# INTERACTING GENERALIZED PÓLYA URN SYSTEMS

GIACOMO ALETTI AND ANDREA GHIGLIETTI



Abstract. We consider a system of interacting *Generalized Pólya Urns* (GPUs) having irreducible mean replacement matrices. The interaction is modeled through the probability to sample the colors from each urn, that is defined as convex combination of the urn proportions in the system. From the weights of these combinations we individuate subsystems of urns evolving with different behaviors. We provide a complete description of the asymptotic properties of urn proportions in each subsystem by establishing limiting proportions, convergence rates and Central Limit Theorems. The main proofs are based on a detailed eigenanalysis and stochastic approximation techniques.

Keywords. *Interacting systems, Generalized Pólya urn models, Central Limit Theorems, Strong Consistency, Stochastic approximation.*

*2010 MSC classification:* 60K35; 62L20; 60F05; 60F15.

## 1. Introduction

The stochastic evolution of systems composed by elements which interact among each other has always been of great interest in several areas of application, e.g. in medicine a tumor growth is the evolution of a system of interacting cells [34], in socio-economics and life sciences a collective phenomenon reflects the result of the interactions among the individuals [27], in physics the concentration of certain molecules within cells varies over time due to interactions between different cells [31]. In the last decade several models have been proposed in which the elements of the system are represented by urns containing balls of different colors, in which the urn proportions reflect the status of the elements, and the evolution of the system is established by studying the dynamics at discrete times of this collection of dependent urn processes. The main reason of this popularity is concerned with the urn dynamics, which is (i) suitable to describe random phenomena in different scientific fields (see e.g. [21]), (ii) flexible to cover a wide range of possible asymptotic behaviors, (iii) intuitive and easy to be implemented in several fields of application.

The dynamics of a single urn typically consists in a sequential repetition of a *sampling phase*, when a ball is sampled from the urn, and a *replacement phase*, when a certain quantity of balls is replaced in the urn. The basic model is the Pólya's urn proposed in [16]: from an urn containing balls of two colors, balls are sequentially sampled and then replaced in the urn with a new ball of the same color. This updating scheme is then iterated generating a sequence of urn proportions whose almost sure limit is random and Beta distributed. Starting from this simple model, several interesting variations have been suggested by considering different distributions in the sampling phase, e.g. [19, 20], or in the replacement phase, e.g. [3, 18, 30]. In a general $K-$colors urn model, the sampled color is usually represented by a vector $\mathbf{X}_n$ such that $X_{k,n} = 1$ when the color is $k \in \{1, .., K\}$, $X_{k,n} = 0$ otherwise; the quantities of balls replaced in the urn are typically defined by a matrix $D_n$ such that $D_{ik,n}$ indicates the number of balls of color $i$ replaced in the urn when the color $k$ is sampled. Considering $\{D_n; n \geq 1\}$ as an i.i.d. sequence, a crucial element to characterize the asymptotic behavior of the urn is the mean replacement matrix $H := \mathbf{E}[D_n]$, typically called *generating matrix*.

The class of urn models considered in this paper is commonly denoted by *Generalized Pòlya's Urn* (GPU), or *Generalized Friedman's Urn* (GFU). The GPU model was introduced in [18] and its extensions and their asymptotic behavior have been studied in several works, see e.g. [4, 5, 6, 33]. The GPU considered in this paper is characterized by a non-negative irreducible generating matrix $H$ with average constant balance, i.e. the columns of $H$ sum up at the same constant $\sum_{i=1}^{K} H_{ik} = c > 0$ for any $k \in \{1, .., K\}$, which implies that its maximum eigenvalues $\lambda_{\max}(H) = c$ has multiplicity one. The irreducibility of $H$ distinguishes the GPU from the *Randomly Reinforced Urn* (RRU) model, which includes the classical Pòlya's Urn, whose replacement matrix is diagonal: when the color $k$ is sampled, the GPU replaces in the urn more colors following the distribution of the $k^{th}$ column of $D_n$ while the RRU only adds balls of colors $k$; hence, the probability to sample color $k$ at next step is reinforced in the RRU, while it may increase or decrease according to the current urn composition in the GPU. As a consequence, the asymptotic behavior is in general very different: in a GPU the urn proportion converges to a deterministic equilibrium identified by $H$ (see e.g. [4, 5, 6, 33]), while in a RRU the limit is random and its distribution depends on the initial composition (see e.g. [1, 2, 15]).

The model proposed in this paper is a collection of $N \geq 1$ GPUs that interact among each other during the sampling phase: the probability to sample a color $k$ in each urn $j$ is a convex combination of the urn proportions of the entire system. Hence, a crucial role to describe the system dynamics is played by the *interacting matrix* $W$ made by the weights of those combinations. Since the properties of the single GPUs are determined by the corresponding generating matrices $\{H^j; 1 \leq j \leq N\}$ and the interaction among them are ruled by $W$, the system dynamics has been studied by defining a new object $\mathbf{Q}$ that merges the information contained in $\{H^j; 1 \leq j \leq N\}$ and $W$. From the analysis of the eigenstructure of $\mathbf{Q}$, we are able to establish the convergence and the second-order asymptotic behavior of the urn proportions in the entire system. Hence, this paper extends the theory on GPU models in the sense that, in the special case of no interaction, i.e. $W = I$, the results presented for the system reduce to the well-known results for a single GPU.

Several interacting urn models have been proposed in the last decade, especially for RRU models. An early work is represented by [29] that considered a collection of two-colors RRU in which the sequence $\{D_n; n \geq 1\}$ is not i.i.d. since the replacements in each urn depend on the colors sampled in the rest of the system. Therefore, in [29] the interaction is modeled through the definition of $D_n$, instead of $\mathbf{X}_n$ as in our model. A complete different updating rule has been used in the two-color urn model proposed in [26], in which sampling color 1 in the urn $j$ increases the composition of color 1 in the urn $j$, while sampling color 2 increases the composition of color 2 in the neighbor urns $i \neq j$ and the urn $j$ comes back to the initial composition. Asymptotic properties for this system have been obtained in [26] where there is no convergence of the urn proportions. Other models in which the interaction enters in the replacement matrices are for instance [11, 10, 8].

Recently there have been more works concerning urn systems in which the interaction is modeled through the sampling probabilities as in our model. They differ from this paper since all of them consider RRUs and the interaction is only modeled as mean-field interaction tuned by a parameter $\alpha \in (0, 1)$, i.e. the urns interact among each other only through the average composition in the entire system. As a consequence, their asymptotic results lead to the *synchronization* property in which all the urn proportions of the system converge to the same random limit. In particular, in [24, 25] the asymptotic behavior of the urn system has been studied for a model that defines the sampling probabilities through the exponential of the urn compositions. In [13, 12] the sampling probabilities are defined directly using the urn compositions and the synchronization property has been proved; moreover, different convergence rates and second-order asymptotic distributions for the urn proportion have been established for different values of the tuning parameter $\alpha$. Since we consider GPU models the asymptotic results established in this paper are totally different from those proved in [13, 12], e.g. our limiting proportion is not random and it does not depend on the initial compositions.

It is also significant to highlight that this work allows a general structure for the urn interaction, which reduces to the mean-filed interaction only for a particular choice of the interacting matrix $W$. Moreover, from the analysis of the structure of $W$ we are able to individuate subsystems of urns evolving with different behaviors: (i) the *leading systems*, whose dynamics is independent of the rest of the system and (ii) the *following systems*, whose dynamics "follows" the evolution of other urns of the system; in the special case of irreducible interacting matrix, which includes the mean-filed interaction considered in [13, 12], there is a unique leading system and no following systems. These two classes of systems have been studied separately, in order to provide an exhaustive description of the asymptotic behavior in any part of the system. In fact, since different systems may converge at different rates, a unique central limit theorem would not been able to characterize the convergence of any urn proportion. Hence, through a careful analysis on the eigen-structure of $\mathbf{Q}$, we individuate the components of the urn processes in the system that actually "lead" or influence the following systems, so that we can establish the right convergence rate and a non-degenerate asymptotic distribution for any subsystem.

A pivotal technique in the proofs consists in revisiting the dynamics of the urn proportions of the system in the stochastic approximation (SA) framework, as suggested for the composition of a single GPU in [23]. To this end, the dynamics of the urn proportions has been properly modified here to embed the processes of the urn proportion into the whole space $\mathbb{R}^K$.

The structure of the paper is the following. In Section 2 the interacting GPU model is described and the main assumptions are presented. Section 2 is also dedicated to analyze the structure of the interacting matrix and hence to define the leading and the following systems. In Section 3 the system dynamics is expressed in the stochastic approximation form and the necessary notation is introduced. The asymptotic results for the leading and following systems are presented in Section 4 and 5, respectively. Section 6 contains a brief discussion on further possible extensions of the interacting GPU model. Finally, the proofs are presented in Section 7.

## 2. Model Setting and main Assumptions

Consider a collection of $N \geq 1$ urns containing balls of $K \geq 1$ different colors. At any time $n \geq 0$ and for any urn $j \in \{1, .., N\}$, let $Y^j_{k,n} > 0$ be the number of balls of color $k \in \{1, .., K\}$, $T^j_n := \sum_{k=1}^K Y^j_{k,n}$ be the total number of balls and let $Z^j_{k,n} := Y^j_{k,n}/T^j_n$ be the proportion of color $k$.

### 2.1. **Model.**

We now describe precisely how the system evolves at any time $n \geq 1$. Denote by $\mathcal{F}_{n-1}$ the $\sigma$-algebra generated by the urn compositions of the entire system up to time $(n-1)$, i.e.

$$\mathcal{F}_{n-1} \ := \ \sigma\left( \ Y^j_{k,n-1}, \ 1 \leq j \leq N, \ 1 \leq k \leq K \ \right).$$

The dynamics of the system is described by two main phases: *sampling* and *replacement*.

*Sampling phase*: for each urn $j \in \{1, .., N\}$, a ball is virtually sampled and its color is represented as follows: $X^j_{k,n} = 1$ indicates that the sampled ball is of color $k$, $X^j_{k,n} = 0$ otherwise. We denote by $\tilde{Z}^j_{k,n-1}$ the probability to sample a ball of color $k$ in the urn $j$ at time $n$, i.e.

$$\tilde{Z}^j_{k,n-1} := \mathbf{E}\left[ \ X^j_{k,n} \mid \mathcal{F}_{n-1} \ \right].$$

Given the sampling probabilities $\{\tilde{Z}^j_{k,n-1}, 1 \leq j \leq N, 1 \leq k \leq K\}$, the colors are sampled independently in all the urns of the system and hence, for any $k \in \{1,..,K\}$, $X^1_{k,n},..,X^N_{k,n}$ are independent conditionally on $\mathcal{F}_{n-1}$. We define the sampling probabilities as convex combinations of the urn proportions of the system. Formally, for any urn $j \in \{1,..,N\}$ we introduce the weights $\{w_{jh}; 1 \leq h \leq N\}$ such that $0 \leq w_{jh} \leq 1$ and $\sum_{h=1}^N w_{jh} = 1$. Thus, the probability to sample the color $k$ in the urn $j$ is defined as follows

$$\tilde{Z}^j_{k,n-1} := \sum_{h=1}^N w_{jh} Z^h_{k,n-1}. \tag{1}$$

*Replacement phase*: after that a ball of color $k$ has been sampled from the urn $j$, we replace $D^j_{ik,n}$ balls of color $i \in \{1,..,K\}$ in the urn $j$. For any urn $j$ we assume that $\{D^j_n; n \geq 1\}$ is a sequence of i.i.d. non-negative random matrices, where $D^j_n := [D^j_{ik,n}]_{ik}$. We will refer to $D^j_n$ as *replacement matrix* and to $H^j := \mathbf{E}[D^j_n]$ as *generating matrix*. Notice that $H^j$ are time-independent since $\{D^j_n; n \geq 1\}$ are identically distributed (see Subsection 6 for possible extensions). Moreover, we assume that at any time $n$ the replacement matrix for the urn $j$, i.e. $D^j_n$, is independent of the sampled colors, i.e. $\{X^j_{k,n}; 1 \leq j \leq N\}$, and independent of the replacement matrices of the other urns of the system, i.e. $D^{j_0}_n$ with $j_0 \neq j$.

In conclusion, the composition of the color $i \in \{1,..,K\}$ in the urn $j \in \{1,..,N\}$ evolves at time $n \geq 1$ as follows:

$$Y^j_{i,n} = Y^j_{i,n-1} + \sum_{k=1}^K D^j_{ik,n} X^j_{k,n}. \tag{2}$$

2.2. **Main assumptions.** We now present the main conditions required to establish the results of the paper. The first assumption is concerned with bounds for the moments of the replacement distributions. Specifically, we require the following condition:

(A1) there exists $\delta > 0$ and a constant $0 < C_\delta < \infty$ such that, for any $j \in \{1,..,N\}$ and any $i,k \in \{1,..,K\}$, $\mathbf{E}[(D^j_{ik,n})^{2+\delta}] < C_\delta$.

Note that $C_\delta$ does not depend on $n$ since $\{D^j_n; n \geq 1\}$ are identically distributed.

The second assumption is the average constant balance of the urns in the system and it is imposed by the following condition on the generating matrices $H^1,..,H^N$:

(A2) for any $j \in \{1,..,N\}$ and $k \in \{1,..,K\}$, there exists a constant $0 < c^j < \infty$ such that $\sum_{i=1}^K H^j_{ik} = c^j$.

Note that (A2) guarantees that the average number of balls replaced in any urn is constant, regardless its composition. Assumption (A2) is essential to obtain the asymptotic configuration of the system, i.e. the limiting urn proportions. The second-order asymptotic properties of the interacting urn system, namely the rate of convergence and the limiting distributions, are obtained by assuming a stricter assumption than (A2). This condition is expressed as follows:

(A'2) for any $j \in \{1,..,N\}$, $k \in \{1,..,K\}$, $\mathbf{P}\left( \sum_{i=1}^K D^j_{ik,n} = c^j \right) = 1$, i.e. each urn is updated with a constant total amount of balls.

Without loss of generality, we may (and do) assume that $c^j = 1$ for all $j \in \{1, .., N\}$. In fact, by defining $\widehat{Y}^j_{k,n} = (c^j)^{-1} Y^j_{k,n}$ and $\widehat{D}^j_{ik,n} = (c^j)^{-1} D^j_{ik,n}$ for all $n \geq 1$, the urn dynamics in (2) can be expressed in the following equivalent form:

$$\widehat{Y}^j_{i,n} \;=\; \widehat{Y}^j_{i,n-1} + \sum_{k=1}^{K} \widehat{D}^j_{ik,n} \cdot X^j_{k,n}, \qquad \widehat{Z}^j_{k,n-1} \;=\; \frac{\widehat{Y}^j_{k,n-1}}{\sum_{k=1}^{K} \widehat{Y}^j_{k,n-1}} \;=\; \frac{Y^j_{k,n-1}}{\sum_{k=1}^{K} Y^j_{k,n-1}} \;=\; Z^j_{k,n-1}.$$

Therefore, from now on we will denote by $Y^j_{k,n}$ and $D^j_{ik,n}$ the normalized quantities $\widehat{Y}^j_{k,n}$ and $\widehat{D}^j_{ik,n}$ and hence (A2) and (A'2) are replaced by the following conditions:

(A2) for any $j \in \{1, .., N\}$ and $k \in \{1, .., K\}$, $\sum_{i=1}^{K} H^j_{ik} \;=\; 1$.

(A'2) for any $j \in \{1, .., N\}$ and $k \in \{1, .., K\}$, $\mathbf{P}\left( \sum_{i=1}^{K} D^j_{ik,n} \;=\; 1 \right) = 1$.

Finally, we consider *Generalized Pòlya urn* (GPU) with irreducible generating matrices, as expressed in the following condition:

(A3) for any $j \in \{1, .., N\}$, $H^j$ is irreducible.

This assumption will guarantee deterministic asymptotic configurations for the urn proportions in the system. Less restrictive conditions to establish deterministic limiting proportion are possible but this analysis is not the focus of this paper.

**Remark 2.1.** *It is worth highlighting that extensions to non-homogeneous generating matrices $\{H_n; n \geq 0\}$ are possible, as discussed in Section 6. In that case, assumption (A2) should be referred to the the limiting matrix $H^j := a.s. - \lim_{n\to\infty} H^j_n$.*

2.3. **A preliminary result.** Assumptions (A2) and (A'2) on the constant balance are essential to obtain the following result on the total number of balls in the urns of the system:

**Theorem 2.1.** *Under assumptions (A1) and (A2), $\{T^j_n - n; n \geq 1\}$ is an $L^2$ martingale and, for any $\alpha < 1/2$,*

$$n^{\alpha} \left( \frac{T^j_n}{n} - 1 \right) \overset{a.s./L^2}{\longrightarrow} 0. \tag{3}$$

*Moreover, under assumption (A'2), $T^j_n = T^j_0 + n$ a.s. and hence* (3) *holds for any $\alpha < 1$.*

2.4. **The interacting matrix.** The interaction among the urns of the system is modeled through the sampling probabilities $\tilde{Z}^j_{k,n-1}$, that are defined in (1) as convex combination of the urn proportions of the system. Formally, we denote by $W$ the $N \times N$ matrix composed by the weights $\{w_{jh}, 1 \leq j, h \leq N\}$ of such linear combinations and we refer to it as *interacting matrix*. We now consider a particular decomposition of $W$ that individuates subsystems of urns evolving with different behaviors. The same decomposition is applied to the transition matrix in the context of discrete-time Markov chains (see [28]) to individuate communicating classes $S^l$, $l \in \mathcal{L}$, and to establish which classes are recurrent, $l \in \mathcal{L}_L$, and which are transient, $l \in \mathcal{L}_F$.

Accordingly, let us denote by $n_L \geq 1$ the multiplicity of $\lambda_{\max}(W) = 1$, and define the integers $n_F \geq 0$ and $1 \leq r^{L_1} < .. < r^{L_{n_L}} < r^{F_1} < .. < r^{F_{n_F}} = N$ such that the interacting matrix can be decomposed as follows:

$$
\begin{aligned}
W &:= \begin{bmatrix} W^{L} & 0 \\ W^{LF} & W^{F} \end{bmatrix}, & W^{L} &:= \begin{bmatrix} W^{L_1} & 0 & ... & 0 \\ 0 & W^{L_2} & ... & ... \\ ... & ... & ... & 0 \\ 0 & 0 & ... & W^{L_{n_L}} \end{bmatrix}, \\
W^{LF} &:= \begin{bmatrix} W^{L_1F_1} & ... & W^{L_{n_L}F_1} \\ ... & ... & ... \\ W^{L_1F_{n_F}} & ... & W^{L_{n_L}F_{n_F}} \end{bmatrix}, & W^{F} &:= \begin{bmatrix} W^{F_1} & 0 & ... & 0 \\ W^{F_1F_2} & W^{F_2} & ... & 0 \\ ... & ... & ... & ... \\ W^{F_1F_{n_F}} & W^{F_2F_{n_F}} & ... & W^{F_{n_F}} \end{bmatrix}.
\end{aligned} \tag{4}
$$

where:

(1) $\mathcal{L} := \mathcal{L}_L \cup \mathcal{L}_F$, $\mathcal{L}_L := \{L^1, .., L^{n_L}\}$ and $\mathcal{L}_F := \{F^1, .., F^{n_F}\}$ are sets of labels that identify subsystems of urns ($\mathcal{L}_F = \emptyset$ when $n_F = 0$);

(2) for any $l \in \mathcal{L}$, $W^l$ is an $s^l \times s^l$ irreducible matrix, where we let $s^l := r^l - r^{l^-}$ and $l^-$ indicates the element in $\mathcal{L}$ that precedes $l$ (by convention $L_1^- \equiv \emptyset$ and $F_1^- \equiv L_{n_L}$);

(3) for any $l_2 \in \mathcal{L}_F$, there is at least an $l_1 \in \mathcal{L}$, $l_1 \neq l_2$, such that $W^{l_1 l_2} \neq 0$; hence, $\lambda_{\max}(W^l) = 1$ if $l \in \mathcal{L}_L$ and $\lambda_{\max}(W^l) < 1$ if $l \in \mathcal{L}_F$.

Naturally, when $n_F = 0$ the elements in $W^{LF}$ and $W^F$ do not exist and we consider $r^{L_{n_L}} = N$. This occurs, for instance, when $W$ is irreducible and hence $n_L = 1$ and $r^1 = N$.

**Remark 2.2.** *It is worth highlighting that extensions to random and time-dependent interacting matrices $\{W_n; n \geq 0\}$ are possible, as discussed in Section 6. In that case, the structure presented in (4) is concerned with the limiting matrix $W := a.s. - \lim_{n\to\infty} W_n$.*

Since the urns of the system interact among each other only through the sampling probabilities $\tilde{Z}^j_{n,k}$, the structure of the matrix $W$ that characterizes such interaction is essential to describe the asymptotic behavior of the system. Specifically, from (4) we individuate

(i) the leading systems $S^l := \{r^{l^-}+1 < j \leq r^l\}$, $l \in \mathcal{L}_L$, that evolve independently with respect to the rest of the system;

(ii) if $n_F \geq 0$, the following systems $S^l := \{r^{l^-}+1 < j \leq r^l\}$, $l \in \mathcal{L}_F$, that evolve depending on the proportions of the urns in the leaders $S^{L_1}, .., S^{L_{n_L}}$ and their upper followers $S^{F_1}, .., S^{l^-}$.

As we will see in the following sections, the asymptotic behaviors of the leading systems and the following systems are quite different. For completeness of the paper, we will present the results for both the types of systems, assuming that $n_F \geq 1$.

## 3. The interacting urn system in the stochastic approximation framework

A crucial technique to characterize the behavior of the interacting urn system consists in revisiting its dynamics into the stochastic approximation (SA) framework. To this end, we need to rewrite the system dynamics expressed in (2) in the classical SA form: given a filtered probability space $(\Omega, \mathcal{A}, (\mathcal{F}_n)_{n\geq 0}, \mathbf{P})$, we consider the following recursive procedure

$$\forall\, n \geq 1, \quad \theta_n = \theta_{n-1} - \frac{1}{n} f(\theta_{n-1}) + \frac{1}{n}\left(\Delta M_n + R_n\right), \tag{5}$$

where $f : \mathbb{R}^d \to \mathbb{R}^d$ is a locally Lipschitz continuous function, $\theta_n$ an $\mathcal{F}_n$-measurable finite random vector and, for every $n \geq 1$, $\Delta M_n$ is an $\mathcal{F}_{n-1}$-martingale increment and $R_n$ is an $\mathcal{F}_n$-adapted remainder term. To this end, we need a compact notation that jointly describes the composition of the urns in the same subsystem $S^l$, $l \in \mathcal{L}$.

3.1. **Notation.** The quantities related to the urn $j \in \{1,..,N\}$ at time $n$ are denoted by:

(1) $Y_n^j = (Y_{1,n}^j,..,Y_{K,n}^j)' \in \mathbb{R}_+^K$,
(2) $X_n^j = (X_{1,n}^j,..,X_{K,n}^j)' \in \{0,1\}^K$,
(3) $Z_n^j = (Z_{1,n}^j,..,Z_{K,n}^j)' \in (0,1)^K$,
(4) $\tilde{Z}_n^j = (\tilde{Z}_{1,n}^j,..,\tilde{Z}_{K,n}^j)' \in (0,1)^K$,

while the corresponding terms of the system $S^l$, $l \in \mathcal{L}$, given by the $s^l$ urns labeled by $\{r^{l^-}+1,..,r^l\}$, are denoted by:

(1) $\mathbf{Y}_n^l := (Y_n^{r^{l^-}+1},..,Y_n^{r^l})' \in \mathbb{R}_+^{s^lK}$,
(2) $\mathbf{X}_n^l := (X_n^{r^{l^-}+1},..,X_n^{r^l})' \in \{0,1\}^{s^lK}$,
(3) $\mathbf{Z}_n^l := (Z_n^{r^{l^-}+1},..,Z_n^{r^l})' \in \mathcal{S}^{s^lK}$, where $\mathcal{S}^{s^lK}$ indicates the composition of $s^l$ simpleces where $Z_n^{r^{l^-}+1},..,Z_n^{r^l}$ are defined.
(4) $\tilde{\mathbf{Z}}_n^l := (\tilde{Z}_n^{r^{l^-}+1},..,\tilde{Z}_n^{r^l})' \in \mathcal{S}^{s^lK}$,
(5) $\mathbf{T}_n^l := (T_n^{r^{l^-}+1}\mathbf{1}_K,..,T_n^{r^l}\mathbf{1}_K)' \in \mathbb{R}_+^{s^lK}$, where $\mathbf{1}_K$ indicates the $K$-vector of all ones.

The replacement matrix for the system $S^l$ is defined by a block diagonal matrix $\mathbf{D}_n^l \in \mathbb{R}_+^{s^lK\times s^lK}$, where the $s^l$ blocks are the replacement matrices of the urns $\{r^{l^-}+1,..,r^l\}$ in $S^l$, i.e. $D_n^{r^{l^-}+1},..,D_n^{r^l}$. Analogously, the generating matrix for $S^l$ is defined by a block diagonal matrix $\mathbf{H}^l \in \mathbb{R}_+^{s^lK\times s^lK}$, where the $s^l$ blocks are $H^{r^{l^-}+1},..,H_n^{r^l}$. The interaction within the system $S^l$ is modeled by the matrix $\mathbf{W}^l \in (0,1)^{s^lK\times s^lK}$ obtained by replacing in $W^l$ (4) the weights $w_{ih}$ with the corresponding diagonal matrix $w_{ih}I_K$, where $I_K$ indicates a $K$-identity matrix. Analogously, the interaction between a following system $S^{l_1}$, $l \in \mathcal{L}_F$, and another system $S^{l_2}$, $l_2 \in \{L_1,..,l_1^-\}$, is modeled by the matrix $\mathbf{W}^{l_1l_2} \in (0,1)^{s^{l_1}K\times s^{l_2}K}$, obtained by replacing in $W^{l_1l_2}$ (4) the weights $w_{ih}$ with the corresponding diagonal matrix $w_{ih}I_K$.

3.2. **The system dynamics in the SA form.** For any system $S^l$, $l \in \mathcal{L}$, the dynamics in (2) can be written, using the notation of Subsection 3.1, as follows:

$$\mathbf{Y}_n^l \;=\; \mathbf{Y}_{n-1}^l \;+\; \mathbf{D}_n^l\mathbf{X}_n^l. \tag{6}$$

We now express (6) in the SA form (5), where the process $\{\theta_n; n \geq 1\}$ is represented by the urn proportions of the system $S^l$, i.e. $\{\mathbf{Z}_n^l; n \geq 1\}$. Since $\mathbf{Y}_n^l = diag(\mathbf{T}_n^l)\mathbf{Z}_n^l$ for any $n \geq 1$, from (6) we have

$$diag(\mathbf{T}_n^l)\mathbf{Z}_n^l \;=\; diag(\mathbf{T}_{n-1}^l)\mathbf{Z}_{n-1}^l \;+\; \mathbf{D}_n^l\mathbf{X}_n^l,$$

that is equivalent to

$$diag(\mathbf{T}_n^l)(\mathbf{Z}_n^l - \mathbf{Z}_{n-1}^l) \;=\; -diag(\mathbf{T}_n^l - \mathbf{T}_{n-1}^l)\mathbf{Z}_{n-1}^l \;+\; \mathbf{D}_n^l\mathbf{X}_n^l. \tag{7}$$

Now, notice that, for any $n \geq 1$,

(1) $\mathbf{E}[diag(\mathbf{T}_n^l - \mathbf{T}_{n-1}^l)|\mathcal{F}_{n-1}] \;=\; \mathbf{I}$ by Theorem 2.1;
(2) $\mathbf{E}[\mathbf{D}_n^l\mathbf{X}_n^l|\mathcal{F}_{n-1}] \;=\; \mathbf{E}[\mathbf{D}_n^l|\mathcal{F}_{n-1}]\mathbf{E}[\mathbf{X}_n^l|\mathcal{F}_{n-1}] \;=\; \mathbf{H}^l\tilde{\mathbf{Z}}_{n-1}^l$, since $\mathbf{D}_n^l$ and $\mathbf{X}_n^l$ are independent conditionally on $\mathcal{F}_{n-1}$.

Hence, defining the martingale increment

$$\Delta \mathbf{M}_n^l \ := \ \mathbf{D}_n^l \mathbf{X}_n^l - \mathbf{H}^l \tilde{\mathbf{Z}}_{n-1}^l \ - \ (diag(\mathbf{T}_n^l - \mathbf{T}_{n-1}^l) - \mathbf{I})\mathbf{Z}_{n-1}^l, \tag{8}$$

we can express (7) as follows:

$$diag(\mathbf{T}_n^l)(\mathbf{Z}_n^l - \mathbf{Z}_{n-1}^l) \ = \ -\mathbf{Z}_{n-1}^l \ + \ \mathbf{H}^l \tilde{\mathbf{Z}}_{n-1}^l \ + \ \Delta \mathbf{M}_n^l. \tag{9}$$

Now, multiplying by $diag(\mathbf{T}_n^l)^{-1}$ and defining the remainder term

$$\mathbf{R}_n^l \ := \ \left(n \cdot diag(\mathbf{T}_n^l)^{-1} - \mathbf{I}\right)\left(-\mathbf{Z}_{n-1}^l \ + \ \mathbf{H}^l \tilde{\mathbf{Z}}_{n-1}^l \ + \ \Delta \mathbf{M}_n^l\right), \tag{10}$$

we can write (9) as follows:

$$\mathbf{Z}_n^l - \mathbf{Z}_{n-1}^l \ = \ -\frac{1}{n}(\mathbf{Z}_{n-1}^l \ - \ \mathbf{H}^l \tilde{\mathbf{Z}}_{n-1}^l) \ + \ \frac{1}{n}\left(\Delta \mathbf{M}_n^l \ + \ \mathbf{R}_n^l\right). \tag{11}$$

The term $(\mathbf{Z}_{n-1}^l - \mathbf{H}^l \tilde{\mathbf{Z}}_{n-1}^l)$ in (11) should represent the function $f$ in (5) in the SA form. However, although in a leader $S^l$, $l \in \mathcal{L}_L$, we have that $\tilde{\mathbf{Z}}_{n-1}^l$ only depends on $\mathbf{Z}_{n-1}^l$, in a follower $S^l$, $l \in \mathcal{L}_F$, the term $\tilde{\mathbf{Z}}_{n-1}^l$ is in general a function of the composition of all the urns of the system, i.e. $\mathbf{Z}_{n-1}^{L_1}, .., \mathbf{Z}_{n-1}^l$. Hence, the dynamics of a leading system can be expressed as in (11), while the dynamics of a following system needs to be incorporated with other systems to be fully described. For this reason, the asymptotic behavior of these two types of systems are studied separately: the leading systems in Section 4 and the following systems in Section 5.

## 4. Leading systems

In this section we present the main asymptotic results concerning the leading systems $S^l$, $l \in \mathcal{L}_L$. We recall that these systems are characterized by irreducible interacting matrices $W^l$ such that $\lambda_{max}(W^l) = 1$ (see (4) in Subsection 2.4). For this reason, their dynamics is independent of the rest of the system and hence, by using $\tilde{\mathbf{Z}}_{n-1}^l = \mathbf{W}^l \mathbf{Z}_{n-1}^l$ in (11), we have

$$\begin{aligned} \mathbf{Z}_n^l - \mathbf{Z}_{n-1}^l \quad &= \ -\frac{1}{n} h^l(\mathbf{Z}_{n-1}^l) \ + \ \frac{1}{n}\left(\Delta \mathbf{M}_n^l \ + \ \mathbf{R}_n^l\right), \\ h^l(\mathbf{x}) \quad &:= (\mathbf{I} - \mathbf{Q}^l)\mathbf{x}, \quad \mathbf{Q}^l \ := \ \mathbf{H}^l \mathbf{W}^l \end{aligned} \tag{12}$$

4.1. **Extension of the urn dynamics to $\mathbb{R}^{s^l K}$.** Since $h^l$ is defined on $\mathbb{R}^{s^l K}$, while the process $\{\mathbf{Z}_n^l; n \geq 0\}$ takes values in the subset $\mathcal{S}^{s^l K}$, then applying theorems based on the SA directly to (12) may lead to improper results for the process $\mathbf{Z}_n^l$. To address this issue, we appropriately modify the dynamics (12) by replacing $h^l$ with a suitable function $f_m^l := h^l + m g^l$, where $m > 0$ is an arbitrary constant and $g^l$ is a function defined in $\mathbb{R}^{s^l K}$ that satisfies the following properties:

- (i) the derivative $\mathcal{D} g^l$ is positive semi-definite and its kernel is $Span\{(x - y) : x, y \in \mathcal{S}^{s^l K}\}$: hence, $g^l$ does not modify the eigen-structure of $\mathcal{D} h^l(\mathbf{x})$ on the subspace $\mathcal{S}^{s^l K}$, where the process $\mathbf{Z}_n^l$ is defined, while the eigen-structure outside $\mathcal{S}^{s^l K}$ can be arbitrary redefined;
- (ii) $g^l(\mathbf{z}) = 0$ for any $\mathbf{z} \in \mathcal{S}^{s^l K}$: hence, since $f_m^l(\mathbf{z}) = h^l(\mathbf{z})$ for any $\mathbf{z} \in \mathcal{S}^{s^l K}$, the modified dynamics restricted to the subset $\mathcal{S}^{s^l K}$ represents the same dynamics as in (12).

Let us now provide an analytic expression of $g^l$. First note that, since by definition of convex combination we always have $W^l \mathbf{1}_{s^l} = \mathbf{1}_{s^l}$, the left eigenvectors of $W^l$ are such that $U_1' \mathbf{1}_{s^l} = 1$ and $U_i' \mathbf{1}_{s^l} = 0$ for all $i \neq 1$. Denote by $Sp(A)$ the set of the eigenvalues of a matrix $A$ and note that, since by $(A2)$ we always have $\mathbf{1}_K' H^j = \mathbf{1}_K'$, then $Sp(W^l) \subset Sp(\mathbf{Q}^l)$ and the $s^l$ left eigenvectors of $\mathbf{Q}^l$ associated to any $\lambda_i \in Sp(W^l) \subset Sp(\mathbf{Q}^l)$, $i \in \{1, .., s^l\}$, present the following structure:

$\mathbf{U}_i := (U_{i1}\mathbf{1}_K, .., U_{is^l}\mathbf{1}_K)'$. As a consequence, for any $\mathbf{z} \in \mathcal{S}^{s^l K}$, we have $\mathbf{U}_1'\mathbf{z} = U_1'\mathbf{1}_{s^l} = 1$ and $\mathbf{U}_i'\mathbf{z} = U_i'\mathbf{1}_{s^l} = 0$ for all $i \in \{2, .., s^l\}$. Hence, denoting by $\mathbb{V}_2$ and $\mathbb{U}_2$ the matrices whose columns are $\mathbf{V}_2, .., \mathbf{V}_{s^l}$ and $\mathbf{U}_2, .., \mathbf{U}_{s^l}$, respectively, we define the function $g^l$ as follows:

$$g^l(\mathbf{x}) \ := \ \mathbf{V}_1 \left(\mathbf{U}_1'\mathbf{x} - 1\right) \ + \ \mathbb{V}_2\mathbb{U}_2'\mathbf{x}, \tag{13}$$

and the dynamics of the process $\mathbf{Z}_n^l$ in (12) can be replaced by the following:

$$\begin{aligned} \mathbf{Z}_n^l - \mathbf{Z}_{n-1}^l \quad &= \ -\frac{1}{n} f_m^l(\mathbf{Z}_{n-1}^l) \ + \ \frac{1}{n}\left(\Delta\mathbf{M}_n^l \ + \ \mathbf{R}_n^l\right), \\ f_m^l(\mathbf{x}) \quad &:= (\mathbf{I} - \mathbf{Q}^l)\mathbf{x} \ + \ m\mathbf{V}_1\left(\mathbf{U}_1'\mathbf{x} - 1\right) \ + \ m\mathbb{V}_2\mathbb{U}_2'\mathbf{x}. \end{aligned} \tag{14}$$

4.2. **First-order asymptotic results.** We now present the main convergence result concerning the limiting proportion of the urns in the leading systems.

**Theorem 4.1.** *Assume* (*A*1), (*A*2) *and* (*A*3). *Thus, for any leading system* $S^l$, $l \in \mathcal{L}_L$, *we have that*

$$\mathbf{Z}_n^l \ \xrightarrow{a.s.} \ \mathbf{Z}_\infty^l \ := \ \mathbf{V}_1, \tag{15}$$

*where* $\mathbf{V}_1$ *indicates the right-eigenvector associated to the simple eigenvalue* $\lambda = 1$ *of the matrix* $\mathbf{Q}^l$, *with* $\sum_i V_{1i} = 1$.

**Remark 4.1.** *Note that when the interacting matrix is the identity matrix, i.e.* $W = I$, $n_L = N$ *and* $n_F = 0$, *each urn represents a leading system and it evolves independently of the rest of the system. In this case,* (15) *expresses the usual result for a single GPU, where the urn proportion converges to the eigenvector associated to the maximum eigenvalue of the generating matrix, see e.g.* [4, 5, 6, 33].

**Remark 4.2.** *In Theorem 4.1, condition* (*A*3) *implies that the maximum eigenvalue* $\lambda = 1$ *of* $\mathbf{Q}^l$ *has multiplicity one, which guarantees* $\mathbf{V}_1$ *to be the unique global attractor for the system* $S^l$. *Without assumption* (*A*3), *there could be multiple attractors and hence the limiting proportions of the system would be a random variable, as in* [13, 12] *where the RRU model is considered.*

4.3. **Second-order asymptotic results.** We now establish the rate of convergence and the asymptotic distribution of the urn proportions in the leading systems $S^l$, $l \in \mathcal{L}_L$. Since to obtain these results we need to apply the Central Limit Theorem of the SA (see Theorem A.2 in Appendix) to the dynamics (14), a crucial role is played by the spectrum of the $Ks^l \times Ks^l$-matrix of the first-order derivative of $f_m^l$ defined as follows: for any $\mathbf{x} \in \mathbb{R}^{Ks^l}$

$$\mathbf{F}_m^l \ := \ \mathcal{D}f_m^l(\mathbf{x}) \ = \ (\mathbf{I} - \mathbf{Q}^l) \ + \ m\mathbf{V}_1\mathbf{U}_1' \ + \ m\mathbb{V}_2\mathbb{U}_2'. \tag{16}$$

Moreover, since the asymptotic variance depends on the second moments of the replacement matrices, we denote by $C^j(k)$ the covariance matrix of the $k^{th}$ column of $D_n^j$, i.e. $C^j(k) := \mathbf{C}ov[D_{\cdot k,n}^j]$, where $D_{\cdot k,n}^j := (D_{1k,n}^j, .., D_{Kk,n}^j)'$; note that (*A*1) ensures the existence of $C^j(k)$. Hence, denoting by $H^j(k) := E[H_{\cdot k}^j (H_{\cdot k}^j)']$ where $H_{\cdot k}^j := (H_{1k}^j, .., H_{Kk}^j)'$, we let

$$G^j \ := \ \sum_{k=1}^K \left(C^j(k) + H^j(k)\right)\tilde{Z}_{k,\infty}^j \ - \ Z_\infty^j (Z_\infty^j)', \tag{17}$$

where $\tilde{Z}_{k,\infty}^j = \sum_{i=1}^N w_{ji} Z_{k,\infty}^i$. Then, for any leading system $S^l$, $l \in \mathcal{L}_L$, we denote by $\mathbf{G}^l$ the block diagonal matrix made by the $s^l$ blocks $G^{r^{l^-}+1}, .., G^{r^l}$.

The following theorem shows the rate of convergence and the limiting distribution of the urn proportions in the leading systems.

**Theorem 4.2.** *Assume (A1), (A′2) and (A3). For any leading system $S^l$, $l \in \mathcal{L}_L$, let $\lambda^{*l}$ be the maximum eigenvalue in $Sp(\mathbf{Q}^l) \setminus Sp(W^l)$. Thus, we have that $\lambda^{*l} \equiv 1 - \min Sp(\mathbf{F}^l_m)$ and*

(a) *if $\lambda^{*l} < 1/2$, then*

$$\sqrt{n}(\mathbf{Z}^l_n - \mathbf{Z}^l_\infty) \xrightarrow{d} \mathcal{N}\left(0, \Sigma^l\right), \qquad \Sigma^l := \lim_{m\to\infty} \int_0^\infty e^{u(\frac{\mathbf{I}}{2} - \mathbf{F}^l_m)} \mathbf{G}^l e^{u(\frac{\mathbf{I}}{2} - \mathbf{F}^l_m)'} du.$$

(b) *if $\lambda^{*l} = 1/2$, then*

$$\sqrt{\frac{n}{\log(n)}}(\mathbf{Z}^l_n - \mathbf{Z}^l_\infty) \xrightarrow{d} \mathcal{N}\left(0, \Sigma^l\right).$$

(c) *if $\lambda^{*l} > 1/2$, then there exists a finite random variable $\psi^l$ such that*

$$n^{1-\lambda^{*l}}(\mathbf{Z}^l_n - \mathbf{Z}^l_\infty) \xrightarrow{a.s.} \psi^l.$$

**Remark 4.3.** *When the interacting matrix $W$ is the identity matrix, each urn represents a leading system and hence $W^l = 1$ and $\mathbf{Q}^l \equiv H^l$. In that case, $\lambda^*$ is the second largest eigenvalue of the generating matrix $H^l$ and hence Theorem 4.2 expresses the usual Central Limit Theorem for a single GPU, see e.g.* [4, 5, 6, 33].

**Remark 4.4.** *The role of $\mathbf{Q}^l$ in Theorem 4.2 shows that the convergence rate of the urns in $S^l$ does not depend only on their generating matrices $\{H^j, r^{l^-} + 1 \leq j \leq r^l\}$ but also on their interaction expressed in $W^l$. For instance, consider two single GPUs whose generating matrices $H^1$ and $H^2$ are such that the convergence rates of the urn proportions $Z^1_n$ and $Z^2_n$ are different. Then, an interaction between these urns with an irreducible $W^l$ would make $Z^1_n$ and $Z^2_n$ converge at the same rate that would depend on the choice of $W^l$.*

## 5. Following systems

In this section we establish asymptotic properties concerning the following systems $S^l$, $l \in \mathcal{L}_F$. Since the dynamics of these systems can be properly expressed in the SA form (5) only through a joint model with the urns in the systems $\{S^{L_1}, .., S^l\}$, we need a special notation to study collections of more systems. In particular, we will replace the label $l$ with $(l)$ whenever an object is referred to the joint system $S^{(l)} := \{S^{L_1}, .., S^l\}$ instead of to the single system $S^l$. For instance, the vector $\mathbf{Y}^{(l)}_n \in \mathbb{R}^{Kr^l}$ indicates $(\mathbf{Y}^{L_1}_n, .., \mathbf{Y}^l_n)'$, and $\mathbf{D}^{(l)}_n$ indicates the block diagonal $(Kr^l \times Kr^l)$-matrix, whose blocks are made by $\mathbf{D}^{L_1}_n, .., \mathbf{D}^l_n$. Then, from (4) we can express the sampling probabilities in the follower $S^l$ as follows:

$$\tilde{\mathbf{Z}}^l_{n-1} = \sum_{i \in \{L_1, .. l^-\}} \mathbf{W}^{il} \mathbf{Z}^i_{n-1} + \mathbf{W}^l \mathbf{Z}^l_{n-1}.$$

Hence, from (11) we obtain

$$(18) \qquad \begin{aligned} \mathbf{Z}^l_n - \mathbf{Z}^l_{n-1} &= -\frac{1}{n} h^l(\mathbf{Z}^{(l^-)}_{n-1}, \mathbf{Z}^l_{n-1}) + \frac{1}{n}\left(\Delta\mathbf{M}^l_n + \mathbf{R}^l_n\right), \\ h^l(\mathbf{x}_1, \mathbf{x}_2) &:= -\mathbf{Q}^{l(l^-)}\mathbf{x}_1 + (\mathbf{I} - \mathbf{Q}^l)\mathbf{x}_2, \\ \mathbf{Q}^{l(l^-)} &:= \left[\mathbf{H}^l \mathbf{W}^{L_1 l} \ldots \mathbf{H}^l \mathbf{W}^{l^- l}\right], \quad \mathbf{Q}^l := \mathbf{H}^l \mathbf{W}^l \end{aligned}$$

Since $h^l$ is not only a function of $\mathbf{Z}^l_{n-1}$, the dynamics in (18) is not already expressed in the SA form (5). To address this issue, we need to consider a joint model for the global system $S^{(l)} = S^{(l^-)} \cup S^l = S^{L_1} \cup .. \cup S^l$ as follows:

$$\begin{aligned} \mathbf{Z}^{(l)}_n - \mathbf{Z}^{(l)}_{n-1} &= -\frac{1}{n} h^{(l)}(\mathbf{Z}^{(l)}_{n-1}) + \frac{1}{n}\left(\Delta \mathbf{M}^{(l)}_n + \mathbf{R}^{(l)}_n\right), \\ h^{(l)}(\mathbf{x}) &:= \left(\mathbf{I} - \mathbf{Q}^{(l)}\right)\mathbf{x}, \end{aligned} \tag{19}$$

where $\mathbf{Q}^{(l)}$ can be recursively defined as follows:

$$\mathbf{Q}^{(l)} := \begin{bmatrix} \mathbf{Q}^{(l^-)} & 0 \\ \mathbf{Q}^{(l^-)l} & \mathbf{Q}^l \end{bmatrix}, \quad \mathbf{Q}^{(L_{n_L})} := \begin{bmatrix} \mathbf{Q}^{L_1} & .. & 0 \\ ... & ... & ... \\ 0 & .. & \mathbf{Q}^{L_{n_L}}, \end{bmatrix}, \tag{20}$$

where by convention $F_1^- = L_{n_L}$.

5.1. **Extension of the urn dynamics to $\mathbb{R}^{r^lK}$.** Since $h^{(l)}$ in (19) is defined in $\mathbb{R}^{r^lK}$, while the process $\{\mathbf{Z}^{(l)}_n; n \geq 0\}$ lies in the subspace $\mathcal{S}^{r^lK}$, then applying theorems based on the SA directly to (18) may lead to improper results for the process $\mathbf{Z}^{(l)}_n$.

For this reason, we replace $h^{(l)}$ in (18) with a suitable function $f^{(l)}_m := h^{(l)} + mg^{(l)}$ such that $m > 0$ is an arbitrary constant and $g^{(l)}$ is a function defined as in (13), where in this case $\{\mathbf{U}_i; 1 \leq i \leq r^l\}$ and $\{\mathbf{V}_i; 1 \leq i \leq r^l\}$, indicate, respectively, the left and right eigenvectors of $\mathbf{Q}^{(l)}$. Hence, the dynamics of the process $\mathbf{Z}^{(l)}_n$ (18) is replaced by the following:

$$\begin{aligned} \mathbf{Z}^{(l)}_n - \mathbf{Z}^{(l)}_{n-1} &= -\frac{1}{n} f^{(l)}_m(\mathbf{Z}^{(l)}_{n-1}) + \frac{1}{n}\left(\Delta \mathbf{M}^{(l)}_n + \mathbf{R}^{(l)}_n\right), \\ f^{(l)}_m(\mathbf{x}) &:= \left(\mathbf{I} - \mathbf{Q}^{(l)}\right)\mathbf{x} + m\mathbf{V}_1\left(\mathbf{U}_1'\mathbf{x} - 1\right) + m\mathbb{V}_2\mathbb{U}_2'\mathbf{x}. \end{aligned} \tag{21}$$

Note that in the joint system $S^{(l)}$ the eigenvalue $\lambda = 1$ of $\mathbf{Q}^{(l)}$ may not have multiplicity one; in that case, $\mathbf{V}_1$ is univocally identified as the right eigenvector of $\mathbf{Q}^{(l)}$ associated to $\lambda = 1$ such that, letting $\mathbf{U}_i := (U_{i1}\mathbf{1}_K, .., U_{ir^l}\mathbf{1}_K)'$ and $U_i'W^{(l)} = \lambda_i U_i'$ for any $i \in \{1, .., r^l\}$, we have $\mathbf{U}_1'\mathbf{V}_1 = U_1'\mathbf{1}_{r^l} = 1$ and $\mathbf{U}_i'\mathbf{V}_1 = U_i'\mathbf{1}_{r^l} = 0$ when $i \neq 1$.

5.2. **Removal of unnecessary components.** The following system $S^l$ may not depend on all the components of $S^{(l^-)}$ and hence the convergence in $S^l$ may be faster than the rate in $S^{(l^-)}$. When this occurs, the asymptotic distribution obtained for the urn proportions in $S^{(l)}$ restricted to the urns in $S^l$ is degenerate. To address this issue and characterize the asymptotic behavior in the following system $S^l$, we need to reduce the dimensionality of $\mathbf{Z}^{(l)}_n$ by deleting those components which do not influence the dynamics of $\mathbf{Z}^l_n$. Since the interaction between $S^l$ and the systems in $S^{(l^-)}$ is expressed by $\mathbf{Q}^{(\mathbf{l})\mathbf{l}^-}$, we exclude the components of $\mathbf{Z}^{(l^-)}_n$ defined on the Kernel of $\mathbf{Q}^{(l^-)l}$. Formally, consider the following decomposition:

$$Sp(\mathbf{Q}^{(\mathbf{l})}) = Sp(\mathbf{Q}^{\mathbf{l}}) \cup Sp(\mathbf{Q}^{(\mathbf{l}^-)}) = \mathcal{A}_{IN} \cup \mathcal{A}_{OUT},$$

where

$$\begin{aligned} \mathcal{A}_{OUT} &:= \left\{ \lambda \in Sp(\mathbf{Q}^{(\mathbf{l}^-)}) : \exists v\{\mathbf{Q}^{(\mathbf{l}^-)}v = \lambda v\} \cap \{\mathbf{Q}^{(\mathbf{l}^-)\mathbf{l}}v = 0\} \right\} \\ \mathcal{A}_{IN} &:= Sp(\mathbf{Q}^{\mathbf{l}}) \cup \left(Sp(\mathbf{Q}^{(\mathbf{l}^-)}) \setminus \mathcal{A}_{OUT}\right). \end{aligned}$$

Then, the eigenspace of $\mathbf{Q^{(l)}}$ associated to $\lambda \in \mathcal{A}_{OUT}$ will be removed from the dynamics in (21). To do this, let us denote by:

(1) $\mathbf{U}_{IN}$ and $\mathbf{V}_{IN}$ the matrices whose columns are the left and right eigenvectors of $\mathbf{Q^{(l)}}$, respectively, associated to eigenvalues in $\mathcal{A}_{IN}$;
(2) $\mathbf{U}_{OUT}$ and $\mathbf{V}_{OUT}$ the matrices whose columns are the left and right eigenvectors of $\mathbf{Q^{(l)}}$, respectively, associated to eigenvalues in $\mathcal{A}_{OUT}$;

Since we do not want to modify the process $\mathbf{Z}_n^{(l)}$ on $S^l$, i.e. $\mathbf{Z}_n^l$, we now construct two conjugate basis in $\mathcal{I}m(\mathbf{U}_{IN})$ and $\mathcal{I}m(\mathbf{V}_{IN})$ that are invariant on $S^l$. Note that, since $Sp(\mathbf{Q^l}) \subset \mathcal{A}_{IN}$, there exists a non-singular matrix $\mathbf{P}$ such that the following decompositions hold:

$$\mathbf{B} := \mathbf{V}_{IN}\mathbf{P} = \begin{bmatrix} \hat{\mathbf{B}} & 0 \\ 0 & \mathbf{I}, \end{bmatrix}, \qquad \mathbf{C} := \mathbf{P}^{-1}\mathbf{U}_{IN}' = \begin{bmatrix} \hat{\mathbf{C}} & 0 \\ 0 & \mathbf{I}, \end{bmatrix}.$$

Since $\hat{\mathbf{C}}'\hat{\mathbf{B}} = \mathbf{I}$ and $\hat{\mathbf{B}}\hat{\mathbf{C}}' = \mathbf{V_{IN}U_{IN}}'$, $\hat{\mathbf{C}}$ and $\hat{\mathbf{B}}$ represent conjugate basis in $\mathcal{I}m(\mathbf{U}_{IN})$ and $\mathcal{I}m(\mathbf{V}_{IN})$, respectively. Thus, for any $\mathbf{x} = (\mathbf{x}^{(l^-)}, \mathbf{x}^l)' \in \mathbb{R}^{Kr^l}$, we have the following decomposition

$$\mathbf{x} \;=\; \mathbf{V_{IN}U_{IN}}'\mathbf{x} \;+\; \mathbf{V_{OUT}U_{OUT}}'\mathbf{x} \;=\; \hat{\mathbf{B}}\hat{\mathbf{x}} \;+\; \mathbf{V_{OUT}x_{OUT}}, \tag{22}$$

where

$$\hat{\mathbf{x}} := \hat{\mathbf{C}}'\mathbf{x} = \begin{bmatrix} \mathbf{C}'\mathbf{x}^{(l^-)} \\ \mathbf{x}^l \end{bmatrix}, \quad \mathbf{x_{OUT}} := \mathbf{U_{OUT}}'\mathbf{x}.$$

In particular, we consider the process $\{\hat{\mathbf{Z}}_n^{(l)}, n \geq 1\}$ defined as follows:

$$\hat{\mathbf{Z}}_n^{(l)} := \hat{\mathbf{C}}'\mathbf{Z}_n^{(l)} = \begin{bmatrix} \mathbf{C}'\mathbf{Z}_n^{(l^-)} \\ \mathbf{Z}_n^l \end{bmatrix}; \tag{23}$$

now, multiplying by $\hat{\mathbf{C}}'$ to (21) and applying the decomposition (22) in (21), since $\hat{\mathbf{C}}'\mathbb{V}_2\mathbb{U}_2'\mathbf{V_{OUT}} = 0$, $\mathbf{U}_1'\mathbf{V_{OUT}} = 0$ and $\hat{\mathbf{C}}'\mathbf{V_{OUT}} = 0$, we have that

$$\begin{aligned} \hat{\mathbf{Z}}_n^{(l)} - \hat{\mathbf{Z}}_{n-1}^{(l)} \quad &= \; -\frac{1}{n}\hat{f}_m^{(l)}(\hat{\mathbf{Z}}_{n-1}^{(l)}) \;+\; \frac{1}{n}\hat{\mathbf{C}}'\left(\Delta\mathbf{M}_n^{(l)} \;+\; \mathbf{R}_n^{(l)}\right), \\ \hat{f}_m^{(l)}(\hat{\mathbf{x}}) \quad &:= \; \left(\mathbf{I} - \hat{\mathbf{C}}'\mathbf{Q}^{(l)}\hat{\mathbf{B}}\right)\hat{\mathbf{x}} \;+\; m\hat{\mathbf{V}}_1\left(\hat{\mathbf{U}}_1'\hat{\mathbf{x}} - 1\right) \;+\; m\hat{\mathbb{V}}_2\hat{\mathbb{U}}_2'\hat{\mathbf{x}}, \end{aligned} \tag{24}$$

where $\hat{\mathbf{U}}_1' := \mathbf{U}_1'\hat{\mathbf{B}}$, $\hat{\mathbb{U}}_2 := \mathbb{U}_2'\hat{\mathbf{B}}$, $\hat{\mathbf{V}}_1 := \hat{\mathbf{C}}'\mathbf{V}_1$ and $\hat{\mathbb{V}}_2 := \hat{\mathbf{C}}'\mathbb{V}_2$ represent the left and right eigenvectors of $\hat{\mathbf{C}}'\mathbf{Q}^{(l)}\hat{\mathbf{B}}$ associated to $\lambda \in Sp(W^{(l)}) \setminus \mathcal{A}_{OUT}$. Since $\hat{f}_m^{(l)}$ is a function of $\hat{\mathbf{Z}}_n^{(l)}$, the dynamics in (24) is now expressed in the SA form (5).

**Remark 5.1.** *It is worth highlighting that the interacting matrix $W$ lonely is not enough to individuate the components of the system that actually influence a following system, but it is necessary to study the eigen-structure of $\mathbf{Q}^{(l)}$, that joins the information of $W$ and of the generating matrices $\{H^j, 1 \leq j \leq r^l\}$ of the urns in $S^{(l)}$. This may be surprising since $W$ is the only element that defines the interaction among the urns in the system. Nevertheless, when $H^j$ is singular, different values of $\tilde{Z}_n^j$ may give the same average replacements, $H^j\tilde{Z}_n^j$, which is equivalent as having singularities in $W$, where different values of $\{Z_n^i; 1 \leq i \leq r^l\}$ may give the same $\tilde{Z}_{k,n}^j$, and hence same $H^j\tilde{Z}_n^j$. For instance, if all the columns of $H^j$ were equal to a given vector $v^j$, the urn $j$ would be updated on average by $v^j$ regardless the value of $\tilde{Z}_{n-1}^j$ and hence the urns in $S^{(l^-)}$ would not play any role in the dynamics of the urn $j$ for any choice of $W$. The eigen-structure of $\mathbf{Q}^{(l)}$ perfectly* *explains this behavior, since in this case the matrix* $\mathbf{Q}^{(l^-)l}$ *would be composed by all zeros and hence* $Sp(\mathbf{Q}^{(l^-)}) \equiv A_{OUT}$ *and* $Sp(\mathbf{Q}^l) \equiv A_{IN}$.

5.3. **First-order asymptotic results.** We now present the convergence result concerning the limiting proportion of the urns in the following systems. The asymptotic behavior of $\mathbf{Z}_n^{(l)}$ is obtained recursively from $\mathbf{Z}_\infty^{(l^-)} := a.s. - \lim_{n\to\infty} \mathbf{Z}_n^{(l-)}$.

**Theorem 5.1.** *Assume* (*A*1), (*A*2) *and* (*A*3). *Thus, for any* $l \in \mathcal{L}_F$, *we have that*

$$\hat{\mathbf{Z}}_n^{(l)} \overset{a.s.}{\longrightarrow} \hat{\mathbf{Z}}_\infty^{(l)} := \hat{\mathbf{V}}_1;$$

*hence, from* (23), *in the following system* $S^l$ *we have that*

$$\mathbf{Z}_n^l \overset{a.s.}{\longrightarrow} \mathbf{Z}_\infty^l := \left(\mathbf{I} - \mathbf{Q}^l\right)^{-1} \mathbf{Q}^{(l^-)l} \mathbf{Z}_\infty^{(l^-)}.$$

5.4. **Second-order asymptotic results.** We now present the results concerning the rate of convergence and the asymptotic distribution of the urn proportions in the following systems. To this end, let us introduce the $Ks^l \times Ks^l$-matrix of the first-order derivative of $\hat{f}_m^l$:

$$\begin{aligned} \hat{\mathbf{F}}_m^{(l)} &:= \hat{\mathbf{C}}'\mathbf{F}_m^{(l)}\hat{\mathbf{B}} \\ &= (\mathbf{I} - \hat{\mathbf{C}}'\mathbf{Q}^{(l)}\hat{\mathbf{B}}) + m\hat{\mathbf{V}}_1\hat{\mathbf{U}}_1' + m\hat{\mathbb{V}}_2\hat{\mathbb{U}}_2'. \end{aligned} \tag{25}$$

Moreover, the asymptotic variance will be based on the quantity $\hat{\mathbf{G}}^{(l)} := \hat{\mathbf{C}}'\mathbf{G}^{(l)}\hat{\mathbf{B}}$, where $\mathbf{G}^{(l)}$ is the block diagonal matrix made by $G^1, .., G^{r^l}$ (see (17)).

The following theorem shows the rate of convergence and the limiting distribution of the urn proportions in the following systems.

**Theorem 5.2.** *Assume* (*A*1), (*A*′2) *and* (*A*3). *For any following system* $S^l$, $l \in \mathcal{L}_F$, *let* $\lambda^{*l}$ *be the maximum eigenvalue in* $Sp(\mathbf{Q}^{(l)}) \setminus (Sp(W^{(l)}) \cup \mathcal{A}_{OUT})$. *Thus, we have that* $\lambda^{*l} \equiv 1 - \min Sp(\hat{\mathbf{F}}_m^{(l)})$ *and*

(a) *if* $\lambda^{*l} < 1/2$, *then*

$$\sqrt{n}(\hat{\mathbf{Z}}_n^{(l)} - \hat{\mathbf{Z}}_\infty^{(l)}) \overset{d}{\longrightarrow} \mathcal{N}\left(0, \hat{\Sigma}^{(l)}\right), \qquad \hat{\Sigma}^{(l)} := \lim_{m\to\infty} \int_0^\infty e^{u(\frac{\mathbf{I}}{2} - \hat{\mathbf{F}}_m^{(l)})} \hat{\mathbf{G}}^{(l)} e^{u(\frac{\mathbf{I}}{2} - \hat{\mathbf{F}}_m^{(l)})'} du.$$

(b) *if* $\lambda^{*l} = 1/2$, *then*

$$\sqrt{\frac{n}{\log(n)}}(\hat{\mathbf{Z}}_n^{(l)} - \hat{\mathbf{Z}}_\infty^{(l)}) \overset{d}{\longrightarrow} \mathcal{N}\left(0, \Sigma^{(l)}\right).$$

(c) *if* $\lambda^{*l} > 1/2$, *then there exists a finite random variable* $\psi^{(l)}$ *such that*

$$n^{1-\lambda^{*l}}(\hat{\mathbf{Z}}_n^{(l)} - \hat{\mathbf{Z}}_\infty^{(l)}) \overset{a.s.}{\longrightarrow} \psi^{(l)}.$$

**Remark 5.2.** *It is worth noticing that, since from* (23) $\hat{\mathbf{Z}}_n^{(l)} = (\mathbf{C}'\mathbf{Z}_n^{(l^-)}, \mathbf{Z}_n^l)'$, *Theorem 5.2 explicitly states the limiting distribution and the asymptotic covariance structure of the urn proportions in any following system* $\mathbf{Z}_n^l$, $l \in \mathcal{L}_F$. *In addition, Theorem 5.2 also determines the correlation between* $\mathbf{Z}_n^l$ *and the components of the urn proportions in the other systems* $S^l$, $l \in \{L_1, .., l^-\}$, *that actually influence the dynamics of* $\mathbf{Z}_n^l$.

**Remark 5.3.** *We highlight that condition* (*A*3), *i.e. irreducibility of the generating matrices* $H^j$, *may be relaxed in Theorems 5.1 and 5.2, by requiring* (*A*3) *only for the urns in the leading systems. In fact, we can note from the proof that* (*A*3) *is not needed for the urns that belong to the following systems.*

## 6. Further extensions

In this section, we discuss some possible extensions of the interacting urn model presented in this paper.

6.1. **Random and time-dependent interacting matrix.** Although we consider a constant interacting matrix $W$, the results of this paper may be extended to a system characterized by a random sequence of interacting matrices $\{W_n; n \geq 0\}$, i.e. $W_n = [w_{jh,n}] \in \mathcal{F}_n$ and $Z^j_{k,n} = \sum_{h=1}^N w_{jh,n} Z^h_{k,n}$ for any $k \in \{1,..,K\}$. In that case, it is essential to assume the existence of a deterministic matrix $W$ such that $W_n \xrightarrow{a.s.} W$, which individuates the leading and the following systems, as in Subsection 2.4.

The dynamics with random and time-dependent interacting matrices could be also expressed in the SA form (5), by including the difference $(W_n - W)$ in the remainder term (10). Naturally, the asymptotic behavior of the urn proportions would depend on the limiting interacting matrix $W$ and on the rate of convergence of the sequence $\{W_n; n \geq 0\}$. Specifically, the convergence of the urn proportions could be obtained with the only assumption $W_n \xrightarrow{a.s.} W$, while extensions for the second-order results presented in would require $n\mathbf{E}[\|W_n - W\|^2] \to 0$.

6.2. **Non-homogeneous generating matrices.** The independence and identically distribution of the replacement matrices is an assumption that could be relaxed by assuming that the sequence of generating matrices $\{H^j_n; n \geq 0\}$, $H^j_{n-1} := \mathbf{E}[D^j_n|\mathcal{F}_{n-1}]$, converges to some deterministic matrix $H^j$. Thus, the urn dynamics could be expressed in the SA form (5), by including the difference $(H^j_n - H^j)$ in the remainder term (10), and the asymptotic behavior would depend on $H^j$ and on the rate of convergence of $H^j_n$. Specifically, the second-order results would require an additional assumption as $n\mathbf{E}[\|H^j_n - H^j\|^2] \to 0$.

## 7. Proofs

The proof of Theorem 2.1 requires the following auxiliary result on the martingale convergence:

**Lemma 7.1.** *Let $\{S_n; n \geq 1\}$, $S_n := \sum_{i=1}^n \Delta S_i$, be a zero-mean martingale with respect to a filtration $\{\mathcal{F}_n; n \geq 1\}$ and let $\{b_n; n \geq 1\}$ be a non-decreasing sequence of positive numbers such that*

$$\sum_{i=1}^{\infty} b_i^{-2}\mathbf{E}[(\Delta S_i)^2|\mathcal{F}_{i-1}] \ < \ \infty, \qquad a.s. \tag{26}$$

*Then, $b_n^{-1}S_n \xrightarrow{a.s.} 0$.*

*Proof.* Let us define the zero-mean martingale $\tilde{S}_n := \sum_{i=1}^n b_i^{-1}\Delta S_i$, where by (26) we have that $\tilde{S}_n$ converges a.s. Thus, the result follows by using Kronacker's Lemma (see Lemma IV.3.2 in [32]). □

*Proof of Theorem 2.1.* By using Lemma 7.1 with $b_n := n^{1-\alpha}$ and $S_n := T^j_n - n$, the proof follows by showing that $T^j_n - n$ is a martingale whose increments have bounded second moments. Now, since

$$T^j_n - T^j_{n-1} \ = \ \sum_{k=1}^K (Y^j_{k,n} - Y^j_{k,n-1}) \ = \ \sum_{k=1}^K \sum_{i=1}^K (D^j_{ki,n} X^j_{i,n}),$$

the result follows by establishing that

(a) $\sup_{n\geq 1} \mathbf{E}\left[\left(\sum_{k=1}^K \sum_{i=1}^K D^j_{ki,n} X^j_{i,n}\right)^2 \ |\mathcal{F}_{n-1}\right] < \infty$;

(b) $\sum_{k=1}^{K}\sum_{i=1}^{K}\mathbf{E}\left[D^j_{ki,n}X^j_{i,n}|\mathcal{F}_{n-1}\right]=1$.

For part (a), by using $|X^j_{i,n}|\leq 1$ and (A1), we have that

$$\sup_{n\geq 1}\mathbf{E}\left[\left(\sum_{k=1}^{K}\sum_{i=1}^{K}(D^j_{n,ki}X^j_{i,n})\right)^2 |\mathcal{F}_{n-1}\right]\ \leq\ K^2\sup_{n\geq 1}\max_{j\in\{1,..,N\}}\max_{i,k\in\{1,..,K\}}\mathbf{E}\left[(D^j_{ki,n})^2\right]\ <\ \infty.$$

For part (b), since $\sum_{k=1}^{K}H^j_{ki}=1$ by (A2) and since $D^j_{ki,n}$ and $X^j_{i,n}$ are independent conditionally on $\mathcal{F}_{n-1}$, we obtain

$$\sum_{k=1}^{K}\sum_{i=1}^{K}\mathbf{E}\left[D^j_{ki,n}X^j_{i,n}\ |\mathcal{F}_{n-1}\right]=\ \sum_{k=1}^{K}\sum_{i=1}^{K}H^j_{ki}\tilde{Z}^j_{i,n-1}\ =\ \sum_{i=1}^{K}\tilde{Z}^j_{i,n-1}\sum_{k=1}^{K}H^j_{ki}\ =\ \sum_{i=1}^{K}\tilde{Z}^j_{i,n-1}.$$

Finally, by the definition of $\tilde{Z}^j_{i,n-1}$ in (1), we have

$$\sum_{i=1}^{K}\tilde{Z}^j_{i,n-1}\ =\ \sum_{i=1}^{K}\sum_{h=1}^{N}w_{jh}Z^h_{i,n-1}\ =\ \sum_{h=1}^{N}w_{jh}\sum_{i=1}^{K}Z^h_{i,n-1}\ =\ \sum_{h=1}^{N}w_{jh}\ =\ 1,$$

which concludes the proof of (3) for $\alpha<1/2$ under assumption (A2).

Concerning the proof of (3) for $\alpha<1$, first note that under assumption (A'2) we have

$$T^j_n-T^j_{n-1}\ =\ \sum_{k=1}^{K}\sum_{i=1}^{K}(D^j_{n,ki}X^j_{i,n})\ =\ \sum_{i=1}^{K}X^j_{i,n}\ =\ 1;$$

hence, $T^j_n=T^j_0+n$ a.s. and, for any $\alpha<1$,

$$n^{\alpha}\left(\frac{T^j_n}{n}-1\right)\ =\ \frac{T^j_0}{n^{1-\alpha}}\ \overset{a.s./L^2}{\longrightarrow}\ 0.$$

□

### 7.1. **Proofs on the leading systems.**

*Proof of Theorem 4.1.* Fix $l\in\mathcal{L}_L$ and consider the leading system $S^l=\{r^{l^-}+1\leq j\leq r^l\}$ with interacting matrix $W^l$. Since the dynamic of the urn proportions $\mathbf{Z}^l_n$ in $S^l$ has been expressed in (14) in the SA form (5), we can establish the convergence result stated in Theorem 4.1 by applying Theorem A.1 in Appendix. To this end, we will show that the assumptions of Theorem A.1 are satisfied by the process $\{\mathbf{Z}^l_n;n\geq 1\}$ of the system $S^l$:

(1) the function $f^l_m$ defined in (14) is a linear transformation and hence locally Lipschitz.
(2) from (8), we have that $\sup_{n\geq 1}\mathbf{E}\left[\left\|\Delta\mathbf{M}^l_n\right\|^2|\mathcal{F}_{n-1}\right]<\infty$ is satisfied by establishing
  (2a) $\sup_{n\geq 1}\mathbf{E}\left[\left\|\mathbf{D}^l_n\mathbf{X}^l_n\right\|^2|\mathcal{F}_{n-1}\right]<\infty$;
  (2b) $\sup_{n\geq 1}\mathbf{E}\left[\left\|diag(\mathbf{T}^l_n-\mathbf{T}^l_{n-1})\mathbf{Z}^l_{n-1}\right\|^2|\mathcal{F}_{n-1}\right]<\infty$.

  Concerning (2a), since $X^j_{k,n}\in\{0,1\}$ a.s., we have that

$$\left\|\mathbf{D}^l_n\mathbf{X}^l_n\right\|^2\ \leq\ \sum_{j\in S^l}\sum_{k=1}^{K}\sum_{i=1}^{K}\left(D^j_{ki,n}\right)^2,\quad a.s.$$

Thus, (2a) follows by assumption (A1), since

$$\sup_{n\geq 1}\mathbf{E}\left[\left\|\mathbf{D}^l_n\mathbf{X}^l_n\right\|^2|\mathcal{F}_{n-1}\right] \;\leq\; \sum_{j\in S^l}\sum_{k=1}^K\sum_{i=1}^K\sup_{n\geq 1}\mathbf{E}\left[\left(D^j_{ki,n}\right)^2\right] \;\leq\; s^lK^2C_\delta.$$

Concerning (2b), since $\sum_{k=1}^K(Z^j_{k,n})^2\leq 1$, we have

$$\left\|diag(\mathbf{T}^l_n-\mathbf{T}^l_{n-1})\mathbf{Z}^l_{n-1}\right\|^2 \;\leq\; \sum_{j\in S^l}(T^j_n-T^j_{n-1})^2,\quad a.s. \tag{27}$$

where we recall that

$$T^j_n-T^j_{n-1} \;=\; \sum_{k=1}^K(Y^j_{k,n}-Y^j_{k,n-1}) \;=\; \sum_{k=1}^K\sum_{i=1}^K(D^j_{ki,n}X^j_{i,n}). \tag{28}$$

Hence, combining (27) and (28), since $X^j_{i,n}\in\{0,1\}$ and $\sum_{i=1}^K X^j_{i,n}=1$ a.s., we obtain that

$$\begin{aligned}\left\|diag(\mathbf{T}^l_n-\mathbf{T}^l_{n-1})\mathbf{Z}^l_{n-1}\right\|^2 &\leq \sum_{j\in S^l}\left(\sum_{k=1}^K\sum_{i=1}^K(D^j_{ki,n}X^j_{i,n})\right)^2\\ &\leq \sum_{j\in S^l}\sum_{i=1}^K\left(\sum_{k=1}^K D^j_{ki,n}\right)^2,\quad a.s.\end{aligned}$$

Finally, using the relation $(\sum_{k=1}^K a_k^2\leq K^2\sum_{k=1}^K a_k^2$, (2b) follows by assumption (A1), since

$$\sup_{n\geq 1}\mathbf{E}\left[\left\|diag(\mathbf{T}^l_n-\mathbf{T}^l_{n-1})\mathbf{Z}^l_{n-1}\right\|^2|\mathcal{F}_{n-1}\right] \;\leq\; \sup_{n\geq 1}\sum_{j\in S^l}\sum_{i=1}^K K^2\sum_{k=1}^K\mathbf{E}\left[\left(D^j_{ki,n}\right)^2\right] \;\leq\; s^lK^4C_\delta.$$

(3) from (10), we show $\|\mathbf{R}^l_n\|\xrightarrow{a.s.}0$ by establishing that, for any $(2+\delta)^{-1}<\alpha<2^{-1}$,
 (3a) $n^\alpha\left\|n\cdot diag(\mathbf{T}^l_n)^{-1}-\mathbf{I}\right\|\xrightarrow{a.s.}0$,
 (3b) $n^{-\alpha}\left\|\mathbf{Z}^l_{n-1}-\mathbf{H}^l\tilde{\mathbf{Z}}^l_{n-1}\right\|\xrightarrow{a.s.}0$,
 (3c) $n^{-\alpha}\left\|\Delta\mathbf{M}^l_n\right\|\xrightarrow{a.s.}0$,

where we recall that $\delta>0$ is defined in assumption (A1) (see Subsection 2.2). Since (3a) follows straightforwardly by Theorem 2.1, consider (3b). For any $\epsilon>0$, using Markov's inequality we obtain

$$\mathbf{P}\left(\left\|\mathbf{Z}^l_{n-1}-\mathbf{H}^l\tilde{\mathbf{Z}}^l_{n-1}\right\|>\epsilon n^\alpha\right) \;\leq\; (\epsilon n^\alpha)^{-(2+\delta)}\mathbf{E}\left[\left\|\mathbf{Z}^l_{n-1}-\mathbf{H}^l\tilde{\mathbf{Z}}^l_{n-1}\right\|^{(2+\delta)}\right].$$

Hence, (3b) follows by Borel-Cantelli Lemma since $\alpha\cdot(2+\delta)>1$ and

$$\sup_{n\geq 0}\mathbf{E}\left[\left\|\mathbf{Z}^l_{n-1}-\mathbf{H}^l\tilde{\mathbf{Z}}^l_{n-1}\right\|^{(2+\delta)}\right] \;\leq\; \sum_{j\in S^l}2^{(2+\delta)} \;<\; \infty.$$

Concerning (3c), we can apply again Markov's inequality and the same arguments of part (3b) since by assumption (A1) we have that

$$\sup_{n\geq 0}\mathbf{E}\left[\left\|\mathbf{D}^l_n\mathbf{X}^l_n-\mathbf{H}^l\tilde{\mathbf{Z}}^l_{n-1}\right\|^{(2+\delta)}\right] \;\leq\; \sup_{n\geq 0}\sum_{j\in S^l}\sum_{k=1}^K\sum_{i=1}^K\mathbf{E}\left[(D^l_{ki,n})^{(2+\delta)}\right] \;<\; \infty.$$

Thus, by applying Theorem A.1 to the dynamics in (14), we have that the limiting values of $\mathbf{Z}_n^l$ are included in the set

$$\left\{ \mathbf{x} \in \mathbb{R}^{Ks_l} \ : \ f_m^l(\mathbf{x}) = 0 \right\}.$$

Now, denote by $\mathbb{V}_3$ and $\mathbb{U}_3$ the matrices whose columns are, respectively, the right and left eigenvectors of $\mathbf{Q}^l$ associated to the eigenvalues $\lambda \in Sp(\mathbf{Q}^l) \setminus Sp(W^l)$. Hence, we have the following decomposition

$$\mathbf{Q}^l \ = \ \mathbf{V}_1\mathbf{U}_1' \ + \ \mathbb{V}_2\mathbf{J}_2\mathbb{U}_2' \ + \ \mathbb{V}_3\mathbf{J}_3\mathbb{U}_3', \tag{29}$$

where $\mathbf{J}_2$ and $\mathbf{J}_3$ represent the corresponding jordan blocks. Since the eigenvectors of $\mathbf{Q}^l$ represent a basis of $\mathbb{R}^{Ks_l}$, for any $\mathbf{x} \in \mathbb{R}^{Ks_l}$ there exists $a \in \mathbb{R}$, $b \in \mathbb{R}^{s^l-1}$ and $c \in \mathbb{R}^{s^l(K-1)}$ such that

$$\mathbf{x} \ = \ \mathbf{V}_1 a \ + \ \mathbb{V}_2 b \ + \ \mathbb{V}_3 c. \tag{30}$$

Hence, by using (29) and (30), we obtain

$$\begin{aligned} h^l(\mathbf{x}) \quad &= \mathbb{V}_2(\mathbf{I} - \mathbf{J}_2)b \ + \ \mathbb{V}_3(\mathbf{I} - \mathbf{J}_3)c, \\ g^l(\mathbf{x}) \quad &= \mathbf{V}_1(a-1) \ + \ \mathbb{V}_2 b, \end{aligned}$$

which, since $f_m^l(\mathbf{x}) = h^l(\mathbf{x}) + mg^l(\mathbf{x})$, it gives us

$$f_m^l(\mathbf{x}) \ = \ m\mathbf{V}_1(a-1) \ + \ \mathbb{V}_2((1+m)\mathbf{I} - \mathbf{J}_2)b \ + \ \mathbb{V}_3(\mathbf{I} - \mathbf{J}_3)c. \tag{31}$$

From the irreducibility of $H^j$ assumed in (A3), for all $\lambda \in Sp(\mathbf{Q}^l) \setminus Sp(W^l)$ we have $\lambda < 1$ and hence $(\mathbf{I} - \mathbf{J}_3)$ is positive definite. Therefore, since $m > 0$, from (31) we have that $f_m^l(\mathbf{x}) = 0$ if and only if $a = 1$ and $b = c = 0$, i.e. $\mathbf{x} = \mathbf{V}_1$.

It just remains to prove that $\mathbf{V}_1$ is a global attractor in $\mathbb{R}^{Ks_l}$. To this end, we will show that $\mathcal{D}f_m^l(\mathbf{x})$ is positive definite for any $\mathbf{x} \in \mathbb{R}^{Ks_l}$. We recall that, from (16) we have

$$\mathbf{F}_m^l \ = \ \mathcal{D}f_m^l(\mathbf{x}) \ = \ m\mathbf{V}_1\mathbf{U}_1' \ + \ \mathbb{V}_2((1+m)\mathbf{I} - \mathbf{J}_2)\mathbb{U}_2' \ + \ \mathbb{V}_3(\mathbf{I} - \mathbf{J}_3)\mathbb{U}_3'. \tag{32}$$

Hence, since $m > 0$ and $(\mathbf{I} - \mathbf{J}_3)$ is positive definite by assumption (A3), we have that $\mathbf{F}_m^l$ is positive definite for any $m > 0$. This concludes the proof. □

*Proof of Theorem 4.2.* The proof consists in showing that the assumptions of the *CLT* for processes in the *SA* form (Theorem A.2 in Appendix) are satisfied by the dynamics in (21) of the urn proportions $\mathbf{Z}_n^l$ in the leading system $S^l$.

First, we show that condition $\{Sp(Df(\theta^*)) > 1/2\}$ in Theorem A.2 is equivalent to $\{\lambda^{*l} < 1/2\}$. Note that the function $f$ of the SA form (5) is represented in our case by $f_m^l$ defined in (14). Similarly, the term $\theta^*$ in Appendix indicates the deterministic limiting proportion $\mathbf{Z}_\infty^l$, while $Dh(\theta^*)$ is represented by $\mathbf{F}_m^l$ defined in (16).

Now, consider the eigen-structure of $\mathbf{Q}^l$ and note that $\mathbf{F}_m^l$ has been expressed in (32) as follows:

$$\mathbf{F}_m^l \ = \ m\mathbf{V}_1\mathbf{U}_1' \ + \ \mathbb{V}_2((1+m)\mathbf{I} - \mathbf{J}_2)\mathbb{U}_2' \ + \ \mathbb{V}_3(\mathbf{I} - \mathbf{J}_3)\mathbb{U}_3',$$

Hence, it is easy to see that the eigenvectors of $\mathbf{F}_m^l$ and $\mathbf{Q}^l$ are the same, since

(1) $\mathbf{F}_m^l \mathbf{V}_1 = m\mathbf{V}_1$,
(2) $\mathbf{F}_m^l \mathbb{V}_2 = \mathbb{V}_2((1+m)\mathbf{I} - \mathbf{J}_2)$,
(3) $\mathbf{F}_m^l \mathbb{V}_3 = \mathbb{V}_3(\mathbf{I} - \mathbf{J}_3)$,

and hence

$$Sp(\mathbf{F}^l_m) = \{m\} \cup \Big\{(1+m) - \lambda, \lambda \in Sp(W^l) \setminus \{1\}\Big\} \cup \Big\{1 - \lambda, \lambda \in Sp(\mathbf{Q}^l) \setminus Sp(W^l)\Big\},$$

which, setting $m > 0$ arbitrary large, implies that $\{Sp(\mathcal{D}f(\theta^*)) > 1/2\} \equiv \{\lambda^{*l} < 1/2\}$.

Then, by following analogous arguments of point (2) in the proof of Theorem 4.1, assumption (A1) implies that condition (36) is satisfied, since

$$\sup_{n\geq 1} \mathbf{E}[\|\Delta \mathbf{M}^l_n\|^{2+\delta}|\mathcal{F}_{n-1}] \ \leq \ K^{2+\delta} \sum_{j=1}^{N}\sum_{i=1}^{K}\sum_{k=1}^{K} \sup_{n\geq 1} \mathbf{E}[(D^j_{ik,n})^{2+\delta}] \ \leq \ NK^{4+\delta}C_\delta;$$

moreover, concerning condition (37), we will show that for any $l \in \mathcal{L}_L$

$$\mathbf{E}[\Delta\mathbf{M}^l_n(\Delta\mathbf{M}^l_n)'|\mathcal{F}_{n-1}] \xrightarrow{a.s.} \mathbf{G}^l, \qquad \mathbf{E}[\Delta\mathbf{M}^{l_1}_n(\Delta\mathbf{M}^{l_2}_n)'] = 0 \quad \forall l_1 \neq l_2.$$

To this end, we first show that, for any urn $j \in S^l$, $\mathbf{E}[\Delta M^j_n(\Delta M^j_n)'|\mathcal{F}_{n-1}] \xrightarrow{a.s.} G^j$. Note that

$$\mathbf{E}[\Delta M^j_n(\Delta M^j_n)'|\mathcal{F}_{n-1}] \ = \ \mathbf{E}[(D^j_nX^j_n)(D^j_nX^j_n)'|\mathcal{F}_{n-1}] \ - \ (H^j\tilde{Z}^j_{n-1})(H^j\tilde{Z}^j_{n-1})';$$

then, concerning the first term, we have that

$$\begin{aligned}\mathbf{E}[(D^j_nX^j_n)(D^j_nX^j_n)'|\mathcal{F}_{n-1}] \quad &= \sum_{k=1}^{K} \mathbf{E}[D^j_{\cdot k,n}(D^j_{\cdot k,n})'|\mathcal{F}_{n-1}]\mathbf{P}(X^j_{k,n} = 1|\mathcal{F}_{n-1}) \\ &= \sum_{k=1}^{K} (C^j(k) + H^j(k))\tilde{Z}^j_{k,n}.\end{aligned}$$

Hence, letting $n$ increase to infinity, from (17) we have

$$\mathbf{E}[\Delta M^j_n(\Delta M^j_n)'|\mathcal{F}_{n-1}] \ \xrightarrow{a.s.} \ \sum_{k=1}^{K}(C^j(k) + H^j(k))\tilde{Z}^j_{k,\infty} \ - \ Z^j_\infty(Z^j_\infty)' \ = \ G^j.$$

Since, for any $j_1 \neq j_2$, $D^{j_1}_nX^{j_1}_n$ and $D^{j_2}_nX^{j_2}_n$ are independent conditionally on $\mathcal{F}_{n-1}$, we have that $\mathbf{E}[\Delta M^{j_1}_n(\Delta M^{j_2}_n)'|\mathcal{F}_{n-1}] = 0$ and hence $\mathbf{E}[\Delta\mathbf{M}^{l_1}_n(\Delta\mathbf{M}^{l_2}_n)'] = 0$ for any $l_1 \neq l_2$.

It remains to check that the remainder sequence $\{\mathbf{R}^l_n; n \geq 1\}$ satisfies (38) for any $\epsilon > 0$, i.e.

$$\mathbf{E}\left[n\|\mathbf{R}^l_n\|^2 \mathbb{1}_{\{\|\mathbf{Z}^l_n - \mathbf{Z}^l_\infty\| \leq \epsilon\}}\right] \longrightarrow 0. \tag{33}$$

Combining (10) and part (3b) in the proof of Theorem 4.1, we can obtain (33) by establishing

$$\mathbf{E}\left[n\left\|n \cdot diag(\mathbf{T}^l_n)^{-1} - \mathbf{I}\right\|^2\right] \longrightarrow 0,$$

that follows by using assumption (A'2) in Theorem 2.1.

Since the assumptions are all satisfied, we can apply Theorem A.2 to any leading system $S^l$, $l \in \mathcal{L}_L$, so obtaining the CLT of Theorem 4.2, with asymptotic variance

$$\Sigma^l \ = \ \lim_{m\to\infty} \int_0^\infty e^{u(\frac{\mathbf{I}}{2} - \mathbf{F}^l_m)}\mathbf{G}^l e^{u(\frac{\mathbf{I}}{2} - \mathbf{F}^l_m)'}du.$$

This concludes the proof. □

### 7.2. Proofs on the following systems.

*Proof of Theorem 5.1.* Consider the joint system $S^{(l)} = \cup_{i\in\{L_1,..l\}} S^i$, for $l \in \mathcal{L}_F$, composed by the leading systems $S^{L_1},..S^{L_{n_L}}$ and the following systems $S^{F_1},..S^l$, where we recall $S^l := \{r^{l^-}+1 \leq j \leq r^l\}$. As explained in Section 5, we focus on the reduced process $\hat{\mathbf{Z}}_n^{(l)} := \hat{\mathbf{C}}'\mathbf{Z}_n^{(l)}$, whose dynamics is expressed in (24) as follows:

$$\begin{aligned}\hat{\mathbf{Z}}_n^{(l)} - \hat{\mathbf{Z}}_{n-1}^{(l)} &= -\frac{1}{n}\hat{f}_m^{(l)}(\hat{\mathbf{Z}}_{n-1}^{(l)}) + \frac{1}{n}\hat{\mathbf{C}}'\left(\Delta\mathbf{M}_n^{(l)} + \mathbf{R}_n^{(l)}\right), \\ \hat{f}_m^{(l)}(\hat{\mathbf{x}}) &:= \left(\mathbf{I} - \hat{\mathbf{C}}'\mathbf{Q}^{(l)}\hat{\mathbf{B}}\right)\hat{\mathbf{x}} + m\hat{\mathbf{V}}_1\left(\hat{\mathbf{U}}_1'\hat{\mathbf{x}} - 1\right) + m\hat{\mathbb{V}}_2\hat{\mathbb{U}}_2'\hat{\mathbf{x}},\end{aligned} \tag{34}$$

where the function $f$ in the SA form (5) is here represented by $\hat{f}_m^{(l)}$ that takes values in $Span\{\mathbf{V}_{IN}\}$.

Analogously to the proof of Theorem 4.1 for the leading systems, one can show that all the assumptions of Theorem A.1 are satisfied by the dynamics in (34) and hence the limiting values of $\hat{\mathbf{Z}}_n^{(l)}$ are represented by those $\mathbf{x} \in Span\{\mathbf{V}_{IN}\}$ such that $\hat{f}_m^{(l)}(\mathbf{x}) = 0$. By using the analogous decompositions of (29) for $\hat{\mathbf{C}}'\mathbf{Q}^{(l)}\hat{\mathbf{B}}$ and (30) for $\mathbf{x} \in Span\{\mathbf{V}_{IN}\}$, we have that

$$\hat{f}_m^l(\mathbf{x}) = m\hat{\mathbf{V}}_1(a-1) + \hat{\mathbb{V}}_2((1+m)\mathbf{I} - \hat{\mathbf{J}}_2)b + \hat{\mathbb{V}}_3(\mathbf{I} - \hat{\mathbf{J}}_3)c, \tag{35}$$

where $\hat{\mathbf{J}}_2 := \hat{\mathbf{C}}'\mathbf{J}_2\hat{\mathbf{B}}$ and $\hat{\mathbf{J}}_3 := \hat{\mathbf{C}}'\mathbf{J}_3\hat{\mathbf{B}}$. From the irreducibility of $H^j$ required in (A3), for all $\lambda \in \mathcal{A}_{IN} \setminus Sp(W^{(l)})$ we have $\lambda < 1$ and hence $(\mathbf{I} - \hat{\mathbf{J}}_3)$ is positive definite. Therefore, since $m > 0$, from (35) we have that $\hat{f}_m^l(\mathbf{x}) = 0$ if and only if $a = 1$ and $b = c = 0$, i.e. $\mathbf{x} = \hat{\mathbf{V}}_1$.

We highlight that, when (A3) does not hold, the matrix $(\mathbf{I} - \mathbf{J}_3)$ in (35) may not be positive definite and hence the solution $\hat{\mathbf{V}}_1$ would not be unique. However, since in the following systems $S^l$, $l \in \mathcal{L}_F$, we have $\lambda_{\max}(W^l) < 1$ and this implies $\lambda_{\max}(\mathbf{Q}^l) < 1$, the irreducibility assumption of $H^j$ required in (A3) is not necessary for the following systems, but it is only essential in the leading systems in which $\lambda_{\max}(W^l) = 1$.

Now, since by definition of $\mathbf{Q}^{(l)}$ in (20) we have that

$$\hat{\mathbf{C}}'\mathbf{Q}^{(l)}\hat{\mathbf{B}} = \begin{bmatrix} \mathbf{C}'\mathbf{Q}^{(l^-)}\mathbf{B} & 0 \\ \mathbf{Q}^{(l^-)l}\mathbf{B} & \mathbf{Q}^l \end{bmatrix},$$

we can express $\hat{\mathbf{V}}_1 = (\hat{\mathbf{V}}_1^{(l^-)}, \hat{\mathbf{V}}_1^l)'$ as follows:

$$\hat{\mathbf{V}}_1 = \begin{bmatrix} \hat{\mathbf{V}}_1^{(l^-)} \\ (\mathbf{I} - \mathbf{Q}^l)^{-1}\mathbf{Q}^{(l^-)l}\mathbf{B}\hat{\mathbf{V}}_1^{(l^-)} \end{bmatrix} = \begin{bmatrix} \mathbf{C}'\mathbf{V}_1^{(l^-)} \\ (\mathbf{I} - \mathbf{Q}^l)^{-1}\mathbf{Q}^{(l^-)l}\mathbf{B}\mathbf{C}'\mathbf{V}_1^{(l^-)} \end{bmatrix}.$$

Then, notice that, since $\mathbf{V}_1^{(l^-)} \in \mathcal{I}m(\mathbf{V}_{IN})$, we have $\mathbf{B}\mathbf{C}'\mathbf{V}_1^{(l^-)} = \mathbf{V}_{IN}\mathbf{U}_{IN}'\mathbf{V}_1^{(l^-)} = \mathbf{V}_1^{(l^-)}$.

Finally, since from (25) $\hat{\mathbf{F}}_m^{(l)} = \hat{\mathbf{C}}'\mathbf{F}_m^{(l)}\hat{\mathbf{B}}$, we have that $Sp(\hat{\mathbf{F}}_m^{(l)}) \subset Sp(\mathbf{F}_m^{(l)})$ and hence $\hat{\mathbf{F}}_m^{(l)}$ is positive definite for any $m > 0$. As a consequence, $\hat{\mathbf{V}}_1$ is a global attractor in $Span\{\mathbf{V}_{IN}\}$ and this concludes the proof. □

*Proof of Theorem 5.2.* Consider the joint system $S^{(l)} = \cup_{i\in\{L_1,..l\}} S^i$, for $l \in \mathcal{L}_F$, composed by the leading systems $S^{L_1},..S^{L_{n_L}}$ and the following systems $S^{F_1},..S^l$, where we recall $S^l := \{r^{l^-}+1 \leq j \leq r^l\}$. As explained in Section 5, we focus on the reduced process $\hat{\mathbf{Z}}_n^{(l)} := \hat{\mathbf{C}}'\mathbf{Z}_n^{(l)}$, whose dynamics is expressed in (24) as follows:

$$\begin{aligned}\hat{\mathbf{Z}}_n^{(l)} - \hat{\mathbf{Z}}_{n-1}^{(l)} &= -\frac{1}{n}\hat{f}_m^{(l)}(\hat{\mathbf{Z}}_{n-1}^{(l)}) + \frac{1}{n}\hat{\mathbf{C}}'\left(\Delta\mathbf{M}_n^{(l)} + \mathbf{R}_n^{(l)}\right),\\ \hat{f}_m^{(l)}(\hat{\mathbf{x}}) &:= \left(\mathbf{I} - \hat{\mathbf{C}}'\mathbf{Q}^{(l)}\hat{\mathbf{B}}\right)\hat{\mathbf{x}} + m\hat{\mathbf{V}}_1\left(\hat{\mathbf{U}}_1'\hat{\mathbf{x}} - 1\right) + m\hat{\mathbb{V}}_2\hat{\mathbb{U}}_2'\hat{\mathbf{x}},\end{aligned}$$

where the function $f$ in the SA form in (5) is here represented by $\hat{f}_m^{(l)}$. The proof will be realized by showing that the assumptions of the $CLT$ for processes in the $SA$ form (Theorem A.2 in Appendix) are satisfied by the process $\hat{\mathbf{Z}}_n^{(l)}$. Hence the results of Theorem 5.2 follow by applying Theorem A.2, where $\theta^*$ indicates the deterministic limiting proportion $\hat{\mathbf{Z}}_\infty^{(l)}$, while $\mathcal{D}f(\theta^*)$ is represented by $\hat{\mathbf{F}}_m^{(l)}$ defined in (25).

First, we show that condition $\{Sp(Df(\theta^*)) > 1/2\}$ in Theorem A.2 is equivalent to $\{\lambda^{*l} < 1/2\}$. To this end, analogously to the proof of Theorem 4.2 for the leading systems, note that

(1) $\hat{\mathbf{F}}_m^{(l)}\hat{\mathbf{V}}_1 = m\hat{\mathbf{V}}_1$,
(2) $\hat{\mathbf{F}}_m^{(l)}\hat{\mathbb{V}}_2 = \hat{\mathbb{V}}_2((1+m)\mathbf{I} - \mathbf{J}_2)$,
(3) $\hat{\mathbf{F}}_m^{(l)}\hat{\mathbb{V}}_3 = \hat{\mathbb{V}}_3(\mathbf{I} - \mathbf{J}_3)$.

Hence, the eigenvectors of $\hat{\mathbf{F}}_m^{(l)}$ and $\hat{\mathbf{C}}'\mathbf{Q}^{(l)}\hat{\mathbf{B}}$ are the same and then

$$Sp(\hat{\mathbf{F}}_m^{(l)}) = \{m\}\cup\Big\{(1+m)-\lambda, \lambda \in Sp(W^{(l)}) \setminus (\{1\} \cup \mathcal{A}_{OUT})\Big\}\cup\Big\{1-\lambda, \lambda \in Sp(\mathbf{Q}^{(l)}) \setminus (Sp(W^{(l)}) \cup \mathcal{A}_{OUT})\Big\},$$

which implies $\{Sp(\mathcal{D}f(\theta^*)) > 1/2\} \equiv \{\lambda^{*l} < 1/2\}$.

Then, by using analogous arguments of the proof of Theorem 4.2 for the leading systems, it is straightforward showing that

$$\mathbf{E}[\hat{\mathbf{C}}'\Delta\mathbf{M}_n^{(l)}(\Delta\mathbf{M}_n^{(l)})'\hat{\mathbf{C}}|\mathcal{F}_{n-1}] \xrightarrow{a.s.} \hat{\mathbf{G}}^{(l)}, \qquad \mathbf{E}[\hat{\mathbf{C}}'\Delta\mathbf{M}_n^{(l_1)}(\Delta\mathbf{M}_n^{(l_2)})'\hat{\mathbf{C}}] = 0 \quad \forall l_1 \neq l_2,$$

and for any $\epsilon > 0$

$$\mathbf{E}\left[n\|\hat{\mathbf{C}}'\mathbf{R}_n^{(l)}\|^2 \mathbb{1}_{\{\|\hat{\mathbf{Z}}_n^{(l)} - \hat{\mathbf{Z}}_\infty^{(l)}\| \leq \epsilon\}}\right] \longrightarrow 0.$$

Since the assumptions are all satisfied, we can apply Theorem A.2 to obtain the CLT with asymptotic variance

$$\hat{\Sigma}^{(l)} := \int_0^\infty e^{u(\frac{\mathbf{I}}{2} - \hat{\mathbf{F}}_m^{(l)})}\hat{\mathbf{G}}^{(l)} e^{u(\frac{\mathbf{I}}{2} - \hat{\mathbf{F}}_m^{(l)})'} du.$$

This concludes the proof. □

# Appendix

## Appendix A. Basic tools of Stochastic Approximation

Consider the recursive procedure define in (5) on a filtered probability space $(\Omega, \mathcal{A}, (\mathcal{F}_n)_{n\geq 0}, \mathbf{P})$, i.e.

$$\forall\, n \geq 1, \quad \theta_n = \theta_{n-1} - \frac{1}{n}f(\theta_{n-1}) + \frac{1}{n}\left(\Delta M_n + R_n\right),$$

where $f : \mathbb{R}^d \to \mathbb{R}^d$ is a locally Lipschitz continuous function, $\theta_n$ an $\mathcal{F}_n$-measurable finite random vector and, for every $n \geq 1$, $\Delta M_n$ is an $\mathcal{F}_{n-1}$-martingale increment and $R_n$ is an $\mathcal{F}_n$-adapted remainder term.

**Theorem A.1.** *(A.s. convergence with ODE method, see e.g.* [9, 14, 22, 17, 7]*). Assume that $f$ is locally Lipschitz, that*

$$R_n \xrightarrow{a.s.} 0 \quad \textit{and} \quad \sup_{n\geq 1} \mathbf{E}\left[\|\Delta M_n\|^2 \mid \mathcal{F}_{n-1}\right] < +\infty \quad a.s..$$

*Then the set $\Theta^\infty$ of its limiting values as $n \to +\infty$ is a.s. a compact connected set, stable by the flow of*

$$ODE_f \equiv \dot{\theta} = -f(\theta).$$

*Furthermore if $\theta^* \in \Theta^\infty$ is a uniformly stable equilibrium on $\Theta^\infty$ of $ODE_f$, then*

$$\theta_n \xrightarrow{a.s.} \theta^*.$$

Comments. By uniformly stable we mean that

$$\sup_{\theta\in\Theta^\infty} |\theta(\theta_0, t) - \theta^*| \longrightarrow 0 \quad \text{as} \quad t \to +\infty,$$

where $\theta(\theta_0, t)_{\theta_0\in\Theta^\infty, t\in\mathbb{R}_+}$ is the flow of $ODE_f$ on $\Theta^\infty$.

We say that the function $f$ is *$\epsilon$-differentiable*, $\epsilon > 0$, at $\theta^*$ if

$$f(\theta) = f(\theta^*) + \mathcal{D}f(\theta^*)(\theta-\theta^*) + o(\|\theta-\theta^*\|^{1+\epsilon}) \quad \text{as} \quad \theta \to \theta^*.$$

**Theorem A.2.** *(Rate of convergence see* [14] *Theorem 3.III.14 p.131 (for CLT see also e.g.* [9, 22]*)). Let $\theta^*$ be an equilibrium point of $\{f = 0\}$. Assume that the function $f$ is differentiable at $\theta^*$ and all the eigenvalues of $\mathcal{D}f(\theta^*)$ have positive real parts. Assume that for some $\delta > 0$,*

$$\sup_{n\geq 1} \mathbf{E}\left[\|\Delta M_n\|^{2+\delta} \mid \mathcal{F}_{n-1}\right] < +\infty \, a.s., \tag{36}$$

*and*

$$\mathbf{E}\left[\Delta M_n \Delta M_n' \mid \mathcal{F}_{n-1}\right] \xrightarrow[n\to+\infty]{a.s.} \Gamma, \tag{37}$$

*where $\Gamma \in \mathcal{S}^+(d,\mathbb{R})$ (deterministic symmetric positive matrix) and for an $\epsilon > 0$,*

$$n\mathbf{E}\left[\|R_n\|^2 \mathbb{1}_{\{\|\theta_{n-1}-\theta^*\|\leq\epsilon\}}\right] \xrightarrow[n\to+\infty]{} 0. \tag{38}$$

*(a) If $\Lambda := \Re e(\lambda_{\min}) > \frac{1}{2}$, where $\lambda_{\min}$ denotes the eigenvalue of $\mathcal{D}f(\theta^*)$ with lowest real part, the above a.s. convergence is ruled on the set $\mathcal{D}f\{\theta_n \to \theta^*\}$ by the following Central Limit Theorem*

$$\sqrt{n}\,(\theta_n - \theta^*) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0,\Sigma) \quad \textit{with} \quad \Sigma := \int_0^{+\infty} e^{(I_d/2-\mathcal{D}f(\theta^*))u}\Gamma e^{(I_d/2-\mathcal{D}f(\theta^*))' u}du.$$

*(b) If $\lambda_{\min} = \frac{1}{2}$, then*

$$\sqrt{\frac{n}{\log n}}\,(\theta_n - \theta^*) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0,\Sigma) \quad \textit{as } n \to +\infty.$$

*(c) If $\lambda_{\min} \in (0, \frac{1}{2})$, then $n^{\lambda_{\min}}(\theta_n - \theta^*)$ a.s. converges as $n \to +\infty$ towards a finite random variable.*

## References


[1] G. Aletti, C. May and P. Secchi. On the distribution of the limit proportion for a two-color, randomly reinforced urn with equal reinforcement distributions. *Adv. Appl. Probab.*, 39, 690:707, 2007.

[2] G. Aletti, C. May and P. Secchi. A Central Limit Theorem, and related results, for two-color randomly reinforced urn. *Adv. Appl. Probab.*, 41, 829:844, 2009.

[3] K. B. Athreya. On a characteristic property of Pólya's urn. *Studia Sci. Math. Hung.* 4:31–35, 1969.

[4] K. B. Athreya and S. Karlin. Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *Ann. Math. Statist.*, 39:1801–1817, 1968.

[5] Z.-D. Bai and F. Hu. Asymptotic theorems for urn models with nonhomogeneous generating matrices. *Stochastic Process. Appl.*, 80(1):87–101, 1999.

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

[7] M. Benaïm. Dynamics of stochastic approximation algorithms. In *Séminaire de Probabilités, XXXIII*, volume 1709 of *Lecture Notes in Math.*, pages 1–68. Springer, Berlin, 1999.

[8] M. Benaïm, I. Benjamini, J. Chen and Y. Lima. A Generalized Pòlya's Urn with Graph Based Interactions. *Random Structures and Algorithms*, 46(4): 614–634, 2015.

[9] A. Benveniste, M. Métivier, and P. Priouret. *Adaptive algorithms and stochastic approximations*, volume 22 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1990. Translated from the French by Stephen S. Wilson.

[10] J. Chen and C. Lucas. A generalized 'Pólyas urn with graph based interactions convergence at linearity. *Electron. Commun. Probab.*, 67(19): 1-13, 2014.

[11] P. Cirillo, P. Gallegati and J. Hüsler. A Pólya lattice model to study leverage dynamics and contagion financial fragility. *Adv. Complex Systems*, 15, 2012.

[12] I. Crimaldi, P. D. Pra and I. G. Minelli. Fluctuation theorems for synchronization of interacting Pòlya's urns. *Stoch. Proc. Appl.*, in press, 2015.

[13] P. D. Pra, P. Y. Louis and I. G. Minelli. Synchronization via interacting reinforcement. *J. Appl. Prob.*, 51:556–568, 2014.

[14] M. Duflo. *Random iterative models*, volume 34 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1997. Translated from the 1990 French original by Stephen S. Wilson and revised by the author.

[15] S. D. Durham, N. Flournoy and W. Li. A sequential design for maximazing the probability of a favourable response. *Can. J. Stat.*, 26(3):479–495, 1998.

[16] F. Effenberger and G. Pólya. Über die Statistik verketteter and Vorgänge. Z. Angewandte *Math. Mech.*, 3,279–289, 1923.

[17] J.-C. Fort and G. Pagès. Convergence of stochastic algorithms: from the Kushner-Clark theorem to the Lyapounov functional method. *Adv. in Appl. Probab.*, 28(4):1072–1094, 1996.

[18] B. Friedman. A simple urn model. *Commun. Pure Appl. Math.*, 36:956–970, 1949.

[19] B. M. Hill, D. Lane and W. Suddert. A strong law for some generalized urn processes. *Ann. Prob.*, 8:214–226, 1980.

[20] B. M. Hill, D. Lane and W. Suddert. Exchangeable urn processes. *Ann. Prob.*, 15:1586–1592, 1987.

[21] N. L. Johnson and S. Kotz. *Urn models and their application*. John Wiley & Sons, New York-London-Sydney, 1977. An approach to modern discrete probability theory, Wiley Series in Probability and Mathematical Statistics.

[22] H. J. Kushner and G. G. Yin. *Stochastic approximation and recursive algorithms and applications*, volume 35 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition, 2003. Stochastic Modelling and Applied Probability.

[23] S. Laruelle and G. Pagès. Randomized urn models revisited using stochastic approximation. *Ann. Appl. Probab.*, 33(4):1409–1436, 2013.

[24] M. Launay. Interacting urn models. Preprint. Available at *http://arxiv.org/abs/1101.1410.* 2011.

[25] M. Launay and V. Limic. Generalized interacting urn models. Preprint. Available at *http://arxiv.org/abs/1207.5635.* 2012.

[26] M. Marsili, and A. Valleriani. Self organization of interacting Pòlya urns. *Europ. Physical J.*, B 3:417–420, 1998.

[27] G. Naldi, L. Pareschi and G. Toscani. *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences*. Springer, 2010.

[28] J. R. Norris *Markov Chains*. Cambridge Series in Statistical and Probabilistic Mathematics, 1998.

[29] A. M. Paganoni and P. Secchi. Interacting reinforced-urn systems. *Adv. in Appl. Probab.*, 36:791–804, 2004.

[30] R. Pemantle. A survey of random processes with reinforcement. *Probab. Surv.*, 3:1549–5787, 2007.

[31] A. Prindle, P. Samayoa, I. Razinkov, T. Danino, L.S. Tsimring and J. Hasty. A sensing array of radically coupled genetic biopixels. *Nature*, 481:39–44, 2011.
[32] A.N. Shiryaev. *Probability*. Springer, New York, 1996.
[33] R. T. Smythe. Central limit theorems for urn models. *Stochastic Process. Appl.*, 65(1):115–137, 1996.
[34] A. Y. Yakovlev, A. D. Tsodikov. *Stochastic Models of Tumor Latency and Their Biostatistical Applications*. World Scientific, 1996.

(G. Aletti) ADAMSS Center & Department of Mathematics, Università degli Studi di Milano, 20133 Milano, Italy
*E-mail address*: giacomo.aletti@unimi.it

(A. Ghiglietti) ADAMSS Center & Department of Mathematics, Università degli Studi di Milano, 20133 Milano, Italy
*E-mail address*: andrea.ghiglietti@unimi.it