The CLT Analogue for Cyclic Urns

Noela S. Müller[*]   Ralph Neininger [*]

Abstract

A cyclic urn is an urn model for balls of types $0, \ldots, m-1$ where in each draw the ball drawn, say of type j, is returned to the urn together with a new ball of type $j + 1 \mod m$. The case $m = 2$ is the well-known Friedman urn. The composition vector, i.e., the vector of the numbers of balls of each type after n steps is, after normalization, known to be asymptotically normal for $2 \le m \le 6$. For $m \ge 7$ the normalized composition vector does not converge. However, there is an almost sure approximation by a periodic random vector. In this paper the asymptotic fluctuations around this periodic random vector are identified. We show that these fluctuations are asymptotically normal for all $m \ge 7$. However, they are of maximal dimension $m - 1$ only when 6 does not divide m. For m being a multiple of 6 the fluctuations are supported by a two-dimensional subspace.

1 Introduction, phenomena and results.

The aim of this extended abstract is to uncover the nature of fluctuations around almost surely oscillating sequences of random variables as they arise in a number of random combinatorial structures, most commonly in random trees. We develop an analysis for the composition vector of cyclic urns and describe new phenomena and characteristics of the fine fluctuations around random oscillating sequences at the example of the cyclic urns.

A cyclic urn is an urn model with a fixed number $m \ge 2$ of possible colours of balls which we call types $0, \ldots, m-1$. Initially, there is one ball of an arbitrary type. In each step we draw a ball from the urn, uniformly from within the balls in the urn and independently of the history of the urn process. If its type is $j \in \{0, \ldots, m-1\}$ it is placed back to the urn together with a new ball of type $j + 1 \mod m$. We denote by $R_n = (R_{n,0}, \ldots, R_{n,m-1})^t$ the (column) vector of the numbers of balls of each type after n steps when starting with one ball of type 0. Hence, we have $R_0 = e_0$ where e_j denotes the j-th unit vector in $\mathbb{R}^m$, indexing the unit vectors by $0, \ldots, m-1$. For fixed $m \ge 2$ we denote the m-th elementary root of unity by

$\omega := \exp(\frac{2\pi i}{m})$. Furthermore, for $0 \le k \le m-1$, we set

$$\lambda_k := \Re(\omega^k) = \cos\left(\frac{2\pi k}{m}\right),$$

$$\mu_k := \Im(\omega^k) = \sin\left(\frac{2\pi k}{m}\right),$$

$$(1.1) \quad v_k := \frac{1}{m}\left(1, \omega^{-k}, \omega^{-2k}, \ldots, \omega^{-(m-1)k}\right)^t \in \mathbb{C}^m.$$

Note that $v_0 = \frac{1}{m}\mathbf{1} := \frac{1}{m}(1, 1, \ldots, 1)^t \in \mathbb{R}^m$.

The asymptotic distributional behavior of the sequence $(R_n)_{n\ge0}$ has been identified in Janson [8, 9, 10], see also Pouyanne [15, 16]. Janson also developed a limit theory for the compositions of rather general urn schemes. For simplicity of presentation we state the case when starting with only one ball (of type 0). However, when starting with one ball of type $j \in \{0, \ldots, m-1\}$, the evolution of the corresponding composition vector $R_n^{[j]}$ is obtained in distribution by the relation

$$(1.2) \qquad (R_n^{[j]})_{n\ge0} \overset{d}{=} \left((\mathcal{A}^t)^j R_n\right)_{n\ge0},$$

$0 \le j \le m-1$, where the replacement matrix $\mathcal{A}$ is defined in (2.4) and $\overset{d}{=}$ denotes equality in distribution. Hence, it is sufficient to consider the cyclic urn process started with one ball of colour 0. An extension to initially having more than one ball is straightforward, see the discussion in [11, p. 1165].

Janson showed that the normalized composition vector R_n of the cyclic urn model converges in distribution towards a multivariate normal distribution for $2 \le m \le 6$, whereas for $m \ge 7$ there is no convergence by a conventionally standardized version of the R_n due to subtle periodicities. For $m \ge 7$ there exists a complex valued random variable Ξ_1 (depending on m) such that almost surely, as $n \to \infty$, we have

$$(1.3) \qquad \frac{R_n - \frac{n}{m}\mathbf{1}}{n^{\lambda_1}} - 2\Re\left(n^{i\mu_1}\Xi_1 v_1\right) \to 0.$$

We now focus on the periodic case $m \ge 7$. According to (1.3) the normalization $n^{-\lambda_1}(R_n - \frac{n}{m}\mathbf{1})$ does not converge but is (strongly) approximated by the oscillating random sequence $(2\Re(n^{i\mu_1}\Xi_1 v_1))_{n\ge0}$. In the present paper we clarify whether it is still possible that the fluctuations of $n^{-\lambda_1}(R_n - \frac{n}{m}\mathbf{1})$ around the periodic

[*]Institute for Mathematics, Goethe University Frankfurt a.M.

sequence $(2\Re(n^{i\mu_1}\Xi_1 v_1))_{n\geq 0}$ do converge although the sequence itself does not converge. Subsequently, we will call the differences in (1.3) residuals.

Our main results stated in Theorems 1.1 and 1.2 show that the nature of the asymptotic behavior of the residuals in (1.3) depends on the number of colours m. For $m \in \{7, 8, 9, 10, 11\}$ there is a direct normalization which implies a multivariate central limit law (CLT) for the residuals. The case $m = 12$ also allows a multivariate CLT with a different scaling. For $m > 12$ the residuals cannot directly by normalized to obtain convergence. However, considering refined residuals allows a multivariate CLT for all $m > 12$. This in fact gives a more refined expansion of the R_n, cf. Theorems 1.1 and 1.2. There is a further subtlety in the nature of the fluctuations of the residuals: If 6 divides m the fluctuations of the residuals are asymptotically supported by a two-dimensional plane, i.e., the covariance matrix of the limit normal distribution has rank 2, whereas for all $m \geq 7$ which are not divisible by 6 this support is a hyperplane (rank $m - 1$).

By $\xrightarrow{d}$ convergence in distribution is denoted, for a symmetric positive semi-definite matrix M by $\mathcal{N}(0, M)$ the centered normal distribution with covariance matrix M. For $v \in \mathbb{C}^m$ we denote by v^* the conjugate transpose of v. Furthermore, $6 \mid m$ and $6 \nmid m$ are short for 6 divides (resp. does not divide) m.

We distinguish the cases $6 \mid m$ and $6 \nmid m$ as follows.

THEOREM 1.1. *Let $m \geq 7$ with $6 \nmid m$ and set $r := \lfloor (m-1)/6 \rfloor$. Then, there exist complex valued random variables $\Xi_1, \ldots, \Xi_r$ such that, as $n \to \infty$, we have*

$$n^{\lambda_1 - 1/2}\left(\frac{R_n - \mathbb{E}[R_n]}{n^{\lambda_1}} - \sum_{k=1}^{r} 2n^{\lambda_k - \lambda_1}\Re\left(n^{i\mu_k}\Xi_k v_k\right)\right)$$
$$\xrightarrow{d} \mathcal{N}\left(0, \Sigma^{(m)}\right).$$

The covariance matrix $\Sigma^{(m)}$ has rank $m-1$ and is given by

$$\Sigma^{(m)} = \sum_{k=1}^{m-1} \frac{1}{|2\lambda_k - 1|} v_k v_k^*.$$

When $6 \mid m$ the normalization requires an additional $\sqrt{\log n}$ factor and the rank of the covariance matrix is reduced to 2:

THEOREM 1.2. *Let $m \geq 7$ with $6 \mid m$ and set $r := \lfloor (m-1)/6 \rfloor$. Then, there exist complex valued random variables $\Xi_1, \ldots, \Xi_r$ such that, as $n \to \infty$, we have*

$$\frac{n^{\lambda_1 - 1/2}}{\sqrt{\log(n)}}\left(\frac{R_n - \mathbb{E}[R_n]}{n^{\lambda_1}} - \sum_{k=1}^{r} 2n^{\lambda_k - \lambda_1}\Re\left(n^{i\mu_k}\Xi_k v_k\right)\right)$$
$$\xrightarrow{d} \mathcal{N}\left(0, \Sigma^{(m)}\right).$$

The covariance matrix $\Sigma^{(m)}$ has rank 2 and is given by

$$\Sigma^{(m)} = v_{m/6} v_{m/6}^* + v_{5m/6} v_{5m/6}^*.$$

The convergences in Theorems 1.1 and 1.2 also hold for all moments. For an expansion of $\mathbb{E}[R_n]$ see (2.6).

We consider Theorems 1.1 and 1.2 as prototypical for a phenomenon which we conjecture to occur frequently in related random combinatorial structures. E.g., we expect similar behavior for other urn models with analog almost sure random periodic behavior (see [9, Theorem 3.24]), for the size of random m-ary search trees, cf. [4, 2, 6], and for the number of leaves in random d-dimensional (point) quadtrees [3]. (For the latter two instances only the case of Theorem 1.1 is expected to occur.)

2 Outline of the proof.

In this section we first recall some known asymptotic behavior of R_n which is used subsequently. Then we state a more refined result on certain projections of residuals in Proposition 2.1 which directly implies Theorems 1.1 and 1.2. Then, an outline of the proof of Proposition 2.1 is given. Technical steps and estimates are then sketched in Section 3. Throughout, we fix an $m \geq 7$.

The cyclic urn with m colours has the $m \times m$ replacement matrix

$$(2.4) \qquad \mathcal{A} := \begin{pmatrix} 0 & 1 & 0 & \cdot & \cdot & 0 & 0 \\ 0 & 0 & 1 & \cdot & \cdot & 0 & 0 \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 & 1 \\ 1 & 0 & 0 & \cdot & \cdot & 0 & 0 \end{pmatrix},$$

where $\mathcal{A}_{ij}$ indicates that after drawing a ball of type i it is placed back together with $\mathcal{A}_{ij}$ balls of type j for all $0 \leq i, j \leq m - 1$. For the urn we consider the initial configuration of one ball of type 0 and write R_n for the composition vector after n steps. The canonical filtration is given by the σ-fields $\mathcal{F}_n = \sigma(R_0, \ldots, R_n)$ for $n \geq 0$. The dynamics of the urn process imply the well-known almost sure relation

$$\mathbb{E}[R_{n+1} \mid \mathcal{F}_n] = \sum_{k=0}^{m-1} \frac{R_{n,k}}{n+1}(R_n + \mathcal{A}^t e_k)$$

$$(2.5) \qquad = \left(\mathrm{Id}_m + \frac{1}{n+1}\mathcal{A}^t\right) R_n$$

for $n \geq 0$. Here, Id_m denotes the $m \times m$ identity matrix and $\mathcal{A}^t$ the transpose of $\mathcal{A}$. The matrices $\mathcal{A}$ and $\mathrm{Id}_m + \frac{1}{n+1}\mathcal{A}^t$ have the same (right) eigenvectors $v_0, \ldots, v_{m-1}$ given in (1.1).

Note that v_0 has the direction of the drift vector $\mathbf{1}$ in Theorems 1.1 and 1.2 and v_1 determines the directions of the a.s. fluctuations around the drift there. By diagonalizing these matrices and using (2.5) one finds explicit expressions for the mean of the R_n, cf. [11, Lemma 6.7]. Let $1 \le k \le r$. With

$$\xi_k := \frac{2}{\Gamma(1+\omega^k)} v_k,$$

these expressions imply the expansion, as $n \to \infty$,

$$(2.6) \quad \mathbb{E}[R_n] = \frac{n+1}{m}\mathbf{1} + \sum_{k=1}^r \Re(n^{i\mu_k}\xi_k)n^{\lambda_k} + \mathrm{O}(\sqrt{n}).$$

It is also known that the variances and covariances of R_n are of the order $n^{2\lambda_1}$ with appropriate periodic prefactors. This explains the normalization $n^{-\lambda_1}(R_n - \frac{n+1}{m}\mathbf{1})$ in Theorems 1.1 and 1.2. The analysis of the asymptotic distribution as stated in (1.3) has been done by different techniques (partly only in a weak sense), by embedding into continuous time multitype branching processes, by (more direct) use of martingale arguments, and by stochastic fixed-point arguments, see [9, 15, 11].

For our further analysis we use a spectral decomposition of the process $(R_n)_{n\ge 0}$. We denote by π_k the projection onto the eigenspace in $\mathbb{C}^m$ spanned by v_k for $0 \le k \le m-1$. Hence, we have

$$R_n = \sum_{k=0}^{m-1} \pi_k(R_n)$$
$$= \pi_0(R_n) + \sum_{k=1}^{\lceil m/2 \rceil -1} (\pi_k + \pi_{m-k})(R_n)$$
$$+ \mathbb{1}_{\{m\ \text{even}\}}\pi_{m/2}(R_n),$$

where $\mathbb{1}$ indicates an indicator. We have deterministically $\pi_0(R_n) = \frac{n+1}{m}\mathbf{1}$. For the other projections $\pi_k(R_n)$ one has similar periodic behavior as for the composition vector R_n, cf. (1.3), as long as we have $\lambda_k > \frac{1}{2}$. We call the projections $\pi_k(R_n)$ large, if $\lambda_k > \frac{1}{2}$, since their magnitudes have orders larger than $\sqrt{n}$. Projections π_k with $\lambda_k \le \frac{1}{2}$ we call small. For the large projections we have for all $1 \le k \le \lfloor m/2 \rfloor$ with $\lambda_k > \frac{1}{2}$ almost surely that

$$Y_{n,k} := \frac{1}{n^{\lambda_k}}(\pi_k + \pi_{m-k})(R_n - \mathbb{E}[R_n])$$
$$- 2\Re\left(n^{i\mu_k}\Xi_k v_k\right)$$
$$(2.7) \qquad \to 0$$

with a complex valued random variable Ξ_k. The small projections $\pi_k(R_n)$ behave differently, see [9, 13]. For those k with $\lambda_k < \frac{1}{2}$ we have

$$X_{n,k} := \frac{1}{\sqrt{n}}(\pi_k + \pi_{m-k})(R_n - \mathbb{E}[R_n])$$
$$(2.8) \qquad \overset{\mathrm{d}}{\longrightarrow} \mathcal{N}(0, \Sigma_k),$$

with an appropriate covariance matrix Σ_k, see (3.16)–(3.18).

If m is even then for $X_{n,m/2} := n^{-1/2}\pi_{m/2}(R_n)$ we have a multivariate CLT as in (2.8).

Finally, if $6 \mid m$, then there is the pair $(\frac{m}{6}, \frac{5m}{6})$ with $\lambda_{m/6} = \lambda_{5m/6} = \frac{1}{2}$. In this case the scaling requires an additional $\sqrt{\log n}$ factor. We have

$$X_{n,m/6} := \frac{1}{\sqrt{n\log n}}(\pi_{m/6} + \pi_{5m/6})(R_n - \mathbb{E}[R_n])$$
$$(2.9) \qquad \overset{\mathrm{d}}{\longrightarrow} \mathcal{N}(0, \Sigma_{m/6}).$$

We identify the orders of the variances and covariances of $Y_{n,k}$ in Section 3.1. These orders imply that an appropriate normalization to study the fluctuations of the large projections is given by

$$(2.10) \qquad X_{n,k} := n^{\lambda_k - \frac{1}{2}}Y_{n,k}.$$

Now, the $X_{n,k}$ are defined for all $1 \le k \le \lfloor m/2 \rfloor$ and describe the normalized fluctuations of all the projections. For the small projections we already know that they are asymptotically normally distributed, see (2.8). As a main contribution of the present paper we show that the residuals of the large projections as normalized in (2.10) are also asymptotically normal. Moreover, we show that all these fluctuations are jointly asymptotically normally distributed and asymptotically independent:

PROPOSITION 2.1. *For the vector* $(X_{n,1}, \ldots, X_{n,\lfloor m/2 \rfloor})$ *defined in (2.8) - (2.10) we have*

$$(X_{n,1}, \ldots, X_{n,\lfloor m/2 \rfloor}) \overset{d}{\longrightarrow} \mathcal{N}(0, \mathrm{diag}(\Sigma_1, \ldots, \Sigma_{\lfloor m/2 \rfloor})),$$

where the blocks Σ_k *of the diagonal block matrix* $\mathrm{diag}(\Sigma_1, \ldots, \Sigma_{\lfloor m/2 \rfloor})$ *are defined in (3.16)–(3.18).*

Proposition 2.1 directly implies Theorems 1.1 and 1.2:

Proof of Theorem 1.1. Let $m \ge 7$ with $6 \nmid m$, set $r = \lfloor (m-1)/6 \rfloor$ and let $\Xi_1, \ldots, \Xi_r$ as in (2.7). Moreover, let $X_{n,1}, \ldots, X_{n,\lfloor m/2 \rfloor}$ be as in Proposition 2.1. Note that $6 \nmid m$ implies that there is no $1 \le k \le m$ with

$\lambda_k = \frac{1}{2}$. We obtain

$$n^{\lambda_1 - 1/2} \left(\frac{R_n - \mathbb{E}[R_n]}{n^{\lambda_1}} - \sum_{k=1}^{r} 2 n^{\lambda_k - \lambda_1} \Re \left(n^{i\mu_k} \Xi_k v_k \right) \right)$$

$$= n^{\lambda_1 - 1/2} \left(n^{-\lambda_1} \sum_{k=1}^{r} \left\{ (\pi_k + \pi_{m-k})(R_n - \mathbb{E}[R_n]) \right. \right.$$
$$\left. - 2 n^{\lambda_k} \Re \left(n^{i\mu_k} \Xi_k v_k \right) \right\}$$
$$+ n^{-\lambda_1} \sum_{r+1}^{\lceil m/2 \rceil - 1} (\pi_k + \pi_{m-k})(R_n - \mathbb{E}[R_n])$$
$$\left. + \mathbb{1}_{\{m \text{ even}\}} n^{-\lambda_1} \pi_{m/2} (R_n - \mathbb{E}[R_n]) \right)$$

$$= X_{n,1} + \cdots + X_{n,\lfloor m/2 \rfloor}$$
$$\xrightarrow{\text{d}} \mathcal{N} \left(0, \Sigma^{(m)} \right),$$

by Proposition 2.1 and the continuous mapping theorem, where $\Sigma^{(m)} = \Sigma_1 + \cdots + \Sigma_{\lfloor m/2 \rfloor}$. That $\Sigma^{(m)}$ has rank $m - 1$ is proven in Theorem 3.1.

Proof of Theorem 1.2. Let $m \geq 7$ with $6 \mid m$ and $\Xi_1, \ldots, \Xi_r$ as in (2.8) and $X_{n,1}, \ldots, X_{n,m/2}$ as in Proposition 2.1. Note that $6 \mid m$ implies that there is the pair $(m/6, 5m/6)$ with $\lambda_{m/6} = \lambda_{5m/6} = \frac{1}{2}$. Rearranging terms as in the proof of Theorem 1.1 we obtain

$$\frac{n^{\lambda_1 - 1/2}}{\sqrt{\log n}} \left(\frac{R_n - \mathbb{E}[R_n]}{n^{\lambda_1}} - \sum_{k=1}^{r} 2 n^{\lambda_k - \lambda_1} \Re \left(n^{i\mu_k} \Xi_k v_k \right) \right)$$

$$= X_{n,m/6} + \frac{1}{\sqrt{\log n}} \sum_{\substack{k=1 \\ k \neq m/6}}^{m/2} X_{n,k}$$

$$\xrightarrow{\text{d}} \mathcal{N} \left(0, \Sigma^{(m)} \right),$$

by Proposition 2.1 and Slutsky's Lemma, where $\Sigma^{(m)} = \Sigma_{m/6}$. That $\Sigma^{(m)}$ has rank 2 is proven in Theorem 3.1.

To prove Proposition 2.1 we first derive moments and mixed moments needed for the normalization in Section 3.1. The ranks of the covariance matrices $\Sigma^{(m)}$ are identified in Section 3.4. In Section 3.2 a pointwise recursive equation for the complex random variables $\Xi_1, \ldots, \Xi_r$ is obtained together with a recurrence for the sequence $(R_n)_{n \geq 0}$ which extends to a recurrence for the residuals in (1.3) as well as to the residuals of the projections of the R_n. Finally, the joint convergence of the normalized residuals of all projections is shown by an application of a stochastic fixed-point argument in the context of the contraction method by use of the Zolotarev metric ζ_3. However, only an indication and a reference are given in Section 3.3.

3 Sketch of the proof of Proposition 2.1.

3.1 Proper normalization of the residuals. Denoting the inner product in $\mathbb{C}^m$ by $\langle \cdot, \cdot \rangle$ we first write the spectral decomposition of the centered composition vector with respect to the orthonormal basis $\{\sqrt{m} v_k : 0 \leq k < m\}$ of the unitary vector space $\mathbb{C}^m$ (denoting coefficients by $u_k(w) := \langle w, v_k \rangle$ for $w \in \mathbb{C}^m$) as

$$R_n - \mathbb{E}[R_n] = \sum_{k=0}^{m-1} \pi_k (R_n - \mathbb{E}[R_n])$$
$$= \sum_{k=0}^{m-1} u_k (R_n - \mathbb{E}[R_n]) v_k.$$

The evolution (2.5) of the process implies that for all $0 \leq k \leq m - 1$ the random variables $M_{0,k} := 0$ and

$$(3.11) \qquad M_{n,k} := \frac{\Gamma(n+1)}{\Gamma(n+1+\omega^k)} u_k (R_n - \mathbb{E}[R_n]),$$

$n \geq 1$, define complex-valued, centered martingales. Note, that the corresponding martingales $M_{n,k}^{[j]}$ when starting with one ball of type $j \in \{0, \ldots, m-1\}$ satisfy

$$M_{n,k}^{[j+1]} \overset{d}{=} \omega^k M_{n,k}^{[j]} \qquad (\text{convention } M_{n,k}^{[m]} := M_{n,k}^{[0]}).$$

It is known, see [9, 10, 15], that for all $k \in \{0, \ldots, m-1\}$ with $\lambda_k = \Re(\omega^k) > 1/2$, there exists a complex random variable Ξ_k such that, as $n \to \infty$, we have

$$(3.12) \qquad M_{n,k} \to \Xi_k \quad \text{almost surely,}$$

where the convergence also holds in L_p for every $p \geq 1$. Note, that the Ξ_k in (3.12) turn out to be identical with the Ξ_k appearing in (2.7) and in Theorems 1.1 and 1.2. The $M_{n,k}$ with $\lambda_k = \Re(\omega^k) \leq 1/2$ are also known to converge, after proper normalization, to normal limit laws.

Our subsequent analysis requires asymptotics for moments of and correlations between the $u_k(R_n)$. Exploiting the dynamic of the urn in (2.5) elementary calculations imply that:

LEMMA 3.1. *For all $k \in \{0, \ldots, m-1\} \setminus \{m/2\}$, we have*

$$\mathbb{E}\left[u_k(R_n)\right] = \sum_{t=0}^{m-1} \omega^{kt} \mathbb{E}[R_{n,t}] = \frac{\Gamma(n+1+\omega^k)}{\Gamma(n+1)\Gamma(1+\omega^k)},$$

while

$$\mathbb{E}\left[u_{m/2}(R_n)\right] = 0.$$

For all $k, \ell \in \{0, \ldots, m-1\}$,

$$
\begin{aligned}
&\mathbb{E}\left[u_k\left(R_n\right) u_\ell\left(R_n\right)\right] \\
&= \prod_{s=1}^n \left(1 + \frac{\omega^k + \omega^\ell}{s}\right) \\
&\quad + \omega^{k+\ell} \sum_{s=1}^n \frac{1}{s} \prod_{t=1}^{s-1}\left(1 + \frac{\omega^{k+\ell}}{t}\right) \prod_{t=s+1}^n \left(1 + \frac{\omega^k + \omega^\ell}{t}\right).
\end{aligned}
$$

From Lemma 3.1 we obtain the L_2-distance of the residuals of the martingales $(M_{n,k})_{n\geq 0}$ with $\lambda_k > \frac{1}{2}$ needed for the proper normalization of these residuals:

LEMMA 3.2. *For $k \geq 1$ such that $\lambda_k > 1/2$, as $n \to \infty$, we have*

$$
\mathbb{E}\left[|M_{n,k} - \Xi_k|^2\right] \sim \frac{1}{2\lambda_k - 1} n^{1-2\lambda_k}.
$$

Lemma 3.2 directly implies the asymptotic covariances of the residuals of the centered projections of the composition vector, which we denote by

$$
\Pi_{n,k} := \begin{cases} \frac{\Gamma(n+1+\omega^k)}{\Gamma(n+1)}\left(M_{n,k} - \Xi_k\right) v_k, & \text{if } \lambda_k > \frac{1}{2}, \\ u_k\left(R_n - \mathbb{E}\left[R_n\right]\right) v_k, & \text{if } \lambda_k \leq \frac{1}{2}. \end{cases}
$$

Note that this notation implies the representation

$$
\begin{aligned}
&\left(R_n - \mathbb{E}[R_n]\right) - \sum_{k \geq 1:\, \lambda_k > 1/2} \frac{\Gamma(n+1+\omega^k)}{\Gamma(n+1)} \Xi_k v_k \\
&= \sum_{k=1}^{m-1} \Pi_{n,k}.
\end{aligned}
$$

Lemma 3.1 and Lemma 3.2 imply:

LEMMA 3.3. *For all $k \in \{1, \ldots, m-1\} \setminus \{\frac{m}{6}, \frac{5m}{6}\}$, as $n \to \infty$, we have*

$$
\text{(3.13)} \qquad \text{Cov}\left(\Pi_{n,k}\right) \sim \frac{1}{|2\lambda_k - 1|} n \cdot v_k v_k^*.
$$

If $6 \mid m$, then

$$
\text{(3.14)} \qquad \text{Cov}\left(\Pi_{n,m/6}\right) \sim n \log(n) \cdot v_{m/6} v_{m/6}^*,
$$

$$
\text{(3.15)} \qquad \text{Cov}\left(\Pi_{n,5m/6}\right) \sim n \log(n) \cdot v_{5m/6} v_{5m/6}^*.
$$

This also determines the covariance matrices Σ_k in Proposition 2.1: We have

(3.16)
$$
\Sigma_k = \frac{1}{|2\lambda_k - 1|} \cdot v_k v_k^* + \frac{1}{|2\lambda_{m-k} - 1|} \cdot v_{m-k} v_{m-k}^*
$$

for $k \in \{1, \ldots, \lceil m/2 \rceil - 1\} \setminus \{\frac{m}{6}\}$ as well as

$$
\text{(3.17)} \qquad \Sigma_{m/6} = v_{m/6} v_{m/6}^* + v_{5m/6} v_{5m/6}^*, \text{ if } 6 \mid m,
$$

$$
\text{(3.18)} \qquad \Sigma_{m/2} = \frac{1}{|2\lambda_{m/2} - 1|} \cdot v_{m/2} v_{m/2}^*, \text{ if } 2 \mid m.
$$

We also need to control correlations of residuals between different eigenspaces. An explicit calculation implies for all $k, \ell \geq 1$ with $k \neq \ell$ and $\lambda_k, \lambda_\ell > \frac{1}{2}$ that

$$
\begin{aligned}
\text{(3.19)} \qquad &\mathbb{E}\left[\left(M_{n,k} - \Xi_k\right)\left(M_{n,\ell} - \Xi_\ell\right)\right] \\
&= \text{O}\left(n^{-1} + n^{\lambda_{k+\ell} - \lambda_k - \lambda_\ell}\right).
\end{aligned}
$$

The bound (3.19) implies:

LEMMA 3.4. *Let $k, \ell \geq 1$ with $k \neq \ell$ and $n \to \infty$. If $\lambda_k, \lambda_\ell > \frac{1}{2}$ or $\lambda_k, \lambda_\ell \leq \frac{1}{2}$ then*

$$
\text{Cov}\left(\Pi_{n,k}, \Pi_{n,\ell}\right) = o(n).
$$

If $\lambda_k > \frac{1}{2}$ and $\lambda_\ell \leq \frac{1}{2}$ then

$$
\text{Cov}\left(\Pi_{n,k}, \Pi_{n,\ell}\right) = 0.
$$

These moments estimates are sufficient to subsequently properly scale the projections of the residuals and to guarantee the finiteness of the Zolotarev metric ζ_3 used.

3.2 Embedding into a random binary search tree. In this section we describe the self-similarity of the martingale limits Ξ_k by deriving an almost sure recursive equation for the Ξ_k and a distributional recurrence for the sequence $(R_n)_{n\geq 0}$ which extends to a recurrence for the residuals in (1.3) as well as to the normalized residuals $X_{n,k}$ of the projections of the R_n.

For this, we embed the cyclic urn process into a random binary search tree. The random binary search tree starts with one external node. In each step one of the external nodes is chosen uniformly at random (and independently from the previous choices) and replaced by one internal node with two children, the children being external nodes attached along a left and right branch. The cyclic urn is embedded into the evolution of the random binary search tree by labeling its external nodes by the types of the balls. The initial external node is labeled by type 0. Whenever an external node of type $j \in \{0, \ldots, m-1\}$ is replaced by an internal node its (new) left child gets label j, its right child gets label $j+1 \mod m$. Note, that the external nodes of the tree correspond to the balls in the urn. A related embedding was exploited in [11, Section 6.3]. Note that the binary search tree starting with one external node labeled 0 decomposes into its left and right subtree starting with external nodes of types 0 and 1, respectively. The size (number of internal nodes) I_n of the left subtree is

uniformly distributed on $\{0, \ldots, n-1\}$. This implies, with $J_n := n - 1 - I_n$, the recurrence

$$(3.20) \quad R_n^{[0]} = R_{I_n}^{[0],(0)} + R_{J_n}^{[1],(1)} = R_{I_n}^{[0],(0)} + \mathcal{A}^t R_{J_n}^{[0],(1)},$$

where the sequences $(R_n^{[0],(0)})_{n\geq 0}$ and $(R_n^{[1],(1)})_{n\geq 0}$ denote the composition vectors of the cyclic urns given by the evolutions of the left and right subtrees of the root of the binary search tree (upper indices (0) and (1) denoting left and right subtree, upper indices $[0]$ and $[1]$ denoting the initial type). They are independent and independent of I_n. Note that the second equation in (3.20) is due to (1.2) where the $R_n^{[0],(1)}$ are chosen appropriately for pointwise equality. Now, applying the transformation and scaling which turns R_n into $M_{n,k}$ to the left and right hand side of (3.20), letting $n \to \infty$ and using the convergence in (3.12) implies the following recursive equation for the Ξ_k:

PROPOSITION 3.1. *For all $k \geq 1$ with $\lambda_k > \frac{1}{2}$ there exist independent random variables U, $\Xi_k^{(0)}$, $\Xi_k^{(1)}$ such that*

$$(3.21) \quad \Xi_k = U^{\omega^k} \Xi_k^{(0)} + \omega^k (1-U)^{\omega^k} \Xi_k^{(1)} + g_k(U),$$

where

$$g_k(u) := \frac{1}{\Gamma(1+\omega^k)} \left(u^{\omega^k} + \omega^k (1-u)^{\omega^k} - 1 \right)$$

and U has the uniform distribution on $[0,1]$ and $\Xi_k^{(0)}$ and $\Xi_k^{(1)}$ have the same distribution as Ξ_k.

Alternatively, the martingale limits Ξ_k can be written explicitly as deterministic functions of the limit of the random binary search tree when interpreting the evolution of the random binary search tree as a transient Markov chain and its limit as a random variable in the Markov chain's Doob-Martin boundary, see [5, 7]. From this representation the self-similarity relation (3.21) can be read off as well. See also [1] for a related explicit construction.

3.3 Proving convergence. Note that the left and right hand sides of (3.20) and (3.21) are linked via the convergence of the $M_{n,k}$ towards Ξ_k. This allows to come up with a recurrence for the vector $(X_{n,1}, \ldots, X_{n,\lfloor m/2 \rfloor})$ in Proposition 2.1. The reader is asked to trust the authors that the techniques developed in [14] for a univariate problem can be extended to the multivariate recurrences for $(X_{n,1}, \ldots, X_{n,\lfloor m/2 \rfloor})$ and that the same type of proof as in [14] based on the Zolotarev metric ζ_3 can be applied.

3.4 The rank of the covariance matrices. The covariance matrices $\Sigma^{(m)}$ in Theorem 1.1 and 1.2 appear as the sums of the covariance matrices in (3.16) and (3.18) if $6 \nmid m$ and as the covariance matrix in (3.17) if $6 \mid m$. We obtain their ranks as follows:

THEOREM 3.1. *For $6 \nmid m$, the matrix*

$$(3.22) \quad \Sigma^{(m)} = \sum_{k=1}^{m-1} \frac{1}{|2\lambda_k - 1|} v_k v_k^*$$

has rank $m-1$, while for $6 \mid m$,

$$(3.23) \quad \Sigma^{(m)} = v_{m/6} v_{m/6}^* + v_{5m/6} v_{5m/6}^*$$

has rank two.

Proof . Note that the matrix-vector product $m v_k v_k^* x$ is the orthogonal projection of $x \in \mathbb{C}^m$ onto the eigenspace spanned by v_k. Hence, we have

$$\mathrm{Id}_m = \sum_{k=0}^{m-1} m v_k v_k^*.$$

The matrix $m\Sigma^{(m)}$ can be interpreted as the orthogonal projection onto $\mathrm{span}\{v_1, \ldots, v_{m-1}\}$ for the case $6 \nmid m$ and onto the subspace $\mathrm{span}\{v_{m/6}, v_{5m/6}\}$ for $6 \mid m$. Hence, we obtain the ranks $m-1$ and 2, respectively.

References

[1] Bindjeme, P. and Fill, J. A. (2012) Exact L^2-Distance from the Limit for QuickSort Key Comparisons (Extended abstract). *DMTCS proc.* **AQ**, 23rd International Meeting on Probabilistic, Combinatorial, and Asymptotic Methods in the Analysis of Algorithms (AofA'12), 339–348.

[2] Chauvin, B. and Pouyanne, N. (2004) m-ary search trees when $m \geq 27$: a strong asymptotics for the space requirements. *Random Structures Algorithms* **24**, 133–154.

[3] Chern, H.-H., Fuchs, M. and Hwang, H.-K. (2007) Phase changes in random point quadtrees. *ACM Trans. Algorithms* **3**, Art. 12, 51 pp.

[4] Chern, H.-H. and Hwang, H.-K. (2001) Phase changes in random m-ary search trees and generalized quicksort. *Random Structures Algorithms* **19**, 316–358.

[5] Evans, S.N., Grübel, R. and Wakolbinger, A. (2012) Trickle-down processes and their boundaries. *Electron. J. Probab.* **17**, 1-58.

[6] Fill, J.A. and Kapur, N. (2004) The Space Requirement of m-ary Search Trees: Distributional Asymptotics for $m \geq 27$. Invited paper, Proceedings of the 7th Iranian Statistical Conference, 2004. Available via `http://www.ams.jhu.edu/~fill/papers/periodic.pdf`

[7] Grübel, R. (2014) Search trees: Metric aspects and strong limit theorems. *Ann. Appl. Probab.* **24**, 1269–1297.

[8] Janson, S. (1983) Limit theorems for certain branching random walks on compact groups and homogeneous spaces. *Ann. Probab.* **11**, 909–930.

[9] Janson, S. (2004) Functional limit theorem for multi-type branching processes and generalized Pólya urns. *Stochastic Process. Appl.* **110**, 177–245.

[10] Janson, S. (2006) Congruence properties of depths in some random trees. *Alea* **1**, 347–366.

[11] Knape, M. and Neininger, R. (2014) Pólya Urns Via the Contraction Method. *Combin. Probab. Comput.* **23**, 1148–1186.

[12] Laruelle, S. and Pagès, G. (2013) Randomized urn models revisited using stochastic approximation. *Ann. Appl. Probab.* **23**, 1409–1436.

[13] Mailler, C. (2014) Describing the asymptotic behaviour of multicolour Pólya urns via smoothing systems analysis.
Preprint available via
`http://arxiv.org/abs/1407.2879`.

[14] Neininger, R. (2015) Refined Quicksort asymptotics. *Random Structures Algorithms* **46**, 346–361.

[15] Pouyanne, N. (2005) Classification of large Pólya-Eggenberger urns with regard to their asymptotics. *2005 International Conference on Analysis of Algorithms*, 275–285 (electronic), Discrete Math. Theor. Comput. Sci. Proc., AD, *Assoc. Discrete Math. Theor. Comput. Sci., Nancy.*

[16] Pouyanne, N. (2008) An algebraic approach to Pólya processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **44**, 293–323.