

STRONG CONVERGENCE FOR URN MODELS WITH REDUCIBLE REPLACEMENT POLICY

R. ABRAHAM,* *Université d'Orléans*

J. S. DHERSIN,** *Université René Descartes*

B. YCART,***, *Université Joseph Fourier*

Abstract

A multitype urn scheme with random replacements is considered. Each time a ball is picked, another ball is added, and its type is chosen according to the transition probabilities of a reducible Markov chain. The vector of frequencies is shown to converge almost surely to a random element of the set of stationary measures of the Markov chain. Its probability distribution is characterized as the solution to a fixed point problem. It is proved to be Dirichlet in the particular case of a single transient state to which no return is possible. This is no longer the case, however, as soon as returns to transient states are allowed.

Keywords: Urn model; Markov chain; strong convergence

2000 Mathematics Subject Classification: Primary 60F15

1. Introduction

In the vast literature devoted to urn models (see [12] as a general reference), a good number of recent papers have been devoted to random replacement policies. Each time a ball is drawn, the types of ball which are added or removed are random variables, whose distribution depends on the type of the ball that has been picked; see, for instance, [1], [2], [3], [10], and [11]. Strong convergence results [9], [10], as well as functional central limit theorems [2], [8], [11], are now available for a vast range of models. However, in all these references, some irreducibility hypothesis is made to ensure that there is only one possible limit for the frequency vector. Our aim here is to answer the natural question: what happens when there is more than one possible limit for the frequency vector?

We study the following simplest possible model: balls are added one by one, where the type of ball added depends only on the type of ball that has been drawn. We believe that our results can be extended to more general schemes, such as those of [11] or [3]. The ball types are numbered from 1 to d . If a ball of type i has been drawn, then a ball of type j is added with probability $p_{i,j}$. The matrix $P = (p_{i,j})$ is a (reducible) stochastic matrix on $\{1, \dots, d\}$. As expected, the distribution of types converges almost surely to a stationary distribution for the matrix P (see Theorem 2.1, below). The proof is based on the classical stochastic algorithm technique [3], [10], [13] and uses the results of [5].

Received 21 August 2006; revision received 4 May 2007.

* Postal address: MAPMO, CNRS UMR 6628, Université d'Orléans, France.

Email address: romain.abraham@univ-orleans.fr

** Postal address: MAP5, CNRS UMR 8145, Université René Descartes, Paris, France.

Email address: jean-stephane.dhersin@univ-paris5.fr

*** Postal address: LJK, CNRS UMR 5224, Université Joseph Fourier, Grenoble, France.

Email address: bernard.ycart@ujf-grenoble.fr

The limit of the frequency vector is a random element of the set of stationary measures, hence, it is a random convex combination of the measures corresponding to irreducible recurrent classes. The question arises of how to characterize its probability law. Theorem 3.1, below, first reduces the problem to computing the d cases, where initially a single ball is present, then characterizes those d distributions as the solution to a fixed point problem. The classical Eggenberger–Pólya model [7] can be seen as a particular case of ours: if P is the identity matrix, it is well known that the vector of frequencies converges to a Dirichlet random vector. In our case, it could seem natural to expect a Dirichlet law for the limit stationary distribution; this would be coherent with the numerous connections between Dirichlet distributions and urn models (see, for instance, [12] and [15]). We prove that it is actually the case if no return to a transient state is allowed; see Proposition 3.1, below. We also show, in Proposition 3.2, that the asymptotic distribution is not Dirichlet if returns to transient states are allowed.

The convergence result is stated and proved in Section 2, and the probability distribution of the limit is studied in Section 3.

2. Almost sure convergence

In this section the model is described then the strong convergence result is stated and proved.

Recall that a transition matrix $P = (p_{i,j})$ on the set of types $\{1, \dots, d\}$ is given. Initially, the number of balls in the urn is n_0 and the distribution of types is X_0 (deterministic or not). At each instant $n > 0$ a ball is added to the urn; hence, the number of balls in the urn at time n is $n_0 + n$. The type of ball added depends on the type of ball drawn with uniform probability. If a ball of type i has been drawn, the probability of adding a ball of type j is $p_{i,j}$. We denote by X_n the distribution of types in the urn at time n : X_n is a d -dimensional vector whose i th coordinate is the frequency of type i after the n th addition. It is a random element of the $(d - 1)$ -dimensional simplex, denoted by $\Delta_d = \{(x_1, \dots, x_d) \in [0, 1]^d, x_1 + \dots + x_d = 1\}$.

We will prove that the frequency distributions, X_n , converge almost surely to a stationary distribution of P . We denote by S their set, i.e. the set of (line) vectors x in Δ_d such that $xP = x$.

Theorem 2.1. *The sequence of random vectors (X_n) converges almost surely to an S -valued random vector.*

2.1. Proof of Theorem 2.1

The proof of Theorem 2.1 is based on the classical technique that consists of expressing (X_n) as a stochastic algorithm; see [6] and [13] as general references. This technique has been used several times for proving strong convergence results in urn schemes by, for instance, [3] and [10].

For $j = 1, \dots, d$, let e_j be the d -dimensional vector with a 1 in the j th coordinate and 0s elsewhere. For $x \in \Delta_d$, let $\epsilon(x)$ denote the probability distribution on $\{e_1, \dots, e_d\}$ such that

$$\epsilon(x)(e_j) = \sum_{i=1}^d x_i p_{i,j}.$$

We can write

$$X_{n+1} = \frac{n + n_0}{n + n_0 + 1} X_n + \frac{1}{n + n_0 + 1} \epsilon_n(X_n), \tag{2.1}$$

where the conditional distribution of $\epsilon_n(X_n)$, knowing $X_0 = x_0, \dots, X_n = x_n$, is $\epsilon(x_n)$.

Denote by η_n the following random vector:

$$\eta_n = \varepsilon_n(X_n) - X_n P.$$

The sequence (η_n) is adapted to the filtration \mathcal{F}_n generated by (X_n) and

$$E[\eta_{n+1} \mid \mathcal{F}_n] = 0.$$

Let us rewrite (2.1) as

$$X_{n+1} = X_n + \frac{1}{n_0 + n + 1} (X_n(P - I) + \eta_n). \tag{2.2}$$

Hence, X_n can be viewed as a Robbins–Monro algorithm. We will use the results of [5]. Equation (2.2) is the same as equation (2) of [5]:

$$X_{n+1} = X_n + \gamma_n h(X_n) + \gamma_n \eta_n,$$

with

$$h(X) = X(P - I), \quad \gamma_n = \frac{1}{n + n_0 + 1}, \quad \text{and} \quad \eta_n = \varepsilon_n(X_n) - X_n P.$$

The main assumption in [5] is the notion of an A -stable algorithm.

Definition 2.1. ([5, Definition 1].) We say that the algorithm is A -stable (in our particular case) if

- it remains in a compact set, and
- the series $\sum \gamma_n \eta_n$ converges almost surely.

The main steps of the proof of Theorem 2.1 are then the following.

Step 1: (X_n) is an A -stable algorithm.

Step 2: The distance from X_n to the set S of stationary measures for P tends to 0 almost surely.

Step 3: The sequence (X_n) converges almost surely; hence, its limit is an element of S .

As X_n remains in a compact subset of \mathbb{R}^d (the simplex of probability vectors), step 1 is proved as soon as we can show that $\sum_{n \geq 0} \gamma_n \eta_n < \infty$. Since the random variables $\gamma_n \eta_n$ are the increments of a martingale, which is bounded in L_2 , this result is true. Hence, it is an A -stable algorithm. A classical method to study this type of stochastic algorithm is to compare its trajectories to the flow of an ordinary differential equation, which in our case is $y' = h(y) = y(P - I)$. It is linear, and the nonnull eigenvalues of its matrix $P - I$ all have a negative real part (since P is a stochastic matrix). Therefore, if $x \in \mathbb{R}^d$ and y_x is the solution such that $y_x(0) = x$, then $\lim_{t \rightarrow \infty} y_x(t)$ exists.

Step 2 is rather standard and can be proved by using, for instance, Theorem 2.2 of [16, p. 2153]: the limiting set of (X_n) is an internally chain recurrent set for the flow of the ordinary differential equation $y' = h(y)$; hence, it is included in S . Since (X_n) takes its values in a compact set, and all possible limits of its subsequences are in S , the distance from X_n to S must tend to 0.

Step 3 is an application of Theorem 2 of [5], which we state as Theorem 2.2, below.

If $x \in \mathbb{R}^d$ and $H \subset \mathbb{R}^d$ we set $d(x, H) = \inf\{\|y - x\|, y \in H\}$, where $\|\cdot\|$ denotes the Euclidean norm on \mathbb{R}^d .

Theorem 2.2. ([5, Theorem 2].) *We assume that the algorithm is A-stable. We say that S satisfies assumption (B) if S is a closed set which has a neighborhood N where h is uniformly Lipschitz, and there exist two uniformly Lipschitz functions π and W, defined on N, taking values in ℝ^d and ℝ, respectively, and such that*

- (a) $|\pi(y(t)) - \pi(y(s))| \leq |W(y(t)) - W(y(s))|$ for any solution $(y(u), s \leq u \leq t)$ of $y' = h(y)$ on N,
- (b) $\pi(x) = x$ if $x \in S$.

If S satisfies assumption (B), if $d(X_n, S)$ tends to 0, and if

$$\sum_{n=0}^{\infty} \gamma_n \left| \sum_{i=n}^{\infty} \gamma_i \eta_i \right| < +\infty,$$

then (X_n) converges almost surely to some point of S.

Let us prove first that S satisfies condition (B) of [5]. Here we will take $N = \mathbb{R}^d$ and $\pi(x) = \lim_{t \rightarrow \infty} y_x(t)$. From the same observation on eigenvalues of $P - I$ as in step 2, it follows that π is Lipschitz. If $(y_x(u), s \leq u \leq t)$ is any solution of $y' = h(y)$, then by the definition of π, $\pi(y_x(s)) = \pi(y_x(t))$ and Theorem 2.2(a) holds with $W = 0$. Of course, if $x \in S$ then $h(x) = 0$ and y_x is constant and equal to x. Thus, $x \in S$ implies that $\pi(x) = x$; hence, Theorem 2.2(b) holds.

It remains to prove that

$$\sum_{n=0}^{\infty} \gamma_n \left| \sum_{i=n}^{\infty} \gamma_i \eta_i \right| < \infty.$$

Let us write

$$\begin{aligned} \mathbb{E} \left[\sum_{n=0}^{\infty} \gamma_n \left| \sum_{i=n}^{\infty} \gamma_i \eta_i \right| \right] &= \sum_{n=0}^{\infty} \gamma_n \mathbb{E} \left[\left| \sum_{i=n}^{\infty} \gamma_i \eta_i \right| \right] \\ &\leq \sum_{n=0}^{\infty} \gamma_n \mathbb{E} \left[\left(\sum_{i=n}^{\infty} \gamma_i \eta_i \right)^2 \right]^{1/2} \\ &= \sum_{n=0}^{\infty} \gamma_n \left[\sum_{i=n}^{\infty} \gamma_i^2 \right]^{1/2} \\ &< \infty. \end{aligned}$$

Hence, step 3 is proved. This completes the proof.

3. Asymptotic distribution

Theorem 2.1 proves that the distribution of types in the urn converges to a random element of the set S of stationary distributions for the transition matrix P. In this section we characterize the probability law of that random element.

Assume that the recurrent classes for the transition matrix P are numbered from 1 to k. For $c = 1, \dots, k$, denote by π_c the unique element of S whose coordinates are positive on class number c and null elsewhere. Any element of S is a convex combination of the π_c s. We will

denote by σ the one to one correspondence between S and Δ_k defined by

$$\sigma^{-1}(\alpha) = \sum_{c=1}^k \alpha_c \pi_c \quad \text{for all } \alpha = (\alpha_1, \dots, \alpha_k) \in \Delta_k.$$

Our goal is to describe the distribution of $\sigma(\lim X_n)$, which depends on the initial state of the urn. We will generically denote by A^{X_0} the Δ_k -valued random variable $\sigma(\lim X_n)$ if the initial state of the urn is the random distribution X_0 . We also set

- A^{x_0} if the initial state of the urn is the deterministic distribution x_0 ,
- $A^{(i)}$ if the urn initially contains a single ball of type i .

Observe that the distribution of X_0 is discrete. Obviously,

$$A^{X_0} \stackrel{D}{=} \sum_{x_0} A^{x_0} \mathbf{1}_{\{X_0=x_0\}},$$

where X_0 and all the A^{x_0} s are mutually independent.

Theorem 3.1, below, reduces the distribution of A^{x_0} to those of the $A^{(i)}$ s, then expresses the $A^{(i)}$ s as a solution of a fixed point problem.

Let $x_0 = (x_0(1), \dots, x_0(d))$ be the initial state and n_0 be the initial number of balls. For $b = 1, \dots, n_0$, let $i(b)$ be the type of ball b so that, for $1 \leq p \leq d$,

$$x_0(p) = \text{card}\{b, i(b) = p\}.$$

Theorem 3.1. (i) Let $Y = (Y^{(1)}, \dots, Y^{(n_0)})$ be a random vector, uniformly distributed on Δ_{n_0} . For $1 \leq b \leq n_0$, let A_b be a copy of $A^{(i(b))}$. Assume that $A_b, b = 1, \dots, n_0$ are mutually independent, and independent from the vector Y . Then

$$A^{x_0} \stackrel{D}{=} \sum_{b=1}^{n_0} Y^{(b)} A_b, \tag{3.1}$$

where A_b represents the distribution of the descendents of ball b (and hence is distributed as $A^{(i(b))}$) and $Y^{(b)}$ represents the asymptotic proportion of balls that descend from ball b .

(ii) For $i = 1, \dots, d$, let $A^{(i)'}$ and $A^{(i)''}$ be independent copies of $A^{(i)}$, $Y^{(i)}$ be uniformly distributed on $[0, 1]$, and U_i have distribution $(p_{i,j})_{j=1,\dots,d}$. Assume that all these random variables are mutually independent. Then

$$A^{(i)} \stackrel{D}{=} \sum_{j=1}^d \mathbf{1}_{\{U_i=j\}} (Y^{(j)} A^{(i)'} + (1 - Y^{(j)}) A^{(i)''}). \tag{3.2}$$

Proof. Assume that the n_0 initial balls are labeled from 1 to n_0 . Assume that at each step the ball that has been added receives the same label as the one that has been drawn. Replacing types by labels, we obtain a standard Eggenberger–Pólya urn [7]. Denote by $Y_n = (Y_n^{(b)})$, $b = 1, \dots, n_0$, the distribution of labels at time n : it converges almost surely to a random vector Y whose distribution is uniform on the simplex Δ_{n_0} . For $b = 1, \dots, n_0$, denote by $Z_n^{(b)}$ the d -dimensional vector of the frequencies of types among the balls with label k at time n . By

Theorem 2.1, $Z_n^{(b)}$ converges almost surely to a random variable $Z^{(b)}$, distributed as if the urn initially had only one ball with label $i(b)$; the distribution of $Z^{(b)}$ is that of $A^{(i(b))}$. Moreover, these random variables are mutually independent. The overall distribution of types at time n decomposes as

$$X_n = \sum_{b=1}^{n_0} Y_n^{(b)} Z_n^{(b)}.$$

As n tends to ∞ , X_n tends to

$$X = \sum_{b=1}^{n_0} Y^{(b)} Z^{(b)};$$

hence, (3.1) follows.

Now assume that, initially, a single ball of type i is present. At time 1, another ball is added, which is of type j with probability $p_{i,j}$. Let us apply (ii) with $n_0 = 2$. If two balls of types i and j are present, then the final distribution is

$$Y^{(i)} A^{(i)'} + (1 - Y^{(i)}) A^{(j)'}$$

The limit starting with one single ball of type i or two balls at time 1 must be the same; hence, (3.2) follows.

Equations (3.1) and (3.2) characterize the distribution of A^{x_0} for any x_0 . This follows from the standard results of [14] and [4]. In practice, finding the actual distribution of A^{x_0} may be rather intricate. We will give two examples with a single transient state: one with no possible return (see Proposition 3.1, below) and the other with possible returns (see Proposition 3.2, below).

Observe that from the point of view of A^{x_0} , the contents of recurrent classes is not relevant: each recurrent class can be aggregated into one single absorbing state. Thus, we can assume, with no loss of generality, that the transition matrix P has k absorbing states and $d - k$ transient states.

Proposition 3.1. *Assume that the matrix P is of the form*

$$P = \begin{pmatrix} 0 & p_2 & p_3 & \cdots & p_d \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix},$$

with $p_2 + \cdots + p_d = 1$. Moreover, assume that initially a single ball of type 1 is present. Then the probability distribution of $A^{(1)}$ is the Dirichlet distribution on Δ_{d-1} , with parameter (p_2, \dots, p_d) .

Proof. We prove this result by the classical method of moments, using a martingale argument. For $u = (u_2, \dots, u_d) \in \Delta_{d-1}$, we set, for every $z \in \mathbb{N}^d$,

$$h_u(z) = \frac{(s(z) - 1)!}{\Gamma(z_2 + p_2) \cdots \Gamma(z_d + p_d)} u_2^{z_2} \cdots u_d^{z_d},$$

where $s(z) = \sum_{i=1}^d z_i$. Then, if e_i denotes the i th vector of the canonical basis of \mathbb{R}^d , we have, for $2 \leq i \leq d$,

$$h_u(z + e_i) = \frac{s(z)}{z_i + p_i} u_i h_u(z)$$

and, hence,

$$\sum_{i=2}^d p(z, z + e_i) h_u(z + e_i) = \sum_{i=2}^d u_i h_u(z) = h_u(z).$$

This shows that h_u is a harmonic function and so the process $(h_u(Z_n))$ is a martingale.

Let $\alpha = (\alpha_2, \dots, \alpha_d) \in \mathbb{R}^{d-1}$ such that $\alpha_i + p_i - 1 \geq 0$. We set

$$g_\alpha(z) = \int_{\Delta_{d-1}} h_u(z) u_2^{\alpha_2} \cdots u_d^{\alpha_d} \lambda_{d-1}(du),$$

where λ_{d-1} denotes the Lebesgue measure on Δ_{d-1} . Note that $(g_\alpha(Z_n))$ is still a martingale.

Let us here suppose that, for every $2 \leq i \leq d$, Z_n^i tends to ∞ . Then, using the fact that

$$\frac{\Gamma(x + h)}{\Gamma(x)} \sim x^h \quad \text{as } x \rightarrow \infty,$$

we have (recall $s(Z_n) = n + 1$)

$$\begin{aligned} g_\alpha(Z_n) &= \frac{\Gamma(n + 1)}{\Gamma(n + s(\alpha) + d)} \frac{\Gamma(Z_n^2 + \alpha_2 + 1)}{\Gamma(Z_n^2 + p_2)} \cdots \frac{\Gamma(Z_n^d + \alpha_d + 1)}{\Gamma(Z_n^d + p_d)} \\ &\sim n^{-(s(\alpha) + d - 1)} (Z_n^2)^{\alpha_2 + 1 - p_2} \cdots (Z_n^d)^{\alpha_d + 1 - p_d} \\ &= \left(\frac{Z_n^2}{n}\right)^{\alpha_2 + 1 - p_2} \cdots \left(\frac{Z_n^d}{n}\right)^{\alpha_d + 1 - p_d} \\ &\rightarrow (A_2^{(1)})^{\alpha_2 + 1 - p_2} \cdots (A_d^{(1)})^{\alpha_d + 1 - p_d}. \end{aligned}$$

Let us add that the same kind of computation shows that $g_\alpha(Z_n)$ tends to 0 if one of the Z_n^i is bounded, so the formula is still true in that case.

This computation also proves that g_α is a continuous function that admits limits at ∞ and hence is bounded. Therefore, $g_\alpha(Z_n)$ is a bounded martingale and the convergence also holds in L^1 .

Consequently, for every integer k_2, \dots, k_d , taking $\alpha_i = p_i + k_i - 1$, we have

$$\mathbb{E} \left[(A_2^{(1)})^{k_2} \cdots (A_d^{(1)})^{k_d} \right] = g_\alpha((1, 0, \dots, 0)) = \frac{1}{\Gamma(s(k) + 1)} \frac{\Gamma(k_2 + p_2)}{\Gamma(p_2)} \cdots \frac{\Gamma(k_d + p_d)}{\Gamma(p_d)}.$$

This is also the moments of the Dirichlet distribution with parameters (p_2, \dots, p_d) on Δ_{d-1} and all these moments characterize the law as its support is compact.

We will now show that the pleasant result of Proposition 3.1 worsens as returns to transient states become possible.

Proposition 3.2. *Let $d = 3$ and*

$$P = \begin{pmatrix} p_1 & p_2 & p_3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

with p_1, p_2 , and p_3 strictly positive. Starting initially with a single ball of type 1, the distribution $A^{(1)}$ charges only the two absorbing states 2 and 3. Let us write $A^{(1)} = (A, 1 - A)$, where A is the asymptotic frequency of type 2. Let φ be the generating function of moments of A ,

$$\varphi(z) = \sum_{n=0}^{\infty} E[A^n]z^n.$$

Then

$$\frac{1}{\varphi(z)} = (1 - z)^{p_2} {}_2F_1(p_2, -p_1, 1 - p_1)(z), \tag{3.3}$$

where ${}_2F_1(p_2, -p_1, 1 - p_1)$ is the hypergeometric function with parameters $(p_2, -p_1)$ and $1 - p_1$.

Having computed the moments of A , it is easy to check that its distribution is not Beta, except for the particular case in which $p_2 = p_3$. Hence, the distribution of $A^{(1)}$ is not Dirichlet.

Proof of Proposition 3.2. From Theorem 3.1(ii), A is equal in distribution to

$$YA' + (1 - Y)(A'' \mathbf{1}_{\{U=1\}} + \mathbf{1}_{\{U=2\}}), \tag{3.4}$$

where A' and A'' are distributed as A , Y is uniformly distributed on $[0, 1]$, U has distribution (p_1, p_2, p_3) , and (A', A'', Y, U) are mutually independent.

Denote by $c_n = E[A^n]$ the n th moment of A . From (3.4), the following induction for c_n is deduced:

$$\begin{aligned} c_n &= \frac{1}{n+1}c_n + \sum_{k=0}^{n-1} \binom{n}{k} E[\rho^k(1 - \rho)^{n-k}] (c_k c_{n-k} p_1 + c_k p_2), \\ \iff nc_n &= p_1 \sum_{k=0}^{n-1} c_k c_{n-k} + p_2 \sum_{k=0}^{n-1} c_k, \\ \iff (n + p_1 + p_2)c_n &= p_1 \sum_{k=0}^n c_k c_{n-k} + p_2 \sum_{k=0}^n c_k. \end{aligned}$$

Multiplying by z^n and summing leads to

$$z\varphi'(z) + (p_1 + p_2)\varphi(z) = p_1\varphi(z)^2 + p_2\frac{\varphi(z)}{1 - z}.$$

Letting $\psi = 1/\varphi$ leads to

$$z(z - 1)\psi'(z) - (p_1(z - 1) + p_2z)\psi(z) = p_1(1 - z),$$

from which (3.3) follows.

The technique of conditioning upon the first drawn ball also permits treatment of the case in which there is a single transient state with possible returns, and more than two absorbing states. By induction on d , we can express the distribution of $A^{(1)}$ using Proposition 3.2.

Acknowledgements

We are indebted to Brigitte Chauvin, Gérard Letac, and Nicolas Pouyanne for helpful comments and hints.

References

- [1] BAI, Z. D. AND HU, F. (2005). Asymptotics in randomized urn models. *Ann. Appl. Prob.* **15**, 914–940.
- [2] BAI, Z. D., HU, F. AND ZHANG, L. X. (2002). Gaussian approximation theorems for urn models and their applications. *Ann. Appl. Prob.* **12**, 1149–1173.
- [3] BENAÏM, M., SCHREIBER, S. AND TARRÈS, P. (2004). Generalized urn models of evolutionary processes. *Ann. Appl. Prob.* **14**, 1455–1478.
- [4] CHAMAYOU, J.-F. AND LETAC, G. (1991). Explicit stationary distributions for compositions of random functions and product of random matrices. *J. Theoret. Prob.* **4**, 3–36.
- [5] DELYON, B. (1996). General results on the convergence of stochastic algorithms. *IEEE Trans. Automatic Control* **41**, 1245–1255.
- [6] DUFLO, M. (1997). *Random Iterative Models*. Springer, New York.
- [7] EGGENBERGER, F. AND PÓLYA, G. (1928). Sur l'interprétation de certaines courbes de fréquence. *C. R. Acad. Sci. Paris* **187**, 870–872.
- [8] GOUET, R. (1993). Martingale functional central limit theorems for a generalized Pólya urn. *Ann. Prob.* **21**, 1624–1639.
- [9] GOUET, R. (1997). Strong convergence of proportions in a multicolor Pólya urn. *J. Appl. Prob.* **34**, 426–435.
- [10] HIGUERAS, I., MOLER, J., PLO, F. AND SAN MIGUEL, M. (2003). Urn models and differential algebraic equations. *J. Appl. Prob.* **40**, 401–412.
- [11] JANSON, S. (2004). Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stoch. Process. Appl.* **110**, 177–245.
- [12] JOHNSON, N. L. AND KOTZ, S. (1977). *Urn Models and Their Application*. John Wiley, New York.
- [13] KUSHNER, H. J. AND YIN, G. G. (1997). *Stochastic Approximation Algorithms and Applications*. Springer, New York.
- [14] LETAC, G. (1986). A contraction principle for certain Markov chains and its applications. *Contemp. Math.* **50**, 263–273.
- [15] PITMAN, J. (1996). Some developments of the Blackwell–MacQueen urn scheme. In *Statistics, Probability and Game Theory* (IMS Lecture Notes Monogr. Ser. **30**), Institute of Mathematical Statistics, Hayward, CA, pp. 245–267.
- [16] SCHREIBER, S. (2001). Urn models, replicator processes and random genetic drift. *SIAM J. Appl. Math.* **61**, 2148–2167.