

CENTRAL LIMIT THEOREMS FOR GENERALIZED PÓLYA URN MODELS

I. HIGUERAS* AND

J. MOLER,** *Universidad Pública de Navarra*

F. PLO*** AND

M. SAN MIGUEL,*** *Universidad de Zaragoza*

Abstract

In this paper we obtain central limit theorems for generalized Pólya urn models with $L \geq 2$ colors where one out of K different replacements (actions) is applied randomly at each step. Each possible action constitutes a row of the replacement matrix, which can be nonsquare and random. The actions are chosen following a probability distribution given by an arbitrary function of the proportions of the balls of different colors present in the urn. Moreover, under the same hypotheses it is proved that the covariance matrix of the asymptotic distribution is the solution of a Lyapunov equation, and a procedure is given to obtain the covariance matrix in an explicit form. Some applications of these results to random trees and adaptive designs in clinical trials are also presented.

Keywords: Generalized Pólya urn model; central limit theorem; Robbins–Monro recurrence equation

2000 Mathematics Subject Classification: Primary 60F05
Secondary 62L20

1. Introduction

In the classical urn model we consider balls of two colors. The replacement policy consists in drawing a ball from the urn, observing its color, and putting it back along with a number of balls of either color. This policy can be described through a replacement matrix $\mathbf{R} = (r_{ij})$ where r_{ij} , $i, j = 1, 2$, denotes the number of balls of color j that must be added to the urn when a ball of color i has been extracted. For instance, the classical Pólya urn has the replacement matrix

$$\mathbf{R} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}.$$

In other words, the rows indicate the different actions that may be taken and the columns represent the number of balls of each color to be added depending on the action taken.

Several generalizations of this model can be found in the probability literature (for a brief account see, for instance, [7]) and have been useful in several applications, such as performing

Received 23 March 2005; revision received 20 September 2006.

* Postal address: Departamento de Matemática e Informática, Universidad Pública de Navarra, Campus de Arrosadía, 31006 Pamplona, Spain.

** Postal address: Departamento de Estadística e Investigación Operativa, Universidad Pública de Navarra, Campus de Arrosadía, 31006 Pamplona, Spain. Email address: jmoler@unavarra.es

*** Postal address: Departamento de Métodos Estadísticos, Universidad de Zaragoza, Pedro Cerbuna 12, 50009 Zaragoza, Spain.

Partially supported by DGA project E22 and MEC project MTM2004-01175.

adaptive design (see, for instance, [18]) or studying the asymptotic behavior of some data structures used in computer science (see, for instance, [21]).

In this paper we consider the following generalized Pólya urn model. An urn initially contains balls of L different colors. The proportion of balls of each color is represented at each step n by the L -dimensional random vector $X_n = (X_{n1}, X_{n2}, \dots, X_{nL})^\top$. We have K different possible actions and at each step one of them is chosen at random with a probability given by some (rather arbitrary) function, G , of the proportions of the balls of different colors that are present in the urn (G will be called the urn function). If the action i , $i = 1, \dots, K$, is applied, we add r_{nij} balls of each color j , $j = 1, \dots, L$. The entries r_{nij} can be random variables. We represent this replacement policy using a rectangular matrix $R_n = (r_{nij})$, called the replacement matrix, where each row represents an action and each column, a color.

The classical Pólya urn model described above turns out to be a particular example of this where the number of actions equals the number of colors, the function G is the identity function, and R is a constant matrix. More complicated examples exist in this framework. For instance, in [22] an urn with two different colors, say black and white, and three possible actions was proposed. The components of the function G in this case are equal to the three probabilities of a double extraction, i.e. extracting two white balls, one white and one black ball, or two black balls. More details are given in Section 4.1 of this paper.

The aim of this paper is to establish conditions that guarantee the existence of a central limit theorem (CLT) for the urn model and to give a procedure to obtain explicit expressions for the covariance matrix of the asymptotic distribution. To the authors' knowledge, there are no CLTs for generalized Pólya urn models with nonsquare replacement matrices. The effect of urn functions has only been treated in [8], where strong laws were obtained for urns with two colors, and in [1], where a CLT for fixed diagonal replacement matrices was obtained.

There exist several papers that provide a CLT when the urn function is the identity function. In [2, Chapter V] CLTs were obtained using the theory of branching processes and, in [6], using martingale theory.

In [20] it was assumed that all the replacement matrices have the same mean value H and that $H\mathbf{1} = s\mathbf{1}$, where s is a positive real number. In this setting, CLTs were obtained under the following assumptions for the eigenvalues and eigenvectors of H : (i) for each nonprincipal eigenvalue λ_j , $\text{Re}(\lambda_j) < s/2$; (ii) all eigenvalues are simple and two distinct complex eigenvalues do not have the same real part, except for conjugate pairs; and (iii) all the eigenvectors are linearly independent.

In [3] the rows of the replacement matrices $\{R_n\}$ were chosen to add to the same value, $s = 1$. CLTs were obtained using martingale techniques when there exists a matrix H such that $\sum_{i=1}^{\infty} \alpha_i / i < \infty$ with $\alpha_i = \|E[R_i] - H\|$ (where $\|\cdot\|$ denotes the matrix norm – see below), and, for each nonprincipal eigenvalue λ_j , $\text{Re}(\lambda_j) < s/2$. Using similar techniques, in [4] the previous results were extended to a wider setting where the matrices $\{R_n\}$ have conditional mean matrices H_n such that all the rows add to the same value, $s = 1$, and $\sum_{i=1}^{\infty} \alpha_i / \sqrt{i} < \infty$ with $\alpha_i = \|H_i - H\|$.

In [9] CLTs were obtained using the theory of branching processes, and a procedure was presented to obtain explicit expressions for the covariance matrix of the asymptotic distribution. Also, in [9, p. 207] a more extensive discussion was presented about the results obtained hitherto in this topic.

In this paper we follow the approach presented in [7]. We prove that the process $\{X_n\}$ evolves as does a Robbins–Monro stochastic approximation scheme (see, for instance, [11, Chapter 5]). Under some extra assumptions, a CLT is obtained for $\{X_n\}$. These assumptions include that the

rows of the replacement matrices $\{\mathbf{R}_n\}$ add up to the same value, $s > 0$, and a second condition which, in the particular case of square matrices and an identity urn function, is equivalent to requiring that, for each nonprincipal eigenvalue λ_j of the matrix limit \mathbf{H} , $\text{Re}(\lambda_j) < s/2$.

The paper is organized as follows. In Section 2 we describe the model and the assumptions in detail, and we state the almost sure (a.s.) convergence results in this particular setting. In Section 3 we obtain our main result, a CLT for the process $\{\mathbf{X}_n\}$. In Section 4 three applications are presented. In the first we obtain a CLT for an urn model with double extraction where G is not the identity function and the replacement matrix is nonsquare. The asymptotic behavior of this model was presented as an open problem in Remark 4.5 of [9]. Moreover, in [22] interesting applications of this model in computer science were presented. The second application illustrates the procedure to obtain the covariance matrix of the asymptotic distribution in a CLT for a random bucket tree presented in [12]. In the third application we present an adaptive design that can be applied in a clinical trial.

2. Generalized Pólya urn model and strong laws

In what follows, the matrix norm $\|\mathbf{R}\| = \sup_i \sum_j |r_{ij}|$ will be used. If $\mathbf{x} \in \mathbb{R}^n$ then $\|\mathbf{x}\| = \sum_{i=1}^n |x_i|$, and the row vector will be denoted by \mathbf{x}^\top . Column vectors of 0s or 1s will be denoted by $\mathbf{0}$ or $\mathbf{1}$, respectively.

We consider an urn that initially contains T_0 balls of L different colors. Each color is identified by a number $i = 1, \dots, L$, and $\mathbf{X}_0 = (X_{01}, \dots, X_{0L})^\top$ indicates the initial proportion of balls of each color. For each n , the random vector \mathbf{X}_n represents the proportion of balls of each color after the n th replacement. Therefore, \mathbf{X}_n takes values in the $(L - 1)$ -simplex set: $\Delta_{L-1} = \{\mathbf{x} \in \mathbb{R}^L : \sum_{i=1}^L x_i = 1, x_i \geq 0\}$. Let T_n be the total number of balls in the urn after the n th replacement. Thus, the number of balls of different colors at step n is given by the vector $T_n \mathbf{X}_n$.

The replacement policy in each step n is associated with a random matrix, $\mathbf{R}_n = (r_{nij})$, which takes values in a set $\mathcal{R}_{K \times L} = \{(r_{ij})_{i=1, \dots, K, j=1, \dots, L} : r_{ij} \in \mathbb{R}, \sum_{j=1}^L r_{ij} > 0\}$, and each row represents an action and each column represents a color. So, the element r_{nij} indicates the number of balls of color j that we add when the action i is applied at the n th replacement.

At each step n , one of K different replacements (actions) is randomly applied. The action applied is represented by a random vector δ_n that takes values in the natural basis of \mathbb{R}^K , denoted $\{\mathbf{e}_i\}_{i=1, \dots, K}$. If $\delta_n = \mathbf{e}_i$, then the i th row of the replacement matrix is applied. Thus, the balls added are given by the vector $\mathbf{R}_n^\top \delta_n$. Consequently, $T_{n+1} \mathbf{X}_{n+1} = T_n \mathbf{X}_n + \mathbf{R}_{n+1}^\top \delta_{n+1}$ and the evolution of the process $\{\mathbf{X}_n\}$ is represented by means of the recurrence equation

$$\mathbf{X}_{n+1} = \mathbf{X}_n + \frac{(\mathbf{I} - \mathbf{X}_n \mathbf{1}^\top) \mathbf{R}_{n+1}^\top \delta_{n+1}}{T_{n+1}}. \tag{2.1}$$

The process $\{U_n\}$, where $U_n = (\mathbf{X}_n, \delta_n, \mathbf{R}_n)$, $n \geq 1$, will be called a generalized Pólya urn model. In what follows we will consider the natural filtration $\{\mathcal{F}_n\}$ with $\mathcal{F}_n = \sigma(U_i, 1 \leq i \leq n)$, $n \geq 1$.

We will assume the following hypotheses to hold in the rest of the paper.

- (A1) There exists a constant $b > 2$ such that $E[\|\mathbf{R}_n\|^b \mid \mathcal{F}_{n-1}] < M$, where M is a positive constant.
- (A2) $\mathbf{R}_n \mathbf{1} = s \mathbf{1}$ for a constant $s > 0$.

(A3) There exists a function $G: \Delta_{L-1} \rightarrow \Delta_{K-1}$ with components $G_i, i = 1, \dots, K$, such that

$$P(\delta_{n+1} = e_i \mid \mathcal{F}_n, \mathbf{R}_{n+1}) = G_i(\mathbf{X}_n), \quad i = 1, \dots, K. \tag{2.2}$$

(A4) $\mathbf{R}_n^\top \delta_n + T_{n-1} \mathbf{X}_{n-1} \geq \mathbf{0}$.

(A5) There exists a matrix $\mathbf{H} \in \mathcal{R}_{K \times L}$ such that

$$\|\mathbf{H}_n - \mathbf{H}\| \rightarrow 0, \quad \text{a.s.},$$

where $\mathbf{H}_n := E[\mathbf{R}_n \mid \mathcal{F}_{n-1}]$.

Remark 2.1. Note that, under Assumption (A2), the process $\{T_n\}$ is deterministic, with $T_n = T_0 + ns$ for each $n \geq 1$. The entries of the replacement matrix can be negative and they can demand impossible removals. Assumption (A4) guarantees that the process does not get stuck.

The process $\{\mathbf{X}_n\}$ can be written within a Robbins–Monro scheme, following the lines of [7].

Theorem 2.1. Consider the generalized Pólya urn model $\{U_n\}$ under assumptions (A1)–(A5). The process $\{\mathbf{X}_n\}$ then satisfies

$$\mathbf{X}_{n+1} = \mathbf{X}_n + \frac{1}{T_{n+1}}(F(\mathbf{X}_n) + \boldsymbol{\varepsilon}_{n+1} + \boldsymbol{\beta}_{n+1}), \tag{2.3}$$

where

1. $F(\mathbf{x}) = (\mathbf{I} - \mathbf{x}\mathbf{1}^\top)\mathbf{H}^\top G(\mathbf{x})$,
2. $\{\boldsymbol{\varepsilon}_n\}$ is a sequence of martingale differences with respect to $\{\mathcal{F}_n\}$, with $\boldsymbol{\varepsilon}_n = \mathbf{R}_n^\top \delta_n - \mathbf{H}_n^\top G(\mathbf{X}_{n-1})$,
3. $\boldsymbol{\beta}_n = (\mathbf{H}_n - \mathbf{H})^\top G(\mathbf{X}_{n-1})$ and $\boldsymbol{\beta}_n \rightarrow \mathbf{0}$, a.s.

Proof. Observe that

$$\mathbf{X}_{n+1} = E[\mathbf{X}_{n+1} \mid \mathcal{F}_n] + \mathbf{X}_{n+1} - E[\mathbf{X}_{n+1} \mid \mathcal{F}_n].$$

From (2.1) and (2.2), we have

$$\begin{aligned} E[\mathbf{X}_{n+1} \mid \mathcal{F}_n] &= \mathbf{X}_n + \frac{1}{T_{n+1}}(\mathbf{I} - \mathbf{X}_n\mathbf{1}^\top)\mathbf{H}_{n+1}^\top G(\mathbf{X}_n) \\ &= \mathbf{X}_n + \frac{1}{T_{n+1}}[(\mathbf{I} - \mathbf{X}_n\mathbf{1}^\top)\mathbf{H}^\top G(\mathbf{X}_n) + (\mathbf{I} - \mathbf{X}_n\mathbf{1}^\top)(\mathbf{H}_{n+1}^\top - \mathbf{H}^\top)G(\mathbf{X}_n)] \end{aligned}$$

and

$$\mathbf{X}_{n+1} - E[\mathbf{X}_{n+1} \mid \mathcal{F}_n] = \frac{1}{T_{n+1}}(\mathbf{I} - \mathbf{X}_n\mathbf{1}^\top)(\mathbf{R}_{n+1}^\top \delta_{n+1} - \mathbf{H}_{n+1}^\top G(\mathbf{X}_n)).$$

Finally, we obtain (2.3) with

$$\begin{aligned} F(\mathbf{x}) &= (\mathbf{I} - \mathbf{x}\mathbf{1}^\top)\mathbf{H}^\top G(\mathbf{x}), \\ \boldsymbol{\varepsilon}_{n+1} &= (\mathbf{I} - \mathbf{X}_n\mathbf{1}^\top)(\mathbf{R}_{n+1}^\top \delta_{n+1} - \mathbf{H}_{n+1}^\top G(\mathbf{X}_n)) = \mathbf{R}_{n+1}^\top \delta_{n+1} - \mathbf{H}_{n+1}^\top G(\mathbf{X}_n), \\ \boldsymbol{\beta}_{n+1} &= (\mathbf{I} - \mathbf{X}_n\mathbf{1}^\top)(\mathbf{H}_{n+1}^\top - \mathbf{H}^\top)G(\mathbf{X}_n) = (\mathbf{H}_{n+1} - \mathbf{H})^\top G(\mathbf{X}_n). \end{aligned}$$

As $G(\mathbf{x}) \in \Delta_{K-1}$, we then see that $\boldsymbol{\beta}_{n+1}$ converges to $\mathbf{0}$ a.s.

Many criteria ensure the almost-sure convergence of $\{X_n\}$ towards a zero of $F(x)$ (see [17] and the references therein). We use [5, Theorem 2.III.4] to obtain sufficient conditions that are suitable for our applications. Assumption (A5) must be strengthened as follows.

(A5') There exist a matrix $H \in \mathcal{R}_{K \times L}$ and a constant $\rho > 0$ such that

$$\|H_n - H\| = O(n^{-\rho}) \quad \text{a.s.,}$$

where $H_n := E[R_n \mid \mathcal{F}_{n-1}]$.

In addition, $F(x)$ in (2.3) must satisfy the next assumption.

(A0) Let $F(x) = (I - x\mathbf{1}^\top)H^\top G(x)$ with $G \in C^1$, and assume that there exist a vector u in Δ_{L-1} and a constant $A > 0$ such that $F(u) = \mathbf{0}$ and $(x - u)^\top F(x) \leq -A\|x - u\|^2$ for $x \in \Delta_{L-1}$.

Theorem 2.2. Consider the generalized Pólya urn model $\{U_n\}$ under assumptions (A1)–(A4), (A5'), and (A0). We have

$$X_n \rightarrow u \quad \text{a.s.}$$

Proof. From Theorem 2.1 it follows that the process follows the Robbins–Monro scheme in (2.3), whose steps are $1/T_{n+1}$ with $T_n = T_0 + ns$. Therefore, $\sum_{n=0}^\infty 1/T_{n+1} = \infty$ and $\sum_{n=0}^\infty 1/T_{n+1}^2 < \infty$. From (A5') we have $\sum_{n=0}^\infty \|\beta_n\|^2/T_n < \infty$. Finally, from (A1) it follows that $E[\|\epsilon_n\|^2 \mid \mathcal{F}_{n-1}]$ is bounded, so we are in the framework of [5, Theorem 2.III.4(b)] with $V(x) = \|x - u\|^2$ and $\Gamma = \Delta_{L-1}$. The gradient of $V(x)$ is $2(x - u)^\top$, and $\|F(x)\|^2 \leq B(1 + V(x))$ in Γ , for a constant B . Therefore, the conditions of [5, Theorem 2.III.4(b)] hold and we see that $\|X_n - u\|^2$ and $\sum_{n=0}^\infty \|X_n - u\|^2/T_{n+1}$ converge a.s. Thus, $X_n \rightarrow u$ a.s.

Remark 2.2. If $H_n = H$ a.s. then $\beta_n = \mathbf{0}$ a.s. in (2.3) and we are in the framework of [5, Theorem 2.III.4(a)]. Condition (A0) can be weakened as follows.

(A0') Let $F(x) = (I - x\mathbf{1}^\top)H^\top G(x)$ with $G \in C^1$, and assume that there exists a vector u in Δ_{L-1} such that $F(u) = \mathbf{0}$ and $(x - u)^\top F(x) < 0$ for $x \in \Delta_{L-1} \setminus \{u\}$.

When $F(x)$ is linear, a different set of conditions for the almost-sure convergence of the process $\{X_n\}$ can be established using the ordinary differential equation method.

Theorem 2.3. Consider the generalized Pólya urn model $\{U_n\}$ under assumptions (A1)–(A5). Assume that $F(x)$ in (2.3) has a unique zero, $u \in \Delta_{L-1}$, and can be written as $F(x) = a + Ax$, $x \in \Delta_{L-1}$, where a is a constant vector and the matrix A has a simple eigenvalue associated with the left eigenvector $\mathbf{1}^\top$, its other eigenvalues having strictly negative real parts. Then

$$X_n \rightarrow u \quad \text{a.s.}$$

Proof. Since $F(u) = \mathbf{0}$, $a = -Au$ and $F(x) = A(x - u)$ for all $x \in \Delta_{L-1}$. Hence, (2.3) can be written as

$$X_{n+1} - u = X_n - u + \frac{1}{T_{n+1}}(A(X_n - u) + \epsilon_{n+1} + \beta_{n+1}). \tag{2.4}$$

Let m be the simple eigenvalue of A associated with the left eigenvector $\mathbf{1}^\top$. Let J be the real canonical Jordan representation of the matrix A such that $P^{-1}JP = A$, where

$$J = \begin{pmatrix} m & \mathbf{0}^\top \\ \mathbf{0} & \tilde{J} \end{pmatrix}.$$

From the assumptions it follows that that \tilde{J} is a stable matrix, that is, its eigenvalues all have negative real parts. The first row of P can be taken to be $\mathbf{1}^\top$. The other rows form an $(L - 1) \times L$ matrix \tilde{P} such that

$$P = \begin{pmatrix} \mathbf{1}^\top \\ \tilde{P} \end{pmatrix}.$$

Observe that, for each n , we have $\mathbf{1}^\top(X_n - \mathbf{u}) = 0$ and $\mathbf{1}^\top \varepsilon_n = 0$, whence $\mathbf{1}^\top \beta_n = 0$. Hence, from (2.4) we have the following recurrence equation for $\tilde{Z}_n := \tilde{P}(X_n - \mathbf{u})$, where $\tilde{\varepsilon}_n := \tilde{P}\varepsilon_n$ and $\tilde{\beta}_n := \tilde{P}\beta_n$:

$$\tilde{Z}_{n+1} = \tilde{Z}_n + \frac{1}{T_{n+1}}(\tilde{J}\tilde{Z}_n + \tilde{\varepsilon}_{n+1} + \tilde{\beta}_{n+1}).$$

As \tilde{J} is stable, $\mathbf{0}$ is a globally asymptotically stable solution to the system of differential equations $\dot{x} = \tilde{J}x$ (see, for instance, [10, Theorem 3.5]). We now invoke [11, Theorem 5.2.3], to conclude that $\tilde{Z}_n \rightarrow \mathbf{0}$ a.s. Since

$$P(X_n - \mathbf{u}) = \begin{pmatrix} 0 \\ \tilde{Z}_n \end{pmatrix},$$

we conclude that $X_n \rightarrow \mathbf{u}$ a.s.

Remark 2.3. The previous theorem is useful in the following setting, provided that assumptions (A1)–(A5) hold. If $K = L$ and $G(x) = x$ then $F(x) = Ax$, where $A = H^\top - sI$. Then, from assumption (A2), we have $\mathbf{1}^\top A = \mathbf{0}^\top$. If H is irreducible and all its off-diagonal entries are nonnegative, from [19, Theorem 2.6] it follows that A has a unique right eigenvector $\mathbf{u} \in \Delta_{L-1}$ associated to the eigenvalue 0, and that its other eigenvalues have strictly negative real parts.

3. A central limit theorem for $\{X_n\}$

Stochastic approximation theory will be the framework we use to obtain central limit theorems for the process $\{X_n\}$. Some extra hypotheses will be needed. First, assumption (A5) must be again strengthened, as follows.

(A5'') There exists a matrix $H \in \mathcal{R}_{K \times L}$ such that, as $n \rightarrow \infty$,

$$\|H_n - H\| \rightarrow 0 \quad \text{a.s.} \quad \text{and} \quad n E[\|H_n - H\|^2] \rightarrow 0.$$

We also make four further assumptions.

(A6) Let R_{jn} be the j th row of the replacement matrix R_n , with $j = 1, \dots, K$. There exist deterministic matrices A_j such that, for $j = 1, \dots, K$,

$$E[R_{jn}^\top R_{jn} \mid \mathcal{F}_{n-1}] \rightarrow A_j \quad \text{a.s.}$$

(A7) Let $F(x) = (I - x\mathbf{1}^\top)H^\top G(x)$ with $G \in C^2$, and assume that \mathbf{u} is the unique vector in Δ_{L-1} such that $F(\mathbf{u}) = \mathbf{0}$. Let m be a simple eigenvalue of the Jacobian matrix evaluated at \mathbf{u} , $JF(\mathbf{u})$, associated with the left eigenvector $\mathbf{1}^\top$. We assume that the other eigenvalues have real parts less than $-s/2$.

(A8) The following matrix has a simple zero eigenvalue:

$$C = \sum_{j=1}^K A_j G_j(\mathbf{u}) - \mathbf{H}^\top G(\mathbf{u}) G(\mathbf{u})^\top \mathbf{H}.$$

(A9) $X_n \rightarrow \mathbf{u}$ a.s.

Remark 3.1. Note that $F(\mathbf{x})$ must be defined in an open set containing Δ_{L-1} in order to obtain the Jacobian matrix $JF(\mathbf{x})$. As $X_n \in \Delta_{L-1}$, $F(\mathbf{x})$ can be extended in any suitable way outside Δ_{L-1} . Although the value of m depends on the extension of $F(\mathbf{x})$, the other eigenvalues (and, hence, the property about them in (A7)) do not depend on the chosen extension. Some examples of this procedure will be seen in the next section.

Remark 3.2. The almost-sure convergence $X_n \rightarrow \mathbf{u}$ can be established under, for instance, the hypotheses of Theorem 2.2 or Theorem 2.3.

Theorem 3.1. Consider the generalized Pólya urn model $\{U_n\}$ under assumptions (A1)–(A4), (A5''), and (A6)–(A9). We have

$$\sqrt{ns}(X_n - \mathbf{u}) \xrightarrow{D} N(\mathbf{0}, \mathbf{V}),$$

where \mathbf{V} is a singular symmetric matrix such that

$$\left(JF(\mathbf{u}) + \frac{s}{2} \mathbf{I} \right) \mathbf{V} + \mathbf{V} \left(JF(\mathbf{u}) + \frac{s}{2} \mathbf{I} \right)^\top = -C$$

and ‘ \xrightarrow{D} ’ denotes convergence in distribution.

Proof. As $G \in C^2$, the Taylor expansion of $F(\mathbf{x})$ can be written as

$$F(\mathbf{x}) = F(\mathbf{u}) + JF(\mathbf{u})(\mathbf{x} - \mathbf{u}) + \theta(\mathbf{x} - \mathbf{u}), \tag{3.1}$$

where $\|\theta(\mathbf{x} - \mathbf{u})\| = O(\|\mathbf{x} - \mathbf{u}\|^2)$. From Theorem 2.1 and (3.1) we have

$$\mathbf{Z}_{n+1} = \mathbf{Z}_n + \frac{1}{T_{n+1}} (JF(\mathbf{u})\mathbf{Z}_n + \boldsymbol{\varepsilon}_{n+1} + \boldsymbol{\beta}_{n+1} + \theta(\mathbf{Z}_n)), \tag{3.2}$$

where $\mathbf{Z}_n := X_n - \mathbf{u}$. Observe that, for each n , we have

$$\mathbf{1}^\top JF(\mathbf{u}) = m\mathbf{1}^\top, \quad \mathbf{1}^\top \mathbf{Z}_n = 0, \quad \mathbf{1}^\top \boldsymbol{\varepsilon}_n = 0, \quad \mathbf{1}^\top \boldsymbol{\beta}_n = 0, \tag{3.3}$$

and, thus, $\mathbf{1}^\top \theta(\mathbf{Z}_n) = 0$.

Let \mathbf{J} be the real canonical Jordan representation of the matrix $JF(\mathbf{u})$ such that $\mathbf{P}^{-1} \mathbf{J} \mathbf{P} = JF(\mathbf{u})$, where

$$\mathbf{J} = \begin{pmatrix} m & \mathbf{0}^\top \\ \mathbf{0} & \tilde{\mathbf{J}} \end{pmatrix}.$$

According to (A7), $\tilde{\mathbf{J}}$ contains the Jordan blocks of the eigenvalues of $JF(\mathbf{u})$ that have negative real parts, so $\tilde{\mathbf{J}}$ is a stable matrix. The first row of \mathbf{P} can be taken to be $\mathbf{1}^\top$. The other rows form a matrix $\tilde{\mathbf{P}}$ such that

$$\mathbf{P} = \begin{pmatrix} \mathbf{1}^\top \\ \tilde{\mathbf{P}} \end{pmatrix}.$$

From (3.2) and (3.3) we obtain the following recurrence equation for $\{\tilde{\mathbf{Z}}_n\} := \{\tilde{\mathbf{P}}\mathbf{Z}_n\}$:

$$\tilde{\mathbf{Z}}_{n+1} = \tilde{\mathbf{Z}}_n + \frac{1}{nS}(\tilde{\mathbf{J}}\tilde{\mathbf{Z}}_n + \tilde{\boldsymbol{\varepsilon}}_{n+1} + \tilde{\mathbf{r}}_{n+1} + \tilde{\boldsymbol{\theta}}(\mathbf{Z}_n)). \tag{3.4}$$

Here, with $a_n := (n - 1)s/T_n$ for each n , we have

$$\tilde{\boldsymbol{\varepsilon}}_{n+1} = a_{n+1}\tilde{\mathbf{P}}\boldsymbol{\varepsilon}_{n+1}, \quad \tilde{\mathbf{r}}_{n+1} = a_{n+1}\tilde{\mathbf{P}}\boldsymbol{\beta}_{n+1} + (a_{n+1} - 1)\tilde{\mathbf{J}}\tilde{\mathbf{Z}}_n, \quad \tilde{\boldsymbol{\theta}}(\mathbf{Z}_n) = a_{n+1}\tilde{\mathbf{P}}\boldsymbol{\theta}(\mathbf{Z}_n).$$

In order to apply [17, Theorem 1] to the recurrence scheme (3.4), we must check that its conditions hold.

As $\{a_n\}$ is a bounded deterministic sequence, we have $\|\tilde{\boldsymbol{\theta}}(\mathbf{Z}_n)\| = O(\|\mathbf{Z}_n\|^2)$. From Theorem 2.1, $\{\boldsymbol{\varepsilon}_n\}$ is a sequence of martingale differences, and from (A1) it follows that, for some $b > 2$,

$$E[\|\tilde{\boldsymbol{\varepsilon}}_{n+1}\|^b \mid \mathcal{F}_n] < \infty.$$

On the other hand, from (A5'') it follows that $n E[\|\tilde{\mathbf{r}}_n\|^2] \rightarrow 0$.

Let $\mathbf{C}_{n+1} := E[\boldsymbol{\varepsilon}_{n+1}\boldsymbol{\varepsilon}_{n+1}^\top \mid \mathcal{F}_n]$. From Theorem 2.1 we have

$$\boldsymbol{\varepsilon}_{n+1} = \mathbf{R}_{n+1}^\top \boldsymbol{\delta}_{n+1} - \mathbf{H}_{n+1}^\top G(\mathbf{X}_n)$$

and, therefore,

$$\begin{aligned} \mathbf{C}_{n+1} &= E[\mathbf{R}_{n+1}^\top \boldsymbol{\delta}_{n+1} \boldsymbol{\delta}_{n+1}^\top \mathbf{R}_{n+1} \mid \mathcal{F}_n] - \mathbf{H}_{n+1}^\top G(\mathbf{X}_n) G(\mathbf{X}_n)^\top \mathbf{H}_{n+1} \\ &= E[\mathbf{R}_{n+1}^\top \text{diag}(G(\mathbf{X}_n)) \mathbf{R}_{n+1} \mid \mathcal{F}_n] - \mathbf{H}_{n+1}^\top G(\mathbf{X}_n) G(\mathbf{X}_n)^\top \mathbf{H}_{n+1} \\ &= \sum_{j=1}^K E[\mathbf{R}_{j,n+1}^\top \mathbf{R}_{j,n+1} \mid \mathcal{F}_n] G_j(\mathbf{X}_n) - \mathbf{H}_{n+1}^\top G(\mathbf{X}_n) G(\mathbf{X}_n)^\top \mathbf{H}_{n+1}, \end{aligned}$$

where $\text{diag}(\mathbf{x})$ denotes the diagonal matrix whose components form the vector \mathbf{x} . From (A5''), (A6), and (A9) we have

$$\mathbf{C}_{n+1} \rightarrow \mathbf{C} = \sum_{j=1}^K A_j G_j(\mathbf{u}) - \mathbf{H}^\top G(\mathbf{u}) G(\mathbf{u})^\top \mathbf{H} \quad \text{a.s.}$$

Now observe that

$$\begin{aligned} \mathbf{P}\mathbf{C}_n\mathbf{P}^\top &= \mathbf{P} E[\boldsymbol{\varepsilon}_n \boldsymbol{\varepsilon}_n^\top \mid \mathcal{F}_{n-1}] \mathbf{P}^\top \\ &= E \left[\begin{pmatrix} 0 \\ \tilde{\mathbf{P}}\boldsymbol{\varepsilon}_n \end{pmatrix} (0 \quad \boldsymbol{\varepsilon}_n^\top \tilde{\mathbf{P}}^\top) \mid \mathcal{F}_{n-1} \right] \\ &= \begin{pmatrix} 0 & \mathbf{0}^\top \\ \mathbf{0} & \tilde{\mathbf{C}}_n \end{pmatrix} \\ &\rightarrow \begin{pmatrix} 0 & \mathbf{0}^\top \\ \mathbf{0} & \tilde{\mathbf{C}} \end{pmatrix} \\ &\equiv \mathbf{P}\mathbf{C}\mathbf{P}^\top \quad \text{a.s.} \end{aligned}$$

As \mathbf{P} is a real, nonsingular matrix, the matrices \mathbf{C} and $\mathbf{P}\mathbf{C}\mathbf{P}^\top$ are congruent and thus have the same number of positive eigenvalues (counting multiplicities). From (A8), \mathbf{C} has 0 as a simple eigenvalue, so $\tilde{\mathbf{C}}$ is positive definite.

Finally, $E[\tilde{\boldsymbol{\epsilon}}_n \tilde{\boldsymbol{\epsilon}}_n^\top \mid \mathcal{F}_{n-1}] = a_{n+1}^2 \tilde{\mathbf{C}}_n$ converges almost surely to $\tilde{\mathbf{C}}$. Therefore, from [17, Theorem 1] it follows that

$$\sqrt{ns} \tilde{\mathbf{Z}}_n \xrightarrow{D} N(\mathbf{0}, \tilde{\mathbf{V}}),$$

where

$$\left(\tilde{\mathbf{J}} + \frac{s}{2} \mathbf{I}\right) \tilde{\mathbf{V}} + \tilde{\mathbf{V}} \left(\tilde{\mathbf{J}} + \frac{s}{2} \mathbf{I}\right)^\top = -\tilde{\mathbf{C}}. \tag{3.5}$$

Moreover, we observe that

$$\sqrt{ns}(X_n - \mathbf{u}) = \mathbf{P}^{-1} \sqrt{ns} \begin{pmatrix} 0 \\ \tilde{\mathbf{Z}}_n \end{pmatrix} \xrightarrow{D} N\left(\mathbf{0}, \mathbf{P}^{-1} \begin{pmatrix} 0 & \mathbf{0}^\top \\ \mathbf{0} & \tilde{\mathbf{V}} \end{pmatrix} (\mathbf{P}^{-1})^\top\right).$$

Let

$$\mathbf{V} := \mathbf{P}^{-1} \begin{pmatrix} 0 & \mathbf{0}^\top \\ \mathbf{0} & \tilde{\mathbf{V}} \end{pmatrix} (\mathbf{P}^{-1})^\top.$$

As

$$\left(\mathbf{J} + \frac{s}{2} \mathbf{I}\right) \begin{pmatrix} 0 & \mathbf{0}^\top \\ \mathbf{0} & \tilde{\mathbf{V}} \end{pmatrix} + \begin{pmatrix} 0 & \mathbf{0}^\top \\ \mathbf{0} & \tilde{\mathbf{V}} \end{pmatrix} \left(\mathbf{J} + \frac{s}{2} \mathbf{I}\right)^\top = -\begin{pmatrix} 0 & \mathbf{0}^\top \\ \mathbf{0} & \tilde{\mathbf{C}} \end{pmatrix}$$

and $\mathbf{J} = \mathbf{P} \mathbf{J} \mathbf{F}(\mathbf{u}) \mathbf{P}^{-1}$, we have

$$\mathbf{P} \left(\mathbf{J} \mathbf{F}(\mathbf{u}) + \frac{s}{2} \mathbf{I}\right) \mathbf{P}^{-1} \begin{pmatrix} 0 & \mathbf{0}^\top \\ \mathbf{0} & \tilde{\mathbf{V}} \end{pmatrix} + \begin{pmatrix} 0 & \mathbf{0}^\top \\ \mathbf{0} & \tilde{\mathbf{V}} \end{pmatrix} (\mathbf{P}^{-1})^\top \left(\mathbf{J} \mathbf{F}(\mathbf{u}) + \frac{s}{2} \mathbf{I}\right)^\top \mathbf{P}^\top = -\begin{pmatrix} 0 & \mathbf{0}^\top \\ \mathbf{0} & \tilde{\mathbf{C}} \end{pmatrix}.$$

Therefore, \mathbf{V} is the solution to the matrix equation

$$\left(\mathbf{J} \mathbf{F}(\mathbf{u}) + \frac{s}{2} \mathbf{I}\right) \mathbf{V} + \mathbf{V} \left(\mathbf{J} \mathbf{F}(\mathbf{u}) + \frac{s}{2} \mathbf{I}\right)^\top = -\mathbf{C}.$$

Remark 3.3. As $\tilde{\mathbf{C}}$ is positive definite and $\tilde{\mathbf{J}} + (s/2)\mathbf{I}$ is a stable matrix, the unique solution to (3.5) is

$$\tilde{\mathbf{V}} = \int_0^\infty \exp\left(\left(\tilde{\mathbf{J}} + \frac{s}{2} \mathbf{I}\right)x\right) \tilde{\mathbf{C}} \exp\left(\left(\tilde{\mathbf{J}} + \frac{s}{2} \mathbf{I}\right)^\top x\right) dx. \tag{3.6}$$

See, for instance, [10, Theorem 3.6].

Before stating a corollary to Theorem 3.1, we introduce one more assumption.

(A8') \mathbf{H} has maximum rank L , $G(\mathbf{u})$ is strictly positive, and \mathbf{R}_n is \mathcal{F}_{n-1} -measurable.

Corollary 3.1. Consider the generalized Pólya urn model $\{U_n\}$ under assumptions (A1)–(A4), (A5''), (A7), (A8'), and (A9). We have

$$\sqrt{ns}(X_n - \mathbf{u}) \xrightarrow{D} N(\mathbf{0}, \mathbf{V}),$$

where \mathbf{V} is a singular symmetric matrix such that

$$\left(\mathbf{J} \mathbf{F}(\mathbf{u}) + \frac{s}{2} \mathbf{I}\right) \mathbf{V} + \mathbf{V} \left(\mathbf{J} \mathbf{F}(\mathbf{u}) + \frac{s}{2} \mathbf{I}\right)^\top = -\mathbf{C}$$

with

$$\mathbf{C} = \mathbf{H}^\top (\text{diag}(G(\mathbf{u})) - G(\mathbf{u})G(\mathbf{u})^\top) \mathbf{H}.$$

Proof. For each n , $\mathbf{H}_n = \mathbf{R}_n$. Then, as in Theorem 3.1,

$$\mathbf{C}_{n+1} = \mathbf{H}_{n+1}^\top (\text{diag}(G(\mathbf{X}_n)) - G(\mathbf{X}_n)G(\mathbf{X}_n)^\top) \mathbf{H}_{n+1}.$$

Therefore, from (A5'') and (A9) we have $\mathbf{C}_n \rightarrow \mathbf{C}$ a.s., where

$$\mathbf{C} = \mathbf{H}^\top (\text{diag}(G(\mathbf{u})) - G(\mathbf{u})G(\mathbf{u})^\top) \mathbf{H}.$$

Let $\mathbf{\Sigma} := \text{diag}(G(\mathbf{u})) - G(\mathbf{u})G(\mathbf{u})^\top$. As $G(\mathbf{u})$ is strictly positive, it follows that $-\mathbf{\Sigma}$ is a symmetric matrix in which all the off-diagonal entries are positive. As all the rows add to the same value, 0, it follows from [19, Theorem 2.6] that $\mathbf{\Sigma}$ has a simple zero eigenvalue and that the other eigenvalues are positive. Then, from (A8'), it follows that $\mathbf{C} = \mathbf{H}^\top \mathbf{\Sigma} \mathbf{H}$ is positive semidefinite with 0 as a simple eigenvalue. Thus, (A8) holds in Theorem 3.1, and the result follows.

Remark 3.4. When $K = L$ and $G(\mathbf{x}) = \mathbf{x}$, we deal with the particular case of a linear function $F(\mathbf{x})$ in the Robbins–Monro scheme. In the framework of Remark 2.3, assumption (A9) is no longer necessary in order to apply Theorem 3.1. Moreover, assumption (A7) is equivalent to the following: \mathbf{u} is the unique left eigenvector of \mathbf{H} in Δ_{L-1} associated with the simple eigenvalue s , and the other eigenvalues of \mathbf{H} have real parts less than $-s/2$.

Remark 3.5. If, in (A7), m is not simple, but there exists a real, nonsingular matrix \mathbf{P} with $\mathbf{1}^\top$ as its first row and such that $\mathbf{P}^{-1} \mathbf{J} \mathbf{P} = \mathbf{J}F(\mathbf{u})$ with

$$\mathbf{J} = \begin{pmatrix} m & \mathbf{0}^\top \\ \mathbf{0} & \tilde{\mathbf{J}} \end{pmatrix},$$

then Theorem 3.1 and Corollary 3.1 still hold, provided that $m < -s/2$.

Remark 3.6. It is worth noting that the proof of Theorem 3.1 provides a simple procedure to calculate the covariance matrix of the asymptotic distribution, provided that all the conditions hold. With the notation used in the proof of Theorem 3.1, the steps are as follows.

1. Obtain the solution, \mathbf{u} , to $F(\mathbf{x}) = \mathbf{0}$, where $\mathbf{x} \in \Delta_{L-1}$.
2. Obtain the eigenvalues of $\mathbf{J}F(\mathbf{u})$ and the real matrix \mathbf{P} such that $\mathbf{P}^{-1} \mathbf{J} \mathbf{P} = \mathbf{J}F(\mathbf{u})$.
3. Consider the matrix $\tilde{\mathbf{P}}$ and obtain $\tilde{\mathbf{V}}$ by solving (3.6).
4. Obtain \mathbf{V} .

4. Applications

4.1. Pólya urn with double extraction

We consider a Pólya urn with balls of two types. The replacement policy consists in drawing and replacing two balls sequentially. Three possible actions can be applied: if we obtain two balls of type 1 then the first row of the replacement matrix is applied; if we obtain one ball of type 1 then the second row is applied; and if we obtain no balls of type 1 then the third row is applied. The replacement matrix,

$$\mathbf{R} = \begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix}, \quad a + b = c + d = e + f = s, \quad eb > 0,$$

is fixed, nonnegative, and has rank two. It follows from the replacement policy that $G(X_n) = (X_{1n}^2, 2X_{1n}X_{2n}, X_{2n}^2)$ and $H_n = \mathbf{R}$, for all n . This urn is of practical use in the study of the asymptotic behavior of a kind of data structure in computation known as a random forest. In [22] a particular case was studied and the open problem of how to deal with Pólya urns with double extraction was presented.

In order to apply Theorem 2.2 we observe that, for any $\mathbf{x} = (x_1, x_2) \in \Delta_1$, we have $F(\mathbf{x}) = (f_1(x_1), -f_1(x_1))$, where $f_1(x_1) = \alpha x_1^2 + \beta x_1 + e$ and $\alpha := a + e - 2c$ and $\beta := 2c - 2e - s$. In order to obtain the vector $\mathbf{u} = (u_1, u_2) \in \Delta_1$ such that $F(\mathbf{u}) = \mathbf{0}$, we follow [7, Example 2.1] and, under the assumption that $eb > 0$, find the following scenarios.

- (S1) If $\alpha = 0$ then the only root of $f_1(x_1)$ in $(0, 1)$ is $u_1 = e/(b + e)$.
- (S2) If $\alpha \neq 0$ then the only root of $f_1(x_1)$ in $(0, 1)$ is $u_1 = (-\beta - \sqrt{\beta^2 - 4\alpha e})/(2\alpha)$.

Now observe that $(\mathbf{x} - \mathbf{u})^\top F(\mathbf{x}) = 2(x_1 - u_1)f_1(x_1)$ for any $\mathbf{x} \in \Delta_1$ and $\mathbf{u} \in \Delta_1$ with $F(\mathbf{u}) = \mathbf{0}$. So, in scenario (S1) this expression is equal to $-2(b + e)(x_1 - u_1)^2$, and (A0) is obviously satisfied. In scenario (S2), $2(x_1 - u_1)f_1(x_1) = 2\alpha(x_1 - u_1)^2(x_1 - r)$, where $r = (-\beta + \sqrt{\beta^2 - 4\alpha e})/(2\alpha) \notin [0, 1]$. Observe that if $\alpha > 0$ then $\beta < 0$ and, so, $r > 1$. If $\alpha < 0$ then $r < 0$. Therefore, there exists a constant $A > 0$ such that $\alpha(x_1 - r) \leq -A$ for all $x_1 \in [0, 1]$ and (A0) thus also holds. As the conditions in Theorem 2.2 hold, it follows that $X_n \rightarrow \mathbf{u}$ a.s.

In order to apply Corollary 3.1, let D be an open set that contains Δ_1 and consider, for all $\mathbf{x} \in D$,

$$G(\mathbf{x}) = (x_1^2, 2x_1x_2, x_2^2)^\top \quad \text{and} \quad F(\mathbf{x}) = (\mathbf{I} - \mathbf{x}\mathbf{1}^\top)\mathbf{R}^\top G(\mathbf{x}).$$

We then have

$$JF(\mathbf{x}) = (\mathbf{R}^\top - s\mathbf{x}\mathbf{1}^\top)JG(\mathbf{x}) - s\mathbf{I} \quad \text{for all } \mathbf{x} \in \Delta_1. \tag{4.1}$$

Observe that, for any $\mathbf{x} \in D$, we have $JG(\mathbf{x})\mathbf{x} = 2G(\mathbf{x})$. From (4.1), it thus follows that

$$JF(\mathbf{u})\mathbf{u} = -s\mathbf{u}, \tag{4.2}$$

where $\mathbf{u} \in \Delta_1$ and $F(\mathbf{u}) = \mathbf{0}$. Expression (4.2) will be useful in simplifying some calculations in this example. We compute $JF(\mathbf{u})$, and obtain

$$JF(\mathbf{u}) = \begin{pmatrix} M - s & N \\ -M & -N - s \end{pmatrix}, \quad M = 2(au_1 + cu_2 - su_1), \quad N = 2(cu_1 + eu_2 - su_1).$$

When $M \neq N$, the eigenvalues are $-s$, associated with the left eigenvector $\mathbf{1}$ (up to constants), and $M - N - s$, associated with the left eigenvector $(M, N)^\top$ (up to constants). Now, in scenarios (S1) and (S2), if $M - N < s/2$ then the conditions of Corollary 3.1 hold and we obtain a central limit theorem for X_n .

We can follow the steps of Remark 3.6 in order to obtain an explicit expression for the covariance matrix of the normal distribution. First we compute \mathbf{P} and \mathbf{P}^{-1} :

$$\mathbf{P} = \begin{pmatrix} 1 & 1 \\ M & N \end{pmatrix}, \quad \mathbf{P}^{-1} = \frac{1}{M - N} \begin{pmatrix} -N & 1 \\ M & -1 \end{pmatrix}.$$

From (4.2) we have $\tilde{\mathbf{P}}\mathbf{u} = \mathbf{0}$, where $\tilde{\mathbf{P}} = (M, N)$. Then, as $F(\mathbf{u}) = \mathbf{0}$, we obtain the relation $\tilde{\mathbf{P}}\mathbf{R}^\top G(\mathbf{u}) = \mathbf{0}$. As we are working under the assumptions of Corollary 3.1, \tilde{C} is a positive

real number that is obtained as follows:

$$\begin{aligned} \tilde{C} &= \tilde{\mathbf{P}}\mathbf{R}^\top(\text{diag}(G(\mathbf{u})) - G(\mathbf{u})G(\mathbf{u})^\top)\mathbf{R}\tilde{\mathbf{P}}^\top \\ &= \tilde{\mathbf{P}}\mathbf{R}^\top\text{diag}(G(\mathbf{u}))\mathbf{R}\tilde{\mathbf{P}}^\top \\ &= (aM + bN)^2u_1^2 + 2(cM + dN)^2u_1u_2 + (eM + fN)^2u_2^2. \end{aligned}$$

From (3.5) we have

$$\tilde{\mathbf{V}} = \frac{-\tilde{C}}{2(M - N - s/2)}.$$

Finally we have $\sqrt{ns}\tilde{\mathbf{P}}(\mathbf{X}_n - \mathbf{u}) \xrightarrow{D} N(0, \tilde{\mathbf{V}})$ and, thus, $\sqrt{ns}(\mathbf{X}_n - \mathbf{u}) \xrightarrow{D} N(\mathbf{0}, \mathbf{V})$ with

$$\mathbf{V} = \mathbf{P}^{-1} \begin{pmatrix} 0 & 0 \\ 0 & \tilde{\mathbf{V}} \end{pmatrix} (\mathbf{P}^{-1})^\top = \frac{\tilde{\mathbf{V}}}{(N - M)^2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

In particular, in scenario (S1) $M - N = a - e$ and

$$\frac{\tilde{\mathbf{V}}}{(M - N)^2} = \frac{(a - e)^2eb}{(b + e)^2(4(b + e) - 2s)}.$$

When $M = N$, from (4.2) we have $M = N = 0$. In scenario (S1) this can only happen when $a = c = e$, and $\{\mathbf{X}_n\}$ is thus a deterministic process. In scenario (S2) we are in the situation described in Remark 3.5, because $-s$ is not a simple eigenvalue of $JF(\mathbf{u})$ but the matrix \mathbf{R} has rank two, $u_1 = c/(b + c) > 0$, and the other conditions of Corollary 3.1 hold. Furthermore, $a \neq c$. If we choose for \mathbf{P} the regular matrix

$$\mathbf{P} = \begin{pmatrix} 1 & 1 \\ -b & c \end{pmatrix},$$

then $\tilde{\mathbf{P}}\mathbf{u} = 0$ and we obtain

$$\mathbf{V} = \frac{c^2(c - a)^2}{(b + c)^2s} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

4.2. Random trees

A random tree is a data structure for storage organization in computers. The knowledge of the asymptotic behavior of a random tree is therefore useful in estimating its memory requirements and the complexity of algorithms that manipulate data, such as searching and sorting algorithms.

In [21] and [12] generalized Pólya urn models were revealed to be indispensable tools in the study of the asymptotic properties of random trees. The generalized Pólya urn model used considers a fixed replacement matrix \mathbf{A} , and the colors of the balls represent different kinds of nodes. It satisfies conditions (A1)–(A5) and the process $\{\mathbf{X}_n\}$ evolves according to the Robbins–Monro recurrence scheme

$$\mathbf{X}_{n+1} = \mathbf{X}_n + \frac{1}{ns + 1}[(\mathbf{A}^\top - s\mathbf{I})\mathbf{X}_n + \mathbf{A}^\top(\delta_{n+1} - \mathbf{X}_n)];$$

we are thus in the framework of Remark 2.2, and Theorem 2.3 can be applied. If conditions (A7) and (A8') hold for \mathbf{A} , then Corollary 3.1 can also be applied. The covariance matrix of the asymptotic distribution is then the solution, \mathbf{V} , to the Lyapunov equation

$$\left(\mathbf{A}^\top - \frac{s}{2}\mathbf{I}\right)\mathbf{V} + \mathbf{V}\left(\mathbf{A} - \frac{s}{2}\mathbf{I}\right) = -\mathbf{C},$$

where $\mathbf{C} = \mathbf{A}^\top(\text{diag}(\mathbf{u}) - \mathbf{u}\mathbf{u}^\top)\mathbf{A}$.

These models are of the standard linear type studied in [2, Chapter V], [9], and [21], and the central limit theorem can be obtained using known theorems.

In [12, Section 8.3] it was indicated that recursive bucket trees present several peculiarities that mean special methods are required to obtain the covariance matrix of the limit distribution explicitly, and its computation is considered rather lengthy. The case with node capacity two was studied in [13]. We consider the case with node capacity three, where the replacement matrix is

$$A = \begin{pmatrix} -1 & 2 & 0 \\ 0 & -2 & 3 \\ 1 & 0 & 0 \end{pmatrix}.$$

From Theorem 2.3 we have $X_n \rightarrow \mathbf{u} = (\frac{3}{11}, \frac{2}{11}, \frac{6}{11})^\top$ a.s., the left eigenvector associated with the eigenvalue $s = 1$ of the matrix A . From Corollary 3.1 we have the following central limit theorem:

$$\sqrt{ns} \tilde{P}(X_n - \mathbf{u}) \xrightarrow{D} N \left(\mathbf{0}, \begin{pmatrix} \frac{936}{605} & \sqrt{\frac{2592}{14641}} \\ \sqrt{\frac{2592}{14641}} & \frac{516}{605} \end{pmatrix} \right).$$

Thus,

$$\sqrt{n}(X_n - \mathbf{u}) \xrightarrow{D} N \left(\mathbf{0}, \begin{pmatrix} \frac{114}{605} & -\frac{12}{121} & -\frac{54}{605} \\ -\frac{12}{121} & \frac{168}{605} & -\frac{108}{605} \\ -\frac{54}{605} & -\frac{108}{605} & \frac{162}{605} \end{pmatrix} \right).$$

Here the covariance matrices are obtained by following Remark 3.6.

4.3. Adaptive designs in clinical trials

Randomized urn models are used to perform adaptive design (see, for instance, [18] and the references therein). The play-the-winner rule is a pioneering adaptive design in clinical trials (see [23]) and the asymptotic behavior of the associated urn model has been widely studied (see, for instance, [14] or [9, Example 7.3]). Several extensions of the classical play-the-winner rule have been studied in the statistical literature. For instance, in [15] an adaptive design with prognostic factors, more than two treatments, and urn functions was studied. In [15, Remark 4.1] a central limit theorem was obtained for the urn model associated with an extension of the play-the-winner rule that includes prognostic factors in the design and an urn function. In all these cases the replacement matrices associated with the urn models have the property that all the rows add to the same value, s , so we are in the framework of this paper.

In [16] an adaptive design was considered and modeled as a generalized Pólya urn such that, for each n , the replacement matrix considered, $R_n = (r_{nij})$, $i, j = 1, \dots, L$, has entries

$$r_{nij} = \begin{cases} Z_{ni} & \text{if } i = j, \\ \hat{\mu}_{n-1,ij}(1 - Z_{ni}) & \text{if } i \neq j, \end{cases}$$

where, for each $i = 1, \dots, L$, $\{Z_{ni}\}$ is a sequence of continuous, positive, independent, and identically distributed random variables bounded by 1, with mean μ_i , $0 < \mu_i < 1$, and variance $\sigma_i^2 > 0$, and $\hat{\mu}_{n-1,ij}$ is an estimator of a function of μ_i that uses the information available up until the replacement of the $(n - 1)$ th patient.

An urn function must be considered in order to satisfy (A3). For instance, let $G(\mathbf{x}) = (G_1(\mathbf{x}), \dots, G_L(\mathbf{x}))$ with $G_i(\mathbf{X}_n) = \alpha + (1 - L\alpha)X_{ni}$ for $i = 1, \dots, L$ and $0 < \alpha < 1/L$. Then $G(\mathbf{x}) = \mathbf{B}\mathbf{x}$ with $\mathbf{B} = \alpha\mathbf{1}\mathbf{1}^\top + (1 - L\alpha)\mathbf{I}$. This allocation rule guarantees that the

probability of allocation of each treatment is not less than the fixed quantity α . We thus compromise between the play-the-winner philosophy, which tends to assign more patients to better treatments, and the classical philosophy, which balances the patients between treatments.

When, for each n , $R_n \mathbf{1} = \mathbf{1}$ and the statistic $\hat{\mu}_{n-1,ij}$ converges a.s. to a constant α_{ij} for $i, j = 1, \dots, L$, we are in the framework of Theorem 2.3 with $F(\mathbf{x}) = (\mathbf{H}^\top \mathbf{B} - \mathbf{I})\mathbf{x}$. Therefore, (A9) holds. Moreover, if $n E[(\hat{\mu}_{n-1,ij} - \alpha_{ij})^2] \rightarrow 0$ for $i, j = 1, \dots, L$, then (A5'') holds. That (A6) holds is immediate. Finally, if (A7) and (A8) hold, then Theorem 3.1 can be applied and a central limit theorem is obtained for the associated urn model.

Acknowledgement

The authors wish to acknowledge the helpful suggestions made by an anonymous referee that greatly improved the contents and the presentation of an initial version of this paper.

References

- [1] ARTUR, B., ERMOL'EV, YU. M. AND KANIOVSKI, YU. M. (1983). A generalized urn problem and its applications. *Cybernetics* **19**, 61–71.
- [2] ATHREYA, K. B. AND NEY, P. E. (1972). *Branching Processes*. Springer, Berlin.
- [3] BAI, Z. D. AND HU, F. (1999). Asymptotic theorems for urn models with nonhomogeneous generating matrices. *Stoch. Process. Appl.* **80**, 87–101.
- [4] BAI, Z. D. AND HU, F. (2005). Asymptotics in randomized urn models. *Ann. Appl. Prob.* **15**, 914–940.
- [5] DUFLO, M. (1996). *Algorithmes Stochastiques*. Springer, Paris.
- [6] GOUET, R. (1993). Martingale functional central limit theorems for a generalized Pólya urn. *Ann. Prob.* **21**, 1624–1639.
- [7] HIGUERAS, I., MOLÉR, J., PLO, F. AND SAN MIGUEL, M. (2003). Urn models and differential algebraic equations. *J. Appl. Prob.* **40**, 401–412.
- [8] HILL, B. M., LANE, D. AND SUDDERTH, W. (1980). A strong law for some generalized urn processes. *Ann. Prob.* **8**, 214–226.
- [9] JANSON, S. (2004). Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stoch. Process. Appl.* **110**, 177–245.
- [10] KHALIL, H. K. (1996). *Nonlinear Systems*, 2nd edn. Prentice-Hall, Upper Saddle River, NJ.
- [11] KUSHNER, H. J. AND YIN, G. G. (1997). *Stochastic Approximation Algorithms and Applications* (Appl. Math. **35**). Springer, Berlin.
- [12] MAHMOUD, H. M. (2002). The size of random bucket trees via urn models. *Acta Informatica* **38**, 813–838.
- [13] MAHMOUD, H. M. AND SMYTHE, R. T. (1995). Probabilistic analysis of bucket recursive trees. *Theoret. Comput. Sci.* **144**, 221–249.
- [14] MATTHEWS, P. C. AND ROSENBERGER, W. F. (1997). Variance in randomized play-the-winner clinical trials. *Statist. Prob. Lett.* **35**, 233–240.
- [15] MOLÉR, J. A., PLO, F. AND SAN MIGUEL, M. (2005). Adaptive designs and Robbins–Monro algorithm. *J. Statist. Planning Infer.* **131**, 161–174.
- [16] MOLÉR, J. A., PLO, F. AND SAN MIGUEL, M. (2006). An adaptive design for clinical trials with non-dichotomous response and prognostic factors. *Statist. Prob. Lett.* **76**, 1940–1946.
- [17] PELLETIER, M. (1998). Weak convergence rates for stochastic approximation with application to multiple targets and simulated annealing. *Ann. Appl. Prob.* **8**, 10–44.
- [18] ROSENBERGER, W. F. (2002). Randomized urn models and sequential designs. *Sequent. Anal.* **21**, 1–28.
- [19] SENETA, E. (1981). *Non-Negative Matrices and Markov Chains*, 2nd edn. Springer, New York.
- [20] SMYTHE, R. T. (1996). Central limit theorems for urn models. *Stoch. Process. Appl.* **65**, 115–137.
- [21] SMYTHE, R. T. AND MAHMOUD, H. M. (1995). A survey of recursive trees. *Theory Prob. Math. Statist.* **51**, 1–27.
- [22] TSUKIJI, T. AND MAHMOUD, H. M. (2001). A limit law for outputs in random recursive circuits. *Algorithmica* **31**, 403–412.
- [23] WEL, L. J. AND DURHAM, S. (1978). The randomized play-the-winner rule in medical trials. *J. Amer. Statist. Soc.* **73**, 840–843.