Electronic Journal of Probability

Electron. J. Probab. **23** (2018), no. 117, 1–20.
ISSN: 1083-6489 https://doi.org/10.1214/18-EJP243

Refined asymptotics for the composition of cyclic urns

Noela Müller* Ralph Neininger†

Abstract

A cyclic urn is an urn model for balls of types $0,\dots,m-1$. The urn starts at time zero with an initial configuration. Then, in each time step, first a ball is drawn from the urn uniformly and independently from the past. If its type is j, it is then 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 is known not to converge. However, there is an almost sure approximation by a periodic random vector.

In the present paper the asymptotic fluctuations around this periodic random vector are identified. We show that these fluctuations are asymptotically normal for all $7\le m\le 12$. For $m\ge 13$ we also find asymptotically normal fluctuations when normalizing in a more refined way. These fluctuations 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.

Keywords: Pólya urn; cyclic urn; cyclic group; periodicities; weak convergence; CLT analogue; probability metric; Zolotarev metric.
AMS MSC 2010: 60F05; 60F15; 60C05; 60J10.
Submitted to EJP on January 30, 2017, final version accepted on November 7, 2018.

1 Introduction and result

A cyclic urn is an urn model with a fixed number $m\ge 2$ of possible colors of balls which we call types $0,\dots,m-1$. We assume that initially there is one ball of type 0. 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 $j\in\{0,\dots,m-1\}$ it is placed back to the urn together with a new ball of type $j+1 \mod m$. These steps are iterated.

*Institut für Mathematik, Goethe Universität, 60054 Frankfurt am Main, Germany.
E-mail: nmueller@math.uni-frankfurt.de
†Institut für Mathematik, Goethe Universität, 60054 Frankfurt am Main, Germany.
E-mail: neiningr@math.uni-frankfurt.de

We denote by $R_n = (R_{n,0}, \dots, R_{n,m-1})^t$ 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 $0, \dots, m-1$. For fixed $m \geq 2$ we denote the m-th elementary root of unity by $\omega := \exp(\frac{2\pi \mathrm{i}}{m})$. Furthermore, we set

$$\lambda_k := \Re(\omega^k) = \cos\left(\frac{2\pi k}{m}\right), \qquad \mu_k := \Im(\omega^k) = \sin\left(\frac{2\pi k}{m}\right),$$
$$v_k := \frac{1}{m}\left(1, \omega^{-k}, \omega^{-2k}, \dots, \omega^{-(m-1)k}\right)^t \in \mathbb{C}^m, \quad 0 \leq k \leq m-1. \tag{1.1}$$

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

The asymptotic distributional behavior of the sequence $(R_n)_{n\geq 0}$ has been identified in Janson [9, 10, 11], see also Pouyanne [18, 19] and, for the case $m = 2$, Freedman [7]. Janson developed a limit theory for the compositions of rather general urn schemes. For the cyclic urns he showed that the normalized composition vector R_n converges in distribution towards a multivariate normal distribution for $2 \leq m \leq 6$, whereas for $m \geq 7$ there is no convergence of a conventionally standardized version of the R_n due to subtle periodicities. Further, for $m \geq 7$, there exists a complex valued random variable Ξ_1 (depending on m) such that almost surely, as $n \to \infty$, we have

$$\frac{R_n - \frac{n+1}{m}\mathbf{1}}{n^{\lambda_1}} - 2\Re\left(n^{i\mu_1}\Xi_1 v_1\right) \to 0. \tag{1.2}$$

We focus mainly on the periodic case $m \geq 7$. In the present paper we study the fluctuations of $n^{-\lambda_1}(R_n - \frac{n+1}{m}\mathbf{1})$ around the periodic sequence $(2\Re(n^{i\mu_1}\Xi_1 v_1))_{n\geq 0}$. We call the differences in (1.2) residuals.

To formulate our results we denote by $\xrightarrow{d}$ convergence in distribution. Further, $\mathcal{N}(0, M)$ denotes the centered normal distribution with covariance matrix M, where M is a symmetric positive semi-definite matrix. For $v \in \mathbb{C}^m$ we write v^* for the conjugate transpose of v and for $z \in \mathbb{C}$, $\bar{z}$ denotes the complex conjugate of z. 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 $m \geq 2$ with $6 \nmid m$ and set $r := \lfloor (m-1)/6 \rfloor$. Then, there exist complex valued random variables $\Xi_1, \dots, \Xi_r$ such that, as $n \to \infty$, we have

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

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

Theorem 1.2. Let $m \geq 2$ with $6 \mid m$ and set $r := \lfloor (m-1)/6 \rfloor$. Then, there exist complex valued random variables $\Xi_1, \dots, \Xi_r$ such that, as $n \to \infty$, we have

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

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

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

Note, that the sum $\sum_{k=1}^{r}$ in Theorem 1.1 is empty for $2 \le m \le 5$, also in Theorem 1.2 for $m = 6$. Hence, for $2 \le m \le 6$ our theorems reduce to the central limit laws of Janson [9, 10, 11]. For $m \in \{7, 8, 9, 10, 11\}$ Theorem 1.1 shows that there is a direct normalization of the residuals which implies a multivariate central limit law (CLT). The case $m = 12$ also admits a multivariate CLT under a different scaling, see Theorem 1.2. For $m > 12$ the residuals cannot directly be normalized to obtain convergence. However, Theorems 1.1 and 1.2 describe refined residuals which satisfy a multivariate CLT for all $m > 12$. These can be considered as asymptotic expansions of the random variables R_n.

The convergences in Theorems 1.1 and 1.2 also hold for all moments. For an expansion of $\mathbb{E}[R_n]$ see (2.3).

We conjecture Theorems 1.1 and 1.2 as being prototypical for a phenomenon 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 Janson [10, Theorem 3.24], further for the size of random m-ary search trees, cf. [4], or 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.)

The remainder of the present paper contains a proof of Theorems 1.1 and 1.2. An outline of the proof is given in Section 2, where also the occurrence of the contributions $\Re(n^{\omega^k}\Xi_k v_k)$ in Theorems 1.1 and 1.2 is explained. Roughly, our proof combines a spectral decomposition of the residuals and estimates of their mixed moments with a recursive decomposition of the urn process and stochastic fixed-point arguments. In work in progress of the first mentioned author of the present paper also an alternative route via martingales is being explored. Within the details of the proofs of the present paper we make mildly use of martingales. However, we could also work out the whole proof without drawing back to any martingale which may provide a useful general technique for related applications where no martingales are directly available.

The results of this paper were announced in the extended abstract [15].

2 Explanation of the result and outline of the proof

In this section we set out our approach towards the proof of Theorems 1.1 and 1.2 and explain the occurrence of the summands $\Re(n^{\omega^k}\Xi_k v_k)$ and the normal fluctuation in the theorems.

We first recall known asymptotic behavior and a spectral decomposition of R_n which are 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. Finally, an outline of the proof of Proposition 2.1 is given. Technical steps and estimates are then carried out in Section 3. Throughout, we fix $m \ge 2$.

For the cyclic urn with m colors we consider an initial configuration of one ball of type 0 and write R_n for the composition vector after n steps. Its dynamics is summarized in the $m \times m$ replacement matrix

$$\mathcal{A} := \begin{pmatrix} 0 & 1 & 0 & \cdot & \cdot & 0 & 0 \\ 0 & 0 & 1 & \cdot & \cdot & 0 & 0 \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 & 1 \\ 1 & 0 & 0 & \cdot & \cdot & 0 & 0 \end{pmatrix}, \tag{2.1}$$

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 $0 \le i, j \le m-1$. The canonical filtration is given by the σ-fields $\mathcal{F}_n = \sigma(R_0, \ldots, R_n)$ for $n \ge 0$. The dynamics of the urn process imply the well-known almost sure relation

$$\mathbb{E}\left[R_{n+1}\,|\,\mathcal{F}_n\right] = \sum_{k=0}^{m-1} \frac{R_{n,k}}{n+1}(R_n + \mathcal{A}^t e_k) = \left(\mathrm{Id}_m + \frac{1}{n+1}\mathcal{A}^t\right) R_n, \qquad n \geq 0. \tag{2.2}$$

Here, Id_m denotes the $m \times m$ identity matrix and $\mathcal{A}^t$ the transpose of $\mathcal{A}$.

Note that v_0 has the direction of the drift vector $\mathbf{1}$ in Theorems 1.1 and 1.2 and, for $m \geq 7$, the vector v_1 determines the direction of the a.s. periodic fluctuations around the drift. By diagonalizing the matrices on the right hand side of (2.2) one finds an exact asymptotic expression for the mean of R_n, cf. [12, Lemma 6.7]. With

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

equation (2.2) implies the expansion, as $n \to \infty$,

$$\mathbb{E}\left[R_n\right] = \frac{n+1}{m}\mathbf{1} + \sum_{k=1}^{r} \Re(n^{i\mu_k}\xi_k) n^{\lambda_k} + \begin{cases} o(\sqrt{n}), & \text{if } 6 \nmid m, \\ \mathrm{O}(\sqrt{n}), & \text{if } 6 \mid m. \end{cases} \tag{2.3}$$

It is also known that the variances and covariances of the numbers of balls of each color are of the order $n^{2\lambda_1}$ when $m \geq 7$, with appropriate periodic prefactors. This explains the normalization $n^{-\lambda_1}(R_n - \frac{n+1}{m}\mathbf{1})$ in (1.2). The analysis of the asymptotic distribution as stated in (1.2) has been carried out 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 [10, 18, 12].

For our further analysis we use a spectral decomposition of the process $(R_n)_{n\geq 0}$. This also leads to an explanation of the terms and fluctuations appearing in Theorems 1.1 and 1.2, see the comments after the proof of Theorem 1.2 in the present section.

We denote by π_k the orthogonal projection onto the eigenspace in $\mathbb{C}^m$ spanned by v_k for $0 \leq k \leq m-1$. Hence, we have

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

where $u_0, \ldots, u_{m-1}$ denotes the basis dual to $v_0, \ldots, v_{m-1}$, as $\mathcal{A}$ is diagonizable. 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 in (1.2), as long as $\lambda_k > \frac{1}{2}$. Commonly, projections $\pi_k(R_n)$ are called *large*, if $\lambda_k > \frac{1}{2}$, since their magnitudes are larger than $\sqrt{n}$. Projections π_k with $\lambda_k \leq \frac{1}{2}$ are called *small*.

For large projections, i.e. for all $k \geq 1$ with $\lambda_k > \frac{1}{2}$, we set

$$X_{n,k} := \frac{1}{\sqrt{n}} \begin{pmatrix} \Re\left(u_k(R_n - \mathbb{E}[R_n]) - n^{\omega^k}\Xi_k\right) \\ \Im\left(u_k(R_n - \mathbb{E}[R_n]) - n^{\omega^k}\Xi_k\right) \end{pmatrix}, \quad n \geq 1, \tag{2.4}$$

with an appropriate complex valued random variable Ξ_k, defined as a martingale limit in (3.2), Section 3.1. The behavior of the small projections $\pi_k(R_n)$ has already been determined, see [10, 14]. For those k with $\lambda_k < \frac{1}{2}$ we have for $n \geq 1$

$$X_{n,k} := \frac{1}{\sqrt{n}} \begin{pmatrix} \Re(u_k(R_n - \mathbb{E}[R_n])) \\ \Im(u_k(R_n - \mathbb{E}[R_n])) \end{pmatrix} \xrightarrow{\mathrm{d}} \mathcal{N}\left(0, \frac{\mathrm{Id}_2}{1-2\lambda_k}\right). \tag{2.5}$$

If m is even, then for $n \geq 1$, $X_{n,m/2} := n^{-1/2} u_{m/2}(R_n) \xrightarrow{\mathrm{d}} \mathcal{N}(0, 1/3)$. For $m = 2$, the last mentioned result has already been established by Freedman [7, Theorem 5.1].

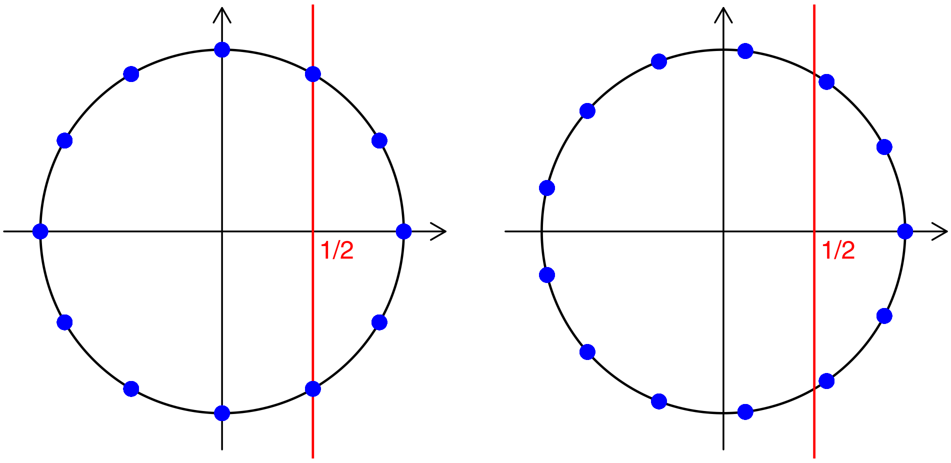

Figure 1: Plots of the eigenvalues of $\mathcal{A}$ for $m = 12$ (left picture) and $m = 13$ (right picture). They correspond to contributions to R_n as follows: The eigenvalue $\lambda_0 = 1$ corresponds to the deterministic drift. All other eigenvalues with $\lambda_k > \frac{1}{2}$ correspond to almost sure periodic contributions with normal fluctuations around the periodic vector. The eigenvalues with $\lambda_k \leq \frac{1}{2}$ correspond to contributions which only consist of a normal fluctuation. All normal fluctuations are of the same order if $6 \nmid m$. They compose an overall fluctuation of rank $m-1$, see Theorem 1.1. If $6 \mid m$ then the eigenspaces with $\lambda_k = \frac{1}{2}$ contribute normal fluctuations of larger orders which dominate the contributions from all other eigenspaces. The overall fluctuations are then just the fluctuations from the two eigenspaces $m/6$ and $5m/6$ with $\lambda_{m/6} = \lambda_{5m/6} = \frac{1}{2}$ and of rank 2, see Theorem 1.2.

Finally, if $6 \mid m$, there are two eigenvalues with real parts equal to $\frac{1}{2}$. Compared to the other small components, the scaling of the associated projections requires an additional $\sqrt{\log n}$ factor for convergence: For $k \in \{m/6, 5m/6\}$ and $n \geq 1$,

$$X_{n,k} := \frac{1}{\sqrt{n \log n}} \begin{pmatrix} \Re(u_k(R_n - \mathbb{E}[R_n])) \\ \Im(u_k(R_n - \mathbb{E}[R_n])) \end{pmatrix} \xrightarrow{\mathrm{d}} \mathcal{N}\left(0, \frac{1}{2}\mathrm{Id}_2\right). \tag{2.6}$$

We prove the convergence of the variances and covariances of all $X_{n,k}$ in Section 3.1. Set $X_{n,0} := u_0(R_n - \mathbb{E}[R_n]) = 0$ and $X_0 := (0, \dots, 0)^t$.

To summarize, $X_{n,0}, \dots, X_{n,m-1}$ describe the normalized fluctuations along the projections. For each pair of complex conjugate eigenvalues, there is one $X_{n,k}$ that captures the behaviour of the corresponding real and imaginary part. Small projections are known to be asymptotically normally distributed, see (2.5). As a main contribution of the present paper we show that residuals of large projections as normalized in (2.4) are also asymptotically normal. Moreover, fluctuations along different projections are asymptotically independent:

Proposition 2.1. Assume that $6 \mid m$. For the vector $Z_n := (X_{n,1}, \dots, X_{n,m/2}) \in \mathbb{R}^{m-1}$ defined for $n \geq 0$ in (2.4)–(2.6) we have, as $n \to \infty$, that

$$Z_n \xrightarrow{d} \mathcal{N}(0, M_m),$$

with

$$M_m := \frac{1}{2}\mathrm{diag}\left(\frac{\mathrm{Id}_2}{|2\lambda_1 - 1|}, \dots, \frac{\mathrm{Id}_2}{|2\lambda_r - 1|}, \mathrm{Id}_2, \frac{\mathrm{Id}_2}{|2\lambda_{r+2} - 1|}, \dots, \frac{\mathrm{Id}_2}{|2\lambda_{m/2-1} - 1|}, \frac{2}{3}\right). \tag{2.7}$$

In the case $6 \nmid m$ Proposition 2.1 holds as well. The only difference is that there is no k with $\lambda_k = \frac{1}{2}$ and thus the matrix corresponding to M_m for the case $6 \nmid m$ does not have the block $\frac{1}{2}\mathrm{Id}_2$. If m is odd, also the last block $\frac{1}{3}$ is not present.

Proposition 2.1 (and its version for $6 \nmid m$) directly imply Theorems 1.1 and 1.2:

Proof of Theorem 1.1. Note that $6 \nmid m$ implies that there is no $0 \le k \le m-1$ with $\lambda_k = \frac{1}{2}$. We obtain

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\left(R_n - \mathbb{E}[R_n] - \sum_{k=1}^{r} 2\Re\left(n^{\omega^k}\Xi_k v_k\right)\right)\\
&= \frac{1}{\sqrt{n}}\left(\sum_{k=1}^{r}\left\{2\Re\left(\left[u_k(R_n - \mathbb{E}[R_n]) - n^{\omega^k}\Xi_k\right]v_k\right)\right\}\right.\\
&\qquad\left. + \sum_{k=r+1}^{\lceil m/2\rceil-1} 2\Re(u_k(R_n - \mathbb{E}[R_n])v_k) + \mathbb{1}_{\{m \text{ even}\}}u_{m/2}(R_n - \mathbb{E}[R_n])v_{m/2}\right)\\
&= 2Z_{n,1}\Re(v_1) - 2Z_{n,2}\Im(v_1) + \cdots + \mathbb{1}_{\{m \text{ even}\}}Z_{n,m-1}v_{m/2}\\
&\xrightarrow{\mathrm{d}} \mathcal{N}\left(0, \Sigma^{(m)}\right),
\end{aligned}$$

by Proposition 2.1 and the continuous mapping theorem, where $\Sigma^{(m)}$ is as in the statement of Theorem 1.1. The image of $\Sigma^{(m)}$ in $\mathbb{R}^m$ is $\mathrm{span}\{\Re(v_1), \Im(v_1), \dots, v_{m/2}\}$, if $2 \mid m$, and $\mathrm{span}\{\Re(v_1), \Im(v_1), \dots, \Im\left(v_{(m-1)/2}\right)\}$ otherwise, hence the rank of $\Sigma^{(m)}$ is $m-1$. ☐

Proof of Theorem 1.2. Note that $6 \mid m$ implies that there is the pair $\lambda_{m/6} = \lambda_{5m/6} = \frac{1}{2}$. Rearranging terms as in the proof of Theorem 1.1 we obtain

$$\begin{aligned}
&\frac{1}{\sqrt{n\log(n)}}\left(R_n - \mathbb{E}[R_n] - \sum_{k=1}^{r} 2\Re\left(n^{\omega^k}\Xi_k v_k\right)\right)\\
&= \frac{1}{\sqrt{\log(n)}}\sum_{k=1, k\neq m/6}^{m/2-1} 2(Z_{n,2k-1}\Re(v_k) - Z_{n,2k}\Im(v_k))\\
&+ 2(Z_{n,m/3-1}\Re(v_{m/6}) - Z_{n,m/3}\Im(v_{m/6})) + \frac{1}{\sqrt{\log(n)}}Z_{n,m-1}v_{m/2} \xrightarrow{\mathrm{d}} \mathcal{N}\left(0, \Sigma^{(m)}\right),
\end{aligned}$$

by Proposition 2.1 and Slutsky's Lemma, where $\Sigma^{(m)}$ is as in Theorem 1.2. Again, it is immediate that the image of $\Sigma^{(m)}$ in $\mathbb{R}^m$ is $\mathrm{span}\{\Re\left(v_{m/6}\right), \Im\left(v_{m/6}\right)\}$, hence its rank is 2. ☐

The proofs of Theorems 1.1 and 1.2 via Proposition 2.1 indicate the role of the terms $\Re(n^{\omega^k}\Xi_k v_k)$ in the overall Gaussian fluctuation, see also Figure 1: All eigenspaces with $\lambda_k > \frac{1}{2}$ (excluding the deterministic drift for $\lambda_k = 1$) contribute two asymptotic components: First, there is the almost sure periodic component

$$\Re(n^{\omega^k}\Xi_k v_k) = n^{\lambda_k}\Re\left(\exp(i\mu_k \log n)\Xi_k v_k\right)$$

of order n^{λ_k} with a random periodic factor, periodic roughly in $\log n$. Second, there is a normal fluctuation (in distribution) of order $\sqrt{n}$. All eigenspaces with $\lambda_k < \frac{1}{2}$ add a contribution of order $\sqrt{n}$ to the normal fluctuation which is the visible order within these eigenspaces. For $6 \mid m$, there are eigenvalues with $\lambda_k = \frac{1}{2}$ and the normal fluctuation is of order $\sqrt{n\log n}$ in the corresponding two eigenspaces. According to Proposition 2.1 all these fluctuations within the eigenspaces are asymptotically independent, which explains the overall asymptotic normal fluctuation. Since this normal fluctuation is of order $\sqrt{n}$ and $\sqrt{n \log n}$, respectively, all the almost sure periodic contributions from the eigenspaces with $\lambda_k > \frac{1}{2}$ are visible as well.

To prove Proposition 2.1 we first derive moments and mixed moments in Section 3.1 needed for the normalization. In Section 3.2 a pointwise recursive equation for the complex random variables $\Xi_1, \dots, \Xi_r$ is obtained together with a recurrence for the sequence $(R_n)_{n \geq 0}$ which extends to a recurrence for the residuals in (1.2) as well as to the residuals Z_n of the projections of the R_n, see equation (3.7) in Section 3.2. Equation (3.7) is then the starting point to show the convergence in Proposition 2.1. For this, a stochastic fixed-point argument in the context of the contraction method within the Zolotarev metric ζ_3, see [17] for general reference, is used. Then, we draw back to an approach to bound the Zolotarev distance and some estimates from [16] where a related, but simpler, (univariate) problem was discussed.

3 Proof of Proposition 2.1

We start with estimates for the covariance matrix of the Z_n appearing in Proposition 2.1 in section 3.1. In section 3.2 we derive the recurrence (3.7) for the Z_n. The use of the Zolotarev metric ζ_3 requires a slightly modified version of recurrence (3.7). This is explained in section 3.3, see in particular the quantities N_n in (3.11) which are the modified versions of the Z_n. Then in section 3.4 asymptotics for the coefficients appearing in the recurrence (3.7) of Z_n and N_n respectively are derived. Based on these asymptotics finally in section 3.5 convergence of the N_n is shown within the Zolotarev metric, which implies convergence in distribution of the Z_n as stated in Proposition 2.1.

Recall that Proposition 2.1 assumes that $6 \mid m$. As mentioned before, the analoguous result for $6 \nmid m$ is true and can be proved along the same lines by some minor modifications.

3.1 Convergence of the covariance matrix

As indicated in Section 2, we study the centered process $(R_n - \mathbb{E}[R_n])_{n \geq 0}$ via its spectral decomposition with respect to the orthogonal basis $\{v_k : 0 \leq k < m\}$ of the unitary vector space $\mathbb{C}^m$, i.e.

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

where $u_k(w) := 1 \cdot w_0 + \omega^k \cdot w_1 + \cdots + \omega^{(m-1)k} \cdot w_{m-1}$ for $w \in \mathbb{C}^m$. The evolution (2.2) of the process implies that for $n \geq 1$, there is a complex normalization

$$M_{k,n} := \frac{\Gamma(n+1)}{\Gamma(n+1+\omega^k)} u_k \left(R_n - \mathbb{E}\left[R_n\right]\right) = \begin{cases} \frac{\Gamma(n+1)}{\Gamma(n+1+\omega^k)} u_k\left(R_n\right) - \frac{1}{\Gamma(1+\omega^k)}, & k \neq m/2, \\ \frac{\Gamma(n+1)}{\Gamma(n+1+\omega^k)} u_k\left(R_n\right), & k = m/2, \end{cases} \tag{3.1}$$

that turns all the eigenspace coefficients, $0 \leq k \leq m-1$, into centered martingales. We set $M_{k,0} := 0$. Depending on λ_k, these martingales are known to exhibit two different kinds of asymptotic behavior, see [10, 11, 18]: For all $k \in \{0, \dots, m-1\}$ with $\lambda_k = \Re\left(\omega^k\right) > 1/2$, there exists a complex valued random variable Ξ_k such that, as $n \to \infty$, we have

$$M_{k,n} \to \Xi_k \ \text{ almost surely,} \tag{3.2}$$

where the convergence also holds in L_p for every $p \geq 1$. Note that the Ξ_k in (3.2) are identical with the Ξ_k in (2.4) and in Theorems 1.1 and 1.2. The $M_{k,n}$ with $\lambda_k = \Re\left(\omega^k\right) \leq 1/2$ are known to converge in distribution, after proper normalization, to normal limit laws.

>From Section 3.2 on, our analysis will also require to start the cyclic urn process with one ball of type $j \in \{0, \dots, m-1\}$. The corresponding composition vector $R_n^{[j]}$ is obtained in distribution by the relation

$$\left(R_n^{[j]}\right)_{n\geq 0} \stackrel{d}{=} \left(\left(\mathcal{A}^t\right)^j R_n\right)_{n\geq 0}, \quad 0 \leq j \leq m-1, \tag{3.3}$$

with the replacement matrix $\mathcal{A}$ from (2.1) and where $\stackrel{d}{=}$ denotes equality in distribution. Similar to the identity (3.3), the corresponding martingales $M_{k,n}^{[j]}$ satisfy

$$M_{k,n}^{[j+1]} \stackrel{d}{=} \omega^k M_{k,n}^{[j]},$$

with convention $M_{k,n}^{[m]} := M_{k,n}^{[0]}$.

Our subsequent analysis requires asymptotics of moments and of correlations between the $u_k(R_n)$. Exploiting the dynamics of the urn in (2.2), elementary calculations imply that:

Lemma 3.1. For $k \in \{0, \dots, m-1\}$, we have

$$\mathbb{E}\left[u_k\left(R_n\right)\right] = \sum_{t=0}^{m-1} \omega^{kt} \mathbb{E}\left[R_{n,t}\right] = \begin{cases} \frac{\Gamma(n+1+\omega^k)}{\Gamma(n+1)\Gamma(1+\omega^k)}, & k \neq m/2, \\ 0, & k = m/2. \end{cases}$$

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

$$\begin{aligned} &\mathbb{E}\left[u_k\left(R_n\right) u_\ell\left(R_n\right)\right] \\ &= \prod_{s=1}^{n} \left(\frac{s+\omega^k+\omega^\ell}{s}\right) + \sum_{s=1}^{n} \frac{\omega^{k+\ell}}{s} \prod_{t=1}^{s-1} \left(\frac{t+\omega^{k+\ell}}{t}\right) \prod_{t=s+1}^{n} \left(\frac{t+\omega^k+\omega^\ell}{t}\right). \end{aligned} \tag{3.4}$$

Proof. The first two identities immediately follow from (2.2). For (3.4), let $k, \ell \in \{0, \dots, m-1\}$ and $n \geq 1$ and note that, almost surely,

$$\mathbb{E}\left[u_k(R_n)u_\ell(R_n)|\mathcal{F}_{n-1}\right] = \left(1 + \frac{\omega^k+\omega^\ell}{n}\right) u_k(R_{n-1})u_\ell(R_{n-1}) + \frac{\omega^{k+\ell}}{n} u_{k+\ell}(R_{n-1}).$$

Here, we use the abbreviation $u_{k+\ell}(R_{n-1}) := u_{(k+\ell) \bmod m}(R_{n-1})$. □

Remark 1. From (3.4) we see that all $\mathbb{E}[|u_k(R_n)|^2]$ with $\lambda_k < 1/2$ are of linear order, all $\mathbb{E}[|u_k(R_n)|^2]$ with $\lambda_k = 1/2$ are of order $n \log n$ and all $\mathbb{E}[|u_k(R_n)|^2]$ with $\lambda_k > 1/2$ have order $n^{2\lambda_k}$. To make this more visible from (3.4), we make some case distinctions.

We first consider the real cases $k = \ell = 0$ and $k = \ell = m/2$ for $2 \mid m$:

$$\mathbb{E}\left[|u_0(R_n)|^2\right] = (n+1)^2$$

and, if $2 \mid m$,

$$\mathbb{E}\left[|u_{m/2}(R_n)|^2\right] = \frac{n+1}{3}.$$

Now, $\omega^k + \omega^\ell = -1$ only if $3 \mid m$ and $\{k, \ell\} = \{m/3, 2m/3\}$. In this case,

$$\mathbb{E}\left[|u_{m/3}(R_n)|^2\right] = \frac{1}{n}\sum_{t=1}^{n} t = \frac{n+1}{2}.$$

On the other hand, $\omega^{k+\ell} = \omega^k + \omega^\ell$ only if $6 \mid m$ and $\{k, \ell\} = \{m/6, 5m/6\}$. In this case, $\omega^k + \omega^\ell = 1$ and

$$\mathbb{E}\left[|u_{m/6}(R_n)|^2\right] = (n+1)\sum_{t=1}^{n+1}\frac{1}{t} \sim n\log n.$$

Thirdly, $\omega^k + \omega^\ell = 0$ if and only if $2 \mid m$ and $\ell = k + m/2 \mod m$, so in this case

$$\mathbb{E}\left[u_k(R_n)u_\ell(R_n)\right] = \begin{cases} 0, & \text{if} \quad \{k,l\} = \{0, m/2\}, \\ \frac{\Gamma(n+1+\omega^{k+\ell})}{\Gamma(n+1)\Gamma(1+\omega^{k+\ell})}, & \text{else}. \end{cases}$$

Finally, $\omega^{k+\ell} = -1$ if $\lambda_k = -\lambda_\ell$ and $\mu_k = \mu_\ell$ and then,

$$\mathbb{E}\left[u_k(R_n)u_\ell(R_n)\right] = \frac{\omega^k + \omega^\ell}{1 + \omega^k + \omega^\ell}\prod_{s=1}^{n}\left(1 + \frac{\omega^k + \omega^\ell}{s}\right) \sim \frac{\omega^k + \omega^\ell}{\Gamma(2 + \omega^k + \omega^\ell)} n^{\omega^k + \omega^\ell}.$$

In all other cases,

$$\mathbb{E}\left[u_k(R_n)u_\ell(R_n)\right] = \frac{1}{\omega^{k+\ell} - \omega^k - \omega^\ell}\left(\frac{\Gamma(n+1+\omega^{k+\ell})}{\Gamma(n+1)\Gamma(\omega^{k+\ell})} - \frac{\Gamma(n+1+\omega^k+\omega^\ell)}{\Gamma(n+1)\Gamma(\omega^k+\omega^\ell)}\right).$$

Remark 2. From (3.4) we obtain the mixed moments of the corresponding real and imaginary parts via the identities

$$\begin{aligned}
\mathbb{E}\left[\Re(u_k(R_n))\Re(u_\ell(R_n))\right] &= \frac{1}{2}\Re\left(\mathbb{E}\left[u_k(R_n)u_\ell(R_n)\right] + \mathbb{E}\left[u_k(R_n)u_{m-\ell}(R_n)\right]\right), \\
\mathbb{E}\left[\Im(u_k(R_n))\Im(u_\ell(R_n))\right] &= \frac{1}{2}\Re\left(\mathbb{E}\left[u_k(R_n)u_{m-\ell}(R_n)\right] - \mathbb{E}\left[u_k(R_n)u_\ell(R_n)\right]\right), \\
\mathbb{E}\left[\Re(u_k(R_n))\Im(u_\ell(R_n))\right] &= \frac{1}{2}\Im\left(\mathbb{E}\left[u_k(R_n)u_\ell(R_n)\right] + \mathbb{E}\left[u_{m-k}(R_n)u_\ell(R_n)\right]\right).
\end{aligned}$$

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

Lemma 3.2. For $k \geq 1$ such that $1/2 < \lambda_k < 1$ and Ξ_k as in (3.2), as $n \to \infty$,

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

and

$$\mathbb{E}\left[(M_{k,n} - \Xi_k)^2\right] \sim \frac{1}{(1 - 2\omega^{-k})\Gamma(2\omega^k)} n^{-1}.$$

In particular,

$$\begin{aligned}
\mathbb{E}\left[\Re\left(M_{k,n} - \Xi_k\right)^2\right] &\sim \frac{1}{2}\frac{1}{2\lambda_k - 1} n^{1-2\lambda_k}, \\
\mathbb{E}\left[\Im\left(M_{k,n} - \Xi_k\right)^2\right] &\sim \frac{1}{2}\frac{1}{2\lambda_k - 1} n^{1-2\lambda_k}, \\
\mathbb{E}\left[\Re\left(M_{k,n} - \Xi_k\right)\Im\left(M_{k,n} - \Xi_k\right)\right] &\sim \frac{1}{2}\Im\left(\frac{1}{(1 - 2\omega^{-k})\Gamma(2\omega^k)}\right) n^{-1}.
\end{aligned}$$

Proof. We show the claim for $\mathbb{E}\left[|M_{k,n}-\Xi_k|^2\right]$ in an exemplary way. Here, we decompose

$$\begin{aligned}
\mathbb{E}\left[|M_{k,n}-\Xi_k|^2\right] &= \sum_{z=n}^{\infty}\mathbb{E}\left[|M_{k,z}-M_{k,z+1}|^2\right]\\
&= \sum_{z=n}^{\infty}\left|\frac{\Gamma(z+2)}{\Gamma(z+2+\omega^k)}\right|^2\mathbb{E}\left[\left|u_k(R_{z+1}-R_z)-\frac{\omega^k}{z+1}u_k(R_z)\right|^2\right]\\
&= \sum_{z=n}^{\infty}\left|\frac{\Gamma(z+2)}{\Gamma(z+2+\omega^k)}\right|^2\left(\mathbb{E}\left[|u_k(R_{z+1}-R_z)|^2\right]-\frac{1}{(z+1)^2}\mathbb{E}\left[|u_k(R_z)|^2\right]\right)\\
&= \sum_{z=n}^{\infty}\left|\frac{\Gamma(z+2)}{\Gamma(z+2+\omega^k)}\right|^2\left(1+\frac{1}{1-2\lambda_k}\frac{1}{(z+1)^2}\left(\frac{\Gamma(z+1+2\lambda_k)}{\Gamma(z+1)\Gamma(2\lambda_k)}-z-1\right)\right)\\
&\sim \sum_{z=n}^{\infty}z^{-2\lambda_k}\sim\frac{1}{2\lambda_k-1}n^{1-2\lambda_k}
\end{aligned}$$

as $n\to\infty$. ☐

The preceding calculations imply that the covariance matrix of Z_n, see Proposition 2.1, converges as $n\to\infty$. Its limit is given by M_m defined in (2.7).

3.2 Embedding and recursions

In this section we briefly explain how to derive an almost sure recurrence for the sequence $(R_n)_{n\geq 0}$ which then extends to the projections. These recursive representations transfer to the martingale limits Ξ_k and thus also to the components of Z_n.

We embed the cyclic urn process into a random binary search tree generated by a sequence $(U_n)_{n\geq 1}$ of i.i.d. random variables, where $U:=U_1$ is uniformly distributed on $[0,1]$. The random binary search tree starts with one external node at time 0, the so-called root. At time $n=1$, the first key U is inserted in this external node, turning it into an internal node. The occupied node then grows two external nodes attached along a left and right branch. We successively insert the following keys, where each key traverses the internal nodes starting at the root, which is occupied by U. Whenever the key traversing is less than the occupying key at a node it moves on to the left child of that node, otherwise to its right child. The first external node visited is occupied by the key, turning it into an internal node with two new external nodes attached. It is easy to see that in each step one of the external nodes is chosen uniformly at random (and independently of the previous choices) and replaced by one internal node with two new external nodes attached. See, e.g., Mahmoud [13], for a detailed description of random binary search trees.

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 $j\in\{0,\dots,m-1\}$ is replaced by an internal node then its new left external node is labeled j (corresponding to returning the chosen ball of type j to the urn) and its new right external node is labeled $(j+1)\mod m$ (corresponding to the addition of a new ball of type $(j+1)\mod m$ to the urn). A related embedding was exploited in [12, Section 6.3], see also [2]. 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 $\{0,\dots,n-1\}$ and, conditional on $U=u$, $u\in(0,1)$, it is binomial $B_{n-1,u}$ distributed. This implies, with $J_n:=n-1-I_n$, the recurrence

$$R_n^{[0]}=R_{I_n}^{[0],(0)}+R_{J_n}^{[1],(1)}=R_{I_n}^{[0],(0)}+\mathcal{A}^tR_{J_n}^{[0],(1)},\quad n\geq 1, \tag{3.5}$$

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 the initial type, upper indices (0) and (1) denoting left and right subtree). They are independent of I_n. We have set $(R_n^{[0],(1)})_{n\geq 0} := (\mathcal{A}R_n^{[1],(1)})_{n\geq 0}$, and note that due to identity (3.3), $(R_n^{[0],(1)})_{n\geq 0}$ is a cyclic urn process started with one ball of type 0 at time 0. Now, applying the transformation and scaling (3.1) which turn R_n into $M_{k,n}$ to the left and right hand side of (3.5), letting $n \to \infty$ and using the convergence in (3.2) yields the following almost sure recursive equation for the Ξ_k:

Proposition 3.3. For all $k \geq 1$ with $\lambda_k > \frac{1}{2}$ there exist random variables $\Xi_k^{(0)}$, $\Xi_k^{(1)}$ such that

$$\Xi_k = U^{\omega^k}\Xi_k^{(0)} + \omega^k(1-U)^{\omega^k}\Xi_k^{(1)} + g_k(U), \tag{3.6}$$

$U, \Xi_k^{(0)}$, $\Xi_k^{(1)}$ are independent, U is uniformly distributed on $[0,1]$ and $\Xi_k^{(0)}$ and $\Xi_k^{(1)}$ have the same distribution as Ξ_k. Here,

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

Here and subsequently, we make no use of the fact that the martingale limits Ξ_k can also 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 [6, 8]. Following this path the Ξ_k become a deterministic function of $(U_n)_{n\geq 1}$ and from this representation the self-similarity relation (3.6) can be read off as well. See [1] for a related explicit construction.

Returning to Z_n, we see that

$$Z_n = \sigma_{I_n}^{-1}\sigma_n Z_{I_n}^{(0)} + \sigma_{J_n}^{-1}\sigma_n \mathcal{D} Z_{J_n}^{(1)} + \sigma_n F_n, \quad n \geq 1, \tag{3.7}$$

where $\sigma_0 := \sigma_1 := \mathrm{Id}_{m-1}$ and $\sigma_k := \frac{1}{\sqrt{k}}\mathrm{diag}\left(1,\dots,1,\frac{1}{\sqrt{\log k}},\frac{1}{\sqrt{\log k}},1,\dots,1\right)$ for $k \geq 2$, where the additional factor of $\sqrt{\log k}$ is needed for the eigenspace $m/6$ (recall that $\lambda_{m/6} = \frac{1}{2}$), the $(m-1)\times(m-1)$ matrix $\mathcal{D}$ is composed of rotation matrices

$$\mathcal{D} = \begin{pmatrix} \cos\left(\frac{2\pi}{m}\right) & -\sin\left(\frac{2\pi}{m}\right) & & & & \\ \sin\left(\frac{2\pi}{m}\right) & \cos\left(\frac{2\pi}{m}\right) & & & & \\ & & \ddots & & & \\ & & & \cos\left(\frac{2\pi(m/2-1)}{m}\right) & -\sin\left(\frac{2\pi(m/2-1)}{m}\right) & \\ & & & \sin\left(\frac{2\pi(m/2-1)}{m}\right) & \cos\left(\frac{2\pi(m/2-1)}{m}\right) & \\ & & & & & -1 \end{pmatrix}$$

and the error term F_n is made up of three components: Setting

$$G_{k,n}(\ell) := \frac{\Gamma(\ell+1+\omega^k)}{\Gamma(\ell+1)\Gamma(1+\omega^k)} + \omega^k\frac{\Gamma((n-1-\ell)+1+\omega^k)}{\Gamma((n-1-\ell)+1)\Gamma(1+\omega^k)} - \frac{\Gamma(n+1+\omega^k)}{\Gamma(n+1)\Gamma(1+\omega^k)} \tag{3.8}$$

for $\ell \in \{0, \dots, n-1\}$, we have $F_n = F_n^{(1)} + F_n^{(2)}$, where

$$F_n^{(1)} := \begin{pmatrix} \Re\left(G_{1,n}(I_n)\right) \\ \Im\left(G_{1,n}(I_n)\right) \\ \vdots \\ \Re\left(G_{r,n}(I_n)\right) \\ \Im\left(G_{r,n}(I_n)\right) \\ \Re\left(G_{r+1,n}(I_n)\right) \\ \Im\left(G_{r+1,n}(I_n)\right) \\ \vdots \\ 0 \end{pmatrix} - \begin{pmatrix} \Re\left(n^{\omega} g_1(U)\right) \\ \Im\left(n^{\omega} g_1(U)\right) \\ \vdots \\ \Re\left(n^{\omega^r} g_r(U)\right) \\ \Im\left(n^{\omega^r} g_r(U)\right) \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

and $F_n^{(2)}$ is given by the sum

$$\begin{pmatrix} \Re\left(\left(I_n^{\omega} - (nU)^{\omega}\right)\Xi_1^{(0)} + \left(J_n^{\omega} - (n(1-U))^{\omega}\right)\omega\Xi_1^{(1)}\right) \\ \Im\left(\left(I_n^{\omega} - (nU)^{\omega}\right)\Xi_1^{(0)} + \left(J_n^{\omega} - (n(1-U))^{\omega}\right)\omega\Xi_1^{(1)}\right) \\ \vdots \\ \Re\left(\left(I_n^{\omega^r} - (nU)^{\omega^r}\right)\Xi_r^{(0)} + \left(J_n^{\omega^r} - (n(1-U))^{\omega^r}\right)\omega^r\Xi_r^{(1)}\right) \\ \Im\left(\left(I_n^{\omega^r} - (nU)^{\omega^r}\right)\Xi_r^{(0)} + \left(J_n^{\omega^r} - (n(1-U))^{\omega^r}\right)\omega^r\Xi_r^{(1)}\right) \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Note that $\mathcal{D}M_m\mathcal{D}^t = M_m$.

3.3 The Zolotarev metric

In the last subsection, we prepared a proof of Proposition 2.1 that is based on the contraction method. To be more precise, weak convergence in Proposition 2.1 is shown by (the stronger) convergence within the Zolotarev metric. The Zolotarev metric has been studied systematically in the context of distributional recurrences in [17]. We only give the definitions of the relevant quantities and properties here.

For $x \in \mathbb{R}^d$, we denote by $\|x\|$ the standard Euclidean norm of x, and for $B \in \mathbb{R}^{d\times d}$, $\|B\|_{\mathrm{op}}$ denotes the corresponding operator norm. For random variables X and $p \geq 1$, we denote by $\|X\|_p$ the L_p-norm of X.

For two $\mathbb{R}^d$ valued random variables X and Y we set

$$\zeta_3(X, Y) := \sup_{f\in\mathcal{F}_3} |\mathbb{E}[f(X) - f(Y)]|,$$

where

$$\mathcal{F}_3 := \left\{f \in C^2(\mathbb{R}^d, \mathbb{R}) : \|D^2 f(x) - D^2 f(y)\|_{\mathrm{op}} \leq \|x - y\|, \;\; x, y \in \mathbb{R}^d\right\}.$$

We call a pair (X, Y) ζ_3-compatible if the expectation and the covariance matrix of X and Y coincide and if both $\|X\|_3, \|Y\|_3 < \infty$. This implies that $\zeta_3(X, Y) < \infty$. A basic property is that ζ_3 is $(3, +)$-ideal, i.e.,

$$\zeta_3(X + Z, Y + Z) \leq \zeta_3(X, Y), \quad \zeta_3(cX, cY) = c^3\zeta_3(X, Y)$$

for random vectors X, Y, Z, where Z is independent of X, Y and $c > 0$. For a linear transformation A of $\mathbb{R}^d$, we have

$$\zeta_3(AX, AY) \leq \|A\|_{\mathrm{op}}^3\zeta_3(X, Y). \tag{3.9}$$

The following lemma will be used in the proof of Proposition 2.1 and can be proved similarly to Lemma 2.1 in [16].

Lemma 3.4. Let V_1, V_2, W_1, W_2 be random variables in $\mathbb{R}^d$ such that (V_1, V_2) and $(V_1 + W_1, V_2 + W_2)$ are ζ_3-compatible. Then we have

$$\zeta_3(V_1 + W_1, V_2 + W_2) \leq \zeta_3(V_1, V_2) + \sum_{i=1}^{2} \left(\|V_i\|_3^2 \|W_i\|_3 + \frac{\|V_i\|_3 \|W_i\|_3^2}{2} + \frac{\|W_i\|_3^3}{2} \right).$$

In order to work with the Zolotarev metric later, it is necessary to adjust the covariance matrix of Z_n. I.e., we need to work with a sequence of random vectors that is sufficiently close to $(Z_n)_{n\geq 0}$ and has fixed covariance matrix M_m to guarantee the finiteness of the corresponding Zolotarev distances ζ_3.

As noted in section 3.1, the covariance matrices $(\mathrm{Cov}(Z_n))_{n\geq 0}$ converge componentwise to M_m, and M_m is invertible. Thus, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $\mathrm{Cov}(Z_n)$ is invertible. Defining

$$\Sigma_n := \mathbb{1}_{\{n<n_0\}} \mathrm{Id}_m + \mathbb{1}_{\{n\geq n_0\}} M_m^{1/2} \mathrm{Cov}(Z_n)^{-1/2}, \tag{3.10}$$

Σ_n is invertible for all $n \geq 0$ and we see that $\Sigma_n Z_n$ has covariance matrix M_m for all $n \geq n_0$. We now set

$$N_n := \Sigma_n Z_n = A_n^{(0)} N_{I_n}^{(0)} + A_n^{(1)} N_{J_n}^{(1)} + b_n, \tag{3.11}$$

where the right hand side is a recursive decomposition of N_n with coefficients

$$A_n^{(0)} := \Sigma_n \sigma_n \sigma_{I_n}^{-1} \Sigma_{I_n}^{-1}, \ \ A_n^{(1)} := \Sigma_n \sigma_n \sigma_{J_n}^{-1} \mathcal{D} \Sigma_{J_n}^{-1}, \ \ b_n := \Sigma_n \sigma_n \left(F_n^{(1)} + F_n^{(2)} \right).$$

3.4 Preparatory lemmata

In this section we collect some technical lemmata needed in the proof of Proposition 2.1 in the next section. We first look at the asymptotics of the coefficients arising in recursion (3.11).

Lemma 3.5. For all $1 \leq p < \infty$, as $n \to \infty$,

$$\left\| A_n^{(0)} - \sqrt{U} \cdot \mathrm{Id}_{m-1} \right\|_p \to 0 \qquad \text{and} \qquad \left\| A_n^{(1)} - \sqrt{1-U} \cdot \mathcal{D} \right\|_p \to 0.$$

Proof. We first check almost sure convergence. Both $\sqrt{I_n/n}, \sqrt{(I_n \log I_n)/(n \log n)} \to \sqrt{U}$ and $\sqrt{J_n/n}, \sqrt{(J_n \log J_n)/(n \log n)} \to \sqrt{1-U}$ a.s. as $n \to \infty$. Also, because $I_n \to \infty$ a.s. as $n \to \infty$, both $\Sigma_n, \Sigma_{I_n}^{-1} \to \mathrm{Id}_{m-1}$. The claim now follows for all $1 \leq p < \infty$ by an application of the dominated convergence theorem. □

Lemma 3.6. Let $k \in \{1, \ldots, r\}$. As $n \to \infty$,

$$\left\| \left(\frac{I_n}{n} \right)^{\omega^k} - U^{\omega^k} \right\|_3 = \mathrm{O}\left(n^{-\lambda_k/2} \right).$$

Proof. The triangle inequality implies

$$\left\| \left(\frac{I_n}{n} \right)^{\omega^k} - U^{\omega^k} \right\|_3 \leq \left\| \left(\frac{I_n}{n} \right)^{\lambda_k} - U^{\lambda_k} \right\|_3 + \mu_k \left\| \left(\frac{I_n}{n} \right)^{\lambda_k} \log\left(\frac{I_n}{nU} \right) \right\|_3. \tag{3.12}$$

We start by considering the first summand in the latter display. Denoting by $B_{n-1,U}$ a mixed binomial distribution with parameters $n-1$ and U, we see that

$$\left\|\left(\frac{I_n}{n}\right)^{\lambda_k} - U^{\lambda_k}\right\|_3 \leq \left\|\left(\frac{I_n}{n}\right) - U\right\|_3^{\lambda_k} = \mathbb{E}\left[\mathbb{E}\left[\left|\frac{B_{n-1,U}}{n} - U\right|^3 |U\right]\right]^{\frac{\lambda_k}{3}}$$

since I_n, conditional on $U=u$, has the $B_{n-1,u}$ distribution. Using the Marcinkiewickz–Zygmund inequality, there exists a constant C independent of $u \in [0,1]$ such that

$$\mathbb{E}\left[|B_{n-1,u} - (n-1)u|^3\right] \leq Cn^{\frac{3}{2}}.$$

This implies $\left\|\left(\frac{I_n}{n}\right)^{\lambda_k} - U^{\lambda_k}\right\|_3 = \mathrm{O}\left(n^{-\lambda_k/2}\right)$. For the analysis of the second summand in (3.12), we also condition on U and write

$$\left\|\left(\frac{I_n}{n}\right)^{\lambda_k} \log\left(\frac{I_n}{nU}\right)\right\|_3^3 = \int_0^1 \mathbb{E}\left[\left|\left(\frac{B_{n-1,u}}{n}\right)^{\lambda_k} \log\left(\frac{B_{n-1,u}}{nu}\right)\right|^3\right] du$$

We divide the integral into two parts. For this purpose, define $\mathrm{E}_u := \{B_{n-1,u} \geq \frac{un}{e}\}$. Chernoff's inequality implies that for $0 \leq t < u(n-1)$

$$\mathbb{P}\left(B_{n-1,u} - u(n-1) < -t\right) \leq \exp(-t^2/(2u(n-1))),$$

so the complement E_u^c of E_u satisfies $\mathbb{P}(\mathrm{E}_u^c) \leq \exp(-C_0 un)$ for some constant $C_0 > 0$. We further denote by $h_{\lambda_k} : [0,\infty) \to \mathbb{R}$ the function $h_{\lambda_k}(x) := x^{\lambda_k}\log(x)$ (convention: $0 \cdot \log 0 := 0$). Then $\sup_{x\in[0,1]} |h_{\lambda_k}(x)| = \frac{1}{\lambda_k e} < \frac{2}{e} < 1$. We can now bound the expectation on E_u^c in the following way:

$$\begin{aligned}\mathbb{E}\left[\left|\left(\frac{B_{n-1,u}}{n}\right)^{\lambda_k} \log\left(\frac{B_{n-1,u}}{nu}\right)\right|^3 \mathbb{1}_{\mathrm{E}_u^c}\right] &= \int_{\mathrm{E}_u^c} u^{3\lambda_k}\left|h_{\lambda_k}\left(\frac{B_{n-1,u}}{un}\right)\right|^3 d\mathbb{P} \\ &\leq u^{3\lambda_k}\exp\left(-C_0 un\right).\end{aligned}$$

On E_u, we apply the mean value theorem to $h_1\left((1+y)^{\lambda_k}\right) - h_1\left(1^{\lambda_k}\right)$ with $y = \frac{B_{n-1,u}-nu}{nu}$. Note that $(\min\{1,1+y\}, \max\{1,1+y\}) \subset [\frac{1}{e}, \frac{1}{u}]$ on E_u and that $|h_1'|$ is nonnegative and increasing on this interval. Thus,

$$\begin{aligned}&\mathbb{E}\left[\left|\left(\frac{B_{n-1,u}}{n}\right)^{\lambda_k} \log\left(\frac{B_{n-1,u}}{nu}\right)\right|^3 \mathbb{1}_{\mathrm{E}_u}\right] \\ &= \int_{\mathrm{E}_u}\left(\frac{1}{\lambda_k}\right)^3 u^{3\lambda_k}\left|h_1\left(\left(1+\frac{B_{n-1,u}-nu}{nu}\right)^{\lambda_k}\right) - h_1\left(1^{\lambda_k}\right)\right|^3 d\mathbb{P} \\ &\leq \int_{\mathrm{E}_u}\left(\frac{1}{\lambda_k}\right)^3 u^{3\lambda_k}\left(\sup_{v\in[\frac{1}{e},\frac{1}{u}]}|h_1'(v)|\right)^3\left|\left(1+\frac{B_{n-1,u}-nu}{nu}\right)^{\lambda_k} - 1^{\lambda_k}\right|^3 d\mathbb{P} \\ &\leq \int_{\mathrm{E}_u}\left(\frac{1}{\lambda_k}\right)^3 u^{3\lambda_k}\left(h_1'\left(\frac{1}{u}\right)\right)^3\left|\frac{B_{n-1,u}-nu}{nu}\right|^{3\lambda_k} d\mathbb{P} \\ &\leq \left(\frac{1}{\lambda_k}\right)^3 n^{-3\lambda_k}(1-\log(u))^3\mathbb{E}\left[|-u+B_{n-1,u}-(n-1)u|^3\right]^{\lambda_k} \\ &\leq C_k\ \frac{(1-\log(u))^3}{n^{3\lambda_k/2}}\end{aligned}$$

for some constant $C_k > 0$. Combining these estimates, we obtain

$$\left\| \left(\frac{I_n}{n}\right)^{\lambda_k} \log\left(\frac{I_n}{nU}\right) \right\|_3^3 \leq \int_0^1 \left(u^{3\lambda_k} \exp\left(-C_0 un\right) + C_k \ \frac{(1-\log(u))^3}{n^{3\lambda_k/2}} \right) du$$
$$= \mathrm{O}\left(\frac{1}{n^{3\lambda_k/2}}\right)$$

as $n \to \infty$. This implies the assertion. □

Lemma 3.7. As $n \to \infty$, we have

$$\|b_n\|_3 \longrightarrow 0.$$

Proof. By the triangle inequality,

$$\|b_n\|_3 \leq \|\Sigma_n\|_{\mathrm{op}} \sum_{j=1}^{2} \left\| \sigma_n F_n^{(j)} \right\|_3 .$$

We have $(I_n, U, \Xi_k^{(0)}) \stackrel{d}{=} (J_n, 1-U, \Xi_k^{(1)})$ with $\Xi_k^{(0)}$ independent of (I_n, U). The triangle inequality implies

$$\left\| \sigma_n F_n^{(2)} \right\|_3 \leq \frac{4}{\sqrt{n}} \sum_{k=1}^{r} n^{\lambda_k} \left\| \Xi_k^{(0)} \right\|_3 \left\| \left(\frac{I_n}{n}\right)^{\omega^k} - U^{\omega^k} \right\|_3$$
$$= \frac{4}{\sqrt{n}} \sum_{k=1}^{r} \mathrm{O}\left(n^{\lambda_k/2}\right) = \mathrm{o}(1)$$

by Lemma (3.6). Also, for $n \to \infty$,

$$\begin{aligned}
\left\| \sigma_n F_n^{(1)} \right\|_3 \leq \frac{2}{\sqrt{n}} & \left(\sum_{k=1}^{r} \left\| G_{k,n}(I_n) - n^{\omega^k} g_k(U) \right\|_3 \right. \\
& \left. + \frac{1}{\sqrt{\log(n)}} \|G_{r+1,n}(I_n)\|_3 + \sum_{k=r+2}^{m/2-1} \|G_{k,n}(I_n)\|_3 \right) \\
\leq & \frac{2}{\sqrt{n}} \left(\sum_{k=1}^{r} \frac{2}{\Gamma(1+\omega^k)} n^{\lambda_k} \left\| \left(\frac{I_n}{n}\right)^{\omega^k} - U^{\omega^k} \right\|_3 \right. \\
& \left. + \frac{1}{\sqrt{\log(n)}} \|G_{r+1,n}(I_n)\|_3 + \sum_{k=r+2}^{m/2-1} \|G_{k,n}(I_n)\|_3 \right) + \mathrm{o}(1) \\
= & \ \mathrm{o}(1)
\end{aligned}$$

as before. Now, the sequence $(\|\Sigma_n\|_{\mathrm{op}})_{n \geq 0}$ is convergent and thus bounded, which implies the claim. □

Finally, we use recursion (3.11) for N_n to show that the sequence $(\|N_n\|_3)_{n \geq 0}$ is bounded.

Lemma 3.8. As $n \to \infty$, we have

$$\|N_n\|_3 = \mathrm{O}(1).$$

Proof. Recall that the composition vector R_n takes only finitely many values, the random variables Ξ_k have finite absolute moments of arbitrary order, see (3.2), and $\|\Sigma_n\|_{\text{op}} \to 1$. Hence, we have $\|N_n\|_3 < \infty$ for all $n \geq 0$.

Recursion (3.11) implies that

$$\|N_n\| \leq \mathcal{Y}^{(0)} + \mathcal{Y}^{(1)} + \|b_n\|,$$

where $\mathcal{Y}^{(0)} := \left\|A_n^{(0)}\right\|_{\text{op}} \left\|N_{I_n}^{(0)}\right\|$ and $\mathcal{Y}^{(1)} := \left\|A_n^{(1)}\right\|_{\text{op}} \left\|N_{J_n}^{(1)}\right\|$. For all $n \geq 0$,

$$\begin{aligned}\mathbb{E}\left[\|N_n\|^3\right] \leq \mathbb{E}\left[\left(\mathcal{Y}^{(0)}\right)^3\right] + \mathbb{E}\left[\left(\mathcal{Y}^{(1)}\right)^3\right] + \mathbb{E}\left[\|b_n\|^3\right] + 3\mathbb{E}\left[\left(\mathcal{Y}^{(0)}\right)^2 \mathcal{Y}^{(1)}\right] \\ + 3\mathbb{E}\left[\left(\mathcal{Y}^{(1)}\right)^2 \mathcal{Y}^{(0)}\right] + 3\mathbb{E}\left[\left(\mathcal{Y}^{(0)}\right)^2 \|b_n\|\right] + 3\mathbb{E}\left[\mathcal{Y}^{(0)}\|b_n\|^2\right] \\ + 3\mathbb{E}\left[\left(\mathcal{Y}^{(1)}\right)^2 \|b_n\|\right] + 3\mathbb{E}\left[\mathcal{Y}^{(1)}\|b_n\|^2\right] + 6\mathbb{E}\left[\mathcal{Y}^{(0)}\mathcal{Y}^{(1)}\|b_n\|\right]. \qquad (3.13)\end{aligned}$$

Set

$$\beta_n := 1 \vee \max_{0\leq k\leq n} \mathbb{E}\left[\|N_k\|^3\right].$$

By Lemma 3.7, $\mathbb{E}\left[\|b_n\|^3\right] \to 0$ as $n \to \infty$. Also,

$$\mathbb{E}\left[\left(\mathcal{Y}^{(j)}\right)^3\right] = \mathbb{E}\left[\left\|A_n^{(j)}\right\|_{\text{op}}^3 \sum_{k=0}^{n-1} \mathbb{1}_{\{I_n=k\}}\mathbb{E}\left[\|N_k\|^3\right]\right] \leq \mathbb{E}\left[\left\|A_n^{(j)}\right\|_{\text{op}}^3\right] \beta_{n-1}$$

for $j = 0, 1$.

To bound the summand $\mathbb{E}\left[\left(\mathcal{Y}^{(0)}\right)^2 \mathcal{Y}^{(1)}\right]$, note that $\left\|A_n^{(0)}\right\|_{\text{op}}$ and $\left\|A_n^{(1)}\right\|_{\text{op}}$ are uniformly bounded in n. This implies that after conditioning on I_n, there is a constant $D > 0$ such that

$$\begin{aligned}\mathbb{E}\left[\left(\mathcal{Y}^{(0)}\right)^2 \mathcal{Y}^{(1)}\right] &\leq D\mathbb{E}\left[\sum_{k=0}^{n-1} \mathbb{1}_{\{I_n=k\}}\mathbb{E}\left[\|N_k\|^2\right]\mathbb{E}\left[\|N_{n-1-k}\|\right]\right] \\ &\leq D\left(\max_{0\leq k\leq n-1} \|N_k\|_2^2\right)\left(\max_{0\leq k\leq n-1} \|N_k\|_1\right).\end{aligned}$$

Now, by construction, $\mathrm{Cov}(N_n) = M_m$ for all $n \geq n_0$, so $\max_{0\leq k\leq n-1} \|N_k\|_2^2 < K$ for some $K > 0$ and hence $\mathbb{E}\left[\left(\mathcal{Y}^{(0)}\right)^2 \mathcal{Y}^{(1)}\right] = \mathrm{O}(1)$. The same applies to $\mathbb{E}\left[\left(\mathcal{Y}^{(1)}\right)^2 \mathcal{Y}^{(0)}\right]$.

All other summands in (3.13) can be bounded using Hölder's inequality. Combining all these bounds leads to the estimate

$$\mathbb{E}\left[\|N_n\|^3\right] \leq \left(\mathbb{E}\left[\left\|A_n^{(0)}\right\|_{\text{op}}^3 + \left\|A_n^{(1)}\right\|_{\text{op}}^3\right] + \mathrm{o}(1)\right)\beta_{n-1} + \mathrm{O}(1).$$

The asymptotics in Lemma 3.5 further imply

$$\mathbb{E}\left[\|N_n\|^3\right] \leq \left(\mathbb{E}\left[U^{3/2} + (1-U)^{3/2}\right] + \mathrm{o}(1)\right)\beta_{n-1} + \mathrm{O}(1) = \left(\frac{4}{5} + \mathrm{o}(1)\right)\beta_{n-1} + \mathrm{O}(1).$$

Since $\beta_n \geq 1$, there exist $J \in \mathbb{N}$ and a constant $0 < E < \infty$ such that for all $n \geq J$, $\mathbb{E}\left[\|N_n\|^3\right] \leq (9/10)\beta_{n-1} + E$. Induction on n gives that for all $n \geq 0$, $\mathbb{E}\left[\|N_n\|^3\right] \leq \max\{\beta_J, 10E\}$. □

3.5 Proof of Proposition 2.1

Proof of Proposition 2.1. Proposition 2.1 claims the convergence $Z_n \stackrel{d}{\longrightarrow} \mathcal{N}$, as $n \to \infty$, where $\mathcal{N} \sim \mathcal{N}(0, M_m)$. In order to establish this convergence, the key point is to show that

$$\zeta_3(N_n, \mathcal{N}) \longrightarrow 0 \quad \text{as} \quad n \to \infty.$$

This is sufficient, as the difference $Z_n - N_n$ tends to 0 in probability and convergence in the Zolotarev metric implies weak convergence of probability measures on $\mathbb{R}^{m-1}$.

Recall that N_n satisfies (3.11) and that $\mathcal{N}(0, M_m)$ is a solution to the distributional recursion

$$\mathcal{N} \stackrel{d}{=} \sqrt{U}\mathcal{N}^{(0)} + \sqrt{1-U}\mathcal{D}\mathcal{N}^{(1)},$$

where $\mathcal{N}^{(0)}, \mathcal{N}^{(1)}$ and U are independent, U is uniform on $[0,1]$ and $\mathcal{N}^{(0)}$ and $\mathcal{N}^{(1)}$ have the same distribution as $\mathcal{N}$.

First, we use recursion (3.11) for N_n to define hybrid random variables that link N_n to $\mathcal{N}(0, M_m)$ as follows: Let $\mathcal{N}^{(0)}$ and $\mathcal{N}^{(1)}$ be defined on the same probability space as $(U_n)_{n\geq 1}$, independent with distribution $\mathcal{N}(0, M_m)$ and also independent of $(U_n)_{n\geq 1}$. We ignore the error term b_n in (3.11) and set

$$Q_n := A_n^{(0)}\left(\mathbb{1}_{\{I_n < n_0\}} N_{I_n}^{(0)} + \mathbb{1}_{\{I_n \geq n_0\}}\mathcal{N}^{(0)}\right) + A_n^{(1)}\left(\mathbb{1}_{\{J_n < n_0\}} N_{J_n}^{(1)} + \mathbb{1}_{\{J_n \geq n_0\}}\mathcal{N}^{(1)}\right)$$

for $n \geq 1$, while $Q_0 := N_0$. Q_n does not necessarily have covariance matrix M_m. However, I_n/n converges to the uniform random variable U almost surely. Together with Lemma 3.5, we obtain

$$\mathrm{Cov}(Q_n) \to M_m.$$

In order to ensure finiteness of the Zolotarev metric, the covariance matrix of Q_n has to be adjusted. However, as $\mathcal{N}^{(0)}$ and $\mathcal{N}^{(1)}$ are independent of $(U_n)_{n\geq 1}$ and have independent components, $\mathrm{Cov}(Q_n)$ has full rank for all $n > n_0$. This implies that we can find a deterministic sequence of matrices $(B_n)_{n\geq 0}$ with $\mathrm{Cov}(B_n Q_n) = M_m$ for all $n > n_0$ and $B_n \to \mathrm{Id}_{m-1}$ componentwise and in operator norm as $n \to \infty$. We write $B_n = \mathrm{Id}_{m-1} + K_n$ with $(K_n)_{n\geq 0}$ tending to the all zero matrix componentwise.

Hence, with $\mathcal{N}$ as before, each pair of N_n, $(\mathrm{Id}_{m-1} + K_n)Q_n$ and $\mathcal{N}$ is ζ_3-compatible for $n > n_0$ and the triangle inequality implies

$$\zeta_3(N_n, \mathcal{N}) \leq \zeta_3(N_n, (\mathrm{Id}_{m-1} + K_n)Q_n) + \zeta_3((\mathrm{Id}_{m-1} + K_n)Q_n, \mathcal{N}), \tag{3.14}$$

which is finite for all $n > n_0$.

First we show that $\zeta_3((\mathrm{Id}_{m-1} + K_n)Q_n, \mathcal{N}) = \mathrm{o}(1)$ by use of an upper bound of ζ_3 by the minimal L_3-metric ℓ_3. The minimal L_3-metric ℓ_3 is given by

$$\ell_3(X, Y) := \ell_3(\mathcal{L}(X), \mathcal{L}(Y)) := \inf\{\|X' - Y'\|_3 : \mathcal{L}(X) = \mathcal{L}(X'), \mathcal{L}(Y) = \mathcal{L}(Y')\}, \tag{3.15}$$

for all random vectors X, Y with $\|X\|_3, \|Y\|_3 < \infty$. For a ζ_3-compatible pair (X, Y), we have the inequality, see [5, Lemma 5.7],

$$\zeta_3(X, Y) \leq \left(\|X\|_3^2 + \|Y\|_3^2\right)\ell_3(X, Y).$$

As $\sup_{n\geq 0}\|Q_n\|_3 < \infty$ by Lemma 3.5 and the properties of the Gaussian distribution, also $\|(\mathrm{Id}_{m-1} + K_n)Q_n)\|_3$ is uniformly bounded in n. So there exists a finite constant $C > 0$

with

$$\zeta_3((\mathrm{Id}_{m-1} + K_n)Q_n, \mathcal{N}) \leq C\ell_3((\mathrm{Id}_{m-1} + K_n)Q_n, \mathcal{N})$$

for all $n > n_0$. In order to upper bound the latter ℓ_3-distance, recall that the random vectors $\mathcal{N}$ and $\sqrt{U}\mathcal{N}^{(0)} + \sqrt{1-U}\mathcal{D}\mathcal{N}^{(1)}$ are identically distributed. Thus for $n \geq n_0$,

$$\begin{aligned}
&\zeta_3((\mathrm{Id}_{m-1} + K_n)Q_n, \mathcal{N}) \leq C\ell_3((\mathrm{Id}_{m-1} + K_n)Q_n, \mathcal{N}) \\
&\leq C \left\| \left((\mathrm{Id}_{m-1} + K_n)A_n^{(0)} \mathbb{1}_{\{I_n \geq n_0\}} - \sqrt{U}\mathrm{Id}_{m-1} \right) \mathcal{N}^{(0)} \right. \\
&\quad \left. + \left((\mathrm{Id}_{m-1} + K_n)A_n^{(1)} \mathbb{1}_{\{J_n \geq n_0\}} - \sqrt{1-U}\mathcal{D} \right) \mathcal{N}^{(1)} \right\|_3 \\
&\quad + C \left\| (\mathrm{Id}_{m-1} + K_n)A_n^{(0)} \mathbb{1}_{\{I_n < n_0\}} N_{I_n}^{(0)} + (\mathrm{Id}_{m-1} + K_n)A_n^{(1)} \mathbb{1}_{\{J_n < n_0\}} N_{J_n}^{(1)} \right\|_3 \\
&\leq C \left(\left\| (\mathrm{Id}_{m-1} + K_n)A_n^{(0)} \mathbb{1}_{\{I_n \geq n_0\}} - \sqrt{U}\mathrm{Id}_{m-1} \right\|_3 \left\| \mathcal{N}^{(0)} \right\|_3 \right. \\
&\quad \left. + \left\| (\mathrm{Id}_{m-1} + K_n)A_n^{(1)} \mathbb{1}_{\{J_n \geq n_0\}} - \sqrt{1-U}\mathcal{D} \right\|_3 \left\| \mathcal{N}^{(1)} \right\|_3 \right) \\
&\quad + C \left\| (\mathrm{Id}_{m-1} + K_n)A_n^{(0)} \mathbb{1}_{\{I_n < n_0\}} N_{I_n}^{(0)} + (\mathrm{Id}_{m-1} + K_n)A_n^{(1)} \mathbb{1}_{\{J_n < n_0\}} N_{J_n}^{(1)} \right\|_3 \stackrel{n\to\infty}{\longrightarrow} 0.
\end{aligned}$$

To bound the first summand in (3.14), we split N_n into two parts and consider the vector

$$\Phi_n := A_n^{(0)} N_{I_n}^{(0)} + A_n^{(1)} N_{J_n}^{(1)}, \quad n \geq 1,$$

with $\Phi_0 := N_0$ such that $N_n = \Phi_n + b_n$. An application of Lemma 3.4 to the sums $N_n = \Phi_n + b_n$ and $(\mathrm{Id}_{m-1} + K_n)Q_n = Q_n + K_n Q_n$ gives for $n > n_0$ that

$$\begin{aligned}
\zeta_3(N_n, (\mathrm{Id}_{m-1} + K_n)Q_n) &\leq \zeta_3(\Phi_n, Q_n) + \|\Phi_n\|_3^2 \|b_n\|_3 + \frac{1}{2}\|\Phi_n\|_3 \|b_n\|_3^2 + \frac{1}{2}\|b_n\|_3^3 \\
&\quad + \left(\|K_n\|_{\mathrm{op}} + \frac{1}{2}\|K_n\|_{\mathrm{op}}^2 + \frac{1}{2}\|K_n\|_{\mathrm{op}}^3 \right) \|Q_n\|_3^3.
\end{aligned}$$

By construction, $\|K_n\|_{\mathrm{op}} \to 0$ and by Lemma 3.7, $\|b_n\|_3 \to 0$. Also, by Lemma 3.8, $\sup_{n\geq 0} \|\Phi_n\|_3 < \infty$ and $\sup_{n\geq 0} \|Q_n\|_3 < \infty$, this yields that, as $n \to \infty$,

$$\zeta_3(N_n, (\mathrm{Id}_{m-1} + K_n)Q_n) \leq \zeta_3(\Phi_n, Q_n) + \mathrm{o}(1).$$

The previous estimates and (3.14) imply that, as $n \to \infty$,

$$\zeta_3(N_n, \mathcal{N}) \leq \zeta_3 (\Phi_n, Q_n) + \mathrm{o}(1). \tag{3.16}$$

Let $\Delta(n) := \zeta_3(N_n, \mathcal{N})$, which is finite for $n \geq n_0$. Note that $\zeta_3 (\Phi_n, Q_n)$ is finite for $n \geq 0$. In the expectations defining the Zolotarev distance, we condition on the value of I_n. With $(N_0^{[0]}, \dots, N_{n-1}^{[0]}),(N_0^{[1]}, \dots, N_{n-1}^{[1]})$ i.i.d. with distribution $\mathcal{L}(N_0, \dots, N_{n-1})$ we make use of independence and the fact that ζ_3 is $(3,+)$-ideal and satisfies (3.9) to get, for $n > 2n_0$,

$$\begin{aligned}\zeta_3\left(\Phi_n, Q_n\right) \leq & \frac{1}{n}\sum_{k=0}^{n_0-1}\zeta_3\left(\Sigma_n\sigma_n\sigma_{n-1-k}\mathcal{D}\Sigma_{n-1-k}^{-1}N_{n-1-k}^{[1]}, \Sigma_n\sigma_n\sigma_{n-1-k}\mathcal{D}\Sigma_{n-1-k}^{-1}\mathcal{N}^{(1)}\right)\\ &+\frac{1}{n}\sum_{k=n-n_0}^{n-1}\zeta_3\left(\Sigma_n\sigma_n\sigma_k\Sigma_k^{-1}N_k^{[0]}, \Sigma_n\sigma_n\sigma_k\Sigma_k^{-1}\mathcal{N}^{(0)}\right)\\ &+\frac{1}{n}\sum_{k=n_0}^{n-n_0-1}\zeta_3\Big(\Sigma_n\sigma_n\sigma_k^{-1}\Sigma_k^{-1}N_k^{[0]}+\Sigma_n\sigma_n\sigma_{n-1-k}\mathcal{D}\Sigma_{n-1-k}^{-1}N_{n-1-k}^{[1]},\\ &\qquad \Sigma_n\sigma_n\sigma_k^{-1}\Sigma_k^{-1}\mathcal{N}^{(0)}+\Sigma_n\sigma_n\sigma_{n-1-k}^{-1}\mathcal{D}\Sigma_{n-1-k}^{-1}\mathcal{N}^{(1)}\Big)\\ &\leq \frac{2}{n}\sum_{k=n-n_0}^{n-1}\|\sigma_n\sigma_k^{-1}\|_{\mathrm{op}}^3\|\Sigma_n\|_{\mathrm{op}}^3\|\Sigma_k^{-1}\|_{\mathrm{op}}^3\zeta_3\left(N_k^{[0]},\mathcal{N}^{(0)}\right)\\ &+\frac{2}{n}\sum_{k=n_0}^{n-n_0}\|\sigma_n\sigma_k^{-1}\|_{\mathrm{op}}^3\|\Sigma_n\|_{\mathrm{op}}^3\|\Sigma_k^{-1}\|_{\mathrm{op}}^3\zeta_3\left(N_k^{[0]},\mathcal{N}^{(0)}\right)\\ &=\frac{2}{n}\sum_{k=n_0}^{n-1}\|\sigma_n\sigma_k^{-1}\|_{\mathrm{op}}^3\|\Sigma_n\|_{\mathrm{op}}^3\|\Sigma_k^{-1}\|_{\mathrm{op}}^3\zeta_3\left(N_k^{[0]},\mathcal{N}^{(0)}\right).\end{aligned}$$

Note that $\|\sigma_n\sigma_{I_n}^{-1}\|_{\mathrm{op}}^3 = \left(\frac{I_n}{n}\right)^{3/2}$ in both cases $6 \mid m$ and $6 \nmid m$. Hence, for $6 \mid m$ and $n > 2n_0$,

$$\Delta(n) \leq 2\mathbb{E}\left[\left(\frac{I_n}{n}\right)^{3/2}\|\Sigma_n\|_{\mathrm{op}}^3\|\Sigma_{I_n}^{-1}\|_{\mathrm{op}}^3\Delta(I_n)\mathbb{1}_{\{I_n\geq n_0\}}\right]+\mathrm{o}(1).$$

Now a standard argument shows that $\zeta_3(N_n, \mathcal{N}) \to 0$ as $n \to \infty$, see [16], for example. □

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. MR-2957340

[2] Chauvin, B., Mailler, C. and Pouyanne, N. (2015) Smoothing equations for large Pólya urns. *J. Theor. Probab.* **28**, 923–957. MR-3413961

[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. MR-2335295

[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. MR-1871558

[5] Drmota, M., Janson, S. and Neininger, R. (2008) A functional limit theorem for the profile of search trees. *Ann. Appl. Probab.* **18**, 288–333. MR-2380900

[6] Evans, S.N., Grübel, R. and Wakolbinger, A. (2012) Trickle-down processes and their boundaries. *Electron. J. Probab.* **17**, 1-58. MR-2869248

[7] Freedman, D. A. (1965) Bernard Friedman's Urn. *Ann. Math. Statist.* **36**, no. 3, 956–970. MR-0177432

[8] Grübel, R. (2014) Search trees: Metric aspects and strong limit theorems. *Ann. Appl. Probab.* **24**, 1269–1297. MR-3199986

[9] Janson, S. (1983) Limit theorems for certain branching random walks on compact groups and homogeneous spaces. *Ann. Probab.* **11**, 909–930. MR-0714955

[10] Janson, S. (2004) Functional limit theorem for multitype branching processes and generalized Pólya urns. *Stochastic Process. Appl.* **110**, 177–245. MR-2040966

[11] Janson, S. (2006) Congruence properties of depths in some random trees. *Alea* **1**, 347–366. MR-2249661

[12] Knape, M. and Neininger, R. (2014) Pólya Urns Via the Contraction Method. *Combin. Probab. Comput.* **23**, 1148–1186. MR-3265841

[13] Mahmoud, H. M. (1992) *Evolution of Random Search Trees*, John Wiley & Sons, New York. MR-1140708

[14] Mailler, C. (2018) Balanced multicolour Pólya urns via smoothing systems analysis. *ALEA - Latin American Journal of Probability and Mathematical Statistics* **XV**, 375–408. MR-3795453

[15] Müller, N. S. and Neininger, R. (2016) The CLT Analogue for Cyclic Urns. *Analytic Algorithmics and Combinatorics (ANALCO)*, 121–127. MR-3480253

[16] Neininger, R. (2015) Refined Quicksort asymptotics. *Random Structures Algorithms* **46**, 346–361. MR-3302901

[17] Neininger, R. and Rüschendorf, L. (2004) A general limit theorem for recursive algorithms and combinatorial structures. *Ann. Appl. Probab.* **14**, 378-418. MR-2023025

[18] 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*. MR-2193125

[19] Pouyanne, N. (2008) An algebraic approach to Pólya processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **44**, 293–323. MR-2446325

Acknowledgments. We thank the anonymous referee and Henning Sulzbach for helpful comments.