Electronic Journal of Probability

Electron. J. Probab. **23** (2018), no. 34, 1–13.
ISSN: 1083-6489 https://doi.org/10.1214/17-EJP80

Moment convergence of balanced Pólya processes

Svante Janson* Nicolas Pouyanne†

Abstract

It is known that in an irreducible small Pólya urn process, the composition of the urn after suitable normalization converges in distribution to a normal distribution. We show that if the urn also is balanced, this normal convergence holds with convergence of all moments, thus giving asymptotics of (central) moments.

Keywords: Pólya urns; Pólya processes; moment convergence.
AMS MSC 2010: 60C05.
Submitted to EJP on June 22, 2016, final version accepted on June 30, 2017.

1 Introduction

A Pólya urn process is defined as follows. Consider an urn containing balls of different colours, with s possible colours which we label $1\dots,s$. At each time step, we draw a ball at random from the urn; we then replace it and, if its colour was i, we add r_{ij} further balls of colour j, for each $j=1,\dots,s$. Here

$$R := (r_{ij})_{i,j=1}^{s} \tag{1.1}$$

is a given matrix, called the *replacement matrix*. The state of the urn at time n is described by a vector $X_n=(X_{n1},\dots,X_{ns})$, where X_{nj} is the number of balls of colour j. We start with some given (deterministic) X_0, and it is clear that X_n evolves according to a Markov process.

As usual, we assume that $r_{ij}\geqslant 0$ when $i\neq j$, but we allow r_{ii} to be negative, meaning removal of balls, provided the urn is *tenable*, i.e., that it is impossible to get stuck. (See (2.2)–(2.3), and see Remark 1.8 for an extension that allows some negative r_{ij}.)

Urn processes of this type have been studied by many different authors, with varying generality, going back to Eggenberger and Pólya [5]; see for example Janson [8], Flajolet, Gabarró and Pekari [6], Pouyanne [14], Mahmoud [12], and the further references given there.

*Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden. E-mail: svante.janson@math.uu.se.

†Laboratoire de Mathématiques de Versailles, UVSQ, CNRS, Université Paris-Saclay, 78035 Versailles, France. E-mail: nicolas.pouyanne@uvsq.fr.

In the present paper we study only the *balanced* case, meaning that the total number of balls added each time is deterministic, i.e., that the row sums of the matrix (1.1) are constant, say m; we assume further that $m > 0$.

We define, for an arbitrary vector $(x_1, \dots, x_n)$, $|(x_1, \dots, x_n)| := \sum_{i=1}^n |x_i|$. In particular, the total number of balls in the urn is $|X_n|$. Note that when the urn is balanced, this number is deterministic, with $|X_n| = |X_0| + nm$.

In the description above, it is implicit that the numbers r_{ij} are integers. However, it has been noted many times that the process is also well-defined for *real* r_{ij}, see e.g. [8, Remark 4.2], [9] and [14] (cf. also [11] for the related case of branching processes); this can be interpreted as an urn containing a certain amount (mass) of each colour, rather than discrete balls. We give a detailed definition of this, more general, version in Section 2, and use it in our results below.

Results on the asymptotic distribution of X_n as $n \to \infty$ have been given by many authors under varying assumptions, using different methods. It is well-known that the asymptotic behaviour of X_n depends on the eigenvalues of R, or equivalently of its transpose $A = R^t$, see e.g. [8, Theorems 3.22–3.24]. By the Perron–Frobenius theory of positive matrices (applied to $R + cI$ for some $c \geqslant 0$), R has a largest real eigenvalue λ_1, and all other eigenvalues λ satisfy $\operatorname{Re}\lambda < \lambda_1$. We say that an eigenvalue λ is *large* if $\operatorname{Re}\lambda > \frac{1}{2}\lambda_1$, *small* if $\operatorname{Re}\lambda \leqslant \frac{1}{2}\lambda_1$ and *strictly small* if $\operatorname{Re}\lambda < \frac{1}{2}\lambda_1$. Similarly, we say that the Pólya process (or urn) is *small* (*strictly small*) if λ_1 is simple and all other eigenvalues are small (strictly small); a process is large whenever it is not small. We call a Pólya process *critically small* if it is small but not strictly small, i.e., if the process is small and R admits an eigenvalue λ such that $\operatorname{Re}\lambda = \lambda_1/2$. We define, letting Λ be the set of eigenvalues,

$$\sigma_2 := \begin{cases} \max\bigl\{\operatorname{Re}\lambda : \lambda \in \Lambda \setminus \{\lambda_1\}\bigr\}, & \lambda_1 \text{ is a simple eigenvalue;} \\ \lambda_1, & \lambda_1 \text{ is not simple.} \end{cases} \tag{1.2}$$

Thus the Pólya urn is strictly small if $\sigma_2 < \lambda_1/2$, critically small if $\sigma_2 = \lambda_1/2$, and large if $\sigma_2 > \lambda_1/2$.

In the main results we assume that the urn is irreducible, i.e., that the matrix R is irreducible. (In other words, every colour is dominating in the sense of [8].) Then, the largest eigenvalue λ_1 is simple. (Thus the second case in (1.2) does not occur.) As said above, we also assume the urn to be balanced, with all row sums of R equal to m, and then $\lambda_1 = m$, with a corresponding right eigenvector $(1, \dots, 1)$. Furthermore, there exists a positive left eigenvector v_1 of R with eigenvalue m; we assume that v_1 is normalized by $|v_1| = 1$, and then v_1 is unique.

If the urn is irreducible and small, then X_n is asymptotically normal [8, Theorems 3.22–3.23]. More precisely, if v_1 is the positive eigenvector of R defined above, and $\nu = 0$ if the urn is strictly small and $\nu \geqslant 1$ is the integer defined in Theorem 1.2 below if the urn is critically small, then, as $n \to \infty$,

$$\frac{X_n - n\lambda_1 v_1}{\sqrt{n \log^\nu n}} \xrightarrow{\mathrm{d}} N(0, \Sigma), \tag{1.3}$$

where the asymptotic covariance matrix Σ can be computed from R. (See e.g. [8, Lemma 5.3 and Lemma 5.4 with (2.15) and (2.17)].) On the other hand, by [8, Theorems 3.24] and, in particular, [14, Theorems 3.5–3.6], if the urn is large, then there exist (complex) random variables W_k, (complex) left eigenvectors v_k of R and an integer $\nu \geqslant 0$ such that, a.s. and in any L^p,

$$X_n = n\lambda_1 v_1 + \sum_{k:\operatorname{Re}\lambda_k = \sigma_2} n^{\lambda_k/\lambda_1} \log^\nu n \, W_k v_k + o\bigl(n^{\sigma_2/\lambda_1} \log^\nu n\bigr). \tag{1.4}$$

In general, there will be oscillations (coming from complex eigenvalues λ_k) and X_n will not converge in distribution (after any non-trivial normalization). Mixed moments of the limit distributions W_k in (1.4) can be computed, see [14]. However, there is in general no explicit description of the limit laws for a large urn. See [2], [4], [3] and Mailler [13] for some recent improvements on these distributions. Note also that (1.4) is valid as soon as the urn is large and λ_1 a simple eigenvalue, the urn being irreducible or not (see [14]).

Results of this type have been proven by several authors, under varying assumptions, using several different methods. The proofs in Janson [8] use an embedding in a continuous-time multi-type branching process, a method that was introduced by Athreya and Karlin [1]. This method leads to general results on convergence in distribution, but not to results on the moments. A different method was developed by Pouyanne [14], where algebraic expressions were obtained for (mixed) moments of various components of X_n, and asymptotics were derived. For large urns, the resulting moment estimates and some simple martingale arguments give the limit results, with convergence a.s. and in L^p, and thus convergence of all moments (after suitable normalization). The method applies also to small urns, and yields limits for the moments. In principle, it should be possible to use the resulting expressions and the method of moments to show (1.3). However, the expressions for the limits are a bit involved, and it seems difficult to do this in general.

The purpose of the present paper is to show moment convergence for small urns by combining these two methods. We use the convergence in distribution (1.3) proven in [8], and we use the estimates of moments proven in [14] to show that any moment of the left-hand side of (1.3) is bounded as $n \to \infty$; these together imply moment convergence in (1.3). (We thus do not have to calculate the limits provided by [14] exactly; it suffices to find bounds of the right order of magnitude.) This yields the following theorems, which are our main results.

All limits and $o(\dots)$ in this paper are as $n \to \infty$.

Theorem 1.1. *Suppose that the urn is balanced, irreducible and strictly small. Then* (1.3) *holds, with $\nu = 0$, with convergence of all moments. In particular, $\mathbb{E}\, X_n = n\lambda_1 v_1 + o\big(n^{1/2}\big)$ and the covariance matrix $\operatorname{Var}(X_n) = n\Sigma + o(n)$.*

Theorem 1.2. *Suppose that the urn is balanced, irreducible and critically small. Let $1 + d$ be the dimension of the largest Jordan block of R corresponding to an eigenvalue λ with $\operatorname{Re}\lambda = \lambda_1/2$ ($d \geqslant 0$). Then* (1.3) *holds, with $\nu = 2d + 1$, with convergence of all moments. In particular, $\mathbb{E}\, X_n = n\lambda_1 v_1 + o\big((n \log^\nu n)^{1/2}\big)$ and the covariance matrix $\operatorname{Var}(X_n) = \big(n \log^\nu n\big)\Sigma + o(n \log^\nu n)$.*

Corollary 1.3. *Suppose that the urn is balanced, irreducible and small, so* (1.3) *holds. Let $w = (w_1, \dots, w_s)$ be any vector in $\mathbb{R}^s$ and let $Y_n := \langle w, X_n\rangle = \sum_{i=1}^s w_i X_{ni}$. Then $\mathbb{E}\, Y_n = n\lambda_1\langle w, v_1\rangle + o\big((n \log^\nu n)^{1/2}\big)$ and $\operatorname{Var} Y_n = \big(\gamma + o(1)\big) n \log^\nu n$, where $\gamma = w^t\Sigma w$. Moreover, if $\gamma \neq 0$, then*

$$\frac{Y_n - \mathbb{E}\, Y_n}{\sqrt{\operatorname{Var} Y_n}} \xrightarrow{\mathrm{d}} N(0,1) \tag{1.5}$$

with convergence of all moments.

The remainder of this section is devoted to remarks and problems that can be skipped on a first reading.

Remark 1.4. For the mean and variance, similar results are also proven in [10] by a related but somewhat different method (under somewhat more general assumptions); that method does not seem to generalise easily to higher moments.

Remark 1.5. If the urn is strictly small, then it can be verified from [8, Lemma 5.4 and (2.13)–(2.15)] that $\gamma = 0$ in Corollary 1.3 only in the trivial case when $w = cu_1 + u_0$ with

$c \in \mathbb{R}$, $u_1 = (1, \dots, 1)$ and $Ru_0 = 0$, which implies that $\langle u_0, X_n \rangle$ is constant and thus $Y_n = \langle w, X_n \rangle = Y_0 + ncm$ is deterministic, see [10, Theorem 3.6].

On the other hand, in the critically small case, the rank of Σ is typically only 1 or 2, and there are non-trivial vectors w such that $\gamma = 0$ and thus $\mathrm{Var}(Y_n) = o(n \log^\nu n)$.

Remark 1.6. More precise error estimates in Theorems 1.1 and 1.2 can be obtained from the proofs below. In particular, for the expectation we have in the strictly small case $\mathbb{E}\, X_n = n\lambda_1 v_1 + O\big(n^{\sigma_2/\lambda_1} \log^{\nu_1} n\big) + O(1)$ for some ν_1. See also [10].

Remark 1.7. It is possible to let balls of different colours have different activities, say $a_i \geqslant 0$ for balls of colour i, with the probability of a ball being drawn proportional to its activity [8]. The condition that the urn is balanced is now that the total activity added each time is a constant. In the case when all activities are positive, this is easily reduced to the standard case $a_i = 1$ by using the real version above; we just multiply the number of balls of colour i by a_i (both in the urn and in the replacement matrix). In general, where there are "dummy balls" of activity 0, which thus never are drawn (see e.g. [8] for the use of such balls), the results above still hold, assuming that the urn is irreducible if dummy balls are ignored. (Note that we get another Pólya process by ignoring dummy balls, and that the non-zero eigenvalues remain the same.) This can be shown by the same proofs as given below; we only have to modify the definitions of balanced in (2.4) and of A and Φ in (2.5) and (2.6) by replacing the vectors ℓ_k used there by $a_k \ell_k$, and note that it is easy to verify that the results in [14] still hold (with the corresponding modification of Φ_∂ defined there).

Remark 1.8. The condition $r_{ij} \geqslant 0$ when $i \neq j$ (and (2.2)–(2.3) below) is customary but can be relaxed if we assume that the urn is tenable for some other reason. (Typically because balls of two different colours always occur together in a fixed proportion, and are added or subtracted together.) See [14, Example 7.2.(5)] for a typical example and [7, Remark 6.3] for another. As remarked in [14, page 295], the results in [14] that we use hold in this case too, and it follows that all moment estimates in the present paper hold. Also (1.3) holds, at least under some supplementary assumptions, see [8, Remark 4.2], and then the results above hold. (In the examples from [14] and [7] just mentioned, (1.3) holds because there is an equivalent urn with random replacements that satisfies the conditions of [8].)

Remark 1.9. It is possible to let the replacement vectors $(r_{ij})_{j=1}^s$ be random, see [8]: with our notations of Section 2, assume that random V-valued increment vectors $W_1, \dots, W_s$ are given and that they admit moments of order p, $p \geq 2$ being an integer or ∞. In this case, the conditional transition probabilities (2.1) keep the same form, and $X_{n+1} = X_n + W_K^{(n)}$, where, given $K = k$, $W_K^{(n)}$ is a copy of W_k, independent of everything that has happened so far. The tenability assumptions (2.2)–(2.3) must be modified: it is sufficient that $\ell_j(W_k) \geqslant 0$ a.s. for all j, k; more generally (2.2) should hold a.s., while for each k either $\ell_k(W_k) \geqslant 0$ a.s. or there exists $d_k > 0$ such that a.s. $\ell_k(W_k) \in \{-d, 0, d, \dots\}$ while $\ell_k(X_0), \ell_k(W_i) \in \{d, 2d, \dots\}$ for $j \neq k$. Assume further that the urn is *almost surely balanced*, which means that (2.4) is a.s. satisfied (replacing w_k by W_k).

Then, our results extend to this case, the moment convergence being valid up to order p.

To see this, note first that in this random replacement context, all results of [8] hold. The techniques developed in [14] and the arguments given in the present paper remain also valid after the following adaptations: the replacement operator (2.5) is now

$$A(v) := \sum_{k=1}^{s} \ell_k(v)\, \mathbb{E}\, W_k \tag{1.6}$$

while the transition operator (2.6), restricted to polynomials f of degree not more than p, becomes

$$\Phi(f)(v) := \sum_{k=1}^{s} \ell_k(v)\, \mathbb{E}\big(f(v+W_k) - f(v)\big). \tag{1.7}$$

Remark 1.10. For an example of applications of the results above on random tree processes (m-ary search trees and preferential attachment trees), one can refer to [7, Remark 3.3].

Problem 1.11. As said above, we consider in this paper only balanced urns. It is a challenging open problem to extend the results to non-balanced urns.

2 Preliminaries

We follow [14] and use the following coordinate-free description of the urn process. It is easily seen to be equivalent to the traditional description in Section 1, with $r_{ij} = \ell_j(w_i)$ and allowing these numbers to be real and not necessarily integers.

Let V be a real vector space of finite dimension $s \geqslant 1$ and let $\ell_1, \dots, \ell_s$ be a basis of the dual space V'; let $V_+ := \{v \in V : \ell_j(v) \geqslant 0, j = 1, \dots, s\} \setminus \{0\}$ be the positive orthant. Let X_0 and $w_1, \dots, w_s$ be given vectors in V, with $X_0 \in V_+$.

Given $X_n \in V_+$, for some $n \geqslant 0$, we let $X_{n+1} := X_n + w_K$, where the random index K is chosen with conditional probability, given X_n,

$$\mathbb{P}(K = k \mid X_n) = \frac{\ell_k(X_n)}{\sum_{j=1}^{s} \ell_j(X_n)}. \tag{2.1}$$

This defines the Pólya process $(X_n)_0^\infty$ (as a Markov process), provided the process is *tenable*, i.e., $X_n \in V_+$ for all n.

The standard sufficient set of conditions for tenability, used by many authors, is in our formulation: for all $j, k = 1, \dots, s$,

$$\ell_j(w_k) \geqslant 0 \quad \text{if } j \neq k, \tag{2.2}$$

$$\ell_k(w_k) \geqslant 0 \quad \text{or} \quad \ell_k(X_0)\mathbb{Z} + \sum_{i=1}^{s} \ell_k(w_i)\mathbb{Z} = \ell_k(w_k)\mathbb{Z}. \tag{2.3}$$

We assume (2.2)–(2.3) for simplicity, but as said in Remark 1.8, the results hold more generally under suitable conditions.

In the present paper, we also assume that the process is *balanced*, which in this context means

$$\sum_{k=1}^{s} \ell_k(w_j) = m, \qquad j = 1, \dots, s, \tag{2.4}$$

for some fixed m. We assume further $m > 0$, and we may without loss of generality assume $m = 1$, since we may divide all X_n and w_k (or, alternatively, all ℓ_j) by m.

We shall also use the following notation from [14], where further details are given.

The replacement matrix R (or rather its transpose) now corresponds to the *replacement operator* $A : V \to V$ defined by

$$A(v) := \sum_{k=1}^{s} \ell_k(v) w_k. \tag{2.5}$$

We choose a basis $(v_k)_1^s$ in the complexification $V_{\mathbb{C}}$ that yields a Jordan block decomposition of A, and let $(u_k)_1^s$ be the corresponding dual basis in $V'_{\mathbb{C}}$. We may assume that these vectors are numbered such that u_1 and v_1 correspond to the eigenvalue $\lambda_1 = m = 1$, and, moreover, for each k either $u_k \circ A = \lambda_k u_k$ (so u_k is an eigenvector of the dual operator A') or $u_k \circ A = \lambda_k u_k + u_{k-1}$, for some eigenvalue λ_k. Since the urn is supposed to be irreducible, $\lambda_1 = 1$ is a simple eigenvalue; furthermore, the balance condition (2.4) (with $m = 1$) implies that $\sum_{j=1}^s \ell_j \in V'$ is an eigenvector of A' with eigenvalue 1; hence we may assume that $u_1 = \sum_{j=1}^s \ell_j$. This means that v_1 is normalized by $\sum_{j=1}^s \ell_j(v_1) = 1$.

Let $\lambda := (\lambda_1, \dots, \lambda_s)$, the vector of eigenvalues of A (repeated according to algebraic multiplicity).

Let π_k denote the projection of $V_{\mathbb{C}}$ onto $\mathbb{C}v_k$ defined by $\pi_k(v) := u_k(v)v_k$. Note that $\sum_{k=1}^s \pi_k = I$.

For a multi-index $\alpha = (\alpha_1, \dots, \alpha_s) \in \mathbb{Z}^s_{\geqslant 0}$, let $\mathbf{u}^\alpha := \prod_{i=1}^s u_i^{\alpha_i}$; this is a homogeneous polynomial function on $V_{\mathbb{C}}^s$. We call such multi-indices α *powers*, and we say that α is a *small power* if only linear forms u_i corresponding to small eigenvalues appear in $\mathbf{u}^\alpha$, i.e., if $\operatorname{Re}\lambda_i \leqslant \frac{1}{2}$ when $\alpha_i > 0$; we define *strictly small power* in the same way.

Let Φ be the linear operator in the space of (complex-valued) functions on V defined by

$$\Phi(f)(v) := \sum_{k=1}^{s} \ell_k(v)\big(f(v + w_k) - f(v)\big). \tag{2.6}$$

Then, using (2.1), $\mathbb{E}\, f(X_{n+1} \mid X_n) = f(X_n) + \sum_{j=1}^s \ell_j(X_n) \cdot \Phi(f)(X_n)$, and thus the expected evolution of any function f of X_n is described by Φ. Note also that Φ is the infinitesimal generator of the Markov branching process defined by $(X_n)_n$ after embedding in continuous time (see [1, 8, 2, 3]).

We order the multi-indices by the *degree-antialphabetic order*, see [14], and define $S_\alpha := \operatorname{span}\{\mathbf{u}^\beta : \beta \leqslant \alpha\}$. Then S_α is a finite-dimensional space of polynomials, and S_α is Φ-stable [14, Proposition 3.1]. Thus S_α has a decomposition into generalized eigenspaces $\ker(\Phi - z)^\infty := \bigcup_n \ker(\Phi - z)^n$, and we define the *reduced polynomial* Q_α as the projection of $\mathbf{u}^\alpha$ onto $\ker(\Phi - \langle\lambda, \alpha\rangle)^\infty$ in this decomposition. Then, for any $\alpha \in \mathbb{Z}^s_{\geqslant 0}$, $\{Q_\beta : \beta \leqslant \alpha\}$ is a basis in S_α [14, Proposition 4.8(2)]. Furthermore, the following statement follows from the more precise [14, Proposition 5.1].

When α is any power, we denote by ν_α the index of nilpotence of Q_α for $\Phi - \langle\lambda, \alpha\rangle$, defined by

$$1 + \nu_\alpha = \min\big\{p \geqslant 1 : \big(\Phi - \langle\lambda, \alpha\rangle\big)^p\big(Q_\alpha\big) = 0\big\}. \tag{2.7}$$

Since Q_α belongs to the generalized eigenspace space $\ker\big(\Phi - \langle\lambda, \alpha\rangle\big)^\infty$, this index is finite. In particular, $\nu_\alpha = 0$ if and only if Q_α is an eigenfunction of Φ.

Proposition 2.1. *For any* $\alpha \in \mathbb{Z}^s_{\geqslant 0}$,

$$\mathbb{E}\, Q_\alpha(X_n) = O\big(n^{\operatorname{Re}\langle\lambda,\alpha\rangle} \log^{\nu_\alpha} n\big), \tag{2.8}$$

where ν_α *is the index of nilpotence of* Q_α *defined in* (2.7).

Our proofs use the whole machinery of [14]. We define a polyhedral cone Σ and, for every power α, a polyhedron A_α (to be precise, the set of integer points in a convex polyhedron). Let δ_j denote the multi-index α with $\alpha_i = \delta_{ij}$, i.e., a single 1 in the j-th place. The cone[1] Σ can be defined by its spanning edges, as the Minkowski sum

$$\Sigma := \sum_{(i,j) \in \{1,\dots s\}^2,\ i \neq j} \mathbb{R}_{\geqslant 0}\,(2\delta_i - \delta_j) \tag{2.9}$$

or equivalently as an intersection of half-spaces:

$$\Sigma := \bigcap_{I \subseteq \{1,\dots,s\}} \{x \in \mathbb{R}^s : \delta_I^*(x) \geqslant 0\} \tag{2.10}$$

[1] There should be no risk of confusion with the covariance matrix Σ in (1.3); we denote this cone too by Σ in order to fit with the notation in [14].

where

$$\delta_I^*(x_1,\dots,x_s)=\sum_{1\leqslant i\leqslant s}x_i+\sum_{i\in I}x_i \tag{2.11}$$

for every subset I of $\{1,\dots,s\}$; the equivalence between the two definitions is proven in [14]. (Moreover, it suffices to consider I with $1\leqslant \#I\leqslant s-1$ in (2.10); these I correspond to the faces of Σ, see [14].)

When $\alpha\in\mathbb{Z}_{\geqslant 0}^s$, the polyhedron A_α is defined as

$$A_\alpha=(\alpha-D_\alpha)\cap\mathbb{Z}_{\geqslant 0}^s \tag{2.12}$$

where $\alpha-D_\alpha$ denotes $\{\alpha-d: d\in D_\alpha\}$ and D_α is[2] the set of $\mathbb{Z}_{\geq 0}$-linear combinations of all vectors $\delta_k-\delta_{k-1}$ such that u_k is not an eigenfunction of A'. Note that for such k, $\lambda_{k-1}=\lambda_k$; hence, if $\alpha'\in A_\alpha$, then

$$\sum_{k:\lambda_k=z}\alpha'_k=\sum_{k:\lambda_k=z}\alpha_k \qquad \text{for every } z\in\mathbb{C}; \tag{2.13}$$

as a consequence, $|\alpha'|=|\alpha|$ and $\langle\lambda,\alpha'\rangle=\langle\lambda,\alpha\rangle$. Note also that always $\alpha\in A_\alpha$, and that if A is diagonalizable, then $D_\alpha=\{0\}$, and thus $A_\alpha=\{\alpha\}$.

We use the following theorem, proven in [14]. It describes more precisely the action of Φ on the generalized eigenspace $\ker\left(\Phi-\langle\lambda,\alpha\rangle\right)^\infty$, which has $\{Q_\beta:\langle\lambda,\beta\rangle=\langle\lambda,\alpha\rangle\}$ as a basis. $A_\alpha-\Sigma$ denotes $\{\alpha'-\sigma:\alpha'\in A_\alpha,\ \sigma\in\Sigma\}$.

Theorem 2.2 ([14, Proposition 4.8(5) and Theorem 4.20])**.** *Let* $\alpha\in\mathbb{Z}_{\geqslant 0}^s$.

(i) $\big(\Phi-\langle\lambda,\alpha\rangle\big)\big(Q_\alpha\big)\in\operatorname{span}\{Q_\beta:\beta<\alpha,\ \langle\lambda,\beta\rangle=\langle\lambda,\alpha\rangle\}$.

(ii) *The subspace*

$$S'_\alpha:=\operatorname{span}\{\mathbf{u}^\beta:\beta\in(A_\alpha-\Sigma)\cap\mathbb{Z}_{\geqslant 0}^s\} \tag{2.14}$$

is Φ*-stable, and*

$$S'_\alpha=\operatorname{span}\{Q_\beta:\beta\in(A_\alpha-\Sigma)\cap\mathbb{Z}_{\geqslant 0}^s\}. \tag{2.15}$$

In particular, $\big(\Phi-\langle\lambda,\alpha\rangle\big)\big(Q_\alpha\big)\in S'_\alpha$.

(iii) *As a consequence,*

$$\big(\Phi-\langle\lambda,\alpha\rangle\big)\big(Q_\alpha\big)\in\operatorname{span}\big\{Q_\beta:\beta\in K_\alpha\big\}, \tag{2.16}$$

where

$$K_\alpha:=\big\{\beta\in(A_\alpha-\Sigma)\cap\mathbb{Z}_{\geqslant 0}^s:\beta<\alpha,\ \langle\lambda,\beta\rangle=\langle\lambda,\alpha\rangle\big\}. \tag{2.17}$$

3 Proofs

Recall that we for convenience, and without loss of generality, assume $\lambda_1=m=1$.

3.1 Powers and nilpotence indices

We begin with the strictly small case, which is rather simple.

Lemma 3.1. *If* α *is a strictly small power, then* $\operatorname{Re}\langle\lambda,\beta\rangle\leqslant|\alpha|/2$ *for any* $\beta\in\mathbb{Z}_{\geqslant 0}^s\cap(A_\alpha-\Sigma)$*, with equality only if* $\beta=c\delta_1$ *with* $c=|\alpha|/2$.

Proof. Let $\alpha'\in A_\alpha$ and $\sigma\in\Sigma$ such that $\beta=\alpha'-\sigma$. Also, let $I:=\{k:\operatorname{Re}\lambda_k\geqslant\frac12\}$ and recall (2.11). Since each $\beta_k\geqslant 0$ and each $\operatorname{Re}\lambda_k\leqslant 1$,

$$\operatorname{Re}\langle\lambda,\beta\rangle=\sum_k\beta_k\operatorname{Re}\lambda_k\leqslant\sum_{k:\operatorname{Re}\lambda_k<\frac12}\tfrac12\beta_k+\sum_{k:\operatorname{Re}\lambda_k\geqslant\frac12}\beta_k \tag{3.1}$$

[2]The definition of D_α corrects a minor error in [14].

$$= \tfrac{1}{2}\delta_I^*(\beta) = \tfrac{1}{2}\delta_I^*(\alpha') - \tfrac{1}{2}\delta_I^*(\sigma). \tag{3.2}$$

Since α is a strictly small power, (2.13) implies that $\alpha' \in A_\alpha$ also is a strictly small power and that $\delta_I^*(\alpha') = |\alpha'| = |\alpha|$. Furthermore, the definition (2.10) of Σ by its faces guarantees that $\delta_I^*(\sigma) \geqslant 0$. Hence, $\operatorname{Re}\langle\lambda, \beta\rangle \leqslant \frac{1}{2}|\alpha|$.

Finally, suppose that equality holds. This implies equality in (3.1), which can hold only if $\beta_k = 0$ when $\operatorname{Re}\lambda_k \neq 1$, which means that $\beta = c\delta_1$ with $c = \beta_1$. Furthermore, then $|\alpha|/2 = \langle\lambda, \beta\rangle = c\langle\lambda, \delta_1\rangle = c\lambda_1 = c$. □

The rest of this subsection is devoted to the critically small case, where we have to pay special attention to eigenvalues λ with $\operatorname{Re}\lambda = \frac{1}{2}$; such eigenvalues are called *critical*. Recall that we have chosen a basis $(v_1, \dots, v_s)$ that yields a Jordan block decomposition of A. A set of indices $J \subseteq \{1, \dots, s\}$ that corresponds to a Jordan block is called a *monogenic block of indices* [14]; if the corresponding eigenvalue is critical, J is called a *critical monogenic block*.

The *support* of a power or another vector $\alpha = (\alpha_1, \dots, \alpha_s) \in \mathbb{Z}^s$ is $\operatorname{supp}(\alpha) := \{k : \alpha_k \neq 0\}$. The power (vector) α is called *critical* if $\alpha_k \neq 0 \implies \operatorname{Re}\lambda_k \in \{1, \frac{1}{2}\}$, and α is called *strictly critical* if $\alpha_k \neq 0 \implies \operatorname{Re}\lambda_k = \frac{1}{2}$. Furthermore, α is called *monogenic* when its support in contained in some monogenic block J, and α is called a *quasi-monogenic power* when $\operatorname{supp}(\alpha) \subseteq \{1\} \cup J$ for some monogenic block J. We consider only critical monogenic blocks, i.e., blocks associated to a critical eigenvalue. (Note that a power $\alpha = c\delta_1$ is critical and quasi-monogenic, and associated to any monogenic block J; otherwise J is determined by α.)

Recall that K_α is the set of powers defined in (2.17).

Lemma 3.2. *Assume that the urn is critically small.*

(i) *Let α be a critical power and let $\beta \in (A_\alpha - \Sigma) \cap \mathbb{Z}^s_{\geqslant 0}$. Then, $\operatorname{Re}\langle\lambda, \beta\rangle \leqslant \operatorname{Re}\langle\lambda, \alpha\rangle$, with equality only if β is critical.*

(ii) *If α is a critical power, then any $\beta \in K_\alpha$ is critical.*

Proof. (i): Let $\beta := \alpha' - \sigma$ with $\alpha' \in A_\alpha$ and $\sigma \in \Sigma$. Then

$$\langle\lambda, \beta\rangle = \langle\lambda, \alpha'\rangle - \langle\lambda, \sigma\rangle = \langle\lambda, \alpha\rangle - \langle\lambda, \sigma\rangle. \tag{3.3}$$

Furthermore, since α is critical, it follows from (2.13) that α' too is critical. Hence for an index k with $\operatorname{Re}\lambda_k < \frac{1}{2}$, we have $\alpha'_k = 0$ and thus $\beta_k = -\sigma_k$ so $\sigma_k \leqslant 0$. Since the urn is critically small, it follows that

$$\begin{aligned}\operatorname{Re}\langle\lambda, \sigma\rangle &= \sigma_1 + \sum_{k:\operatorname{Re}\lambda_k < \frac{1}{2}} \sigma_k \operatorname{Re}\lambda_k + \sum_{k:\operatorname{Re}\lambda_k = \frac{1}{2}} \tfrac{1}{2}\sigma_k \\ &\geqslant \sigma_1 + \sum_{k:\operatorname{Re}\lambda_k < \frac{1}{2}} \tfrac{1}{2}\sigma_k + \sum_{k:\operatorname{Re}\lambda_k = \frac{1}{2}} \tfrac{1}{2}\sigma_k = \frac{1}{2}\delta^*_{\{1\}}(\sigma) \geqslant 0,\end{aligned} \tag{3.4}$$

where the last inequality comes from (2.10). Hence, (3.3) yields $\operatorname{Re}\langle\lambda, \beta\rangle \leqslant \operatorname{Re}\langle\lambda, \alpha\rangle$; moreover, equality holds only if $\operatorname{Re}\lambda_k < \frac{1}{2}$ implies $\sigma_k = 0$ and thus $\beta_k = \alpha'_k = 0$, i.e., β is critical. (Equality also requires $\delta^*_{\{1\}}(\sigma) = 0$.)

(ii): Let α be a critical power. If $\beta \in K_\alpha$, then $\beta \in (A_\alpha - \Sigma) \cap \mathbb{Z}^s_{\geqslant 0}$ and equality holds in (i). Then β is critical. □

As a consequence of Lemma 3.2 and Theorem 2.2, the space $\mathcal{C}$ of polynomial functions on V defined by

$$\mathcal{C} := \operatorname{span}\{Q_\alpha : \alpha \text{ critical}\} \tag{3.5}$$

is Φ-stable; thus, when α is a critical power, ν_α is also the index of Q_α for the nilpotent endomorphism induced by $\Phi - \langle\lambda,\alpha\rangle$ on $\mathcal{C}$. This property is the basic fact that allows us to prove Proposition 3.3 which constitutes the key argument of Theorem 1.2.

Proposition 3.3. *Assume that the urn is critically small. If α is a quasi-monogenic critical power associated with a Jordan block of size $1+r$, $r \geqslant 0$, then $\nu_\alpha \leqslant \big(r+\frac{1}{2}\big)|\alpha|$.*

The remainder of this section is devoted to the proof of Proposition 3.3. We assume that α is a critical power with $\operatorname{supp}(\alpha) \subseteq \{1\} \cup J$ for some monogenic block J, and we may without loss of generality assume that $J = \{2,\dots,r+2\}$ for some $r \geqslant 0$, since we otherwise may permute the Jordan blocks of the chosen basis. In this case, we define for vectors γ with $\operatorname{supp}(\gamma) \subseteq \{1\} \cup J$,

$$M(\gamma) := \sum_{k=1}^{r+2} k\gamma_k - 2\sum_{k=1}^{r+2}\gamma_k + \operatorname{Re}\langle\lambda,\gamma\rangle = \sum_{k=2}^{r+2}\big(k-\tfrac{3}{2}\big)\gamma_k. \tag{3.6}$$

Note that $M(\gamma)$ is a linear function of γ.

Lemma 3.4. *Assume that α is a quasi-monogenic critical power with monogenic block $J = \{2,\dots,r+2\}$, $r \geqslant 0$. Let $\alpha' \in A_\alpha \setminus \{\alpha\}$. Then, α' is also a critical quasi-monogenic power with monogenic block J and $M(\alpha') \leqslant M(\alpha) - 1$.*

Proof. By (2.12) and (2.13), only the inequality is non-trivial. Furthermore, (2.12) implies that α' can be written, with $J' := \{k : k, k-1 \in J\} = \{3,\dots,r+2\}$,

$$\alpha' = \alpha - \sum_{k\in J'} \varepsilon_k\big(\delta_k - \delta_{k-1}\big) \tag{3.7}$$

where the ε_k are nonnegative integers, not all 0 since $\alpha \neq \alpha'$. Then, since $M\big(\delta_k - \delta_{k-1}\big) = 1$ for $k \in J'$,

$$M(\alpha') = M(\alpha) - \sum_{k\in J'} \varepsilon_k M\big(\delta_k - \delta_{k-1}\big) = M(\alpha) - \sum_{k\in J'}\varepsilon_k \leqslant M(\alpha) - 1. \tag{3.8}$$

□

Lemma 3.5. *Assume that the urn is critically small. Let α be a quasi-monogenic critical power with monogenic block $J = \{2,\dots,r+2\}$, $r \geqslant 0$. Assume that $\beta \in \big(\alpha - \Sigma\big) \cap \mathbb{Z}^s_{\geqslant 0}$ satisfies $\operatorname{Re}\langle\lambda,\beta\rangle = \operatorname{Re}\langle\lambda,\alpha\rangle$ and $\beta \neq \alpha$. Then, β is also a critical quasi-monogenic power with monogenic block J and $M(\beta) \leqslant M(\alpha) - 1$.*

Proof. When $i, j \in \{1,\dots,s\}$ are distinct, denote by $\delta_{(i,j)}$ the s-dimensional vector $\delta_{(i,j)} = 2\delta_i - \delta_j$. These vectors span Σ, see (2.9). We divide the proof into three steps.

① *Let i, j be distinct indices in $\{1,\dots,s\}$. Then $\delta^*_{\{1\}}(\delta_{(i,j)}) \geqslant 0$ with equality if and only if $j = 1$.*

Indeed, by (2.11), $\delta^*_{\{1\}}(\delta_{(i,j)}) = 2 + 2\delta_{i1} - 1 - \delta_{j1}$ and the result follows.

② *Let $\sigma = \alpha - \beta \in \Sigma$. Then, σ is a linear combination of $\delta_{(k,1)}$, $k \in J$, with nonnegative coefficients.*

Indeed, Lemma 3.2 guarantees that β is critical, so that σ is also critical. Consequently, by (2.11), $\delta^*_{\{1\}}(\sigma) = 2\operatorname{Re}\langle\lambda,\sigma\rangle$. Furthermore, by the assumption, $\operatorname{Re}\langle\lambda,\sigma\rangle = \operatorname{Re}\langle\lambda,\alpha\rangle - \operatorname{Re}\langle\lambda,\beta\rangle = 0$. Hence, $\delta^*_{\{1\}}(\sigma) = 0$.

Since σ is a linear combination of vectors $\delta_{(i,j)}$ with nonnegative coefficients (definition (2.9) of Σ by edges), ① proves that all j that appear are equal to 1. Thus

$$\sigma = \sum_{k=2}^{s} \varepsilon_k \delta_{(k,1)} \tag{3.9}$$

where the ε_k are nonnegative (real) numbers. Furthermore, if $k \geqslant 2$ and $k \notin J$, then $0 = \alpha_k \geqslant \alpha_k - \beta_k = \sigma_k = 2\varepsilon_k \geqslant 0$ and thus $\varepsilon_k = 0$.

③ It follows from ② that $\operatorname{supp}(\sigma) \subseteq \{1\} \cup J$, and thus this is also true for β, proving the assertion that β is critical and quasi-monogenic with monogenic block J. Furthermore, by (3.9) and (3.6),

$$M(\sigma) = \sum_{k=2}^{r+2} \varepsilon_k M(\delta_{(k,1)}) = \sum_{k=2}^{r+2} \varepsilon_k (2k-3) \geqslant \sum_{k=2}^{r+2} \varepsilon_k = -\sigma_1 \geqslant 1 \tag{3.10}$$

since σ_1 is an integer and the sum is nonnegative and nonzero (because $\beta \neq \alpha$). Consequently, $M(\beta) = M(\alpha) - M(\sigma) \leqslant M(\alpha) - 1$. □

Lemma 3.6. *Assume that the urn is critically small. Let α be a quasi-monogenic critical power with monogenic block $\{2, \dots, r+2\}$, $r \geqslant 0$. Then $\nu_\alpha \leqslant M(\alpha)$.*

Proof. Let $J = \{2, \dots, r+2\}$ be a critical monogenic block and fix $\ell \in \frac{1}{2}\mathbb{Z}_{\geqslant 0}$. Let

$$I_\ell := \big\{\alpha \in \mathbb{Z}^s_{\geqslant 0} : \operatorname{supp}(\alpha) \subseteq \{1\} \cup J,\ \operatorname{Re}\langle \lambda, \alpha \rangle = \ell \big\}. \tag{3.11}$$

We show by induction on α (using the degree-antialphabetical order) that the inequality $\nu_\alpha \leqslant M(\alpha)$ is true for every $\alpha \in I_\ell$. Note that I_ℓ is finite and thus well-ordered.

Take any $\alpha \in I_\ell$ and suppose by induction that $\nu_\beta \leqslant M(\beta)$ for any $\beta \in I_\ell$ such that $\beta < \alpha$. By Theorem 2.2, (2.16)–(2.17) hold. In particular, by the definition of the index of nilpotence,

$$\nu_\alpha \leqslant \begin{cases} 0, & K_\alpha = \emptyset, \\ 1 + \max\{\nu_\beta : \beta \in K_\alpha\}, & K_\alpha \neq \emptyset. \end{cases} \tag{3.12}$$

In particular, if $K_\alpha = \emptyset$, then $\nu_\alpha = 0 \leqslant M(\alpha)$.

Assume $K_\alpha \neq \emptyset$ and let $\beta \in K_\alpha$. Then $\beta = \alpha' - \sigma$ with $\alpha' \in A_\alpha$ and $\sigma \in \Sigma$. By Lemmas 3.4 and 3.5, α' and β are also critical quasi-monogenic powers with monogenic block J. Thus $\beta \in I_\ell$. Furthermore, if $\alpha' \neq \alpha$, then Lemmas 3.4 and 3.5 also yield $M(\beta) \leqslant M(\alpha') \leqslant M(\alpha) - 1$, while if $\alpha' = \alpha$, then Lemma 3.5 yields $M(\beta) \leqslant M(\alpha) - 1$. Hence, in any case, $M(\beta) \leqslant M(\alpha) - 1$. By the inductive assumption, we thus have $\nu_\beta \leqslant M(\beta) \leqslant M(\alpha) - 1$.

Consequently, (3.12) shows that if $K_\alpha \neq \emptyset$, then $\nu_\alpha \leqslant 1 + (M(\alpha) - 1) = M(\alpha)$, which completes the induction. □

Remark 3.7. Since ν_α is an integer, in fact, $\nu_\alpha \leqslant \lfloor M(\alpha) \rfloor$. Strict inequality is possible. For example, if $\lambda_2 = \frac{1}{2} + \mathrm{i}t$ is a critical eigenvalue with $t \neq 0$, then $Q_{2\delta_2}$ is an eigenfunction of Φ and thus $\nu_{2\delta_2} = 0$.

Proof of Proposition 3.3. Let J be a Jordan block of size $1 + r$ associated to α. As said above, we may assume that $J = \{2, \dots, r+2\}$. Then, by Lemma 3.6 and (3.6),

$$\nu_\alpha \leqslant M(\alpha) = \sum_{k=2}^{r+2} \big(k - \tfrac{3}{2}\big)\alpha_k \leqslant \big(r + \tfrac{1}{2}\big)|\alpha|. \tag{3.13}$$

□

Remark 3.8. The upper bound in Proposition 3.3 is reached only for $\alpha = |\alpha|\delta_{\max J}$ where J is a critical Jordan block. Moreover, it is reached only when $|\alpha|$ is even, explaining why the odd moments of X_n are asymptotically negligible after normalization.

3.2 Moments

Lemma 3.9. *If α is a strictly small power, then $\mathbb{E}\,\mathbf{u}^\alpha(X_n) = O\big(n^{|\alpha|/2}\big)$.*

Proof. Since $\mathbf{u}^\alpha \in S'_\alpha$ by definition (2.14), it follows from (2.15) that we have a decomposition

$$\mathbf{u}^\alpha = \sum_{\beta\in\mathbb{Z}^s_{\geqslant 0}\cap(A_\alpha-\Sigma)} q_{\alpha,\beta}Q_\beta \tag{3.14}$$

for some constants $q_{\alpha,\beta}$.

If $\beta \in \mathbb{Z}^s_{\geqslant 0} \cap (A_\alpha - \Sigma)$ and $\beta \neq (|\alpha|/2)\delta_1$, then $\operatorname{Re}\langle\lambda,\beta\rangle < |\alpha|/2$ by Lemma 3.1. Furthermore, by [14, Proposition 5.1], for some $\nu_\beta \geqslant 0$,

$$\mathbb{E}\,Q_\beta(X_n) = O\big(n^{\operatorname{Re}\langle\lambda,\beta\rangle}\log^{\nu_\beta} n\big) = o\big(n^{|\alpha|/2}\big). \tag{3.15}$$

On the other hand, if $\beta = (|\alpha|/2)\delta_1$ (and thus $|\alpha|$ is even), then Q_β is an eigenfunction of Φ and by [14, Proposition 5.1(1)], (3.15) holds with $\nu_\beta = 0$, so

$$\mathbb{E}\,Q_\beta(X_n) = O\big(n^{\langle\lambda,\beta\rangle}\big) = O\big(n^{|\alpha|/2}\big). \tag{3.16}$$

In fact, in this case $Q_\beta = u_1(u_1+1)\cdots(u_1+|\alpha|/2-1)$ so $Q_b(X_n)$ is deterministic, and a polynomial in n of degree $|\alpha|/2$, see [14, Remark 4.10]. ☐

Lemma 3.10. *Assume that the urn is critically small. Let, as in Theorem 1.2, $1+d$ be the largest dimension of a critical Jordan block of the replacement matrix R. Then, if α is a strictly critical power α,*

$$\mathbb{E}\,\mathbf{u}^\alpha(X_n) = O\big(n\log^{2d+1} n\big)^{|\alpha|/2}. \tag{3.17}$$

Proof. Decomposing $\mathbf{u}^\alpha = \mathbf{u}^{\alpha_1}\dots\mathbf{u}^{\alpha_t}$ where the α_k are monogenic critical powers, thanks to the Cauchy–Schwarz inequality applied $t-1$ times, it suffices to show the lemma when α is strictly critical and monogenic.

Suppose thus that α is strictly critical and monogenic. Note that, since α is strictly critical, $\operatorname{Re}\langle\lambda,\alpha\rangle = |\alpha|/2$. As above, we use the decomposition (3.14) of $\mathbf{u}^\alpha$; we now split it as

$$\mathbf{u}^\alpha = \sum_{\beta\in A_\alpha-\Sigma,\operatorname{Re}\langle\lambda,\beta\rangle=\operatorname{Re}\langle\lambda,\alpha\rangle} q_{\alpha,\beta}Q_\beta + \sum_{\beta:\operatorname{Re}\langle\lambda,\beta\rangle<\operatorname{Re}\langle\lambda,\alpha\rangle} q_{\alpha,\beta}Q_\beta. \tag{3.18}$$

When $\operatorname{Re}\langle\lambda,\beta\rangle < \operatorname{Re}\langle\lambda,\alpha\rangle$, Proposition 2.1 yields $\mathbb{E}\,Q_\beta(X_n) = o\big(n^{|\alpha|/2}\big)$. To deal with the first sum in (3.18), suppose that $\beta \in A_\alpha - \Sigma$ satisfies $\operatorname{Re}\langle\lambda,\beta\rangle = \operatorname{Re}\langle\lambda,\alpha\rangle$. Then, thanks to Lemmas 3.4 and 3.5, β is also critical and quasi-monogenic so that Proposition 3.3 asserts that $\nu_\beta \le (d+\frac{1}{2})|\alpha|$. Thus Proposition 2.1 yields

$$\mathbb{E}\,Q_\beta(X_n) = O\big(n^{\operatorname{Re}\langle\lambda,\beta\rangle}\log^{(d+\frac{1}{2})|\alpha|} n\big) = O\big(n^{\frac{1}{2}|\alpha|}\log^{(d+\frac{1}{2})|\alpha|} n\big). \tag{3.19}$$

Putting the small o and the big O together, one gets the result. ☐

3.3 Proofs of Theorems 1.1 and 1.2, and of Corollary 1.3

Proof of Theorems 1.1 and 1.2. Assume that the urn is small. Let $P_I := \sum_{k:\operatorname{Re}\lambda_k<\frac{1}{2}}\pi_k$ and $P_{II} := \sum_{k:\operatorname{Re}\lambda_k=\frac{1}{2}}\pi_k$, so that $\mathrm{id}_{\mathbb{C}^s} = \pi_1 + P_I + P_{II}$. Remember that $\pi_k(v) = u_k(v)v_k$.

• We first deal with P_I. Let $J_I := \{k : \operatorname{Re}\lambda_k < \frac{1}{2}\}$. Then, for any $v \in \mathbb{C}^s$,

$$\big|P_I(v)\big|^2 = \Big|\sum_{k\in J_I} u_k(v)v_k\Big|^2 = \sum_{k,j\in J_I}\langle v_k,v_j\rangle u_k(v)\overline{u_j(v)}. \tag{3.20}$$

Taking the ℓ-th power and expanding, we see that for any $\ell \geqslant 1$, there exists a set of strictly small powers β with $|\beta| = 2\ell$, and constants c_β, such that, for all v,

$$\big|P_I(v)\big|^{2\ell} = \sum_\beta c_\beta \mathbf{u}^\beta(v). \tag{3.21}$$

Hence, Lemma 3.9 yields

$$\mathbb{E}\big|P_I(X_n)\big|^{2\ell} = \sum_\beta c_\beta \, \mathbb{E}\, \mathbf{u}^\beta(X_n) = O\big(n^\ell\big). \tag{3.22}$$

• For P_{II}, we argue as in (3.20) and obtain an identity similar to (3.21), now for a set of strictly critical powers β with $|\beta| = 2\ell$. Hence, Lemma 3.10 yields

$$\mathbb{E}\big|P_{II}(X_n)\big|^{2\ell} = \sum_\beta c'_\beta \, \mathbb{E}\, \mathbf{u}^\beta(X_n) = O\big(n \log^{2d+1} n\big)^\ell. \tag{3.23}$$

• Finally, because of the balance assumption (2.4) (with $m = 1$), $\pi_1(X_n)$ is nonrandom and

$$\pi_1(X_n) = u_1(X_n)v_1 = \big(u_1(X_0) + n\big)v_1 = nv_1 + O(1). \tag{3.24}$$

When the urn is strictly small (Theorem 1.1), $P_{II} = 0$ and thus

$$X_n = \pi_1(X_n) + P_I(X_n) = nv_1 + P_I(X_n) + O(1), \tag{3.25}$$

and (3.22) implies

$$\mathbb{E}\,|X_n - nv_1|^{2\ell} = O\big(n^\ell\big). \tag{3.26}$$

When the urn is critically small (Theorem 1.2), we instead have

$$X_n = \pi_1(X_n) + P_I(X_n) + P_{II}(X_n) = nv_1 + P_I(X_n) + P_{II}(X_n) + O(1), \tag{3.27}$$

so that (3.22) and (3.23) imply

$$\mathbb{E}\,|X_n - nv_1|^{2\ell} = O\big(n \log^{2d+1} n\big)^\ell. \tag{3.28}$$

In other words, if $\widetilde{X}_n$ denotes $\widetilde{X}_n := (X_n - nv_1)/n^{1/2}$ when the urn is strictly small and $\widetilde{X}_n := (X_n - nv_1)/\sqrt{n \log^{2d+1} n}$ when the urn is critically small, then $\mathbb{E}\,|\widetilde{X}_n|^{2\ell} = O(1)$, for every positive integer ℓ. Consequently, if $0 \leqslant p < 2\ell$, then the sequence $\mathbb{E}\,|\widetilde{X}_n|^p$ is uniformly integrable. Since ℓ is arbitrary, this sequence is uniformly integrable for every fixed $p \geqslant 0$. Furthermore, by [8, Theorems 3.22 and 3.23], $\widetilde{X}_n \xrightarrow{\mathrm{d}} N(0; \Sigma)$, for some covariance matrix Σ. The uniform integrability just shown implies that any mixed moment $\mathbb{E}\,\widetilde{X}_n^\alpha$ converges to the corresponding moment of $N(0, \Sigma)$. □

Proof of Corollary 1.3. The estimates for $\mathbb{E}\,Y_n$ and $\operatorname{Var} Y_n$ follow directly from the results for $\mathbb{E}\,X_n$ and $\operatorname{Var}(X_n)$ in Theorem 1.1 or 1.2. Furthermore, (1.3) yields

$$\frac{Y_n - n\lambda_1 \langle w, v_1 \rangle}{\sqrt{n \log^\nu n}} \xrightarrow{\mathrm{d}} N(0, \gamma), \tag{3.29}$$

and (1.5) follows when $\gamma \neq 0$. Moreover, the moment convergence in (1.3) asserted in Theorems 1.1 and 1.2 implies moment convergence in (3.29), and thus in (1.5). □

References

[1] Krishna B. Athreya and Samuel Karlin, Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *Ann. Math. Statist.* **39** (1968), 1801–1817. MR-0232455

[2] Brigitte Chauvin, Nicolas Pouyanne and Réda Sahnoun, Limit distributions for large Pólya urns. *Ann. Appl. Probab.* **21** (2011), no. 1, 1–32. MR-2759195

[3] Brigitte Chauvin, Quansheng Liu and Nicolas Pouyanne, Limit distributions for multitype branching processes of m-ary search trees, *Ann. Inst. Henri Poincaré Probab. Stat.* **50** (2014), no. 2, 628–654. MR-3189087

[4] Brigitte Chauvin, Cécile Mailler and Nicolas Pouyanne, Smoothing equations for large Pólya urns, *J. Theor. Probab.* **28** (2015), 923–957. MR-3413961

[5] F. Eggenberger and G. Pólya, Über die Statistik verketteter Vorgänge. *Zeitschrift Angew. Math. Mech.* **3** (1923), 279–289.

[6] Philippe Flajolet, Joaquim Gabarró and Helmut Pekari, Analytic urns. *Ann. Probab.* **33** (2005), no. 3, 1200–1233. MR-2135318

[7] Cecilia Holmgren, Svante Janson and Matas Šileikis, Multivariate normal limit laws for the numbers of fringe subtrees in m-ary search trees and preferential attachment trees. Preprint, 2016. arXiv:1603.08125 MR-3626585

[8] Svante Janson, Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stoch. Process. Appl.* **110** (2004), 177–245. MR-2040966

[9] Svante Janson, Limit theorems for triangular urn schemes. *Probab. Theory Rel. Fields* **134** (2005), 417–452. MR-2226887

[10] Svante Janson, Mean and variance of balanced Pólya urns. Preprint, 2016. arXiv:1602.06203

[11] Miloslav Jiřina, Stochastic branching processes with continuous state space. *Czechoslovak Math. J.* **8 (83)** (1958), 292–313. MR-0101554

[12] Hosam M. Mahmoud, *Pólya urn models.* CRC Press, Boca Raton, FL, 2009. MR-2435823

[13] Cécile Mailler, Describing the asymptotic behaviour of multicolour Pólya urns via smoothing systems analysis. Preprint, 2014. arXiv:1407.2879

[14] Nicolas Pouyanne, An algebraic approach to Pólya processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** (2008), no. 2, 293–323. MR-2446325

Acknowledgments. Svante Janson partly supported by the Knut and Alice Wallenberg Foundation.