Strong convergence for urn models with reducible replacement policy

R. Abraham*, J.S. Dhersin**, and B. Ycart***

Running title: Urn models with reducible replacements

Addresses: * MAPMO, CNRS UMR 6628, Université d'Orléans, France
** MAP5, CNRS UMR 8145, Université René Descartes, Paris, France
*** LJK, CNRS UMR 5224, Université Joseph Fourier, Grenoble, France

E-mail addresses: * `Romain.Abraham@univ-orleans.fr`
** `Jean-Stephane.Dhersin@univ-paris5.fr`
*** `Bernard.Ycart@ujf-grenoble.fr`

Corresponding author : J.-S. Dhersin

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 more the case as soon as returns to transient states are allowed.

Keywords: urn model, Markov chain, strong convergence

AMS Subject Classification: 60F15

1 Introduction

In the vast literature devoted to urn models (see Johnson and Kotz [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 balls 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, 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 are more than one?

We study the simplest possible model: balls are added one by one, the type of the added ball only depends on the type of the one that has been drawn. We believe that our results can be extended to more general schemes, such as those of Janson [11] or Benaïm et al. [3]. The 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 (Theorem 2.1). The proof is based on the classical stochastic algorithm technique [3, 10, 13], and uses the results of Delyon [5].

The limit is a random element of the set of stationary measures, hence a random convex combination of the measures corresponding to irreducible recurrent classes. The question arises to characterize its probability law. Theorem 3.1 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, 15]). We prove that it is actually the case if no return to a transient state is allowed (Proposition 3.2). We also show in Proposition 3.3 that the asymptotic distribution is not Dirichlet if returns to transient states are allowed.

The convergence result is stated and proved in Section 2, 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 the ball which is added depends on that of a ball drawn with uniform probability. If a ball of type i has been drawn, the probability to add 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 Δ_d.

$$\Delta_d = \{ (x_1, \ldots, x_d) \in [0,1]^d ,\ x_1 + \cdots + x_d = 1 \} .$$

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 a S-valued random vector.*

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

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

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

One can write:

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

where the conditional distribution of $\varepsilon_n(X_n)$ knowing $X_0 = x_0, \ldots, X_n = x_n$ 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

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

Let us rewrite (2.1) as:

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

Hence X_n can be seen as a Robbins-Monro algorithm. We shall use the results of Delyon [5]. Equation (2.2) is the same as equation (2) in [5]:

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

with

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

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

Definition 2.2 *([5], Definition 1)*
We say that the algorithm is A-stable (in our particular case) if

- *It remains in a compact set.*
- *The serie* $\sum \gamma_n \eta_n$ *converges a.s.*

The main steps of the proof 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 a.s.

Step 3 The sequence (X_n) converges a.s., 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 non-null 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 $\lim_{t\to+\infty} y_x(t)$ exists.

Step 2 is rather standard and can be proved by using for instance Theorem 2.2 p. 2153 of [17]: the limiting set of (X_n) is an internally chain recurrent set for the flow of the ODE $y' = h(y)$, 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 in [5]:

Theorem 2.3 *([5], Theorem 2)*
We assume that the algorithm is A-stable.
If $\mathcal{S}$ *satisfies assumption* (B)*:*

> $\mathcal{S}$ *is a closed set which has a neighbourhood* N *where* h *is uniformly Lipschitz and there exist two uniformly Lipschitz functions* π, W, *defined on* N, *taking values in* $\mathbb{R}^d$ *and* $\mathbb{R}$ *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 $d(X_n, \mathcal{S})$ *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 a.s. to some point of* $\mathcal{S}$.

Let us prove first that S satisfies condition (B) of [5]. Here we shall take $N = \mathbb{R}^d$ and $\pi(x) = \lim_{t \to +\infty} y_x(t)$. $>$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 definition of π, $\pi(y_x(s)) = \pi(y_x(t))$ and (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 $\pi(x) = x$, hence (b).

There 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, which ends the proof of Theorem 2.1. □

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 the recurrent classes for the transition matrix P are numbered from 1 to k. For $c = 1, \ldots, k$, 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 shall denote by σ the one to one correspondence between S and Δ_k defined by:

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

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^* the Δ_k-valued random variable $\sigma(\lim X_n)$ for $X_0 = *$:

- A^{X_0} if the initial state of the urn is the random distribution X_0,

- 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,

$$A^{X_0} \stackrel{(d)}{=} \sum_{x_0} A^{x_0}\, \mathbb{I}_{X_0=x_0} \ ,$$

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

$$x_0(p) = \mathrm{Card}\{b,\ i(b) = p\}.$$

Theorem 3.1

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

$$A^{x_0} \stackrel{(d)}{=} \sum_{b=1}^{n_0} Y^{(b)}\, A_b \ . \tag{3.3}$$

   (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)

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

$$A^{(i)} \stackrel{(d)}{=} \sum_{j=1}^{d} \mathbb{I}_{U_i=j} \left(Y^{(j)}\, A^{(i)\prime} + (1 - Y^{(j)})\, A^{(j)\prime\prime} \right) \ . \tag{3.4}$$

Proof: Assume the n_0 initial balls are labelled 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, one gets a standard Eggenberger-Pólya urn [7]. Denote by $Y_n = (Y_n^{(b)})$, $b = 1, \ldots, n_0$ the distribution of labels at time n: it converges a.s. to a random vector Y whose distribution is uniform on the simplex Δ_{n_0}. For $b = 1, \ldots, n_0$, 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 a.s. to a random variable $Z^{(b)}$, distributed as if initially the urn only had 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 infinity, X_n tends to

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

hence equation (3.3).

Assume now 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 point *1* with $n_0 = 2$: if two balls of types i and j are present, then the final distribution is that of:

$$Y^{(i)}\, A^{(i)\prime} + (1 - Y^{(i)})\, A^{(j)\prime\prime}.$$

The limit starting with one single ball of type i or the two balls of time 1 must be the same, hence equation (3.4). $\square$

Equations (3.3) and (3.4) characterize the distribution of A^{x_0}, for any x_0. This follows from standard results of Letac [14] and Chamayou and Letac [4]. In practice, finding the actual distribution of A^{x_0} may be rather intricate. We shall give two examples with a single transient state, one with no possible return (Proposition 3.2), the other with possible returns (Proposition 3.3).

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 one can assume with no loss of generality that the transition matrix P has k absorbing states and $d - k$ transient states.

Proposition 3.2 *Assume the matrix P is the following*

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

with $p_2 + \ldots + p_d = 1$. Assume moreover that initially a single ball of type 1 is present.

The probability distribution of $A^{(1)}$ is the Dirichlet distribution on Δ_{d-1}, with parameter $(p_2, \ldots, p_d)$.

Proof: We prove this result by the classical method of moments, using a martingale argument.

For $u = (u_2, \ldots, u_d) \in \Delta_{d-1}$, 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 ith vector of the canonical basis of $\mathbb{R}^d$, we have, for $2\le i\le d$

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

and hence

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

This shows that h_u is an 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 $\alpha_i+p_i-1\ge 0$. 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}. Remark that $(g_\alpha(Z_n))$ is still a martingale.

Let us here suppose that for every $2\le i\le d$, Z_n^i tends to infinity. Then, using that

$$\frac{\Gamma(x+h)}{\Gamma(x)}\underset{x\to+\infty}{\sim} x^h,$$

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

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

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 infinity and hence is bounded. Therefore $g_\alpha(Z_n)$ is a bounded martingale and the convergence also holds in L^1.

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

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

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. $\square$

We will now show that the pleasant result of Proposition 3.2 worsens as returns to transient states become possible.

Proposition 3.3 *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, p_3 strictly positive. Starting initially with a single ball of type 1, the distribution $A^{(1)}$ only charges 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.

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

Then

$$\frac{1}{\varphi(z)} = (1 - z)^{p_2}\, {}_2F_1(p_2, -p_1, 1 - p_1)(z) , \tag{3.5}$$

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

Proof: >From point *2* of Theorem 3.1, A is equal in distribution to:

$$Y A' + (1 - Y)(A''\mathbb{I}_{U=1} + \mathbb{I}_{U=2}) , \tag{3.6}$$

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 $c_n = \mathbb{E}[A^n]$ the n-th moment of A. From (3.6), the following induction for c_n is deduced.

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

Multiplying by z^n and summing leads to

$$z\varphi'(z) + (p_1 + p_2)\varphi(z) = p_1 \varphi(z)^2 + p_2 \frac{\varphi(z)}{1 - z} . \tag{3.7}$$

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

$$z(z - 1)\psi'(z) - (p_1(z - 1) + p_2 z)\psi(z) = p_1(1 - z) , \tag{3.8}$$

from which (3.5) follows. □

The technique of conditioning upon the first drawn ball also permits to treat the case where there is a single transient state with possible return, and more than 2 absorbing ones. By induction on d, one can express the distribution of $A^{(1)}$ using Proposition 3.3.

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

References

[1] Z.D. Bai and F. Hu. Asymptotics in randomized urn models. *Ann. Appl. Probab*, 15(1B):914–940, 2005.

[2] Z.D. Bai, F. Hu, and L.X. Zhang. Gaussian approximation theorems for urn models and their applications. *Ann. Appl. Probab*, 12(4):1149–1173, 2002.

[3] M. Benaïm, S. Schreiber, and P. Tarrès. Generalized urn models of evolutionary processes. *Ann. Appl. Probab.*, 14(3):1455–1478, 2004.

[4] J.-F. Chamayou and G. Letac. Explicit stationary distributions for compositions of random functions and product of random matrices. *J. Theor. Probab.*, 4:3–36, 1991.

[5] B. Delyon. General results on the convergence of stochastic algorithms. *IEEE Trans. Automatic Control*, 41:1245–1255, 1996.

[6] M. Duflo. *Random iterative models*. Springer, New York, 1997.

[7] F. Eggenberger and G. Pólya. Sur l'interprétation de certaines courbes de fréquence. *C. R. Acad. Sci. Paris*, 187:870–872, 1928.

[8] R. Gouet. Martingale functional central limit theorems for a generalized Pólya urn. *Ann. Probab.*, 21(3):1624–1639, 1993.

[9] R. Gouet. Strong convergence of proportions in a multicolor Pólya urn. *J. Appl. Probab.*, 34(2):426–435, 1997.

[10] I. Higueras, J. Moler, F. Plo, and M. San Miguel. Urn models and differential algebraic equations. *J. Appl. Probab.*, 40(2):401–412, 2003.

[11] S. Janson. Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stoch. Proc. Appl.*, 110(2):177–245, 2004.

[12] N.L. Johnson and S. Kotz. *Urn models and their application.* Wiley, New York, 1977.

[13] H.J. Kushner and G.G. Yin. *Stochastic approximation algorithms and applications.* Springer, New York, 1997.

[14] G. Letac. A contraction principle for certain Markov chains and its applications. *Contemp. Math.*, 50:263–273, 1986.

[15] J. Pitman. Some developments of the Blackwell-MacQueen urn scheme. In T.S. Ferguson et al., editor, *Statistics, Probability and Game theory; Papers in honor of David Blackwell*, volume 30, pages 245–267. IMS, Hayward, 1996.

[16] N. Pouyanne. Classification of large Pólya-Eggenberger urns with regard to their asymptotics. *Disc. Math. Theor. Comp. Science*, 2006. to appear.

[17] S. Schreiber. Urn models, replicator processes and random genetic drift. *SIAM J. Appl. Math.*, 61(6):2148–2167, 2001.