 2B|I||J|(2|I| - 1)\sum_i a_i^3$$
$$\le 4|I||J||B|\|a\|^3.$$

By symmetry, we have

$$II = 2A(-|I||J| + 2|I||J|^2)\sum_i a_i^3 = 2A|I||J|(2|J| - 1)\sum_i a_i^3 \le 4|I||J||A|\|a\|^3.$$

Finally, by the same reasoning as above, we get

$$IV = -4AB|I||J|\sum_i a_i^2 \le 4|AB||I||J|\|a\|^2,$$

and the claim follows from putting together the four quantities. □

Lemma 6. *Let $l, N \in \mathbb{N}$ and let $\lambda_1, \dots, \lambda_\ell$ be different, nonzero complex numbers. For $(i, j) \in \mathbb{N}_0 \times [\ell]$ we define $f_{i,j} : \mathbb{Z} \to \mathbb{C}$ by $f_{i,j}(k) = k^i \lambda_j^k$. Then the collection of functions $\{f_{i,j} : (i, j) \in \mathbb{N}_0 \times [\ell]\}$ is linearly independent. In particular, if, for any $i \in \mathbb{N}_0$ and $j \le \ell$, $p_{j,i}$ is a polynomial of degree i, then the collection of functions $\{k \mapsto \lambda_j^k p_{j,i}(k) : (i, j) \in \mathbb{N}_0 \times [\ell]\}$ is also linearly independent.*

Proof. Let $h = \sum_{i,j} c_{i,j} f_{i,j}$ be a finite linear combination of the functions $f_{i,j}$. Our aim is to show that

$$\text{if } h(k) = 0 \text{ for } k \ge N, \text{ then } c_{i,j} = 0 \text{ for all } i, j. \tag{33}$$

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press

By $d_j(h)$ we denote the maximal power i such that the element $f_{i-1,j}$ appears in the combination of h. Furthermore, we set $d(h) = d_1(h) + \ldots d_\ell(h)$. We prove the claim by induction on $d(h)$. If $d(h) = 1$ then $h(k) = \lambda_j^k$ for some j and the conclusion follows. Now suppose that (33) holds for any h with $d(h) = n$ and take h with $d(h) = n + 1$. If $d_j(h) \leq 1$ for all $j \leq \ell$, then $h(k) = \sum_{m=1}^{n+1} c_{j_m} \lambda_{j_m}^k$ for some $j_1, \ldots, j_{n+1} \leq n + 1$. Since $\left(\lambda_{j_m}^{-N+1} f_{0,j_m}(k)\right)_{1 \leq k,m \leq n+1}$ forms a Vandermonde matrix, this implies (33).

Therefore, we may now assume that for some $j_0 \leq \ell$ we have $d_{j_0}(h) \geq 2$. We denote by ∇ the difference operator defined by $\nabla f(k) = f(k+1) - f(k)$, and m_{j_0} is defined by $m_{j_0} f(k) = \lambda_{j_0}^k f(k)$. We now define a linear operator $\nabla_{j_0} = m_{j_0} \nabla m_{j_0}^{-1}$. Clearly ∇_{j_0} acts on the linear combinations g of $f_{i,j}$ with $d_j(\nabla_{j_0} g) \leq d_j(g)$ and also $d_{j_0}(\nabla_{j_0} f_{i,j_0}) = d_{j_0}(f_{i,j_0}) - 1$. In particular, $1 \leq \nabla_{j_0} h \leq n$, and hence by the induction hypothesis $\nabla_{j_0} h(k) \neq 0$ for some $k \geq N$, which finally implies that $h(k) \neq 0$ or $h(k+1) \neq 0$, thus proving (33). □

Acknowledgements

We are very grateful to the two anonymous referees for their suggestions and positive criticism, which substantially improved the quality and the presentation of the paper.

Funding information

This research was funded in part by the Austrian Science Fund (FWF) 10.55776/P34129. For open access purposes, the authors have applied a CC BY public copyright license to any author-accepted manuscript version arising from this submission.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] ALDOUS, D. J. AND EAGLESON, G. K. (1978). On mixing and stability of limit theorems. *Ann. Prob.* **6**, 325–331.

[2] ATHREYA, K. B. AND KARLIN, S. (1968). Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *Ann. Math. Statist.* **39**, 1801–1817.

[3] BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, 2nd edn. John Wiley, New York.

[4] BURKHOLDER, D., DAVIS, B. AND GUNDY, R. (1972). Integral inequalities for convex functions of operators on martingales. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 2, University of California Press, Berkeley, CA, pp. 223–240.

[5] JANSON, S. (2004). Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stoch. Process. Appl.* **110**, 177–245.

[6] JANSON, S., MAILLER, C. AND VILLEMONAIS, D. (2023). Fluctuations of balanced urns with infinitely many colours. *Electron. J. Prob.* 28, article no. 82.

[7] KESTEN, H. AND STIGUM, B. P. (1966). A limit theorem for multidimensional Galton–Watson processes. *Ann. Math. Statist.* **37**, 1211–1223.

[8] KOLESKO, K. AND SAVA-HUSS, E. (2023). Limit theorems for discrete multitype branching processes counted with a characteristic. *Stoch. Process. Appl.* **162**, 49–75.

[9] MÜLLER, N. AND NEININGER, R. (2018). Refined asymptotics for the composition of cyclic urns. *Electron. J. Prob.* 23, article no. 117.

[10] PÓLYA, G. (1930). Sur quelques points de la théorie des probabilités. *Ann. Inst. H. Poincaré Prob. Statist.* **1**, 117–161.

[11] POUYANNE, N. (2008). An algebraic approach to Pólya processes. *Ann. Inst. H. Poincaré Prob. Statist.* **44**, 293–323.
[12] SMYTHE, R. T. (1996). Central limit theorems for urn models. *Stoch. Process. Appl.* **65**, 115–137.
[13] TOPCHIĬ V. A. AND VATUTIN, V. A. (1997). The maximum of critical Galton–Watson processes, and left-continuous random walks. *Teor. Veroyat. Primen.* **42**, 21–34.

https://doi.org/10.1017/apr.2024.37 Published online by Cambridge University Press