Multiple drawing multi-colour urns by stochastic approximation

Nabil Lasmar[*], Cécile Mailler[†] and Olfa Selmi[‡]

October 3, 2017

Abstract

A classical Pólya urn scheme is a Markov process whose evolution is encoded by a replacement matrix $(R_{i,j})_{1\leqslant i,j\leqslant d}$. At every discrete time-step, we draw a ball uniformly at random, denote its colour c, and replace it in the urn together with $R_{c,j}$ balls of colour j (for all $1 \leqslant j \leqslant d$).

We study multi-drawing Pólya urns, where the replacement rule depends on the random drawing of a set of m balls from the urn (with or without replacement). Many particular examples of this situation have been studied in the literature, but the only general results are by Kuba & Mahmoud (arXiv:1503.09069 and 1509.09053). These authors prove second order asymptotic results in the 2-colour case, under the so-called *balance* and *affinity* assumptions, the latter being somewhat artificial.

The main idea of this work is to apply stochastic approximation methods to this problem, which enables us to prove analogous results to Kuba & Mahmoud, but without the artificial *affinity* hypothesis, and, for the first time in the literature, in the d-colour case ($d \geqslant 3$). We also give some partial results in the two-colour non-balanced case, the novelty here being that the only results for this case currently in the literature are for particular examples.

Keywords: multiple-drawing Pólya urns, limit theorems, stochastic approximation; reinforced processes; discrete-time martingales

AMS subjects: 60F05; 60F15; 62L20

Contents

1 Introduction **2**

1.1 Classical Pólya urns . . . 2

1.2 Multi-drawing Pólya urns . . . 3

1.3 The main contributions of this paper . . . 3

1.4 Definition of the model and main assumptions . . . 3

1.5 Main result for d-colour balanced urns . . . 4

1.6 Main results for the two-colour case . . . 6

1.7 Plan of the article . . . 7

1.8 Tenability . . . 7

2 Law of large numbers and central limit theorem for balanced urn schemes **8**

2.1 This model can be described as a stochastic algorithm . . . 8

2.2 Strong law of large numbers (Proof of Theorem 1(*a*)) . . . 9

[*]Département des Mathématiques, Institut Préparatoire aux Études d'Ingénieur, Monastir, Tunisia (`nabillasmar@yahoo.fr`).
[†]University of Bath, Department of Mathematical Sciences, Claverton Down, BA2 7AY Bath, UK (`c.mailler@bath.ac.uk`).
[‡]Département des Mathématiques, Faculté des Sciences de Monastir, Monastir, Tunisia (`selmiolfa3@yahoo.fr`).

2.3 Central limit theorem (Proof of Theorem 1(b)) . . . 10
2.4 The diagonal case (proof of Theorem 2) . . . 12

3 The two-colour particular case (proof of Corollary 1) 14

4 Examples 14
4.1 Two-colour examples . . . 14
4.2 Three-colour examples . . . 16

5 The two-colour non-balanced case (proof of Theorem 3) 19
5.1 Estimates for the total number of balls in the urn . . . 20
5.2 The associated stochastic algorithm . . . 20
5.3 Law of large numbers . . . 21
5.4 Central limit theorem . . . 21
5.5 Examples . . . 22

1 Introduction

1.1 Classical Pólya urns

Pólya urn schemes are the simplest example of stochastic process with reinforcement. Understanding these objects is thus the first step in understanding many more intricate models such as reinforced random walks, preferential attachment networks, interacting urn models, etc. Furthermore, reinforced stochastic processes appear in a wide range of applications; in biology (e.g. ants walks, reinforced branching processes), in finance and clinical trials, and in computer science (preferential attachment networks are used to model the complex networks such as the internet, the World Wide Web, or social networks). We refer the reader to Pemantle's survey [Pem07] for an overview on both mathematics and applications of reinforced processes.

The classical Pólya urn, first introduced by Eggenberger & Pólya [EP23], is described as follows: an urn contains initially one white ball and one black ball. At each discrete time step, one picks a ball uniformly at random and replaces it in the urn together with another ball of the same colour. Many generalisations of this model have been studied in the literature, namely urns with more than two colours, or with different, possibly random, replacement rules.

The methods used to study Pólya urns are quite varied: since the seminal work of Athreya & Karlin [AK68], one successful approach is to embed the urn process in continuous time using exponential clocks and then apply martingale arguments. This method can be very effective - see, for example Janson [Jan04, Jan06]; the main difficulty is then to pull the results back into the discrete-time framework.

Since the work of Flajolet, Dumas & Puyhaubert [FDP06], analytic combinatorics have been used to study Pólya's urn schemes (see Morcrette [Mor12] and Morcrette & Mahmoud [MM12]). The main advantage of this method is that, when successful, it gives results for fixed finite time and not only asymptotically when time goes to infinity as the embedding-in-continuous-time method does. The major drawback is that this method is often non-tractable as it relies on solving a non-linear differential system.

Finally, stochastic approximation (or stochastic algorithms) provides a powerful toolbox to prove strong laws of large numbers and central limit theorems for the composition of a Pólya urn. This method relies on discrete time martingale methods; a good introduction to stochastic approximation is the book by Duflo [Duf97]. The literature on stochastic algorithms is very wide; in the urn context, stochastic approximation has already been used, for example by Laruelle & Pagès [LP13] for the study of Pólya urns with random replacement rules. In this article, we apply the stochastic approximation results of Zhang [Zha16] and Renlund [Ren11] to a model of multi-drawing Pólya urns.

1.2 Multi-drawing Pólya urns

In this article, we focus on the multi-drawing generalisation of Pólya urns: instead of choosing the replacement rule according to the random drawing of one ball in the urn, one picks at random a handful of balls (more precisely a fixed number m of balls), and the replacement rule then depends on the colours of those m balls. This model has numerous applications, such as degrees in increasing trees and preferential attachment networks (see Kuba & Sulzbach [KS16]), leaves in random circuits (see [KS16]) and tournaments *à la rock-paper-scissors* (see Laslier & Laslier [LL13]).

Many particular examples have been studied in the literature, mostly in the two-colour case: see, for example, Chen & Wei [CW05], Chen & Kuba [CK13] and Kuba, Mahmoud & Panholzer [MKP13] and Aguech, Lasmar & Selmi [ALS].

In recent work, Kuba & Mahmoud [KM17] and Kuba & Sulzbach [KS16] considered the general case of two-colour Pólya urns with multiple drawing and proved third-order asymptotics of the composition of the urn under two hypothesis. The standard *balance* hypothesis makes the total number in the urn deterministic. The *affinity* hypothesis is the assumption that, if one denotes by W_n the number of white balls in the urn at time n, then there exists two deterministic sequences $(\alpha_n)_{n\geqslant 0}$ and $(\beta_n)_{n\geqslant 0}$ such that, for all integers n, $\mathbb{E}[W_{n+1}|\mathcal{F}_n] = \alpha_n W_n + \beta_n$.

1.3 The main contributions of this paper

The main idea of this paper is to apply the stochastic approximation methods to the multi-drawing problem. Our main motivation for doing this was the successful application by Laruelle & Pagès [LP13] of these techniques to the classical Pólya urn model with random replacement matrices.

Using these methods we prove results in the d-colour ($d \geqslant 2$), multi-drawing case. In the case $d \geqslant 3$, these are the first general results in the literature. In the case $d \geqslant 2$, our results are analogous to Kuba & Mahmoud [KM17], but importantly do not require the somewhat artificial affinity hypothesis.

Furthermore, we prove partial results on the two-colour *non-balanced* case; these are the first results in the literature for this situation, apart from the study of particular examples in, for example, [ALS].

1.4 Definition of the model and main assumptions

The model we study is defined as follows. An urn contains balls of $d \geqslant 2$ different colours. At time 0, the urn contains $U_{0,i} \geqslant 0$ balls of colour i, for all $1 \leqslant i \leqslant d$. We assume that the urn is originally non empty, namely $\sum_{i=1}^d U_{0,i} > 0$. We fix an integer $m \geqslant 1$ and a replacement rule $R : \Sigma_m^{(d)} \to \mathbb{Z}^d$, where

$$\Sigma_m^{(d)} = \{(v_1, \dots, v_d) \in \mathbb{N}^d \,:\, \sum_{i=1}^d v_i = m\}.$$

We denote by $R_1, \dots, R_d$ the coefficient functions of R. At each (discrete) time step $n \geqslant 1$, we draw m balls in the urn *with or without replacement*, and denote by $\xi_{n,i}$ the number of balls of colour i among those m balls. Let $\xi_n = (\xi_{n,1}, \dots, \xi_{n,d})$. Conditionally on $\xi_n = v$, we then add into the urn $R_i(v)$ balls of type i into the urn, for all $1 \leqslant i \leqslant d$. More precisely, we have $U_n = U_{n-1} + R(\xi_n)$, for all $n \geqslant 1$, where $U_{n,i}$ is the number of balls of colour i in the urn at time n, and $U_n = (U_{n,1}, \dots, U_{n,d})$.

Let us denote by T_n the total number of balls and $Z_{n,i} = {}^{U_{n,i}}/_{T_n}$ the proportion of balls of type i in the urn at time n. Note that, with these notations, in the with-replacement case,

$$P_n(v) := \mathbb{P}(\xi_{n+1} = v|\mathcal{F}_n) = \binom{m}{v_1, \dots, v_d} \prod_{i=1}^d Z_{n,i}^{v_i},$$

while in the without-replacement case,

$$P_n(v) := \mathbb{P}(\xi_{n+1} = v|\mathcal{F}_n) = \frac{1}{\binom{T_n}{m}} \prod_{i=1}^d \binom{U_{n,i}}{v_i}.$$

In the literature, it is often assumed that the urn is *balanced*, meaning that, for all $v \in \Sigma_m^{(d)}$, $r(v) := \sum_{i=1}^d R_i(v) = S$, where S is a positive integer. Without this assumption, one needs to control the speed of convergence of T_n/n to its limit, and this is not yet understood for $m \geqslant 2$ (i.e. in the multiple-drawing case). It is also standard to assume that the urn is *tenable*, meaning that it is never asked to remove from the urn balls that are not in the urn: we give necessary and sufficient conditions for the tenability to be achieved. We give labels to these two assumptions since they will be used all along the article:

(B) For all $v \in \Sigma_m^{(d)}$, $r(v) := \sum_{i=1}^d R_i(v) = S$.

(T) The urn scheme is tenable.

Although the tenability assumption is natural (an alternative would be to work conditionally on the event that no impossible configuration happens), we believe it is an interesting and challenging open question to remove the balance hypothesis altogether. Note that in the classical $m = 1$ case, strong results can be proved without the balance assumption (see Janson [Jan04]); the reason is that the urn scheme embedded in continuous time is a multi-type Galton-Watson process, which is no longer the case when $m \geqslant 2$. Although our main results for d-colour urns require the balance assumption, we are able to prove partial results for two-colour non-balanced urns.

The following function from $\Sigma^{(d)} = \{(x_1, \dots, x_d) \in [0,1]^d \colon \sum_{i=1}^d x_i = 1\}$ onto $\{(x_1, \dots, x_d) \in \mathbb{R}^d \colon \sum_{i=1}^d x_i = 0\}$ (two $(d-1)$-dimensional spaces) and its zeros will play a crucial role in the article (recall that $r(v) := \sum_{i=1}^d R_i(v)$):

$$h(x) = \sum_{v \in \Sigma_m^{(d)}} \binom{m}{v_1, \dots, v_d} \left(\prod_{i=1}^d x_i^{v_i}\right) \Big(R(v) - r(v)\, x\Big).$$

We will especially focus on the stable zeros of h, i.e. the zeros at which all eigenvalues of ∇h (the Jacobian matrix of h) have a negative real part: we denote by $\mathtt{Stable}(h)$ the set of stable zeros of h.

In this article, we focus on the renormalised composition vector $Z_n := (Z_{n,1}, \dots, Z_{n,d})$. In view of the definition above, the process $(Z_n)_{n \geqslant 0}$ is a Markov process that depends on two parameters, the initial composition Z_0 and the replacement function R. Before stating them in full detail in the rest of this introduction, let us summarise our main results. We prove the following results for balanced urns:

- The *limit set* (see Definition 1) of the renormalised composition vector Z_n is almost surely a compact connected set of $\Sigma^{(d)}$ stable by flow of the differential equation $\dot{x} = h(x)$ (see Theorem 1(a)). Note that, given a function h, it is a non-trivial question to determine the compact connected sets stable by the flow of $\dot{x} = h(x)$ (we give some examples in Section 4). Also, the limit set of Z_n a priori depends on the initial composition vector Z_0. Reducing the number of colours to $d = 2$ makes the matter much simpler and, in this particular case, we state almost sure convergence to a constant vector θ as long as the function h is not constant equal to zero (see Corollary 1(a)).
- Assuming that Z_n converges almost surely to a stable zero of h, we prove convergence of the fluctuations around this almost sure limit (see Theorem 1(b)).
- When $h \equiv 0$ (we call this case the *diagonal case*), we prove almost sure convergence of the renormalised composition vector to a random vector Z_∞ (see Theorem 2).

In the two-colour non-balanced case, we prove almost sure convergence to a zero of h (as long as $h \not\equiv 0$), and some partial result for the fluctuations around this almost sure limit (see Theorem 3).

1.5 Main result for d-colour balanced urns

Definition 1 (See, e.g., Pemantle [Pem07, Definition 2.11]). *Given a stochastic process $(\Pi_n)_{n \geqslant 0}$, we define its limit set as*

$$L(\Pi) := \bigcap_{n \geqslant 0} \overline{\bigcup_{m \geqslant n} \Pi_m}.$$

Theorem 1 (Balanced and tenable d-colour urns). *Under Assumptions (B) and (T), we have:*

(a) *The limit set of the renormalised composition vector $Z_n = (Z_{n,1}, \dots, Z_{n,d})$ is almost surely a compact connected set of $\Sigma^{(d)}$ stable by the flow of the differential equation $\dot{x} = h(x)$. If there exists $\theta \in \Sigma^{(d)}$ such that $h(\theta) = 0$ and, for all $n \geqslant 0$, $\langle h(Z_n), Z_n - \theta\rangle < 0$, then Z_n converges almost surely to θ.*

(b) *Assume that there exists a stable zero θ of h such that Z_n converges almost surely to θ when n goes to infinity. Let Λ be the eigenvalue of $-\nabla h(\theta)$ with the smallest real part, let*

$$\Gamma = \frac{1}{S^2} \sum_{v \in \Sigma_m^{(d)}} \binom{m}{v_1, \dots, v_d} \left(\prod_{i=1}^d \theta_i^{v_i}\right) (R(v) - S\theta)(R(v) - S\theta)^t.$$

Then,

- *if* $\mathtt{Re}(\Lambda) > S/2$, *then* $\sqrt{n}(Z_n - \theta) \to \mathcal{N}(0, \Sigma)$, *in distribution when* $n \to \infty$, *where*

$$\Sigma = \int_0^{+\infty} \left(\exp\left(\left(\frac{\nabla h(\theta)}{S} + \frac{\mathtt{Id}}{2}\right)u\right)\right)^t \Gamma \exp\left(\left(\frac{\nabla h(\theta)}{S} + \frac{\mathtt{Id}}{2}\right)u\right) du.$$

Assume additionally that all Jordan blocks of $\nabla h(\theta)$ associated to Λ are of size 1. Then,

- *if* $\mathtt{Re}(\Lambda) = S/2$, *then* $\sqrt{n/\log n}(Z_n - \theta) \to \mathcal{N}(0, \Sigma)$, *in distribution when* $n \to \infty$, *where*

$$\Sigma = \lim_{n\to\infty} \frac{1}{\log n} \int_0^{\log n} \left(\exp\left(\left(\frac{\nabla h(\theta)}{S} + \frac{\mathtt{Id}}{2}\right)u\right)\right)^t \Gamma \exp\left(\left(\frac{\nabla h(\theta)}{S} + \frac{\mathtt{Id}}{2}\right)u\right) du;$$

- *if* $\mathtt{Re}(\Lambda) < S/2$, *then* $n^{\mathtt{Re}(\Lambda)/S}(Z_n - \theta)$ *converges almost surely to a random variable.*

Remark: The fact that both Z_n and θ are in $\Sigma^{(d)}$ implies that the matrix Σ satisfies $\Sigma \cdot (1, \dots, 1)^t = 0$.

Remark: When applying this theorem to particular examples, the difficulty is to understand the flow of the differential equation $\dot{x} = h(x)$. It is in general non-trivial to describe the limit set of $(Z_n)_{n\geqslant 0}$ (see, e.g., Laslier & Laslier [LL13] where this is carried out for one particular example), and this limit set is a priori not a point and depends on the initial composition Z_0. The strength of our result is that, if one can prove, for a particular replacement rule and for one particular initial composition vector Z_0, that the renormalised composition vector Z_n converges almost surely to a stable zero of h, then the fluctuations are given relatively easily by Theorem 1(b). We give numerous examples in Section 4.

To discuss our main result, let us first apply it to the well-known particular case $m = 1$. In that case, the replacement rule is encoded by a replacement matrix $R = (R_{i,j})_{1\leqslant i,j\leqslant d}$. At every time step, we pick a ball uniformly at random in the urn, denote by i its colour, and put it back in the urn together with $R_{i,j}$ balls of colour j, for all $1 \leqslant i, j \leqslant d$. Note that this translates into our framework as $R_j(\boldsymbol{e}_i) = R_{i,j}$ for all $1 \leqslant i, j \leqslant d$ (where $\boldsymbol{e}_i$ is the vector whose coordinates are all equal to 0 except the ith, which is equal to 1). We thus have $h(x) = (R - S\mathtt{Id}_d)x$. If we assume that R is irreducible and that the urn is tenable, the Perron-Frobenius theorem implies that S is the spectral radius of R, the multiplicity of S as eigenvalue of R is one, and there exists an eigenvector ϖ associated to S whose coordinates are all non-negative and such that $\sum_{i=1}^d \varpi_i = 1$.

In particular, ϖ is the unique zero of h on $\Sigma^{(d)}$, and the eigenvalues of $R - S\mathtt{Id}_d$ restricted to $\Sigma^{(d)}$ all have negative real parts. Thus ϖ is a stable zero of h, and, for all $x \neq \varpi$, $\langle h(x), x - \varpi\rangle = \langle R - S\mathtt{Id}_d)(x - \varpi), x - \varpi\rangle < 0$. Theorem 1$(a)$ applies and gives almost sure convergence of Z_n to ϖ and Theorem 1(b) applies straightforwardly, reproving some standard results from the literature (see Janson [Jan04]).

By analogy to the classical $m = 1$ case, we can say that the cases for which Z_n does not converge to a constant vector or converges to a constant vector that depends on Z_0 are the *non-irreducible* cases. Unfortunately, we are not yet able to give a nice characterisation of those cases in terms of the replacement function R. Theorem 1(a) gives a sufficient condition; namely, if there exists a zero θ of h such that $\langle h(x), x - \theta\rangle < 0$ for all $x \in \Sigma^{(d)}$, then Z_n converges almost surely to θ (for all choices of Z_0 satisfying (T)). Note that, in the classical setting $m = 1$, the non-irreducible cases are also more intricate; see Janson [Jan06] for two-colour triangular urns and Bose, Dasgupta & Maulik [BDM09] for the d-colour balanced and triangular case.

In the case when $h \equiv 0$, which we call *diagonal* case (although [KM17, KS16] call it *triangular* - we choose to change the terminology so that it is coherent with the $m = 1$ case), we prove the following:

Theorem 2 (Diagonal balanced case)**.** *Under Assumption* (B) *and* (T)*, if $h \equiv 0$, then Z_n converges almost surely to a random vector Z_∞ when n tends to infinity.*

1.6 Main results for the two-colour case

In the two-colour case, the renormalised composition vector Z_n of the urn at time n is characterised by its first coordinate since $Z_{n,1} + Z_{n,2} = 1$. Thus, it is enough to study the first coordinate and the problem becomes unidimensional and easier. Let us introduce simplified notations for this special case.

In the two-colour case, we say that the urn contains white balls and black balls. We denote by W_n and $X_n = {}^{W_n}/_{T_n}$ respectively the number and the proportion of white balls in the urn at time n. The replacement function can be seen as a replacement matrix

$$R = \begin{pmatrix} a_0 & b_0 \\ a_1 & b_1 \\ \vdots & \vdots \\ a_m & b_m \end{pmatrix},$$

where the a_i's, and b_i's are integers (such that $a_i + b_i = S$ for all $1 \leqslant i \leqslant m$ under Assumption (B)) . At each (discrete) time step n, we draw m balls in the urn *with*- or *without*-replacement, and denote by ζ_n the number of white balls among those m. Conditionally on $\zeta_n = k$, we then add into the urn a_{m-k} white balls and b_{m-k} black balls (we keep the notations of [KM17, KS16]).

Note that, we these notations, in the *with*-replacement case,

$$\mathbb{P}(\zeta_{n+1} = k | \mathcal{F}_n) = \binom{m}{k} X_n^k (1 - X_n)^{m-k};$$

and in the *without*-replacement case,

$$\mathbb{P}(\zeta_{n+1} = k | \mathcal{F}_n) = \binom{m}{k} \frac{(W_n)_k (T_n - W_n)_{m-k}}{(T_n)_m},$$

where $(n)_q = n(n-1)\cdots(n-q+1)$ for all integers q and n.

The following result is a corollary of Theorem 1.

Corollary 1 (Two-colour balanced urns)**.** *Assume $d = 2$,* (B) *and* (T) *and let*

$$g(x) := \sum_{k=0}^{m} \binom{m}{k} x^k (1-x)^{m-k} (a_{m-k} - Sx).$$

(a) If $g \not\equiv 0$, then there exists a random variable $\theta_\star \in [0,1]$ such that the proportion of white balls in the urn at time n, denoted by X_n, converges almost surely to $\theta_\star$. Furthermore, we have $g(\theta_\star) = 0$ and $g'(\theta_\star) \leqslant 0$ almost surely.

(b) Furthermore, conditionally on $\theta_\star$, the following holds with $\Lambda = -{}^{g'(\theta_\star)}/_S$ and

$$\Gamma = \frac{1}{S^2} \sum_{k=0}^{m} \binom{m}{k} \theta_\star^k (1-\theta_\star)^{m-k} (a_{m-k} - S\theta_\star)^2 :$$

- *If $\Lambda > {}^1/_2$, then $\sqrt{n}(X_n - \theta_\star) \to \mathcal{N}\big(0, \frac{\Gamma}{2\Lambda - 1}\big)$ in distribution when $n \to \infty$.*
- *If $\Lambda = {}^1/_2$, then $\sqrt{{}^n/_{\log n}}(X_n - \theta_\star) \to \mathcal{N}(0, \Gamma)$ in distribution when $n \to \infty$.*
- *If $\Lambda < {}^1/_2$, then $n^\Lambda (X_n - \theta_\star)$ converges almost surely to a finite random variable.*

In the two–colour case, we are also able to get some partial results about the non-balanced case. The following result is not a corollary of Theorem 1, which only applies to balanced urn schemes. Note that in the non-balance case, we need a stronger assumption than the tenability assumption: we are only able to study urns under the assumption that the total number of balls in the urn grows linearly in n, i.e. $\liminf_{n\to\infty} {}^{T_n}/_n > 0$. A way to ensure that this is true is for example to assume that $\min_{0 \leqslant k \leqslant m}(a_k + b_k) \geqslant 1$.

Theorem 3 (Non-balanced two-colour case)**.** *Assume that* $\liminf_{n\to\infty} T_n/n > 0$ *and let*

$$\tilde{g}(x) := \sum_{k=0}^{m} \binom{m}{k} x^k (1-x)^{m-k} \big((1-x)a_{m-k} - xb_{m-k}\big).$$

(a) If $\tilde{g} \not\equiv 0$, then there exists a random variable $\theta_\star \in [0,1]$ such that the proportion of white balls in the urn, denoted by X_n, converges almost surely to $\theta_\star$. Furthermore, we have $\tilde{g}(\theta_\star) = 0$ and $\tilde{g}'(\theta_\star) \leqslant 0$ almost surely.

(b) Conditionally on $\theta_\star$, let

$$\omega := \sum_{k=0}^{m} \binom{m}{k} \theta_\star^k (1-\theta_\star)^{m-k} c_{m-k},$$

where $c_i := a_i + b_i$ for all $0 \leqslant i \leqslant m$,

$$H(x) := \sum_{k=0}^{m} \binom{m}{k} x^k (1-x)^{m-k} \big((1-x)a_{m-k} - xb_{m-k}\big)^2,$$

and

$$\lambda := \frac{|h'(\theta_\star)|}{\omega} \quad \text{and} \quad \sigma^2 := \frac{H(\theta_\star)}{\omega^2}.$$

Assume that $\sigma^2 > 0$. Then, if $\lambda > 1/2$,

$$\sqrt{n}(Z_n - \theta_\star) \to \mathcal{N}\Big(0, \frac{\sigma^2}{2\lambda - 1}\Big) \quad \text{in distribution when } n \to \infty.$$

Remark: Note that in both Corollary 1 and Theorem 3, Statement (a) gives almost sure convergence to a random variable $\theta_\star$ which belongs to the set of zeros of the function g (resp. $\tilde{g}$). The fact that g (resp. $\tilde{g}$) is a non-zero polynomial function ensures that this set is a set of at most m isolated points of $[0,1]$. If g admits a unique zero θ_0 in $[0,1]$ such that $g'(\theta_0) \leqslant 0$, then the proportion of white balls converges almost surely to this zero independently of the initial composition of the urn; we give examples in Section 4. This particular case is what we called the *irreducible* case when discussing Theorem 1(a). If g admits at least two such zeros, then the almost sure limit of X_n depends on the initial composition X_0; we refer the reader to Example 4.1.5 for such an example.

1.7 Plan of the article

The end of this introduction section (see Section 1.8) is devoted to stating an algebraic necessary and sufficient condition for the urn to be tenable: it is a straightforward generalisation of the two-colour case studied by Kuba and Sulzbach [KS16, Lemma 1]. Section 2 contains the proofs of Theorems 1 and 2. Section 3 treats the two-colour balanced case and contains the proof of Corollary 1. We give in Section 4 numerous two-colour and three-colour examples and show how to apply our main results to them. Finally, Section 5 treats the two-colour non-balanced case and contains the proof of Theorem 3 as well as some examples.

Acknowledgements – We are very grateful to Rafik Aguech for making this collaboration possible, and for carefully proof-reading the manuscript. A preliminary version of Theorem 1(a) was inaccurate in a previous version of this article; we warmly thank Benoît Laslier and Jean-François Laslier for noticing it. CM would also like to thank the EPSRC for support through the grant EP/K016075/1.

1.8 Tenability

We here state a lemma giving necessary and sufficient conditions for tenability. This lemma and its proof are very similar to those stated in Kuba & Sulzbach's article [KS16] in the case of two-colour urns. The proof in the d-colour case would be very similar to the $d = 2$ case and would not give additional insight. We therefore do not detail it.

Lemma 1. *Consider the urn process with initial composition U_0 and replacement function R. For all $1 \leqslant i \leqslant d$, we denote by ν_i the greatest common divisor of $\{R_i(v) \colon v \in \Sigma_m^{(d)} \backslash \{m\boldsymbol{e}_i\}\}$. For all integers ℓ, we denote by $[\ell]_i$ the remainder of division of ℓ by ν_i.*

- *In the with-replacement case, the urn scheme is tenable if and only if, for all $1 \leqslant i \leqslant d$, $R_i(v) \geqslant 0$ for all $v \neq m\boldsymbol{e}_i$, and $R_i(m\boldsymbol{e}_i) \geqslant 0$ or $-R_i(m\boldsymbol{e}_i)$ is a divisor of $U_{0,i}$ and ν_i.*

- *In the without-replacement case, the urn scheme is tenable if and only if, for all $1 \leqslant i \leqslant d$, $R_i(v) \geqslant -v_i$ for all $v \neq m\boldsymbol{e}_i$ and*

$$R_i(m\boldsymbol{e}_i) \in [-m, \infty] \cup \Big([-m-\nu_i+1, -m) \cap \Big\{\ell \in -\mathbb{N} \colon [U_{0,i}]_i \in \big\{[-\ell]_i, [-\ell+1]_i, \dots, [m+\nu_i-1]_i\big\}\Big\}\Big).$$

2 Law of large numbers and central limit theorem for balanced urn schemes

2.1 This model can be described as a stochastic algorithm

Lemma 2. *Under Assumption* (B) *and* (T)*, the renormalised composition vector of the urn at time n, denoted by Z_n, satisfies the following recursion:*

$$Z_{n+1} = Z_n + \frac{1}{T_{n+1}}\big(h(Z_n) + \Delta M_{n+1} + \varepsilon_{n+1}\big), \tag{1}$$

where

$$h(x) = \sum_{v \in \Sigma_m^{(d)}} \binom{m}{v_1, \dots, v_d}\left(\prod_{i=1}^{d} x_i^{v_i}\right)\big(R(v) - r(v)\,x\big) \quad \text{and} \quad \Delta M_{n+1} = Y_{n+1} - \mathbb{E}[Y_{n+1}|\mathcal{F}_n]$$

with

$$Y_{n+1} = R(\xi_{n+1}) - r(\xi_{n+1})Z_n,$$

where ξ_{n+1} stands for the (random) vector of balls drawn at time $n+1$, and finally $\varepsilon_{n+1} = 0$ in the with-*replacement case and ε_{n+1} is a $\mathcal{F}_{n+1}$-adapted term satisfying $\varepsilon_n \to 0$ almost surely when n tends to infinity in the* without-*replacement case.*

Proof. Recall that $U_{n+1} = U_n + R(\xi_{n+1})$ and $T_{n+1} = T_n + r(\xi_{n+1})$, implying that

$$\begin{aligned} Z_{n+1} - Z_n &= \frac{U_n + R(\xi_{n+1})}{T_{n+1}} - \frac{U_n}{T_n} = \frac{1}{T_{n+1}}\left(U_n + R(\xi_{n+1}) - \frac{T_n + r(\xi_{n+1})}{T_n}\,U_n\right) \\ &= \frac{1}{T_{n+1}}\,\big(R(\xi_{n+1}) - r(\xi_{n+1})\,Z_n\big) = \frac{1}{T_{n+1}}\big(\mathbb{E}[Y_{n+1}|\mathcal{F}_n] + \Delta M_{n+1}\big). \end{aligned}$$

Note that, in the *with*-replacement case,

$$\mathbb{E}[Y_{n+1}|\mathcal{F}_n] = \sum_{v \in \Sigma_m^{(d)}} P_n(v)(R(v) - r(v)\,Z_n) = h(Z_n),$$

since $P_n(v) = \binom{m}{v_1,\dots,v_d}\prod_{i=1}^d Z_{n,i}^{v_i}$, which concludes the proof in the *with*-replacement case. In the *without*-replacement case,

$$\mathbb{E}[Y_{n+1}|\mathcal{F}_n] = \sum_{v \in \Sigma_m^{(d)}} \frac{\prod_{i=1}^d \binom{U_{n,i}}{v_i}}{\binom{T_n}{m}}\big(R(v) - r(v)\,Z_n\big) = h(Z_n) + \varepsilon_{n+1},$$

with

$$\begin{aligned} \varepsilon_{n+1} &:= \sum_{v \in \Sigma_m^{(d)}} \left(\frac{\prod_{i=1}^d \binom{U_{n,i}}{v_i}}{\binom{T_n}{m}} - \binom{m}{v_1, \dots, v_d}\prod_{i=1}^{d} Z_{n,i}^{v_i}\right)\big(R(v) - r(v)Z_n\big) \\ &= \sum_{v \in \Sigma_m^{(d)}} \binom{m}{v_1, \dots, v_d}\big(R(v) - r(v)Z_n\big)\left(\frac{\prod_{i=1}^d \prod_{j=0}^{v_i-1}(Z_{n,i} - {}^{j}\!/_{T_n})}{\prod_{j=0}^{m-1}(1 - {}^{j}\!/_{T_n})} - \prod_{i=1}^{d} Z_{n,i}^{v_i}\right). \end{aligned} \tag{2}$$

For any two sequences of real-valued random variables $(A_n)_{n\geqslant 0}$ and $(B_n)_{n\geqslant 0}$, let us denote $A_n = \mathcal{O}(B_n)$ if and only if there exists a deterministic constant M such that, for all $n \geqslant 1$, $|A_n| \leqslant MB_n$ almost surely. Using the fact that $Z_{n,i} \leqslant 1$ almost surely for all $1 \leqslant i \leqslant d$ and $n \geqslant 0$, and that $T_n = T_0 + nS \to \infty$ when n goes to infinity, we have, for all $v \in \Sigma_m^{(d)}$,

$$\frac{\prod_{i=1}^d \prod_{j=0}^{v_i-1}(Z_{n,i} - j/T_n)}{\prod_{j=0}^{m-1}(1 - j/T_n)} = \prod_{i=1}^d Z_{n,i}^{v_i} + \mathcal{O}(1/T_n),$$

where the deterministic constant in the $\mathcal{O}$-term depends on v; since there is a finite number of elements in $\Sigma_m^{(d)}$, we can take the maximum of these v-dependent constants so that the constant in the $\mathcal{O}$-term above does not depend on v. We thus have, using in particular the fact that $\|Z_n\|$ is bounded by 1 since $Z_n \in \Sigma^{(d)}$ (we use the L^2-norm, i.e. for all $u \in \mathbb{R}^d$, $\|u\|^2 := \sum_{i=1}^d u_i^2$),

$$\begin{aligned}\|\varepsilon_{n+1}\| &\leqslant \sum_{v\in\Sigma_m^{(d)}} \binom{m}{v_1,\dots,v_d} \big(\|R(v)\| + |r(v)|\,\|Z_n\|\big) \left|\prod_{i=1}^d Z_{n,i}^{v_i} + \mathcal{O}(1/T_n) - \prod_{i=1}^d Z_{n,i}^{v_i}\right| \\ &\leqslant d^m \max_{v\in\Sigma_m^{(d)}} \big(\|R(v)\| + |r(v)|\big)\,\mathcal{O}(1/T_n) = \mathcal{O}(1/T_n), \qquad (3)\end{aligned}$$

almost surely when n goes to infinity. We get, in particular, that $\varepsilon_{n+1} \to 0$ almost surely when n goes to infinity, which concludes the proof in the without-replacement case. □

2.2 Strong law of large numbers (Proof of Theorem 1(a))

As mentioned in Pemantle's survey [Pem07], there are few results for the almost sure convergence of d-dimensional ($d \geqslant 2$) stochastic algorithms. Section 2.5 of [Pem07] is dedicated to such higher-dimension stochastic algorithms. Different versions of the following theorem (originally due to Robbins & Monro [RM51]) exist in the literature; the version we state here is a combination of Duflo [Duf97, Theorem 1.4.26] and Laruelle & Pagès [LP13, Appendix Theorem A.1], which are both stated in $d \geqslant 2$ dimension. We also refer the reader to [Pem07, Corollary 2.15], also stated in the multi-dimensional case.

Theorem 4. *Consider the sequence of random vectors $(\theta_n)_{n\geqslant 0}$ defined by the following recursion:*

$$\forall n \geqslant n_0; \quad \theta_{n+1} = \theta_n + \gamma_{n+1} f(\theta_n) + \gamma_{n+1}(\Delta\hat{M}_{n+1} + \hat{\varepsilon}_{n+1}), \qquad (4)$$

where $f : \mathbb{R}^d \to \mathbb{R}^d$ is a differentiable non-null function, θ_0 is a deterministic vector, for all $n \geqslant n_0$, $\Delta\hat{M}_n$ is an $\mathcal{F}_n$-increment martingale and $\hat{\varepsilon}_n$ is an $\mathcal{F}_n$-adapted remainder term.

Assume in addition that

$$\hat{\varepsilon}_n \xrightarrow{a.s.} 0, \quad \sup_{n\geqslant 0} \mathbb{E}[\|\Delta\hat{M}_{n+1}\|^2|\mathcal{F}_n] < +\infty,$$

and that $(\gamma_n)_{n\geqslant 1}$ is a sequence of positive real numbers satisfying

$$\sum_{n\geqslant 0} \gamma_n = \infty \quad \text{and} \quad \sum_{n\geqslant 0} \gamma_n^2 < +\infty.$$

Then the limit set of $(\theta_n)_{n\geqslant 0}$ (see Definition 1) is almost surely a compact connected set stable by the flow of the differential equation $\dot{x} = f(x)$.

If, in addition, f is uniformly bounded, and admits a zero θ such that, for all n, $\langle f(\theta_n), \theta_n - \theta\rangle < 0$, then θ_n converges almost surely to θ.

This theorem applies to our particular case, in the balanced case only (since otherwise, the sequence $(\gamma_n)_n$ is random and not deterministic), and gives the proof of Theorem 1(a):

Proof of Theorem 1(a). In view of Lemma 2, and since we assume that the urn is balanced, the vector Z_n satisfies the recursion (4) with $\gamma_n = \frac{1}{T_n} = \frac{1}{T_0+nS}$, and $f = h$ a differentiable function, $\Delta\hat{M}_{n+1} = \Delta M_{n+1}$ a $\mathcal{F}_n$-martingale difference and $\hat{\varepsilon}_{n+1} = \varepsilon_{n+1}$ a remainder term tending to zero when n tends to infinity.

It only remains to check that $\sup_{n\geqslant 0}\mathbb{E}[\|\Delta M_{n+1}\|^2|\mathcal{F}_n] < +\infty$. Recall that $\Delta M_{n+1} = Y_{n+1} - \mathbb{E}[Y_{n+1}|\mathcal{F}_n]$, where $Y_{n+1} = R(\xi_{n+1}) - SZ_n$. If we let $\|R\| := \max_{v\in\Sigma_m^{(d)}}\|R(v)\|$, then, for all $n \geqslant 0$, $\|Y_{n+1}\| \leqslant \|R\| + S$, implying that $\|\Delta M_{n+1}\| \leqslant 2(\|R\| + S)$ almost surely. In particular, we have that $\sup_{n\geqslant 0}\mathbb{E}[\|\Delta M_{n+1}\|^2|\mathcal{F}_n] < +\infty$.

Finally, note that $f = h$ is uniformly bounded by $\|R\| + S$. We can thus apply Theorem 4, which concludes the proof. $\square$

2.3 Central limit theorem (Proof of Theorem 1(b))

To prove Theorem 1(b), the idea is again to apply standard theorems from the stochastic algorithms literature. We state here a weak version of a result by Zhang [Zha16]. We also refer the reader to Laruelle & Pagès [LP13, Appendix Theorem A.2][1].

Theorem 5. *Assume that θ_n satisfies the recursion (4) with $\gamma_n = 1/n$ and that there exists $\theta \in \mathbb{R}^d$ a stable zero of f such that θ_n converges to θ with positive probability. Also assume that, for some $\delta > 0$,*

$$\sup_{n\geqslant 0}\mathbb{E}(\|\Delta\hat{M}_{n+1}\|^{2+\delta}|\mathcal{F}_n) < +\infty, \quad \text{and} \quad \mathbb{E}(\Delta\hat{M}_{n+1}\Delta\hat{M}_{n+1}^t|\mathcal{F}_n) \xrightarrow{n\to\infty} \hat{\Gamma} \quad \text{almost surely,}$$

where $\hat{\Gamma}$ is a deterministic symmetric positive semi-definite matrix and for some $\eta > 0$

$$n^{3/2}\,\mathbb{E}\big[\|\hat{\varepsilon}_{n+1}\|^2 \mathbb{1}_{\|\theta_n-\theta\|\leqslant\eta}|\mathcal{F}_n\big] \xrightarrow{n\to\infty} 0. \tag{5}$$

Let $\hat{\Lambda}$ be the eigenvalue of $-\nabla f(\theta)$ with the largest real part. If we assume that θ_n converges almost surely to some deterministic limit θ, then:

- *If $\mathtt{Re}(\hat{\Lambda}) > 1/2$, then $\sqrt{n}(\theta_n - \theta) \xrightarrow{n\to\infty} \mathcal{N}(0,\hat{\Sigma})$ in distribution, where*

$$\hat{\Sigma} = \int_0^{+\infty}\Big(\mathrm{e}^{(\nabla f(\theta)+\frac{\mathtt{Id}}{2})u}\Big)^t\hat{\Gamma}\,\mathrm{e}^{(\nabla f(\theta)+\frac{\mathtt{Id}}{2})u}\,du.$$

Assume additionally that f is twice differentiable, and that all Jordan blocks of $\nabla f(\theta)$ associated to $\hat{\Lambda}$ have size 1. Then:

- *If $\mathtt{Re}(\hat{\Lambda}) = 1/2$, then $\sqrt{\dfrac{n}{\log n}}(\theta_n - \theta) \xrightarrow{n\to\infty} \mathcal{N}(0,\hat{\Sigma})$, in distribution, where*

$$\hat{\Sigma} = \lim_{n\to\infty}\frac{1}{\log n}\int_0^{\log n}\Big(\mathrm{e}^{(\nabla f(\theta)+\frac{\mathtt{Id}}{2})u}\Big)^t\hat{\Gamma}\,\mathrm{e}^{(\nabla f(\theta)+\frac{\mathtt{Id}}{2})u}\,du.$$

- *If $\mathtt{Re}(\hat{\Lambda}) < 1/2$, then $n^{\mathtt{Re}(\hat{\Lambda})}(\theta_n - \theta)$ converges almost surely to a finite random variable.*

Remark: Note that Assumption (5) is not stated as such in [LP13] and [Zha16], in which the assumption depends on the values of $\mathtt{Re}(\hat{\Lambda})$. One can check that (5) is stronger than the assumptions of [LP13] and [Zha16], and since it holds in our particular case, we only state this weaker version.

Applying this result to our framework gives the proof of the second claim of Theorem 1:

Proof of Theorem 1(b). Recall that (see Lemma 2 for details)

$$Z_{n+1} - Z_n = \frac{1}{T_{n+1}}\big(h(Z_n) + \Delta M_{n+1} + \varepsilon_{n+1}\big).$$

We also have $T_n = T_0 + nS$ (by Assumption ($\mathtt{B}$)), which gives

$$\begin{aligned} Z_{n+1} - Z_n &= \frac{1}{(n+1)S}\big(h(Z_n) + \Delta M_{n+1} + \varepsilon_{n+1}\big) + \frac{1}{(n+1)S}\left(\frac{(n+1)S}{T_0 + (n+1)S} - 1\right)\big(h(Z_n) + \Delta M_{n+1} + \varepsilon_{n+1}\big) \\ &= \frac{1}{n+1}\big(f(Z_n) + \Delta\hat{M}_{n+1} + \hat{\varepsilon}_{n+1}\big), \end{aligned}$$

[1] There are different versions of this paper. We refer the reader to the ArXiV version `ArXiV:1101.2786`, which has been regularly updated by the authors.

where $f = {}^{h}/_{S}$ (note that this function is infinitely differentiable), $\Delta\hat{M}_{n+1} = {}^{\Delta M_{n+1}}/_{S}$ and

$$\hat{\varepsilon}_{n+1} = \frac{\varepsilon_{n+1}}{S} + \left(\frac{1}{1+\frac{T_0}{(n+1)S}} - 1\right)\big(f(Z_n) + \Delta\hat{M}_{n+1} + {}^{\varepsilon_{n+1}}/_{S}\big).$$

This last equality implies that, almost surely when n tends to infinity,

$$\|\hat{\varepsilon}_{n+1}\| \leqslant \frac{\|\varepsilon_{n+1}\|}{S} + \mathcal{O}({}^{1}/_{n}).$$

Recall that in the *with*-replacement case, $\varepsilon_n = 0$ for all integers n. In the *without*-replacement case, we have already proved that $\|\varepsilon_n\| = \mathcal{O}({}^{1}/_{T_n}) = \mathcal{O}({}^{1}/_{n})$ (see Equation (3)), implying that $\|\hat{\varepsilon}_n\| = \mathcal{O}({}^{1}/_{n})$, and

$$n^{3/2}\,\mathbb{E}[\|\hat{\varepsilon}_n\|^2|\mathbb{1}_{|\theta_n-\theta|\leqslant\eta}|\mathcal{F}_{n-1}] = \mathcal{O}(n^{-1/2}) \to 0,$$

almost surely when n tends to infinity, for all $\eta > 0$.

We already mentioned that ΔM_{n+1} is almost surely bounded by $2(\|R\|+S)$, and thus $\sup_{n\geqslant 0}\mathbb{E}[\|\hat{\Delta}M_{n+1}\|^{2+\delta}|\mathcal{F}_n] < +\infty$ almost surely for all $\delta > 0$.

Finally, note that

$$\begin{aligned}
&\mathbb{E}[\Delta M_{n+1}\Delta M_{n+1}^t|\mathcal{F}_n]\\
&\qquad = \mathbb{E}[Y_{n+1}Y_{n+1}^t|\mathcal{F}_n] - \mathbb{E}[Y_{n+1}|\mathcal{F}_n]\mathbb{E}[Y_{n+1}^t|\mathcal{F}_n]\\
&\qquad = \sum_{v\in\Sigma_m^{(d)}} P_n(v)(R(v)-SZ_n)(R(v)-SZ_n)^t - \sum_{v,w\in\Sigma_m^{(d)}} P_n(v)P_n(w)(R(v)-SZ_n)(R(w)-SZ_n)^t.
\end{aligned}$$

Since $Z_n \to \theta$ almost surely, we have, in both the with- and the without-replacement cases that $P_n(v) \to \binom{m}{v_1,\dots,v_d}\prod_{i=1}^d \theta_i^{v_i}$ almost surely, implying that

$$\mathbb{E}[\Delta\hat{M}_{n+1}\Delta\hat{M}_{n+1}^t|\mathcal{F}_n] = \frac{1}{S^2}\,\mathbb{E}[\Delta M_{n+1}\Delta M_{n+1}^t|\mathcal{F}_n] \xrightarrow{a.s.} \Gamma,$$

where

$$\begin{aligned}
S^2\,\Gamma &= \sum_{v\in\Sigma_m^{(d)}} \binom{m}{v_1,\dots,v_d}\left(\prod_{i=1}^d \theta_i^{v_i}\right)(R(v)-S\theta)(R(v)-S\theta)^t\\
&\quad - \sum_{v,w\in\Sigma_m^{(d)}} \binom{m}{v_1,\dots,v_d}\binom{m}{w_1,\dots,w_d}\left(\prod_{i=1}^d \theta_i^{v_i+w_i}\right)(R(v)-S\theta)(R(w)-S\theta)^t\\
&= \sum_{v\in\Sigma_m^{(d)}} \binom{m}{v_1,\dots,v_d}\left(\prod_{i=1}^d \theta_i^{v_i}\right)(R(v)-S\theta)(R(v)-S\theta)^t - h(\theta)h(\theta)^t\\
&= \sum_{v\in\Sigma_m^{(d)}} \binom{m}{v_1,\dots,v_d}\left(\prod_{i=1}^d \theta_i^{v_i}\right)(R(v)-S\theta)(R(v)-S\theta)^t,
\end{aligned}$$

since $h(\theta) = 0$. Note that Γ is the limit of a sequence of symmetric positive, semi-definite matrices, and, as such, it is also symmetric positive and semi-definite. Therefore, Theorem 5 applies, which concludes the proof. $\square$

Remark: The result obtained in [Zha16] also permits to treat the cases $\mathrm{Re}(\hat{\Lambda}) \leqslant {}^{S}/_{2}$ when the Jordan blocks associated $\hat{\Lambda}$ are not all of size one. The asymptotic renormalisation then depends on the size of the largest blocks of $\nabla f(\theta)$ associated to $\hat{\Lambda}$. Applying this stronger version to our framework would give a generalisation of Theorem 1 to these cases. Since this generalisation is quite technical to state and since most of the examples fall under Theorem 1 as it is, we do not give more details.

2.4 The diagonal case (proof of Theorem 2)

Before proving Theorem 2, we give a simple characterisation of the diagonal balanced urn schemes:

Lemma 3. *Under assumption* (B)*, the function h is identically null on $\Sigma^{(d)}$ if, and only if, there exists an integer σ such that $S = m\sigma$ and*

$$R(v) = \sigma v \quad \text{for all} \quad v \in \Sigma_m^{(d)}. \tag{6}$$

Proof. First note that, straightforwardly, the fact that there exists an integer σ such that $S = m\sigma$ and such that Equation (6) holds implies that $h \equiv 0$.

The reverse implication is less straightforward. First note that, for all $1 \leqslant i \leqslant d$, $h(\boldsymbol{e}_i) = R(m\boldsymbol{e}_i) - S\boldsymbol{e}_i = 0$, implying that $R(m\boldsymbol{e}_i) = S\boldsymbol{e}_i$ for all $1 \leqslant i \leqslant d$.

We reason by induction and prove that there exists σ such that $R(v) = \sigma v$ for all vectors $v \in \Sigma_m^{(d)}$ having at most k non-null coordinates. Note that if this result is true, it implies that $\sigma = {}^{S}\!/\!{}_{m}$ and thus that m divides S (take $v = \boldsymbol{e}_1 + (m-1)\boldsymbol{e}_2$ and recall that $R_1(v) = \sigma$ is an integer).

Assume that this is true for some integer $k < m$, $R(v) = \sigma v$ for all vector $v \in \Sigma_m^{(d)}$ having at most k non-null coordinates. Let us prove that this property extends to vectors with at most $k+1$ non-zero coordinates. Without loss of generality, we can focus on vectors $v \in \Sigma_m^{(d)}$ such that the last $d-k-1$ coordinates are equal to zero, which is actually the set $\Sigma_m^{(k+1)}$ (with a slight abuse of notation since these vectors are still d-dimensional). Let $x = x_1\boldsymbol{e}_1 + \ldots + x_{k+1}\boldsymbol{e}_{k+1} \in \Sigma^{(d)}$, we have, by assumption,

$$\sum_{v\in\Sigma_m^{(k+1)}} \binom{m}{v_1,\ldots,v_{k+1}} \left(\prod_{i=1}^{k+1} x_i^{v_i}\right) R(v) = Sx. \tag{7}$$

First note that for all $k+2 \leqslant j \leqslant d$, we have

$$\sum_{v\in\Sigma_m^{(k+1)}} \binom{m}{v_1,\ldots,v_{k+1}} \left(\prod_{i=1}^{k+1} x_i^{v_i}\right) R_j(v) = 0.$$

Note that the tenability assumption implies that $R_j(v) \geqslant 0$ for all $v \notin \{m\boldsymbol{e}_1, \ldots, m\boldsymbol{e}_d\}$ (see Lemma 1). Moreover, we have proved that for all $1 \leqslant i \leqslant d$, $R(m\boldsymbol{e}_i) = S\boldsymbol{e}_i$, implying in particular that $R_j(m\boldsymbol{e}_i) \geqslant 0$ for all $1 \leqslant j \leqslant d$. Therefore, the sum in the above display is a sum of non-negative terms. The fact that it is equal to zero thus implies that all its terms are null, and thus that $R_j(v) = 0$ $(= \sigma v_j)$, for all $k+2 \leqslant j \leqslant d$. It remains to prove that the same equality, namely $R_j(v) = \sigma v_j$ holds for all $1 \leqslant j \leqslant k+1$.

Note that

$$\sum_{v\in\Sigma_m^{(k+1)}} \binom{m}{v_1,\ldots,v_{k+1}} \left(\prod_{i=1}^{k+1} x_i^{v_i}\right) (R(v)-\sigma v) = \sum_{v\in\Sigma_m^{(k+1)}} \binom{m}{v_1,\ldots,v_{k+1}} \left(\prod_{i=1}^{k+1} x_i^{v_i}\mathbb{1}_{v_i\neq 0}\right) \big(R(v)-\sigma v\big),$$

since all $v \in \Sigma_m^{(k+1)}$ having at least one null coordinate satisfy $R(v) = \sigma v$ by the induction hypothesis. Therefore, Equation (7) implies that

$$Sx = \sum_{v\in\Sigma_m^{(k+1)}} \binom{m}{v_1,\ldots,v_{k+1}} \left(\prod_{i=1}^{k+1} x_i^{v_i}\,\mathbb{1}_{v_i\neq 0}\right) (R(v)-\sigma v) + \sigma \sum_{v\in\Sigma_m^{(k+1)}} \binom{m}{v_1,\ldots,v_{k+1}} \left(\prod_{i=1}^{k+1} x_i^{v_i}\right) v. \tag{8}$$

Note that, for all $1 \leqslant j \leqslant k+1$,

$$x_j \frac{\partial}{\partial x_j}\left(x_1+\cdots+x_{k+1}\right)^m = x_j \frac{\partial}{\partial x_j}\left(\sum_{v\in\Sigma_m^{(k+1)}} \binom{m}{v_1,\ldots,v_{k+1}} \left(\prod_{i=1}^{k+1} x_i^{v_i}\right)\right) = \sum_{v\in\Sigma_m^{(k+1)}} \binom{m}{v_1,\ldots,v_{k+1}} \left(\prod_{i=1}^{k+1} x_i^{v_i}\right) v_j,$$

implying that

$$mx = \sum_{v\in\Sigma_m^{(k+1)}} \binom{m}{v_1,\ldots,v_{k+1}} \left(\prod_{i=1}^{k+1} x_i^{v_i}\right) v_j.$$

Therefore, Equation (8) implies

$$Sx = \sum_{v \in \Sigma_m^{(k+1)}} \binom{m}{v_1, \dots, v_{k+1}} \left(\prod_{i=1}^{k+1} x_i^{v_i} \mathbb{1}_{v_i \neq 0} \right) (R(v) - \sigma v) + \sigma m x,$$

and thus, for all $x = x_1 \boldsymbol{e}_1 + \cdots + x_{k+1} \boldsymbol{e}_{k+1} \in \Sigma^{(d)}$,

$$\sum_{v \in \Sigma_m^{(k+1)}} \binom{m}{v_1, \dots, v_{k+1}} \left(\prod_{i=1}^{k+1} x_i^{v_i} \mathbb{1}_{v_i \neq 0} \right) (R(v) - \sigma v) = 0.$$

Take $x_p = t \in (0,1)$ and $x_i = (1-t)/k$ for all $i \in \{1, \dots, k+1\} \backslash \{p\}$. To all $v \in \Sigma_m^{(k+1)}$, we associate bijectively the couple (v_p, v_p^*), where v_p is the p-th coordinate of v and where $v_p^* \in \Sigma_{m-v_p}^{(k)}$ is the vector whose k coordinates are $v_1, \dots, v_{p-1}, v_{p+1}, \dots, v_{k+1}$. With these notations, we have

$$\begin{aligned} 0 &= \sum_{v \in \Sigma_m^{(k+1)}} \binom{m}{v_1, \dots, v_{k+1}} (t/k)^{v_p} (1-t)^{m-v_p} (R_p(v) - \sigma v_p) \\ &= \sum_{v_p=1}^{m} \sum_{v_p^* \in \Sigma_{m-v_p}^{(k)}} \binom{m}{v_1, \dots, v_{k+1}} (t/k)^{v_p} (R_p(v) - \sigma v_p) \sum_{\ell=0}^{m-v_p} \binom{m-v_p}{\ell} (-t)^{\ell} \\ &= \sum_{u=1}^{m} t^u \sum_{v_p=1}^{u} \sum_{v_p^* \in \Sigma_{m-v_p}^{(k)}} \binom{m}{v_1, \dots, v_{k+1}} \binom{m-v_p}{u-v_p} k^{-v_p} (R_p(v) - \sigma v_p) (-1)^{u-v_p}, \end{aligned}$$

implying that, for all $1 \leqslant u \leqslant m$,

$$\sum_{v_p=1}^{u} \binom{m-v_p}{u-v_p} k^{-v_p} (-1)^{u-v_p} \sum_{v_p^* \in \Sigma_{m-v_p}^{(k)}} \binom{m}{v_1, \dots, v_{k+1}} (R_p(v) - \sigma v_p) = 0.$$

This equation for $u = 1$ gives that $R_p(v) = \sigma$ for all vector v such that $v_p = 1$. Using the above equation for $u = 2$, one can then induce that $R_p(v) = 2\sigma$ for all vector v such that $v_p = 2$, and, inductively, prove that $R_p(v) = \sigma v_p$ for all $1 \leqslant p \leqslant k+1$.

In total, for all $v \in \Sigma_m^{(k+1)}$, we have $R(v) = \sigma v$, which concludes the induction argument. ☐

Proof of Theorem 2. We have that (since we assume that the urn is balanced)

$$Z_{n+1} = Z_n + \frac{1}{T_0 + (n+1)S} \big(\Delta M_{n+1} + \varepsilon_{n+1} \big), \tag{9}$$

where we recall that ΔM_{n+1} is the increment of a martingale and $\|\varepsilon_{n+1}\| = \mathcal{O}(1/n)$ almost surely when n tends to infinity. We infer that, for all $1 \leqslant k \leqslant d$,

$$Z_{n,k} = Z_{0,k} + \sum_{i=0}^{n-1} \frac{\Delta M_{i+1,k}}{T_0 + (i+1)S} + \sum_{i=0}^{n-1} \frac{\varepsilon_{i+1,k}}{T_0 + (i+1)S}. \tag{10}$$

Note that the last sum of this last equation converges almost surely since either $\varepsilon_i = 0$ for all i, or $|\varepsilon_{i,k}| \leqslant \|\varepsilon_i\| = \mathcal{O}(1/i)$ almost surely. The first sum in Equation (10) is a martingale and its quadratic variation is given by

$$\texttt{Hook}_n := \sum_{i=0}^{n-1} \frac{\mathbb{E}[\Delta M_{i+1,k}^2 | \mathcal{F}_i]}{(T_0 + (i+1)S)^2}.$$

Recall that $\Delta M_{i+1,k}^2$ is almost surely bounded by $4(\|R\| + S)^2$ implying that $\texttt{Hook}_n$ is almost surely convergent when n goes to infinity. Therefore, the martingale itself, i.e. the first sum in Equation (10) converges almost surely to a finite random variable. In total, Z_n converges almost surely to a random vector Z_∞. ☐

3 The two-colour particular case (proof of Corollary 1)

In the two-colour balanced case, the renormalised composition vector Z_n of the urn at time n is actually characterised by its first coordinate since $Z_{n,1} + Z_{n,2} = 1$. Thus, it is enough to study this first coordinate, and the problem becomes unidimensional and thus slightly easier. Although Corollary 1 could be deduced from Theorem 1, we give here a stand-alone and much simpler proof: we apply directly a result about one-dimensional stochastic algorithms (see [Pem07, Corollary 2.7 and Theorem 2.9]). Let

$$g(x) := \sum_{k=0}^{m} \binom{m}{k} x^k (1-x)^{m-k} (a_{m-k} - Sx),$$

then

$$X_{n+1} - X_n = \frac{1}{n} \left(\frac{g(X_n)}{S} + \Delta M_{n+1} + \varepsilon_{n+1} \right),$$

where

$$\Delta M_{n+1} = \frac{1}{S}\big(Y_{n+1} - \mathbb{E}[Y_{n+1}|\mathcal{F}_n]\big),$$

with $Y_{n+1} := a_{m-\zeta_{n+1}} - SZ_n$, and where $\varepsilon_{n+1} = 0$ in the *with*-replacement case and ε_{n+1} satisfies $|\varepsilon_n| = \mathcal{O}(1/n)$ almost surely when n tends to infinity in the *without*-replacement case.

Since $g(x)$ is a polynomial, either it is identically null, or it has isolated zeros. The case $g \equiv 0$ is called the *diagonal* case and is treated separately. If $g \not\equiv 0$, then g has at most m zeros on $[0,1]$. In view of Corollary 2.7 and Theorem 2.9 in [Pem07], we can conclude that X_n converges almost surely to one of the zeros of g where g' is non-positive, depending on the initial composition of the urn. This proves Corollary 1(a).

If we assume that X_n converges almost surely to some stable zero $\theta_\star$ of g, then we can apply Theorem 5. We have

$$\mathbb{E}[(\Delta M_{n+1})^2|\mathcal{F}_n] \to \Gamma := \frac{1}{S^2} \sum_{k=0}^{m} \binom{m}{k} \theta_\star^k (1-\theta_\star)^{m-k} (a_{m-k} - S\theta_\star)^2.$$

Finally let $\Lambda = -g'(\theta_\star)/S$, and the following central limit theorem holds:

- If $\Lambda > 1/2$, then $\sqrt{n}(X_n - \theta_\star) \overset{n\to\infty}{\longrightarrow} \mathcal{N}\big(0, \frac{\Gamma}{2\Lambda-1}\big)$, in distribution.
- If $\Lambda = 1/2$, then $\sqrt{n/\log n}(X_n - \theta_\star) \overset{n\to\infty}{\longrightarrow} \mathcal{N}(0,\Gamma)$, in distribution.
- If $\Lambda < 1/2$, then $n^\Lambda(X_n - \theta_\star)$ converges almost almost surely to a random variable.

This concludes the proof of Corollary 1.

Diagonal case. In the two-colour case, assuming both that the urn is balanced and that $g \equiv 0$ implies that there exists an integer q such that

$$R = \begin{pmatrix} mq & 0 \\ (m-1)q & q \\ \vdots & \vdots \\ 0 & mq \end{pmatrix}.$$

This case is treated by Kuba and Mahmoud [KM17] and Kuba and Sulzbach [KS16].

4 Examples

4.1 Two-colour examples

The first four examples are two-colour examples: we thus apply Corollary 1 to them. We choose to focus on non-affine examples to which the results of Kuba & Mahmoud [KM17] do not apply; note that affine models correspond to the case when g is linear.

Example 4.1.1: Take

$$R = \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

In that case, $g(x) = (1 + x)(1 - 2x)$, whose single zero in $[0, 1]$ is $1/2$. Thus, almost surely when n goes to infinity, the proportion of white balls in the urn at time n converges to $1/2$. In this particular case, $S = 3$, and $\Lambda = -g'(1/2)/S = 1 > 1/2$. One can check that $\Gamma = 1/36$, and since $2\Lambda - 1 = 1$, we get, by Corollary $1(b)$,

$$\sqrt{n}(X_n - 1/2) \to \mathcal{N}(1, 1/36),$$

in distribution when n tends to infinity.

Example 4.1.2: Take

$$R = \begin{pmatrix} 4 & 0 \\ 1 & 3 \\ 1 & 3 \end{pmatrix}.$$

In that case, $g(x) = (1 - x)(1 - 3x)$, whose two zeros in $[0, 1]$ are 1 and $1/3$. Note that $g'(1) = 2$ and $g'(1/3) = -2$, thus, almost surely when n tends to infinity, the proportion of white balls in the urn converges to $1/3$. Note that $\Lambda = -g'(1/3)/S = 1/2$. Also, one can check that $\Gamma = 1/18$, which implies, by Corollary $1(b)$,

$$\sqrt{n/\log n}(X_n - 1/3) \to \mathcal{N}(1, 1/18),$$

in distribution when n tends to infinity.

Example 4.1.3: Take

$$R = \begin{pmatrix} 7 & 1 \\ 3 & 5 \\ 1 & 7 \end{pmatrix}.$$

Then $g(x) = 2x^2 - 4x + 1$ and this polynomial has a unique root in $[0, 1]$, given by $\theta_\star = 1 - \sqrt{2}/2$. Applying Theorem $1(a)$, we get that the proportion of white balls X_n converges almost surely to $1 - \sqrt{2}/2$. Moreover, $\Lambda = -g'(1/2)/8 = \sqrt{2}/4 < 1/2$. Thus, by Corollary $1(b)$, there exists a random variable Ψ such that

$$n^{\sqrt{2}/4}(X_n - 1/2) \to \Psi,$$

almost surely when n tends to infinity.

Example 4.1.4: Take

$$R = \begin{pmatrix} 6 & 0 \\ 3 & 3 \\ 1 & 5 \end{pmatrix}.$$

In this case, $g(x) = (x - 1)^2$, implying that 1 is the unique zero of g in $[0, 1]$. Thus, the proportion of white balls in the urn converges almost surely to 1. We have $g'(1) = 0$, which makes us unable to apply Corollary $1(b)$: we have no information about the speed of convergence of the proportion of white balls to its limit.

Example 4.1.5: Let us consider an example for which $m = 3$:

$$R = \begin{pmatrix} 82 & 9 \\ 91 & 0 \\ 0 & 91 \\ 9 & 82 \end{pmatrix}.$$

In that case, $S = 91$, and $g(x) = -200(x - 1/10)(x - 1/2)(x - 9/10)$ admits three zeros on $[0, 1]$, namely $1/10, 1/2$, and $9/10$. Note that $g'(1/2) > 0$ while $1/10$ and $9/10$ are stable zeros of g. Therefore, the proportion of white balls in the urn converges to some random variable $\theta_\star$, and $\theta_\star \in \{1/10, 9/10\}$ almost surely (see Figure 1). Note that $g'(1/3) = g'(2/3) = -64$ implying that $\Lambda = 64/91 > 1/2$. Therefore, $\sqrt{n}(X_n - \theta_\star)$ converges in distribution to a standard Gaussian.

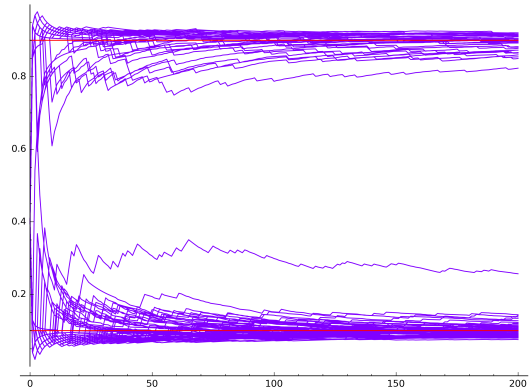

Figure 1: A hundred realisations of the urn process of Example 4.1.5, all starting at $(4, 6)$ and run for 200 steps. One can see that each trajectory converges to one of the two stable zeros of g, namely $1/10$ and $9/10$.

4.2 Three-colour examples

For all three-colour examples, we use a computer algebra system to help with the calculations. Note that it is in practise harder to apply Theorem 1(a) than Theorem 1(b); the strategy is to first find the zeros of the function h, then calculate the Jacobian of h at these zeros, extract their spectrum and isolate the zeros whose Jacobians only have negative eigenvalues. In all the examples below, there is at most one such zero and it is our candidate for the limit of the normalised composition vector Z_n; if there is no such zero, we cannot say more about the considered urn (see example 4.2.5) except that its limit set is a compact connected set stable by the flow of $\dot{x} = h(x)$.

Once we have a candidate θ to be the limit of Z_n, we need to check that $\langle h(Z_n), Z_n - \theta \rangle < 0$ for all n. The way we do it is by proving that $\langle h(x), x - \theta \rangle < 0$ for all $x \in \Sigma^{(d)}$ that is not a zero of h. This is the most intricate part. All the examples below are given for $m = 2$, since it implies that $\langle h(x), x - \theta \rangle$ is a polynomial of order at most 3 and the set where it is negative can be calculated exactly (we do it with a computer algebra system).

Finally, we prove that Z_n never reaches the other zeros of h, which proves convergence to θ by Theorem 1(a). Applying Theorem 1(b) to get the fluctuations is then straightforward.

Example 4.2.1: Let consider the three-colour urn scheme defined by the following replacement rule:

$$\begin{aligned} R(2,0,0) &= (1,0,0); & R(0,1,1) &= (1,0,0);\\ R(0,2,0) &= (0,1,0); & R(1,0,1) &= (0,1,0);\\ R(0,0,2) &= (0,0,1); & R(1,1,0) &= (0,0,1). \end{aligned}$$

In that case,

$$h(x) = \begin{pmatrix} x_1^2 + 2x_2x_3 - x_1 \\ x_2^2 + 2x_1x_3 - x_2 \\ x_3^2 + 2x_2x_3 - x_3 \end{pmatrix}.$$

The function $h(x)$ admits four zeros on the simplex $\Sigma_2^{(3)}$ being given by $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ and $(1/3, 1/3, 1/3)$.

We recall that the function h is a function from $\Sigma^{(d)}$ onto $\{(x_1, x_2, x_3) \in \mathbb{R}^d \colon x_1 + x_2 + x_3 = 0\}$, two spaces of dimension 2, therefore, the Jacobian matrix of h at any point of $\Sigma^{(d)}$ is a 2×2 matrix. The eigenvalues of the Jacobian matrix of h at the first three zeros mentioned above are 1 and -3. The eigenvalues of $\nabla h(1/3, 1/3, 1/3)$ are -1 (with multiplicity 2), thus this zero is stable and is our candidate to be the almost sure limit of Z_n. We actually have $\nabla h(1/3, 1/3, 1/3) = -\mathtt{Id}_2$.

One can check that the only solutions of $\langle h(x), x - \theta \rangle = 0$ on $\Sigma^{(3)}$, where $\theta = (1/3, 1/3, 1/3)$, are the four zeros of h. Since θ is a stable zero of h, $\langle h(x), x - \theta \rangle < 0$ in a neighbourhood of θ, and thus, by continuity, $\langle h(x), x - \theta \rangle < 0$ for all x in $\Sigma^{(3)}$ such that $h(x) \neq 0$.

Finally, note that $Z_0 \neq (1,0,0)$ implies that $Z_n \neq (1,0,0)$ for all $n \geqslant 0$, and similarly for $(0,1,0)$ and $(0,0,1)$. Thus, Theorem 1(a) applies and gives that Z_n converges almost surely to $\theta = (1/3, 1/3, 1/3)$ (if the urn contained balls of more than one colour at time 0).

Let us now apply Theorem 1(b). One can check that

$$\Gamma = \frac{1}{9}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} \tag{11}$$

and

$$\Sigma = \int_0^{+\infty} (\mathtt{e}^{-u/2}\mathtt{Id})^t\,\Gamma(\mathtt{e}^{-u/2}\mathtt{Id})\,du = \Gamma.$$

Finally note that the eigenvalue of $-\nabla h(\theta)$ with the smallest real part is $\Lambda = 1 > S/2 = 1/2$. We have that $2\mathtt{Re}(\Lambda)/S - 1 = 1$, implying that, in distribution when n tends to infinity,

$$\sqrt{n}(Z_n - \mathbf{1}/3) \to \mathcal{N}(0, \Gamma),$$

where $\mathbf{1} = (1,1,1)$ and Γ given by Equation (11).

Since Z_n and $\mathbf{1}/3$ are in the simplex $\Sigma^{(d)}$, we know that $\Sigma = \Gamma$ are of rank at most 2, and that, in particular,

$$\Sigma \cdot (1,1,1)^t = \Gamma \cdot (1,1,1)^t = 0,$$

which can indeed be checked using (11).

Example 4.2.2: Consider the three-colour urn scheme defined by the following replacement rule:

$$\begin{aligned} R(2,0,0) &= (2,0,0); & R(0,1,1) &= (0,1,1); \\ R(0,2,0) &= (1,0,1); & R(1,0,1) &= (0,2,0); \\ R(0,0,2) &= (1,1,0); & R(1,1,0) &= (0,0,2). \end{aligned}$$

In that case,

$$h(x) = \begin{pmatrix} 2x_1^2 + x_2^2 + x_3^3 - 2x_1 \\ x_3(4x_1 + 2x_2 + x_3) - 2x_2 \\ x_2(4x_1 + x_2 + 2x_3) - 2x_3 \end{pmatrix}.$$

One can check that $h(x)$ admits two zeros on the simplex $\Sigma_2^{(3)}$ being given by $(1,0,0)$, $(1/5, 2/5, 2/5)$. The eigenvalues of the Jacobian matrix of h at $(1,0,0)$ are 2 and -6. The eigenvalues of $\nabla h(1/5, 2/5, 2/5)$ are -2 and $-18/5$, thus this zero is stable and is our candidate to be the almost sure limit of Z_n.

Using a computer algebra software, we can also check that the only solutions of $\langle h(x), x - \theta\rangle = 0$ are $(1,0,0)$ and $\theta = (1/5, 2/5, 2/5)$. Since θ is a stable zero, $\langle h(x), x - \theta\rangle$ is negative in a neighbourhood of θ, and, by continuity, it is negative on $\Sigma^{(3)} \setminus \{(1,0,0), \theta\}$.

Note that if $Z_0 \neq (1,0,0)$, then $Z_n \neq (1,0,0)$ for all $n \geqslant 0$. We can thus apply Theorem 1(a) and conclude that, if $Z_0 \neq (1,0,0)$, then Z_n converges almost surely to $\theta = (1/5, 2/5, 2/5)$.

Let us now apply Theorem 1(b) One can calculate that

$$\Gamma = \frac{1}{25}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 3 & -2 \\ -1 & -2 & 3 \end{pmatrix}, \quad \text{and} \quad \Sigma = \frac{1}{25}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 19 & -6 \\ -1 & -6 & 19 \end{pmatrix}. \tag{12}$$

Finally note that the eigenvalue of $-\nabla h(\theta)$ with the smallest real part is $\Lambda := 2 > S/2 = 1$, implying that in distribution when n tends to infinity, since $2\mathtt{Re}(\Lambda)/S - 1 = 1$,

$$\sqrt{n}(Z_n - (1/5, 2/5, 2/5)^t) \to \mathcal{N}(0, \Sigma),$$

where Σ is given by Equation (12). (Again, one can check that, as expected, $\Sigma \cdot (1,1,1)^t = 0$.)

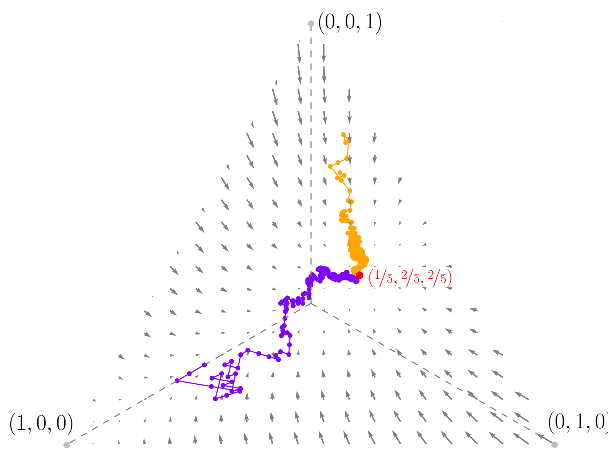

Figure 2: Two realisations (purple and orange) of the urn of Example 4.2.2 starting respectively from $(10,3,3)$ and $(2,6,20)$, and run for 400 steps. The gray arrows are the vector field associated to h on the simplex $\Sigma^{(3)}$; one can in particular see that $(1,0,0)$ is indeed an unstable zero while $(1/5, 2/5, 2/5)$ is stable.

Example 4.2.3: Consider the three-colour urn scheme defined by the following replacement rule:

$$\begin{aligned} R(2,0,0) &= (3,0,0); \qquad & R(0,1,1) &= (3,0,0);\\ R(0,2,0) &= (0,3,0); \qquad & R(1,0,1) &= (1,1,1);\\ R(0,0,2) &= (0,0,3); \qquad & R(1,1,0) &= (1,1,1). \end{aligned}$$

In that case,

$$h(x) = \begin{pmatrix} 3x_1^2 + 2x_1x_2 + 2x_1x_3 + 6x_2x_3 - 3x_1 \\ 3x_2^2 + 2x_1x_2 + 2x_1x_3 - 3x_2 \\ 3x_3^2 + 2x_1x_2 + 2x_1x_3 - 3x_3 \end{pmatrix}$$

one can check that $h(x)$ admits four zeros on the simplex $\Sigma_2^{(3)}$, given by $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ and $(3/5, 1/5, 1/5)$. The eigenvalues of ∇h at the first three zeros are $-2-\sqrt{13}$ and $-2+\sqrt{13}$. The eigenvalues of $\nabla h(3/5, 1/5, 1/5)$ are -1 and $-9/5$. Thus $\theta = (3/5, 1/5, 1/5)$ is a stable zero of h.

One can also check that the only solutions of $\langle h(x), x-\theta\rangle = 0$ on $\Sigma^{(3)}$ are the zeros of h. Using the fact that $\langle h(x), x-\theta\rangle < 0$ on a neighbourhood of θ, thus on the whole domain except at the zeros of h, and using the fact that $Z_0 \neq (1,0,0)$ implies $Z_n \neq (1,0,0)$ for all $n \geqslant 0$ and similarly for $(0,1,0)$ and $(0,0,1)$, we can apply Theorem 1(a) and get that the renormalised composition of the urn Z_n converges almost surely to $\theta = (3/5, 1/5, 1/5)$ (if the urn contained balls of more than one colour at time 0).

We can now apply Theorem 1(b). Note that, in that case, $\Lambda = 1 < S/2 = 3/2$, and the Jordan block of $\nabla h(\theta)$ associated to 1 has size 1 (since ∇h has two distinct eigenvalues given by -1 and $-9/5$), implying that $n^{1/3}(Z_n - (3/5, 1/5, 1/5)^t)$ converges almost surely to a random variable Ψ.

Example 4.2.4: Consider the three colour urn scheme defined by the following replacement rule:

$$\begin{aligned} R(2,0,0) &= (0,0,2); \qquad & R(0,1,1) &= (0,1,1);\\ R(0,2,0) &= (0,0,2); \qquad & R(1,0,1) &= (1,0,1);\\ R(0,0,2) &= (0,0,2); \qquad & R(1,1,0) &= (1,1,0). \end{aligned}$$

We have

$$h(x) = 2\begin{pmatrix} x_1x_2 + x_1x_3 - x_1 \\ x_1x_2 + x_2x_3 - x_2 \\ x_1^2 + x_2^2 + x_3^2 + x_1x_3 + x_2x_3 - x_3 \end{pmatrix}.$$

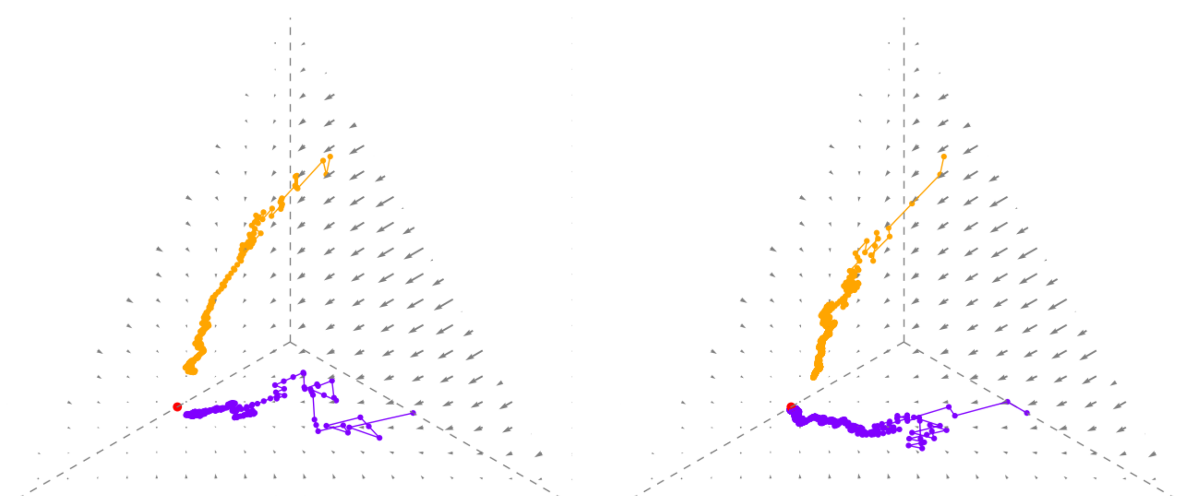

Figure 3: Two realisations of the urn of Example 4.2.3 starting respectively from $(3,10,3)$ and $(2,5,20)$, run for 400 steps (left) and 4000 steps (right). Both realisation converge to the stable zero $(1/5, 2/5, 2/5)$; one can compare the speed of convergence with the faster Example 4.2.1 displayed in Figure 2.

The function h admits a unique zero on the simplex $\Sigma_2^{(3)}$ given by $(0, 0, 1)$. The eigenvalues of $\nabla h(0, 0, 1)$ are 0 (with multiplicity 2).

We can solve explicitly $\langle h(x), x-(0, 0, 1)\rangle$ on $\Sigma^{(3)}$ and check that $(0, 0, 1)$ is the only solution. Also, $\langle h(1, 0, 0), (1, 0, 0)-(0, 0, 1)\rangle = -4$. By continuity, $\langle h(x), x - (0, 0, 1)\rangle < 0$ for all $x \in \Sigma^{(3)}\backslash\{(0, 0, 1)\}$. Thus Theorem 1($a$) applies, and the renormalised composition vector Z_n converges almost surely to $(0, 0, 1)$. Nothing can be said about the rate of convergence since $\nabla h(0, 0, 1)$ admits zero as an eigenvalue (and Theorem 1(b) thus does not apply).

Example 4.2.5: Rock-paper-scissor game [LL13]. Consider the three colour urn scheme defined by the following replacement rule:

$$\begin{aligned} R(2,0,0) &= (1,0,0); & R(0,1,1) &= (0,1,0);\\ R(0,2,0) &= (0,1,0); & R(1,0,1) &= (0,0,1);\\ R(0,0,2) &= (0,0,1); & R(1,1,0) &= (1,0,0). \end{aligned}$$

This urn scheme is studied in Laslier & Laslier [LL13] as a model for the famous *rock-paper-scissors* game; the first colour represents *scissors*, the second *paper* and the third *rock*, and the replacement rule encodes the standard rules of the game (scissors wins against paper, paper wins against rock and rock wins against scissors). We have

$$h(x) = \begin{pmatrix} x_1(x_1 + 2x_2 - 1)\\ x_2(x_2 + 2x_3 - 1)\\ x_3(x_3 + 2x_1 - 1) \end{pmatrix}.$$

The function h admits four zeros, being $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ and $(1/3, 1/3, 1/3)$, but none of them is stable.

In [LL13, Theorem 7] it is proved that the composition vector does not converge when n goes to infinity, but concentrate on a cycle instead. Therefore, our Theorem 1 does not apply further than saying that the limit set of $(Z_n)_{n\geqslant 0}$ is a compact connected set.

The authors also study what the call a *three-alternative* scheme, in which a sample of three balls is sampled at every time step, and the replacement rule depends on the order of sampling of those balls: if one samples *(scissors, rock, paper)*, we get

$$(\textit{scissors} \text{ against } (\textit{rock} \text{ against } \textit{paper})) = \textit{scissors} \text{ against } \textit{paper} = \textit{scissors},$$

while the sample *(paper, scissors, rock)* gives

$$(\textit{paper} \text{ against } (\textit{scissors} \text{ against } \textit{rock})) = \textit{paper} \text{ against } \textit{rock} = \textit{paper}.$$

This rock-paper-scissor game shows that an interesting extension of our model is to sample an ordered set of balls (or rather a sequence of balls) at each step and make the replacement rule depend on this sequence of m randomly sampled balls.

5 The two-colour non-balanced case (proof of Theorem 3)

We have shown in this article how stochastic algorithms can be used as a powerful toolbox to study multi-drawing Pólya urns. Our main achievement is to remove the *affinity* assumption that was previously made in the literature (see [KM17]) and thus considerably increase the number of urn scheme to which the strong law of large numbers and the central limit theorem apply.

The main open question is to generalise these results to the non-balanced case: we show in this section how to prove Theorem 3.

There are two main difficulties occurring when treating the unbalanced case. Firstly, there are only very few results in the literature on stochastic algorithms that allow the sequence $(\gamma_n)_{n\geqslant 0}$ to be random (which is the case when the urn is non-balanced since $\gamma_n = 1/T_n$). Secondly, we need to control the speed of convergence of the total number of balls dividing by n, i.e. T_n/n to its limit.

To prove Theorem 3, we apply the work of Renlund [Ren11]. In this work, a strong law of large numbers and a central limit theorem are proved, but only in the case when the stochastic algorithm $(\theta_n)_{n\geqslant 0}$ is one-dimensional, which is why this result can only be applied to two-colour Pólya urn schemes.

5.1 Estimates for the total number of balls in the urn

Proposition 1. *If X_n converges almost surely to a limit $\theta_\star$, then the sequence ${}^{T_n}\!/_{n}$ converges almost surely to*

$$\omega := \sum_{k=0}^{m} \binom{m}{k} \theta_\star^k (1-\theta_\star)^{m-k}(a_{m-k}+b_{m-k}).$$

Proof. In this proof, we use the notation $\mathbb{E}_i$ for the expectation conditionally on the canonical filtration $\mathcal{F}_i$. Write

$$\frac{T_{n+1}}{n+1} = \frac{T_n}{n+1} + \frac{c_{m-\zeta_{n+1}}}{n+1} = \frac{T_0}{n+1} + \frac{1}{n+1}\sum_{i=0}^{n}(c_{m-\zeta_{i+1}} - \mathbb{E}_i c_{m-\zeta_{i+1}}) + \frac{1}{n+1}\sum_{i=0}^{n}\mathbb{E}_i c_{m-\zeta_{i+1}}, \tag{13}$$

where we recall that ζ_{i+1} is the number of white balls sampled at time $i+1$. Note that

$$\mathbb{E}_i c_{m-\zeta_{i+1}} = \sum_{k=0}^{m}\binom{k}{i} X_i^k (1-X_i)^{m-k} c_{m-k} \overset{i\to\infty}{\longrightarrow} \sum_{k=0}^{m}\binom{k}{i}\theta_\star^k(1-\theta_\star)^{m-k}c_{m-k} = \omega,$$

and thus, by Cesáro's lemma, almost surely,

$$\frac{1}{n+1}\sum_{i=0}^{n}\mathbb{E}_i c_{m-\zeta_{i+1}} \overset{n\to\infty}{\longrightarrow} \omega.$$

Moreover, $\Delta D_{i+1} = c_{m-\zeta_{i+1}} - \mathbb{E}_i c_{m-\zeta_{i+1}}$ is a martingale difference sequence such that

$$\frac{\langle D\rangle_n}{n} = \frac{1}{n}\sum_{i=0}^{n}\mathbb{E}_i(\Delta D_{i+1})^2 = \frac{1}{n}\sum_{i=0}^{n}\Big[\mathbb{E}_i c^2_{m-\zeta_{i+1}} - (\mathbb{E}_i c_{m-\zeta_{i+1}})^2\Big].$$

Note that

$$\begin{aligned}\mathbb{E}_i c^2_{m-\zeta_{i+1}} - (\mathbb{E}_i c_{m-\zeta_{i+1}})^2 &= \sum_{k=0}^{m}\binom{m}{k} c^2_{m-k} X_i^k(1-X_i)^{m-k} - \Big(\sum_{k=0}^{m}\binom{m}{k} c_{m-k}X_i^k(1-X_i)^{m-k}\Big)^2 \\ &\overset{i\to\infty}{\longrightarrow} \sum_{k=0}^{m}\binom{m}{k} c^2_{m-k}\theta_\star^k(1-\theta_\star)^{m-k} - \Big(\sum_{k=0}^{m}\binom{m}{k} c_{m-k}\theta_\star^k(1-\theta_\star)^{m-k}\Big)^2,\end{aligned}$$

thus, by Cesáro's lemma, ${}^{\langle D\rangle_n}\!/_{n}$ converges to the same limit, and, by [Duf97, Theorem 1.3.17], we get that ${}^{D_n}\!/_{n} \to 0$ almost surely when n tends to infinity. Therefore, by Equation (13), ${}^{T_n}\!/_{n}$ converges almost surely to ω. ☐

5.2 The associated stochastic algorithm

As already mentioned in Section 3, in the two-colour case, we can reduce the study to the first colour since $B_n + W_n = T_n$, and the associated stochastic algorithm is thus one-dimensional.

Lemma 4. *The proportion of white balls in the urn at time n, denoted by $X_n = {}^{W_n}\!/_{T_n}$ satisfies the following recursion:*

$$X_{n+1} = X_n + \frac{1}{T_{n+1}}\big(\tilde{g}(X_n) + \Delta M_{n+1} + \varepsilon_{n+1}\big),$$

where ε_{n+1} is $\mathcal{F}_n$-measurable and goes almost surely to zero as n goes to infinity,

$$\tilde{g}(x) := (1-x)\sum_{k=0}^{m}\binom{m}{k} a_{m-k}x^k(1-x)^{m-k} - x\sum_{k=0}^{m}\binom{m}{k} b_{m-k}x^k(1-x)^{m-k},$$

and

$$\Delta M_{n+1} = Y_{n+1} - \mathbb{E}[Y_{n+1}|\mathcal{F}_n],$$

with

$$Y_{n+1} := a_{m-\zeta_{n+1}} - Z_n(a_{m-\zeta_{n+1}} + b_{m-\zeta_{n+1}}).$$

The proof of this lemma is skipped since it is very similar to the proof of Lemma 2.

5.3 Law of large numbers

In view of Lemma 4, one can apply the following result by Renlund [Ren11]:

Theorem 6. *Assume that the sequence $(X_n)_{n\geqslant 0}$ satisfies*

$$X_{n+1} - X_n = \gamma_{n+1}\big(f(X_n) + V_{n+1}\big), \tag{14}$$

where $(\gamma_n)_{n\geqslant 1}$ and $(V_n)_{n\geqslant 1}$ are two $\mathcal{F}_n$-measurable sequences of random variables and f is a continuous function from $[0,1]$ onto $\mathbb{R}$ such that $f(0)\geqslant 0$, $f(1)\leqslant 1$ and $f\not\equiv 0$. Assume that, almost surely,

$$c_1/n \leqslant \gamma_n \leqslant c_2/n, \qquad |V_n|\leqslant K_v \qquad |f(X_n)|\leqslant K_f, \quad \text{and} \quad |\mathbb{E}[\gamma_{n+1}V_{n+1}|\mathcal{F}_n]|\leqslant K_e\gamma_n^2,$$

where the constants c_1, c_2, K_v, K_f, and K_e are positive real numbers. Then, $\theta_\star := \lim_{n\to+\infty} X_n$ exists almost surely, $f(\theta_\star) = 0$ and $f'(\theta_\star)\leqslant 0$.

This result allows us to prove Theorem 3(a):

Proof of Theorem 3(a). In view of Lemma 4, the proportion X_n of white balls in the urn satisfies

$$X_{n+1} = X_n + \frac{1}{T_{n+1}}\big(\tilde{g}(X_n) + \Delta M_{n+1} + \varepsilon_{n+1}\big),$$

which is the same as Equation (14) with $f=\tilde{g}$, $V_{n+1} = \Delta M_{n+1} + \varepsilon_{n+1}$, and $\gamma_{n+1} = 1/T_{n+1}$. Note that, by the tenability assumption (see Lemma 1), $\tilde{g}(0) = a_m \geqslant 0$ and $\tilde{g}(1) = -b_0\leqslant 0$ as requested in Lemma 4.

By assumption, we have that $\liminf_{n\to\infty} T_n/n > 0$, implying that there exists a constant c such that $T_n\geqslant cn$ for all $n\geqslant 0$. If we let $C = \max_{1\leqslant k\leqslant m} c_k$, then we have almost surely $cn\leqslant T_n\leqslant T_0 + Cn$, implying that there exists two constants c_1 and c_2 such that $c_1/n\leqslant\gamma_n\leqslant c_2/n$ almost surely for all $n\geqslant 0$.

Recall that $\Delta M_{n+1} = Y_{n+1} - \mathbb{E}[Y_{n+1}|\mathcal{F}_n]$, where $Y_{n+1} = \tilde{g}(X_n)$. Note that the function $\tilde{g}$ is a continuous and thus bounded on $[0,1]$, implying that $|\Delta M_{n+1}|\leqslant 2\max_{[0,1]}\tilde{g}$. Also, $\varepsilon_{n+1}\to 0$ almost surely when n tend to infinity, implying that this sequence is bounded. Thus there exists K_v such that $|V_n| = |\Delta M_n + \varepsilon_n|\leqslant K_v$ for all $n\geqslant 1$.

For all $n\geqslant 0$, we have $|\tilde{g}(X_n)|\leqslant\max_{[0,1]}\tilde{g}$, and finally, since $T_{n+1}\geqslant c(n+1)$ almost surely, we have

$$|\mathbb{E}[\gamma_{n+1}V_{n+1}|\mathcal{F}_n]| = \left|\frac{1}{T_n}\mathbb{E}\left[\left(1-\frac{c_{\xi_{n+1}}}{T_{n+1}}\right)V_{n+1}\middle|\mathcal{F}_n\right]\right| = \frac{\mathbb{E}[|\varepsilon_{n+1}||\mathcal{F}_n]}{T_n} + \frac{1}{T_n}\mathbb{E}\left[\left|\frac{c_{\xi_{n+1}}V_{n+1}}{T_{n+1}}\right|\middle|\mathcal{F}_n\right],$$

where we have used that $\mathbb{E}[\Delta M_{n+1}|\mathcal{F}_n] = 0$. Recall that $\varepsilon_n = 0$ in the with-replacement case, and that $|\varepsilon_n| = \mathcal{O}(1/T_n)$ in the without-replacement case. Also recall that, by assumption, $T_0 + cn\leqslant T_n\leqslant T_0 + Cn$, implying that there exists a constant K_v such that $|\mathbb{E}[\gamma_{n+1}V_{n+1}|\mathcal{F}_n]|\leqslant K_v\gamma_n^2$ as requested in Lemma 4.

We can thus apply Lemma 4, which proves Theorem 3(a). □

5.4 Central limit theorem

Our aim in this section is to apply Renlund's central limit theorem for stochastic algorithms. This result is expressed as follows:

Theorem 7 ([Ren11])**.** *Let $(X_n)_{n\geqslant 0}$ satisfying Equation (14) and $\theta_\star$ a stable zero of f. Conditionally on the event $\{X_n\to\theta_\star\}$, let $\hat{\gamma}_n := n\gamma_n\hat{f}(X_{n-1})$ where $\hat{f}(x) = -f(x)/x-\theta_\star$, and assume that $\hat{\gamma}_n$ converges almost surely to some deterministic limit $\hat{\gamma}$. Then:*

(i) If $\hat{\gamma} > 1/2$ and if $\mathbb{E}[(n\gamma_nV_n)^2|\mathcal{F}_{n-1}]\to\sigma^2 > 0$, then $\sqrt{n}(X_n-\theta_\star)\stackrel{n\to\infty}{\longrightarrow}\mathcal{N}\Big(0,\frac{\sigma^2}{2\hat{\gamma}-1}\Big)$, in distribution.

(ii) If $\hat{\gamma} = 1/2$, $(\ln n)(1/2-\hat{\gamma}_n)\to 0$, and $\mathbb{E}[(n\gamma_nV_n)^2|\mathcal{F}_{n-1}]\xrightarrow{a.s.}\sigma^2 > 0$ when $n\to\infty$, then

$$\sqrt{\frac{n}{\ln n}}(X_n-\theta_\star)\stackrel{n\to\infty}{\longrightarrow}\mathcal{N}(0,\sigma^2),\ \textit{in distribution.}$$

Renlund also proves the following result:

Theorem 8 ([Ren11])**.** *Let $(X_n)_{n\geqslant 0}$ be a sequence of random variables that satisfies Equation* (14)*, and $\theta_\star$ a stable zero of f. Conditionally on $\{X_n \to \theta_\star\}$, assume that there exists $\hat{\gamma} \in (0, 1/2)$ such that, almost surely when n tends to infinity,*

$$\hat{\gamma}_n - \hat{\gamma} = \mathcal{O}(|X_n - \theta_\star| + 1/n).$$

Then $n^{\hat{\gamma}}(X_n - \theta_\star)$ converges almost surely to a finite random variable when n tends to infinity.

Theorem 6(a) applies to our framework and allows us to prove Theorem 3(b):

Proof of Theorem 3(b). The idea of the proof is to apply Renlund's result to the stochastic algorithm of Lemma 4. In that case, $\hat{\gamma}_n = \frac{n}{T_n}\frac{\tilde{g}(X_{n-1})}{\theta_\star - X_{n-1}}$, $f = \tilde{g}$ and $V_n = \Delta M_n + \varepsilon_n$. First note that, since $\tilde{g}$ is a polynomial, it is in particular differentiable in $\theta_\star$, and thus, conditionally on $\{X_n \to \theta_\star\}$, we have

$$\frac{\tilde{g}(X_{n-1})}{\theta_\star - X_{n-1}} \to -\tilde{g}'(\theta_\star).$$

Moreover, Lemma 1 gives that $\hat{\gamma}_n \to {-\tilde{g}'(\theta_\star)}/{\omega} =: \lambda$. Let us now have look at

$$\mathbb{E}[(n\gamma_n V_n)^2|\mathcal{F}_{n-1}] = \mathbb{E}\Big[\Big(\frac{n(\Delta M_n + \varepsilon_n)}{T_n}\Big)^2\Big|\mathcal{F}_{n-1}\Big].$$

Note that ${T_n}/{n} \to \omega$ almost surely, $\varepsilon_n \to 0$ almost surely, and $\Delta M_n = Y_{n+1} - \mathbb{E}[Y_{n+1}|\mathcal{F}_{n-1}]$. A simple calculation (using $\tilde{g}(\theta_\star) = 0$) then leads to

$$\mathbb{E}\Big[\Big(\frac{n(\Delta M_n + \varepsilon_n)}{T_n}\Big)^2\Big|\mathcal{F}_{n-1}\Big] \to \frac{1}{\omega^2}H(\theta_\star),$$

almost surely when n goes to infinity. We can thus apply Theorem 7, which concludes the proof of Theorem 3(b). □

As one can now understand, we need information about the speed of convergence of ${T_n}/{n}$ to its limit in order to solve the two remaining cases: $\lambda = 1/2$ and $\lambda < 1/2$. And, in view of Theorem 8, it would be enough to prove that

$$\left|\frac{T_n}{n} - \omega\right| = \mathcal{O}\big(|Z_n - \theta_\star| + 1/n\big).$$

Unfortunately, we are not yet able to prove this statement.

5.5 Examples

Theorem 3 already permits to treat some particular cases of non-balanced Pólya urn processes. We give in this section examples to which our result applies, but also some examples that do not fall into our framework, either because $\lambda \leqslant 1/2$ or because $\sigma^2 = 0$.

Note that in all examples given below, $\min_{0\leqslant k\leqslant m} c_k \geqslant 1$ (recall that $c_k = a_k + b_k$), implying in particular that $\liminf_{n\to\infty} {T_n}/{n} > 0$, as required in Theorem 3.

Example 5.1: Take

$$R = \begin{pmatrix} 2 & 1 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

In that case, $\tilde{g}(x) = (1-x)^3 - x^3$, and this function admits a unique zero $\theta_\star = 1/2$. This particular case thus falls under Theorem 3 and X_n converges almost surely to $1/2$. We also have $\omega = 5/2$, and $\tilde{g}'(1/2) = -3/2$, implying that $\lambda = 3/5 > 1/2$. Finally, note that $H(1/2) = 1/8$, $\sigma^2 = 1/50$, and $2\lambda - 1 = 1/5$ meaning that, in distribution,

$$\sqrt{n}(X_n - 1/2) \xrightarrow{n\to\infty} \mathcal{N}(0, 1/10).$$

Example 5.2: Let a and b be two distinct integers and let $a_{m-k} = a(m-k)$ and $b_{m-k} = bk$ for all $0 \leqslant k \leqslant m$. This example is studied in [ALS], where it is proved that X_n converges almost surely to $\frac{\sqrt{a}}{\sqrt{a}+\sqrt{b}}$, and that, almost surely,

$$\sqrt{n}\Big(X_n - \frac{\sqrt{a}}{\sqrt{a}+\sqrt{b}}\Big) \xrightarrow{n\to\infty} \mathcal{N}\Big(0, \frac{\sqrt{ab}}{3m(\sqrt{a}+\sqrt{b})^2}\Big). \tag{15}$$

This result can be re-proved as a corollary of our main results. One can check that

$$\tilde{g}(x) = m[a(1-x)^2 - bx^2],$$

whose unique root in $[0,1]$ is $\theta_\star := \frac{\sqrt{a}}{\sqrt{a}+\sqrt{b}}$. Thus theorem 3($a$) gives the almost sure convergence of Z_∞ to $\theta_\star$. Moreover, $\tilde{g}'(\theta_\star) = -2m\sqrt{ab}$, and $\omega = m\sqrt{ab}$, implying that $\lambda = 2$. Finally,

$$H(\theta_\star) = \frac{m(ab)^{3/2}}{(\sqrt{a}+\sqrt{b})^2}, \quad \sigma^2 = \frac{\sqrt{ab}}{m(\sqrt{a}+\sqrt{b})^2} \quad \text{and} \quad 2\lambda - 1 = 3,$$

which, applying Theorem 3(b), gives Equation (15).

Example 5.3: Let a and b be two distinct integers and let $a_{m-k} = ak$ and $b_{m-k} = b(m-k)$ for all $0 \leqslant k \leqslant m$. This example is studied in [ALS], where it is proved that, if $a < b$ (resp. $a > b$) X_n converges almost surely to 0 (resp. 1) when n tends to infinity, and that, almost surely

$$n^{-a/b} X_n \to \Psi,$$

where Ψ is an absolutely continuous random variable (which first and second moments are calculated).

Let us try to apply our results: one can satisfy that

$$\tilde{g}(x) = m(a-b)x(1-x).$$

This polynomial admits two zeros, 0 and 1. Note that $\tilde{g}'(0) = -\tilde{g}'(1) = m(a-b)$. Thus, by Theorem 3($a$), if $a < b$, then X_n converges almost surely to 0, and if $a > b$, X_n converges almost surely to one. In both cases,

$$\omega = \max(a,b)m,$$

and thus, $\lambda = \frac{|a-b|}{\max(a,b)}$. Also note that in both cases, $H(\theta_\star) = 0$, which means that $\sigma^2 = 0$. Even in the cases when $\lambda > 1/2$, this example does not fall into our framework since $\sigma^2 = 0$.

References

[AK68] K. B. Athreya and S. Karlin. Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *Annals of Mathematical Statistics*, 39:1801–1817, 1968.

[ALS] R. Aguech, N. Lassmar, and O. Selmi. A generalized urn model with multiple drawing and random addition. (preprint).

[BDM09] A. Bose, A. Dasgupta, and K. Maulik. Strong laws for balanced triangular urns. *Journal of Applied Probability*, 46:571–584, 2009.

[CK13] M. R Chen and M. Kuba. On generalized Pólya urn models. *Journal of Applied Probabilty Theory*, 50(4):909–1216, 2013.

[CW05] M. R Chen and C. Z Wei. A new urn model. *Journal of Applied Probability*, 42, 2005.

[Duf97] M. Duflo. *Random Iterative Models.* Springer-Verlag Berlin, 1997.

[EP23] F. Eggenberger and G. Pólya. Über die statistik verketetter vorgäge. *Zeitschrift für Angewandte Mathematik und Mechanik*, 1:279–289, 1923.

[FDP06] P. Flajolet, P. Dumas, and V. Puyhaubert. Some exactly solvable models of urn process theory. *DMTCS Proceedings*, AG:59–118, 2006.

[Jan04] S. Janson. Functional limit theorem for multitype branching processes and generalized Pólya urns. *Stochastic Processes and their Applications*, 110:177–245, 2004.

[Jan06] S. Janson. Limit theorems for triangular urn schemes. *Probability Theory and Related Fields*, 134(3):417–452, 2006.

[KM17] M. Kuba and H. M. Mahmoud. Two-colour balanced affine urn models with multiple drawings. *Advances in Applied Mathematics*, 90:1–26, 2017.

[KS16] M. Kuba and H. Sulzbach. On martingale tail sums in affine two-color urn models with multiple drawings. *Journal of Applied Probability*, 2016. to appear.

[LL13] B. Laslier and J.-F. Laslier. Reinforcement learning from comparisons: Three alternatives is enough, two is not. 2013. `ArXiV:1301.5734`.

[LP13] S. Laruelle and G. Pagès. Randomized urn models revisited using stochastic approximation. *Annals of Applied Probability*, 23(4):1409–1436, 2013.

[MKP13] H. Mahmoud M. Kuba and A. Panholzer. Analysis of a generalized Friedman's urn with multiple drawings. *Discrete Applied Mathematics*, 161:2968–2984, 2013.

[MM12] B. Morcrette and H. Mahmoud. Exactly solvable balanced tenable urns with random entries via the analytic methodology. In *23rd International Meeting on Probabilistic, Combinatorial, and Asymptotic Methods for the Analysis of Algorithms (AofA)*, volume AQ, pages 219–232. DMTCS, 2012.

[Mor12] B. Morcrette. Fully analyzing an algebraic Pólya urn model. In *Latin American Symposium on Theoretical Informatics (LATIN)*, volume 7256, pages 568–581. LNCS, 2012.

[Pem07] R. Pemantle. A survey of random processes with reinforcement. *Probability surveys*, 4(1-79):25, 2007.

[Ren10] H. Renlund. Generalized pólya urns via stochastic approximation. 2010. `ArXiV:1002.3716v1`.

[Ren11] H. Renlund. Limit theorems for stochastic approximation algorithms. 2011. `ArXiV:1102.4741v1`.

[RM51] H. Robbins and S. Monro. A stochastic approximation method. *Annals of mathematical statistics*, 22:400–407, 1951.

[Zha16] L.-X. Zhang. Central limit theorems of a recursive stochastic algorithm with applications to adaptive design. *Annals of Applied Probability*, 26(6):3630–3658, 2016.