



This is a repository copy of *Phase transitions in non-linear urns with interacting types*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/193934/>

Version: Accepted Version

---

**Article:**

Costa, M. and Jordan, J. [orcid.org/0000-0003-4686-5440](https://orcid.org/0000-0003-4686-5440) (2022) Phase transitions in non-linear urns with interacting types. *Bernoulli*, 28 (4). pp. 2546-2562. ISSN 1350-7265

<https://doi.org/10.3150/21-BEJ1428>

---

© 2022 ISI/BS. This is an author-produced version of a paper subsequently published in *Bernoulli*. Uploaded in accordance with the publisher's self-archiving policy.

**Reuse**

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

# Phase transitions in non-linear urns with interacting types

Marcelo Costa

Universidad de Buenos Aires \*

Jonathan Jordan

University of Sheffield †

November 1, 2021

## Abstract

We investigate reinforced non-linear urns with interacting types, and show that where there are three interacting types there are phenomena which do not occur with two types. In a model with three types where the interactions between the types are symmetric, we show the existence of a double phase transition with three phases: as well as a phase with an almost sure limit where each of the three colours is equally represented and a phase with almost sure convergence to an asymmetric limit, which both occur with two types, there is also an intermediate phase where both symmetric and asymmetric limits are possible. In a model with anti-symmetric interactions between the types, we show the existence of a phase where the proportions of the three colours cycle and do not converge to a limit, alongside a phase where the proportions of the three colours can converge to a limit where each of the three is equally represented.

---

\*mrocha@dm.uba.ar

†jonathan.jordan@sheffield.ac.uk

# 1 Introduction and definitions

In general terms, an urn model is a system containing a number of particles of different types (often regarded as balls of different colours, for ease of visualisation). At each time step, a set of particles is sampled from the system, whose contents are then altered depending on the sample which was drawn. Pemantle [10] surveys several ways to approach this model framework.

This paper is limited to models with a single urn from which a single ball is drawn, its colour is noted and it is then returned to the urn along with one new ball of that same colour. In addition, we introduce a graph-based interaction according to which the probability of choosing a ball of a given colour is reinforced not only by its own proportion in the urn, but also by the proportions of balls of other colours. Therefore, the interaction arises among balls of different colours, as opposed to the so-called interacting urn models consisting of systems of multiple urns (e.g. Benaïm et al [2], Launay and Limic [7]) in which different urns (each containing balls of different colours) interact with one another. Our model is also different from the graph-based competition described in van der Hofstad et al [11], where the colours correspond to edges of the graph, which compete with, as opposed to being reinforced by, other edges incident on the same vertices.

We now formally define our model. Consider an urn containing balls of  $d$  colours. The vector  $x(n) = (x_1(n), \dots, x_d(n)) \in \mathbb{N}^d$  denotes the number of balls of each colour at time  $n = 0, 1, 2, \dots$ . The strength of the reinforcement is given by a positive real number  $\beta > 0$  and we denote by  $x^\beta(n)$  the coordinate-wise  $\beta$  power of the column vector  $x(n)$ . The interaction is defined as follows. Given a non-negative matrix  $A = (a_{ij})_{i,j=1}^d$ , define the column vector

$$u(n) := Ax^\beta(n) \tag{1}$$

Let  $\mathcal{F}_n$  be the  $\sigma$ -algebra generated by the  $x(m)$  for  $0 \leq m \leq n$  and write  $u_i(n)$  for the  $i$ -th component of  $u(n)$ . The transition probabilities are then

$$\mathbb{P}(x(n+1) - x(n) = \mathbf{e}_i \mid \mathcal{F}_n) = \frac{u_i(n)}{\sum_{j=1}^d u_j(n)}, \quad i = 1, \dots, d, \tag{2}$$

where  $\mathbf{e}_i$  is the unit vector in direction  $i$ . That is, one ball is added to the urn at each time step, and the right hand side of (2) gives the probability that it is of colour  $i$ .

Now, let  $n_0 > 0$  be the initial number of balls so that at time  $n$  the urn contains  $n + n_0$  balls. Then, the proportion of balls of each colour is a process in the  $(d-1)$ -dimensional simplex  $\Delta^{d-1}$

given by the vector

$$\bar{x}(n) = x(n)/(n + n_0). \quad (3)$$

When  $A$  is a multiple of the identity matrix (therefore,  $\bar{x}(n)$  having no interaction) it is well-known (see Oliveira [8]) that the process  $\bar{x}$  undergoes a phase transition as follows. For  $\beta < 1$ , the process converges almost surely to the ‘centre’ of the simplex, that is, the asymptotic proportion of balls of each colour are all the same. For  $\beta = 1$ , commonly referred to as the Pólya urn model, the process converges almost surely to a non-trivial random variable supported in the interior of the simplex. For  $\beta > 1$ , the process converges almost surely to one of the corners of the simplex. In this case that a single type dominates was proved by Khanin and Khanin [5] following on from the two-type case which can be covered using Rubin’s Theorem in Davis [4].

For a two-colour urn model  $d = 2$  and symmetric interaction,

$$A = \begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix}, \quad a > 0,$$

it was proved by the first author in Theorem 2.2.1, [3], that there was a phase transition as follows.

- (i) if  $\left(\frac{1-a}{1+a}\right)\beta \leq 1$ , then  $\bar{x}(n) \rightarrow (\frac{1}{2}, \frac{1}{2})$  *a.s.*
- (ii) if  $\left(\frac{1-a}{1+a}\right)\beta > 1$ , then  $\bar{x}(n) \rightarrow \Psi$  *a.s.*,

where  $\Psi$  is a random vector supported on  $\left\{ \left(\frac{1}{1+r}, \frac{r}{1+r}\right), \left(\frac{r}{1+r}, \frac{1}{1+r}\right) \right\}$  and  $r := r(a, \beta)$  is the unique root in  $(0, 1)$  of  $\mathcal{P}_{a,\beta}(z) = az^{\beta+1} - z^\beta + z - a = 0$ . In case (ii),  $\mathbb{P}[\bar{x}(n) \rightarrow (\frac{1}{2}, \frac{1}{2})] = 0$ . Note that for  $\beta = 1$ , the process  $(u(n))_{n \geq 0}$  is a Friedman’s urn model and statement (i) yields  $u(n)/(u_1(n) + u_2(n)) \rightarrow (\frac{1}{2}, \frac{1}{2})$  *a.s.* as expected.

Some similar results appear in Laruelle and Pagès [6]. The definition of the model there allows for a more general *skewing function* than  $x^\beta$  as above, but in the  $d = 2$  case gives phase transitions similar to those in [3], and relates them to the eigenvalues of the *generating matrix*  $H$ , which is related to our matrix  $A$ . Furthermore they show that, for any  $d$ , for a concave skewing function (corresponding to  $\beta < 1$  in our model) and a bi-stochastic generating matrix there is almost sure convergence to the centre of the simplex, and they give conditions, related to the eigenvalues of the generating matrix, under which for a convex skewing function (our  $\beta > 1$ ) there is probability zero of converging to the centre of the simplex.

In this paper, we follow up on the results for  $d = 2$  in [3] and those in [6], with the aim to generalise from  $d = 2$  to larger values of  $d$  and to see whether more types of behaviour emerge when this is done. We show that this is indeed the case when  $d = 3$ , where we consider two

particular choices of  $A$ . Our results can be seen as extensions of those of [6] in certain specific cases.

First of all, we consider a choice of  $A$  with a symmetric interaction of the same strength  $a$  for each pair of colours. The following theorem shows that in this system there are three phases, as opposed to two when  $d = 2$ ; there are phases where there is almost sure convergence to a symmetric limit and where there is almost sure convergence to one of a number of asymmetric limits, which are analogues of the phases when  $d = 2$ , but there is also an intermediate phase where both symmetric and asymmetric limits are possible.

**Theorem 1.1.** *Let  $A$  be the matrix*

$$\begin{pmatrix} 1 & a & a \\ a & 1 & a \\ a & a & 1 \end{pmatrix}, \quad a > 0. \quad (4)$$

(i) *Fix  $a < 1$ . Then there exists  $\beta_1(a)$  satisfying  $1 < \beta_1(a) < \frac{1+2a}{1-a}$ , with  $\beta_1(a)$  an increasing function of  $a$  satisfying  $\beta_1(a) \rightarrow \infty$  as  $a \rightarrow 1$ , and we have the following three phases.*

- (a) *Symmetric limit almost surely. If  $\beta < \beta_1(a)$ , then almost surely  $\bar{x}(n) \rightarrow (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ .*
- (b) *Symmetric or asymmetric limit. If  $\beta_1(a) < \beta < \frac{1+2a}{1-a}$  then there exists  $r_2 > 1$  such that almost surely  $\bar{x}(n)$  converges to one of the four points in  $\Delta^2$  given by  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ ,  $(\frac{r_2}{2+r_2}, \frac{1}{2+r_2}, \frac{1}{2+r_2})$ ,  $(\frac{1}{2+r_2}, \frac{r_2}{2+r_2}, \frac{1}{2+r_2})$  and  $(\frac{1}{2+r_2}, \frac{1}{2+r_2}, \frac{r_2}{2+r_2})$ . All of these points have positive probability of being limits.*
- (c) *Asymmetric limit almost surely. If  $\beta > \frac{1+2a}{1-a}$  then there exists  $r_+ > 1$  such that almost surely  $\bar{x}(n)$  converges to one of the three points in  $\Delta^2$  given by  $(\frac{r_+}{2+r_+}, \frac{1}{2+r_+}, \frac{1}{2+r_+})$ ,  $(\frac{1}{2+r_+}, \frac{r_+}{2+r_+}, \frac{1}{2+r_+})$  and  $(\frac{1}{2+r_+}, \frac{1}{2+r_+}, \frac{r_+}{2+r_+})$ .*

(ii) *Fix  $a \geq 1$ . Then almost surely  $\bar{x}(n) \rightarrow (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ .*

Theorem 1.1 presents the results in terms of phase transitions in  $\beta$  with  $a$  fixed. However, because both  $\beta_1(a)$  and  $\frac{1+2a}{1-a}$  are increasing functions of  $a$  which converge to 1 as  $a \rightarrow 0$  and to  $\infty$  as  $a \rightarrow 1$ , it is also possible to see them as phase transitions in  $a$  with  $\beta > 1$  fixed: if  $a < \frac{\beta-1}{2+\beta}$  then we will be in case (c), if  $\frac{\beta-1}{2+\beta} < a < \beta_1^{-1}(\beta)$  then we will be in case (b), and if  $a > \beta_1^{-1}(\beta)$  we will be in case (a).

Theorem 1.1 can be seen as an extension of the results of Proposition 2.15 of Laruelle and

Pagès [6] in the case where the matrix  $H = \frac{1}{1+2a} \begin{pmatrix} 1 & a & a \\ a & 1 & a \\ a & a & 1 \end{pmatrix}$  in that their result shows the non-convergence to  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  in the case (i)(c), as the second largest eigenvalue of  $H$  is  $\frac{1-a}{1+2a}$ . We can also state the condition  $\beta > \frac{1+2a}{1-a}$  in terms of the eigenvalues of  $A$ : the right hand side can be seen as the ratio of the two largest eigenvalues. It might be reasonable to conjecture that an extension to  $d > 3$  of Theorem 1.1 might involve a similar condition on the eigenvalues; however we note that the other phase transition does not appear to be related to the eigenvalues of  $A$ . We discuss the question of what happens with  $d > 3$  further in Section 5.4.

We also consider a system where each colour is reinforced by itself and by one other, in a cyclic way. For this system, the following theorem shows the existence of a phase transition between a phase with convergence with positive probability to a symmetric limit and a phase where there is no convergence to a limit and there is cycling behaviour.

**Theorem 1.2.** *Let  $A$  be the matrix*

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

- *When  $\beta < 4$  there is positive probability that  $\bar{X}(n) \rightarrow (1/3, 1/3, 1/3)$ .*
- *When  $\beta > 4$ , almost surely  $\bar{X}(n)$  fails to converge, and the limit set is either a periodic orbit or a connected union of periodic orbits.*

In Section 2 we discuss the stochastic approximation methods we use in the proofs, while the proofs themselves are in Section 3 for Theorem 1.1 and Section 4 for Theorem 1.2. In the final Section 5, we illustrate the results with some examples and simulations, including some examples beyond those covered by Theorems 1.1 and 1.2.

## 2 Stochastic approximation approach

In this section we introduce some of the stochastic approximation ideas which appear in our proofs.

The type of stochastic approximation process we will be interested in is a Robbins-Monro algorithm in  $\mathbb{R}^d$ , following section 4.2 of Benaïm [1]. This is a stochastic process  $(y(n))_{n \in \mathbb{N}}$ , with

natural filtration  $(\mathcal{F}_n)$ , taking values in  $\mathbb{R}^d$  which satisfies

$$y(n+1) - y(n) = \gamma_n(F(y(n)) + \xi_{n+1}), \quad (5)$$

where  $F : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is a deterministic vector field,  $\gamma_n$  is a step size satisfying certain conditions, and  $\mathbb{E}(\xi_{n+1}|\mathcal{F}_n) = 0$  with  $\xi_n$  being  $\mathcal{F}_n$ -measurable. For most results it is required that  $\gamma_n \rightarrow 0$  as  $n \rightarrow \infty$  but that  $\sum_{n=1}^{\infty} \gamma_n = \infty$ , and some conditions are also needed on the noise term  $\xi_{n+1}$ , typically that it is bounded in  $L^q$  for some  $q$  which depends on the  $\gamma_n$ : see for example Proposition 4.2 of Benaïm [1].

The above sequence can be thought as a numerical approximating method with varying step size  $\gamma_n$  for solving the ODE  $dx/dt = F(x)$ . Under the conditions discussed above, the asymptotic behavior of  $(\bar{x}_n)_{n \in \mathbb{N}}$  and the underlying ODE are closely connected: as described in [1], define an interpolated version  $(X(t))_{t \geq 0}$  by defining  $\tau_0 = 0$  and for  $n \in \mathbb{N}$   $\tau_n = \sum_{i=1}^n \gamma_i$ , then defining  $X(\tau_n + s) = \bar{x}_n + \frac{\bar{x}_{n+1} - \bar{x}_n}{\gamma_{n+1}} s$  for  $0 \leq s < \gamma_{n+1}$ . Then the interpolated process  $(X(t))_{t \geq 0}$  is an *asymptotic pseudotrajectory* for the ODE. This is called the *ODE method* or the *dynamical system approach*, which alongside some probabilistic techniques, is applied to examine almost sure dynamics of stochastic approximation processes.

Let  $\Phi : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  be the semiflow induced by the vector field  $F$ , so that  $(\Phi(t, x))_{t \geq 0}$  is the trajectory of  $F$  started at  $x$ . A useful situation for analysis of stochastic approximation processes is where there is a *Lyapunov function* for the vector field  $F$ : a function  $V : \mathbb{R}^d \rightarrow \mathbb{R}^d$  such that on a trajectory  $(\Phi(t, x))_{t \geq 0}$  of the vector field,  $V(\Phi(t, x))$  is strictly decreasing in  $t$  except where  $\Phi(t, x)$  is a stationary point of  $F$ . If this holds, then under mild conditions then the Robbins-Monro algorithm will converge almost surely to a (possibly random) limit, which will be a stationary point of  $F$ . See section 6.2 of Benaïm [1], and in particular Corollary 6.6 therein.

We now show that our process  $(\bar{x}_n)$  can be put into Robbins-Monro form, and that the assumptions of Proposition 4.2 of Benaïm [1] are satisfied, allowing the theory of asymptotic pseudotrajectories to be used. For a general matrix  $A$  and a given configuration of balls  $x(n)$  at time  $n$ , let  $i_{n+1} \in \{1, \dots, d\}$  be the random colour of the ball to be added in the urn at time  $n+1$ . Then, note that

$$\begin{aligned} \bar{x}(n+1) &= \frac{x(n) + \mathbf{e}_{i_{n+1}}}{n_0 + n + 1} = \frac{(n_0 + n)\bar{x}(n) + \mathbf{e}_{i_{n+1}}}{n_0 + n + 1} \\ &= \left(1 - \frac{1}{n_0 + n + 1}\right) \bar{x}(n) + \frac{\mathbf{e}_{i_{n+1}}}{n_0 + n + 1}, \end{aligned} \quad (6)$$

implying

$$\bar{x}(n+1) - \bar{x}(n) = \frac{1}{n_0 + n + 1} (\mathbf{e}_{i_{n+1}} - \bar{x}(n)). \quad (7)$$

To put (7) into Robbins-Monro form, we rearrange the right-hand side into a deterministic part and a zero mean “noise”. More specifically, let

$$F(\bar{x}(n)) := \mathbb{E}[\mathbf{e}_{i_{n+1}} | \mathcal{F}_n] - \bar{x}(n), \quad (8)$$

and

$$\xi_{n+1} := \mathbf{e}_{i_{n+1}} - \mathbb{E}[\mathbf{e}_{i_{n+1}} | \mathcal{F}_n]. \quad (9)$$

By setting  $\gamma_n = 1/(n_0 + n + 1)$ , we obtain

$$\bar{x}(n+1) - \bar{x}(n) = \gamma_n(F(\bar{x}(n)) + \xi_{n+1}). \quad (10)$$

We note that the components of  $\xi_{n+1}$  are uniformly bounded in modulus by 2, meaning that the conditions of Proposition 4.2 of Benaïm [1] are satisfied, meaning that the interpolated process  $(X(t))_{t \geq 0}$  is indeed an asymptotic pseudotrajectory of  $F$ .

We can write the components of  $F$  as

$$F_i(x) = \frac{u_i}{\sum_{j=1}^d u_j} - x_i, \quad i = 1, \dots, d,$$

where  $u_i$  has the same relationship to  $x$  as  $u_i(n)$  to  $x(n)$ .

For a stochastic approximation heuristic, and to apply some of the results we use, it is necessary to classify the stationary points, for which we use the following terminology.

**Definition 1.** *Consider a stationary point  $p$  of  $F$ . We will say that  $p$  is stable if it is an attractor for the vector field  $F$ , meaning that there exists a neighbourhood  $W$  of  $p$  where for  $x \in W$  the trajectory  $(\Phi(t, x))_{t \geq 0}$  satisfies  $\Phi(x, t) \rightarrow p$  as  $t \rightarrow \infty$ , uniformly in  $W$ . Furthermore, if all eigenvalues of  $DF(p)$  have negative real part,  $p$  is said to be linearly stable, while if some eigenvalue has positive real part,  $p$  is said to be linearly unstable. If all eigenvalues have positive real part, then  $p$  is said to be a source.*

Typically convergence happens with positive probability to a stable stationary point but with probability zero to a linearly unstable one. For the former, Theorem 7.3 of Benaïm [1], which holds under the assumptions on the noise of Proposition 4.2, shows that a stable stationary point has positive probability of being a limit as long as its basin of attraction  $W$  contains points which are *attainable* in the sense that the process has positive probability of being indefinitely close to them at indefinitely large times, a condition which is usually satisfied for all points in the simplex for urn type processes under mild conditions; see Example 7.2 of [1].

To show that linearly unstable stationary points have zero probability of being limits it is usually necessary to check that there is expectation bounded away from zero of the positive

part of the component of the noise in any given direction; see Pemantle [9], Theorem 1. This condition ensures that the process has zero probability of being trapped on a stable manifold and converging to the unstable point. Theorem 1 of [9] also requires the noise term  $\xi_{n+1}$  (in our notation) to be uniformly bounded; as noted above this is satisfied in our process.

### 3 Proofs for the symmetric case

In this section we prove Theorem 1.1.

Throughout this section we let  $A$  be the matrix

$$\begin{pmatrix} 1 & a & a \\ a & 1 & a \\ a & a & 1 \end{pmatrix}, \quad a > 0. \quad (11)$$

In this case the vector field  $F$  is given by

$$F_1(x_1, x_2, x_3) = \frac{x_1^\beta + ax_2^\beta + ax_3^\beta}{(1+2a)(x_1^\beta + x_2^\beta + x_3^\beta)} - x_1 \quad (12)$$

$$F_2(x_1, x_2, x_3) = \frac{ax_1^\beta + x_2^\beta + ax_3^\beta}{(1+2a)(x_1^\beta + x_2^\beta + x_3^\beta)} - x_2 \quad (13)$$

$$F_3(x_1, x_2, x_3) = \frac{ax_1^\beta + ax_2^\beta + x_3^\beta}{(1+2a)(x_1^\beta + x_2^\beta + x_3^\beta)} - x_3 \quad (14)$$

The stochastic approximation approach indicates that the possible limits of our process will be stationary points of  $F$ , so we start by identifying these. Noting that the lines  $x_1 = x_2$ ,  $x_1 = x_3$  and  $x_2 = x_3$  are each invariant under  $F$ , define the function

$$\mathcal{P}_{a,\beta}(z) = az^{\beta+1} - z^\beta + (1+a)z - 2a, \quad (15)$$

which we will see is related to the dynamics restricted to one of these lines. The following result shows that all stationary points of  $F$  are located on at least one of these lines and expresses them in terms of solutions to  $\mathcal{P}_{a,\beta}(z) = 0$ .

**Proposition 3.1.** *All stationary points  $(x_1, x_2, x_3)$  of  $F$  have at least two of  $x_1, x_2, x_3$  equal, and are of one of the forms  $(\frac{r}{r+2}, \frac{1}{r+2}, \frac{1}{r+2})$ ,  $(\frac{1}{r+2}, \frac{r}{r+2}, \frac{1}{r+2})$  or  $(\frac{1}{r+2}, \frac{1}{r+2}, \frac{r}{r+2})$ , with  $r$  a solution of  $\mathcal{P}_{a,\beta}(z) = 0$  in  $\mathbb{R}^+$ .*

Furthermore, there are at most three possible values of  $r$ , one of which is always 1, corresponding to the stationary point  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ .

*Proof.* We start off by showing that any stationary point has at least two co-ordinates equal. We do this by writing the stationary point in the form  $(x, rx, sx)$  and showing that one of  $r = 1$ ,  $s = 1$  or  $r = s$  must hold.

Rearranging (12), (13) and (14) at  $(x, rx, sx)$  gives

$$x = \frac{1 + a(r^\beta + s^\beta)}{(2a + 1)(1 + r^\beta + s^\beta)} \quad (16)$$

$$rx = \frac{r^\beta + a(1 + s^\beta)}{(2a + 1)(1 + r^\beta + s^\beta)} \quad (17)$$

$$sx = \frac{s^\beta + a(1 + r^\beta)}{(2a + 1)(1 + r^\beta + s^\beta)}. \quad (18)$$

It follows that

$$r^\beta + a(1 + s^\beta) = r(1 + a(r^\beta + s^\beta)) \quad (19)$$

$$s^\beta + a(1 + r^\beta) = s(1 + a(r^\beta + s^\beta)). \quad (20)$$

Take the linear combination  $(s + \frac{1}{a}) \times (19) - (r + 1) \times (20)$ . This eliminates  $s^\beta$  and  $s^{\beta+1}$ , giving

$$(r^\beta + a)s - ar(1 + r^\beta) = (ar^\beta + 1)s - \left( a(1 + r^\beta) + \frac{r}{a} - \frac{r^\beta}{a} + r^{\beta+1} - 1 \right), \quad (21)$$

which can be rearranged to give

$$s(r^\beta - 1)(a - 1) = (r^{\beta+1} - 1)(1 - a) + (r^\beta - r) \left( a - \frac{1}{a} \right). \quad (22)$$

Assuming  $a \neq 1$ , (22) gives  $r = 1$  or

$$s = \frac{a(r^\beta + 1)(1 - r) + r^\beta - r}{a(r^\beta - 1)} = \frac{\mathcal{P}_{a,\beta}(r) - ar^\beta + a}{a(1 - r^\beta)}. \quad (23)$$

Using this form for  $s$  in (20) gives (if  $r \neq 1$ )

$$s^\beta = \frac{-ar^{\beta+1} + r^\beta + a - r}{a(r - 1)} = \frac{\mathcal{P}_{a,\beta}(r)}{a(1 - r)} + 1. \quad (24)$$

Combining (23) and (24) tells us that either  $r = 1$  or  $s = 1$  or

$$\frac{s^\beta - 1}{s - 1} = \frac{r^\beta - 1}{r - 1},$$

and the latter case implies  $r = s$ . Hence any stationary point has two co-ordinates equal.

We now assume, without loss of generality, that the stationary point is of the form  $(x, x, rx)$  or equivalently  $\left(\frac{1}{r+2}, \frac{1}{r+2}, \frac{r}{r+2}\right)$ . That the stationary point equations for a point of this form imply  $\mathcal{P}_{a,\beta}(r) = 0$  is easy to check, and it is also easy to check that  $\mathcal{P}_{a,\beta}(1) = 0$  for any  $a, \beta > 0$ .

The function  $\mathcal{P}_{a,\beta}$  satisfies  $\mathcal{P}_{a,\beta}(0) < 0$  and  $\mathcal{P}_{a,\beta}(z) \rightarrow \infty$  as  $z \rightarrow \infty$ ; furthermore it is concave for  $z < \frac{\beta-1}{a(\beta+1)}$  and convex for  $z > \frac{\beta-1}{a(\beta+1)}$ , which indicates that it has either one root or three in  $\mathbb{R}^+$ , counting multiplicity. This completes the proof. □

We now show that the process will indeed, almost surely, converge to one of the stationary points identified in Proposition 3.1.

**Lemma 3.2.** *The limit set of the process  $(\bar{x}(n))_{n \in \mathbb{N}}$ , defined in (3) with given matrix (11), will, almost surely, be a single point which is a stationary point of  $F$ .*

*Proof.* Let

$$L(x_1, x_2, x_3) = (x_1 + x_2 + x_3) - \frac{1}{2a+1} \left[ a \log(x_1 x_2 x_3) - \frac{1}{\beta} (a-1) \log(x_1^\beta + x_2^\beta + x_3^\beta) \right]. \quad (25)$$

Then  $L$  is a strict Lyapunov function for  $F$ . In fact, straightforward differentiation gives

$$\frac{\partial L}{\partial x_i} = -\frac{1}{x_i} F_i.$$

Then, denoting an trajectory of  $F$  by  $x(t) = (x_1(t), x_2(t), x_3(t))$ , we obtain

$$\frac{d(L \circ x)}{dt} = \sum_{i=1}^3 \frac{\partial L}{\partial x_i} \frac{dx_i}{dt} = - \sum_{i=1}^3 x_i \left( \frac{\partial L}{\partial x_i} \right)^2 \leq 0,$$

where the equality holds in the above inequality if and only if  $F(x) = 0$ . Hence  $L(x(t))$  is decreasing in  $t$ , and strictly decreasing except where  $x(t)$  is a stationary point of  $F$ .

We now apply Corollary 6.6 of Benaïm [1] to show that the limit set of  $(\bar{x}(n))_{n \in \mathbb{N}}$  will almost surely be a stationary point of  $F$ . As  $\Delta^2$  is precompact and we have identified a strict Lyapunov function, the only condition which needs to be checked is that  $F$  has countably many stationary points, which follows from Proposition 3.1. In fact, Proposition 3.1 shows that  $F$  has no connected sets of stationary points other than single points, so the limit set must be a single point, which is one of the stationary points of  $F$ .  $\square$

Note that the Lyapunov function  $L$  generalises in an obvious way to more than three types, as long as all off-diagonal entries are equal.

**Proposition 3.3.** *Consider  $\mathcal{P}_{a,\beta}(z) = az^{\beta+1} - z^\beta + (1+a)z - 2a$  for  $z \in \mathbb{R}^+$ , with  $a, \beta > 0$ .*

- (i) *For a given value of  $a > 1$ ,  $\mathcal{P}_{a,\beta}$  has only one root at  $z = 1$ .*
- (ii) *For a given value of  $a < 1$ , there exists  $\beta_1(a)$  satisfying  $\frac{1+2a}{1-a} > \beta_1(a) > 1$  such that*
  - (a) *If  $\beta > \frac{1+2a}{1-a}$  then  $\mathcal{P}'_{a,\beta}(1) < 0$  and we have that  $\mathcal{P}_{a,\beta}$  has three roots in  $\mathbb{R}^+$ ,  $1$ ,  $r_-$  and  $r_+$ , labelled so that  $r_- < 1 < r_+$ . As functions of  $\beta$  for fixed  $a$ ,  $r_+$  is increasing and  $r_-$  is decreasing.*
  - (b) *If  $\beta_1(a) < \beta < \frac{1+2a}{1-a}$  then  $\mathcal{P}'_{a,\beta}(1) > 0$  and  $\mathcal{P}_{a,\beta}$  has three roots in  $\mathbb{R}^+$ ,  $1$ ,  $r_1$  and  $r_2$ , labelled so that  $1 < r_1 < r_2$ . As functions of  $\beta$  for fixed  $a$ ,  $r_2$  is increasing and  $r_1$  is decreasing.*
  - (c) *If  $\beta < \beta_1(a)$  then  $\mathcal{P}'_{a,\beta}(1) > 0$  and the only root of  $\mathcal{P}_{a,\beta}$  in  $\mathbb{R}^+$  is  $1$ .*

*Furthermore  $\beta_1(a)$  is an increasing function of  $a$  with  $\beta_1(a) \rightarrow \infty$  as  $a \rightarrow 1$ .*

*Proof.* We start by observing that  $\mathcal{P}_{a,\beta}(0) = -2a$  and that  $\mathcal{P}_{a,\beta}(z) \rightarrow \infty$  as  $z \rightarrow \infty$ , indicating that  $\mathcal{P}_{a,\beta}$  has an odd number of roots in  $\mathbb{R}^+$ , counting multiplicity. Differentiating with respect to  $z$ , we have

$$\mathcal{P}'_{a,\beta}(z) = a(\beta + 1)z^\beta - \beta z^{\beta-1} + 1 + a \quad (26)$$

and

$$\mathcal{P}''_{a,\beta}(z) = z^{\beta-2}(a\beta(\beta + 1)z - \beta(\beta - 1)). \quad (27)$$

Because  $\mathcal{P}''_{a,\beta}(z)$  is increasing on  $z \in \mathbb{R}^+$  there cannot be more than three roots of  $\mathcal{P}_{a,\beta}$  in  $\mathbb{R}^+$ .

(i) First, if  $\beta < 1$  we have that  $\mathcal{P}''_{a,\beta}(z) > 0$  in  $\mathbb{R}^+$  implying that  $\mathcal{P}'_{a,\beta}(z)$  is strictly increasing. Moreover,  $\mathcal{P}'_{a,\beta}(z)$  goes from  $-\infty$  to  $+\infty$  when  $z$  ranges from  $0$  to  $+\infty$  and  $\mathcal{P}'_{a,\beta}(1) > 0$ . Then  $\mathcal{P}'_{a,\beta}(z)$  changes sign only once at some  $z^* < 1$ . Now, since  $\mathcal{P}_{a,\beta}(0) = -2a < 0$  and  $\mathcal{P}_{a,\beta}(z)$  is decreasing for  $z < z^* < 1$  and increasing otherwise, it follows that  $\mathcal{P}_{a,\beta}(z)$  crosses the  $z = 0$  line

only once at  $z = 1$ . Second, the same happens for  $\beta > 1$  since  $\mathcal{P}'_{a,\beta}(z) > 0$  in  $\mathbb{R}^+$  and so  $\mathcal{P}_{a,\beta}(z)$  is strictly increasing. The case  $\beta = 1$  is trivial.

(ii) If  $\beta > \frac{1+2a}{1-a}$  we have  $\mathcal{P}'_{a,\beta}(1) = 1 + 2a - \beta(1-a) < 0$ , indicating that in this case  $\mathcal{P}_{a,\beta}$  must have three roots. Note that if  $\beta = \frac{1+2a}{1-a}$  then  $\mathcal{P}'_{a,\beta}(1) = 0$  but that  $\mathcal{P}''_{a,\beta}(1) < 0$ , showing that this is a double root, not a triple root, and so there must be another root in that case for larger  $z$ . If  $\beta < \frac{1+2a}{1-a}$  then  $\mathcal{P}'_{a,\beta}(1) > 0$  and  $\mathcal{P}_{a,\beta}$  has no root less than 1. Then, there must be either none, one double, or two distinct additional roots greater than 1.

The derivative of  $\mathcal{P}_{a,\beta}(z)$  with respect to  $\beta$ , for fixed  $a$  and  $z$ , is  $(az - 1)(\log z)z^\beta$ . Thus, for any  $z \in (1, 1/a)$ ,  $\mathcal{P}_{a,\beta}(z)$  is decreasing in  $\beta$ , and it follows that if there are roots of  $\mathcal{P}_{a,\beta}$  in this range for a particular value of  $\beta$  there must also be for any larger  $\beta$ . As  $\mathcal{P}_{a,\beta}(z) > 0$  if  $z \geq 1/a$ , this shows that there exists  $\beta_1(a) \in [1, \frac{1+2a}{1-a}]$  such that there is one root of  $\mathcal{P}_{a,\beta}$  when  $\beta < \beta_1(a)$  and three when  $\beta > \beta_1(a)$ .

Let  $\beta_0(a) > \frac{2}{1-a} - 1 > 1$  be the unique solution to

$$1 + a = \left( \left( 1 - \frac{2}{\beta + 1} \right) \frac{1}{a} \right)^{\beta - 1}. \quad (28)$$

(It can be seen that (28) has a unique solution for fixed  $a < 1$ , as in that case the right hand side is increasing in  $\beta$  if the right hand side is greater than 1, is equal to 1 at  $\beta = 1$  and tends to  $\infty$  as  $\beta \rightarrow \infty$ . That  $\beta_0(a) > \frac{2}{1-a} - 1$  can be seen by noting that if  $\beta$  is a solution of (28) we must have  $\left( 1 - \frac{2}{\beta + 1} \right) \frac{1}{a} > 1$ .) Then if  $\beta < \beta_0(a)$  we have  $\mathcal{P}'_{a,\beta}(z) > 0$  for all  $z$  and hence  $\mathcal{P}_{a,\beta}$  is increasing in  $z$  and so  $z = 1$  is the only root. This shows that  $\beta_1(a) \geq \beta_0(a) > 1$ .

Now, the fact that as mentioned above there is a root greater than 1 when  $\beta = \frac{1+2a}{1-a}$  together with the continuity of  $\mathcal{P}_{a,\beta}(z)$  in  $\beta$  ensures that there remains a root greater than 1 for  $\beta \in \left( \frac{1+2a}{1-a} - \epsilon, \frac{1+2a}{1-a} \right)$  for some  $\epsilon > 0$ , so  $\beta_1(a) < \frac{1+2a}{1-a}$ .

The claims that  $r_+$  and  $r_2$  are increasing functions of  $\beta$  and that  $r_1$  is a decreasing function of  $\beta$  also follow from the negative derivative of  $\mathcal{P}_{a,\beta}(z)$  with respect to  $\beta$  for  $z \in (1, 1/a)$ . Similarly the claim that  $r_-$  is a decreasing function of  $\beta$  follows from the derivative of  $\mathcal{P}_{a,\beta}(z)$  with respect to  $\beta$  being positive on  $(0, 1)$ .

To see that  $\beta_1(a)$  is an increasing function of  $a$ , note that the derivative of  $\mathcal{P}_{a,\beta}(z)$  with respect to  $a$  is  $z^{\beta+1} + z - 2$ , and for fixed  $z > 1$  this is positive, meaning that if we are in case (c) for particular choices of  $a$  and  $\beta$  we will also be in case (c) for the same value of  $\beta$  and any larger value of  $a$ . That  $\beta_1(a) \rightarrow \infty$  as  $a \rightarrow 1$  follows from  $\beta_0(a) > \frac{2}{1-a} - 1$ .  $\square$

We now investigate the stability of these roots, for which recall the terminology in Definition 1.

**Proposition 3.4.** *If a stationary point for  $F$  is of the form  $(x, x, rx)$  or  $(x, rx, x)$  or  $(rx, x, x)$ , then it is linearly stable if  $\mathcal{P}'_{a,\beta}(r) > 0$  and  $\frac{r^\beta+2}{r+2} > \frac{\beta(1-a)}{2a+1}$ , and linearly unstable if either  $\mathcal{P}'_{a,\beta}(r) < 0$  or  $\frac{r^\beta+2}{r+2} < \frac{\beta(1-a)}{2a+1}$ .*

*Proof.* Without loss of generality we focus on the case  $(x, x, rx)$ .

We note that the differential equation driven by  $F$  keeps the line  $x_1 = x_2$  invariant, so we consider it restricted to this line; the equation for  $F_3$  gives

$$F_3 \left( \frac{1-x_3}{2}, \frac{1-x_3}{2}, x_3 \right) = \frac{2a \left( \frac{1-x_3}{2} \right)^\beta + x_3^\beta}{(1+2a) \left( 2 \left( \frac{1-x_3}{2} \right)^\beta + x_3^\beta \right)} - x_3.$$

Let  $x_3 = z/(z+2)$  so that  $x_1 = x_2 = 1/(z+2)$ . Then

$$F_3 \left( \frac{1-x_3}{2}, \frac{1-x_3}{2}, x_3 \right) = \frac{-2\mathcal{P}_{a,\beta}(z)}{(1+2a)(2+z^\beta)},$$

and so is positive when  $\mathcal{P}_{a,\beta}(z)$  is negative and vice versa. Hence a stationary point  $(x, x, rx)$  is stable in this direction if  $\mathcal{P}'_{a,\beta}(r) > 0$  and linearly unstable in this direction if  $\mathcal{P}'_{a,\beta}(r) < 0$ .

Because  $F$  is symmetric in  $x_1$  and  $x_2$ , the other direction in which we need to consider stability will be perpendicular to this one. Hence we consider

$$\begin{aligned} F_1(x + \epsilon, x - \epsilon, rx) &= \frac{(x + \epsilon)^\beta + (x - \epsilon)^\beta + ar^\beta x^\beta}{(2a + 1)((x + \epsilon)^\beta + (x - \epsilon)^\beta + r^\beta x^\beta)} - x - \epsilon \\ &= F_1(x, x, rx) + \epsilon \left( -1 + \frac{\beta(1-a)}{(2a+1)(2+r^\beta)x} \right) + o(\epsilon) \\ &= F_1(x, x, rx) + \epsilon \left( -1 + \frac{\beta(1-a)(r+2)}{(2a+1)(2+r^\beta)} \right) + o(\epsilon). \end{aligned} \tag{29}$$

It follows that  $(x, x, rx)$  is a stable stationary point in the direction perpendicular to the line  $x_1 = x_2$  if  $\frac{r^\beta+2}{r+2} > \frac{\beta(1-a)}{2a+1}$  and linearly unstable in that direction if the reverse inequality applies.  $\square$

We shall henceforth restrict ourselves to the case  $a < 1$  since Propositions 3.1, 3.3(1) and 3.4 imply that if  $a > 1$ ,  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  is the only stable stationary point for  $F$  and by Lemma 3.2,  $(\bar{x}(n))_{n \in \mathbb{N}}$  must converge to it. The case  $a = 1$  has the probabilities of each colour being  $1/3$  regardless of  $\bar{x}(n)$  and so it is easily seen that  $\bar{x}(n) \rightarrow (1/3, 1/3, 1/3)$  almost surely.

**Corollary 3.5.** *Assume  $a < 1$ .*

- (i) *If  $\beta < \beta_1(a)$ , then the stationary point  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  is stable, and is the limit with probability 1.*
- (ii) *If  $\beta_1(a) < \beta < \frac{1+2a}{1-a}$ , then the stationary points  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  and  $(\frac{1}{r_2+2}, \frac{1}{r_2+2}, \frac{r_2}{r_2+2})$  (and its permutations) are stable, while the stationary point  $(\frac{1}{r_1+2}, \frac{1}{r_1+2}, \frac{r_1}{r_1+2})$  and its permutations are linearly unstable.*
- (iii) *If  $\beta > \frac{1+2a}{1-a}$ , then there are three stationary points of  $F$  of the form  $(x, x, rx)$  corresponding to the three solutions  $r_- < 1 < r_+$  of  $\mathcal{P}_{a,\beta}(z) = 0$  in  $\mathbb{R}^+$ . The stationary points  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  and  $(\frac{1}{r_-+2}, \frac{1}{r_-+2}, \frac{2}{r_-+2})$  (and its permutations) are linearly unstable, while  $(\frac{1}{r_++2}, \frac{1}{r_++2}, \frac{r_+}{r_++2})$  and its permutations are stable.*

*Proof.* (i) Stability follows from Proposition 3.4, and almost sure convergence from Lemma 3.2.

(ii) The shape of  $\mathcal{P}_{a,\beta}$  as discussed in the proof of Proposition 3.3 ensures that  $r_1, r_2 > 1$  and that  $\mathcal{P}'_{a,\beta}(r_1) < 0$  and  $\mathcal{P}'_{a,\beta}(r_2) > 0$ , showing that  $(\frac{1}{r_1+2}, \frac{1}{r_1+2}, \frac{r_1}{r_1+2})$  is linearly unstable, and that for the other two stationary points we just need to check the stability perpendicular to the line  $x_1 = x_2$ . But  $\frac{r_2^\beta+2}{r_2+2} > 1 > \frac{\beta(1-a)}{2a+1}$  by our assumption on  $\beta$ , so the condition from Proposition 3.4 is satisfied and so  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  and  $(\frac{1}{r_2+2}, \frac{1}{r_2+2}, \frac{r_2}{r_2+2})$  are stable. (iii) That  $\mathcal{P}_{a,\beta}(z) = 0$  has three solutions follows from Proposition 3.3. As  $\mathcal{P}'_{a,\beta}(1) < 0$  it follows from Proposition 3.4 that  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  is linearly unstable, and as  $r_- < 1$  we have  $\frac{r_-^\beta+2}{r_-+2} < 1 < \frac{\beta(1-a)}{2a+1}$ , so  $(\frac{1}{r_-+2}, \frac{1}{r_-+2}, \frac{r_-}{r_-+2})$  is also linearly unstable. The global minimum of the Lyapunov function on  $\Delta^2$  must be a stable stationary point of  $F$ , so the remaining stationary points,  $(\frac{1}{r_++2}, \frac{1}{r_++2}, \frac{r_+}{r_++2})$  and its permutations, must be stable.  $\square$

The following result completes the proof of Theorem 1.1.

**Proposition 3.6.** *Let  $l(\bar{x})$  denote the limit set of the process  $(\bar{x}(n))_{n \in \mathbb{N}}$  defined in (10) and recall the stability criteria in Proposition 3.4. Then we have*

- (i)  $\mathbb{P}[l(\bar{x}) = \{p\}] > 0$  for stable points  $p$  of  $F$ .
- (ii)  $\mathbb{P}[l(\bar{x}) = \{p\}] = 0$  for linearly unstable points  $p$  of  $F$ .

*Proof.* Without loss of generality we focus on the case  $(x, x, rx)$ .

(i) Let us now show that the process  $\bar{x}$  in fact converges with positive probability toward a given stable fixed point. Of course, it is necessary that the process has positive probability of being arbitrarily close to the attractor at arbitrarily large times. That is, a point  $p$  is said to be attainable by a process  $X$  if for each  $t > 0$  we have that  $\mathbb{P}[\exists s \geq t : X(s) \in N_p] > 0$  for every neighborhood  $N_p$  of  $p$ . It turns out that if the function  $F + Id$  associated with an urn process  $X$  maps the simplex into its interior, it follows that every point of the simplex is attainable by  $X$ . This is indeed the case for our process  $\bar{x}$ . Finally, Benaïm [1] Theorem 7.3 ensures that a given attainable attractor  $p$  with non-empty basin of attraction is such that  $\mathbb{P}[l(\bar{x}) = \{p\}] > 0$ .

(ii) Let  $p$  be a linearly unstable critical point in the interior of  $\Delta^2$  and  $\mathcal{N}_p \subset \Delta^2$  a neighborhood of  $p$ . The simplex is considered as a differential manifold by identifying its tangent space at any point with the linear subspace  $T\Delta^2 = \{x \in \mathbb{R}^3 : \sum_i x_i = 0\}$ . We need to check Pemantle's non-convergence criteria (Theorem 1 in [9]). As we have  $\gamma_n = 1/(n_0 + n + 1)$  and we have bounded noise, the only condition we need to check is condition (6): that the expectation of the positive part of the component of the noise in any given direction is uniformly bounded away from zero. Formally, we need that whenever  $\bar{x}_n \in \mathcal{N}_p$ , there is a constant  $\kappa$  such that  $\mathbb{E}[\max\{\xi_{n+1} \cdot \theta, 0\} | \mathcal{F}_n] \geq \kappa$  for every unit vector  $\theta = (\theta_1, \theta_2, \theta_3) \in T\Delta^2$ , where  $\xi_{n+1}$  is the noise term from (9). For notational simplicity, write  $\tilde{u}_i = u_i(n)/(\sum_j u_j(n))$  and note that

$$\begin{aligned} \mathbb{E}[\max\{\xi_{n+1} \cdot \theta, 0\} | \mathcal{F}_n] &= \tilde{u}_1 \max\{\theta_1(1 - \tilde{u}_1) - \theta_2\tilde{u}_2 - \theta_3\tilde{u}_3, 0\} + \\ &\quad \tilde{u}_2 \max\{-\theta_1\tilde{u}_1 + \theta_2(1 - \tilde{u}_2) - \theta_3\tilde{u}_3, 0\} + \\ &\quad \tilde{u}_3 \max\{\theta_1\tilde{u}_1 - \theta_2\tilde{u}_2 + \theta_3(1 - \tilde{u}_3), 0\}. \end{aligned} \quad (30)$$

Now, write  $\theta = (\theta_1, \theta_2, \theta_3)$  with  $\theta_1 + \theta_2 + \theta_3 = 0$  and  $\theta_1^2 + \theta_2^2 + \theta_3^2 = 1$ , and suppose that  $\theta_i < 0$  for exactly two coordinates, without loss of generality  $i = 2, 3$ . Then  $\min \theta_2, \theta_3 \leq -\frac{1}{\sqrt{6}}$ , as otherwise  $(-\theta_2 - \theta_3)^2 + \theta_2^2 + \theta_3^2 < 1$ . Then we can write

$$\theta_1(1 - \tilde{u}_1) - \theta_2\tilde{u}_2 - \theta_3\tilde{u}_3 = -\theta_2(1 - \tilde{u}_1 + \tilde{u}_2) - \theta_3(1 - \tilde{u}_1 + \tilde{u}_3) \geq \frac{2a}{(1+2a)\sqrt{6}},$$

as  $\tilde{u}_1 \leq \frac{1}{1+2a}$ , meaning that the first term on the right hand side of (30) is at least  $\frac{2a^2}{(1+2a)^2\sqrt{6}}$ . On the other hand, suppose that  $\theta_i < 0$  for exactly one coordinate; without loss of generality assume  $\theta_3 < 0$ . Then if  $\theta_1 > \theta_2$  we can write

$$\theta_1(1 - \tilde{u}_1) - \theta_2\tilde{u}_2 - \theta_3\tilde{u}_3 = \theta_1(1 - \tilde{u}_1 + \tilde{u}_3) - \theta_2(\tilde{u}_2 - \tilde{u}_3).$$

If  $\tilde{u}_3 \geq \tilde{u}_2$  then this is at least  $\theta_1(1 - \tilde{u}_1 + \tilde{u}_3) \geq \frac{2a}{(1+2a)\sqrt{6}}$ , while if  $\tilde{u}_3 < \tilde{u}_2$  it is at least  $\theta_1(1 - \tilde{u}_1 - \tilde{u}_2 + 2\tilde{u}_3)$ , which is equal to  $3\theta_1\tilde{u}_3$  because  $\tilde{u}_1 + \tilde{u}_2 + \tilde{u}_3 = 1$ . Hence in that case

$\theta_1(1 - \tilde{u}_1) - \theta_2\tilde{u}_2 - \theta_3\tilde{u}_3 \geq \frac{3a}{(1+2a)\sqrt{6}}$ , so that again the first term on the right hand side of (30) is at least  $\frac{3a^2}{(1+2a)^2\sqrt{6}}$ . If  $\theta_2 \geq \theta_1$  similar arguments show the the same but for the second term. Hence we can set  $\kappa = \frac{2a^2}{(1+2a)^2\sqrt{6}} > 0$ , and apply Theorem 1 of [9] to conclude the result.  $\square$

## 4 Proofs for the cyclic case

### 4.1 Introduction

In this section we prove Theorem 1.2. Here  $A$  is the matrix

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix},$$

and we have

$$\begin{aligned} F_1(x_1, x_2, x_3) &= \frac{x_1^\beta + x_2^\beta}{2(x_1^\beta + x_2^\beta + x_3^\beta)} - x_1 \\ F_2(x_1, x_2, x_3) &= \frac{x_2^\beta + x_3^\beta}{2(x_1^\beta + x_2^\beta + x_3^\beta)} - x_2 \\ F_3(x_1, x_2, x_3) &= \frac{x_3^\beta + x_1^\beta}{2(x_1^\beta + x_2^\beta + x_3^\beta)} - x_3 \end{aligned}$$

First, we note that for any choice of  $\beta$ ,  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  is a stationary point of  $F$ . The following two results give information on its stability and show that it is in fact the only stationary point.

**Lemma 4.1.** *For the vector field  $F$ , the stationary point at  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  is stable if  $\beta < 4$ , and a linearly unstable source if  $\beta > 4$ .*

*Proof.* As we are working with  $x \in \Delta^2$ , write  $x_3 = 1 - x_1 - x_2$ . Routine calculus then shows that the Jacobian matrix at  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  is

$$\begin{pmatrix} \frac{\beta}{2} - 1 & \frac{\beta}{2} \\ -\frac{\beta}{2} & -1 \end{pmatrix}.$$

The eigenvalues of this Jacobian are then the roots  $\lambda$  of

$$\lambda^2 + \lambda \left(2 - \frac{\beta}{2}\right) - \left(\frac{\beta}{2} - 1\right) + \left(\frac{\beta}{2}\right)^2$$

which are

$$\left(\frac{\beta}{4} - 1\right) \pm \frac{i\sqrt{3}\beta}{4}.$$

Since the real part of both eigenvalues is positive if  $\beta > 4$  and negative if  $\beta < 4$ , the result follows.  $\square$

**Lemma 4.2.** *The only stationary point of  $F$  in  $\Delta^2$  is  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ .*

*Proof.* For  $x = (x_1, x_2, x_3) \in \Delta^2$  with  $F_1(x) = F_2(x) = F_3(x) = 0$ , we have  $x_1 = \frac{x_1^\beta + x_2^\beta}{x_2^\beta + x_3^\beta} x_2$  and similarly  $x_2 = \frac{x_2^\beta + x_3^\beta}{x_3^\beta + x_1^\beta} x_3$  and  $x_3 = \frac{x_3^\beta + x_1^\beta}{x_1^\beta + x_2^\beta} x_1$ . Using this,

$$x_1 - x_2 = x_2 \frac{x_1^\beta - x_3^\beta}{x_2^\beta + x_3^\beta},$$

indicating that (if  $x_2 > 0$ ) if  $x_1 > x_2$  then also  $x_1 > x_3$ , while if  $x_1 < x_2$  then  $x_1 < x_3$ . Similarly, if  $x_3 > 0$  then the signs of  $x_2 - x_3$  and  $x_2 - x_1$  are the same, and if  $x_1 > 0$  then the signs of  $x_3 - x_1$  and  $x_3 - x_2$  are the same. Hence the only stationary point of  $F$  in the interior of  $\Delta^2$  is  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ .

It is also easy to check that if  $x_1 = 0$  then  $x_2 = 0$ , and similarly that if  $x_2 = 0$  then  $x_3 = 0$  and if  $x_3 = 0$  then  $x_1 = 0$ . Hence there are no stationary points of  $F$  on the boundary of  $\Delta^2$ .  $\square$

We can now complete the proof of Theorem 1.2.

That we only have one stationary point, and that it is never a saddle, restricts the possibilities for chain transitive sets. In two dimensions Theorem 6.12 of Benaïm [1] states that chain transitive sets must be unions of stationary points, periodic orbits and cyclic orbit chains. However, with only one stationary point which is not a saddle cyclic orbit chains are impossible. We can thus conclude that the limit set must be a connected union of periodic orbits and stationary points.

By Lemma 4.1, if  $\beta < 4$  then by Lemma 4.1 the stationary point  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  is stable, and hence is an attractor for the flow given by  $F$ . As in the proof of Proposition 3.6, by Theorem 7.3 of Benaïm [1], to show that there is positive probability of convergence to  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  it is enough

to show that it is an attainable point, that is that we have that  $\mathbb{P}[\exists s \geq t : X(s) \in N] > 0$  for every neighborhood  $N$  of  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ . This is straightforward to show: for any  $\epsilon > 0$  there will be points of the form  $(\frac{n_1}{n}, \frac{n_2}{n}, \frac{n_3}{n})$  with  $n_1, n_2, n_3$  integers satisfying  $n_1 + n_2 + n_3 = n$  and  $\max\{\frac{n_1}{n} - \frac{1}{3}, \frac{n_2}{n} - \frac{1}{3}, \frac{n_3}{n} - \frac{1}{3}\} < \epsilon$  for arbitrarily large  $n$ . Because even if only one colour is present in the urn initially, there will be positive probability that a second colour is chosen at the first time step, and once two colours are present in the urn all three colours have positive probability of being chosen at each step, there will be positive probability of any such point being reached, so  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  is indeed attainable.

If  $\beta > 4$ , then by Lemma 4.1  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  is linearly unstable, so as in Proposition 3.6 it will follow that it is a limit with probability zero if we have positive expectation of the positive part of the component of the noise in any given direction. In fact, the same argument as in Proposition 3.6 will work here, except that the explicit bounds for the  $\tilde{u}_i$  must be replaced by the fact that in a neighbourhood of  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  there will exist  $\delta$  such that  $\delta < \tilde{u}_i < 1 - \delta$  for each  $i$ , giving  $\kappa = \delta^2/\sqrt{6}$ . So we can conclude that  $\bar{x}(n)$  has probability zero of converging to a stationary point, and it follows that the limit set must be a periodic orbit or a connected union of periodic orbits, completing the proof of Theorem 1.2.

## 5 Examples and simulations

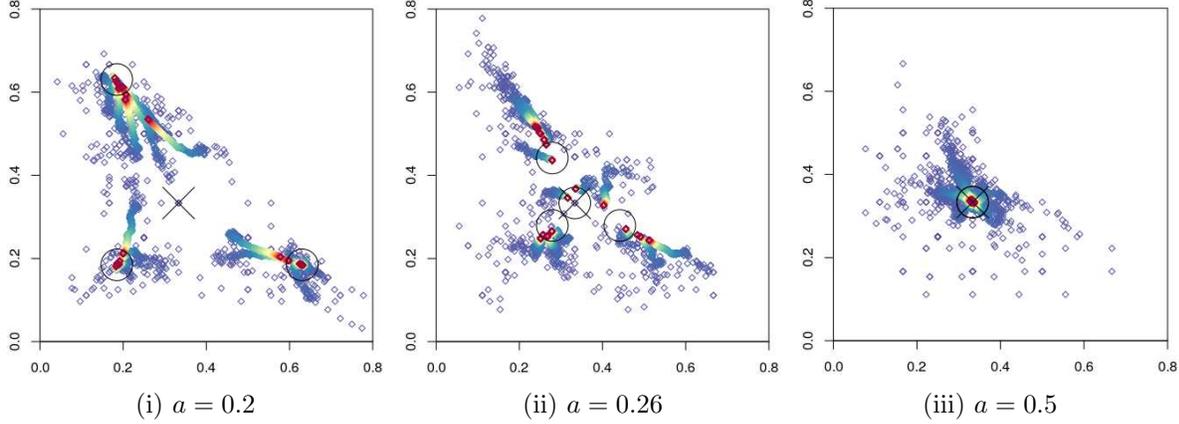
In this section we consider some examples, including some where exact calculations are possible, and some simulations. We also consider some examples which go beyond the cases covered by Theorems 1.1 and 1.2.

### 5.1 The symmetric case with $\beta = 2$

In the case of Theorem 1.1 with  $\beta = 2$ , the possible limits and the phase transitions can be explicitly identified. We find that  $\mathcal{P}_{a,2}(z) = (z - 1)(az^2 + (a - 1) + 2a)$ , with roots given by  $z = 1$  and  $z = \frac{1-a \pm \sqrt{1-2a-7a^2}}{2a}$ . If  $a < \frac{\sqrt{8}-1}{7}$ , then these are real and positive, and letting  $\lambda_1 = \frac{3a+1-\sqrt{1-2a-7a^2}}{4(2a+1)}$ ,  $\lambda_2 = \frac{3a+1+\sqrt{1-2a-7a^2}}{4(2a+1)}$ ,  $\lambda_3 = \frac{a+1-\sqrt{1-2a-7a^2}}{2(2a+1)}$  and  $\lambda_4 = \frac{a+1+\sqrt{1-2a-7a^2}}{2(2a+1)}$ , we obtain linearly stable stationary points  $(\lambda_1, \lambda_1, \lambda_4)$ ,  $(\lambda_1, \lambda_4, \lambda_1)$  and  $(\lambda_4, \lambda_1, \lambda_1)$ , and linearly unstable stationary (except at  $a = \frac{1}{4}$ ) points  $(\lambda_2, \lambda_2, \lambda_3)$ ,  $(\lambda_2, \lambda_3, \lambda_2)$ ,  $(\lambda_3, \lambda_2, \lambda_2)$ . In addition, the stationary point  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  is linearly stable if  $a > \frac{1}{4}$ , and linearly unstable if  $a < \frac{1}{4}$ .

If  $a > \frac{\sqrt{8}-1}{7}$ , then  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  is the only stationary point, and is stable. In the notation of Theorem 1.1, we have  $\beta_1\left(\frac{\sqrt{8}-1}{7}\right) = 2$ , and the three phases are as follows:

Figure 1: 20 simulations of the symmetric model for  $\beta = 2$



- When  $a < \frac{1}{4}$ ,  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  is linearly unstable; there are three other stationary points,  $(\lambda_1, \lambda_1, \lambda_4)$  and permutations, placed symmetrically, which are stable. For example when  $a = \frac{1}{5}$ ,  $(0.1847, 0.1847, 0.6306)$  and permutations are stable. Almost surely, one of these three points will be the limit. A simulation of 20 trajectories of the stochastic process in this case appears in Figure 1i.
- For  $\frac{1}{4} < a < \frac{\sqrt{8}-1}{7}$ ,  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  is now stable but there are also stable stationary points elsewhere, near  $(\frac{1}{4}, \frac{1}{4}, \frac{1}{2})$ . In this case, both symmetric and asymmetric limits have positive probability. For example, at  $a = 0.26$ , there are stable stationary points at  $(0.2792, 0.2792, 0.4416)$  and permutations as well as  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ . A simulation of 20 trajectories of the stochastic process in this case appears in Figure 1ii.
- For  $a > \frac{\sqrt{8}-1}{7}$ ,  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  is the only stationary point, and is stable, and will be the limit almost surely. A simulation of 20 trajectories of the stochastic process in the  $a = \frac{1}{2}$  case appears in Figure 1iii.

At the critical value  $a = \frac{1}{4}$ ,  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}) = (\lambda_2, \lambda_2, \lambda_3)$  has zeros as eigenvalues of its Jacobian and so is neither linearly stable nor linearly unstable, while there are stable stationary points at  $(\frac{1}{4}, \frac{1}{4}, \frac{1}{2})$  and permutations; similarly at the critical value  $a = \frac{\sqrt{8}-1}{7}$  the stationary point  $(\lambda_2, \lambda_2, \lambda_3) = (\lambda_1, \lambda_1, \lambda_4)$  is neither linearly stable nor linearly unstable.

## 5.2 The symmetric case with $\beta = 3$

It is also possible to do some explicit calculations when  $\beta = 3$ . In this case  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  is linearly stable when  $a > \frac{2}{5}$  and linearly unstable when  $a < \frac{2}{5}$ , and we have

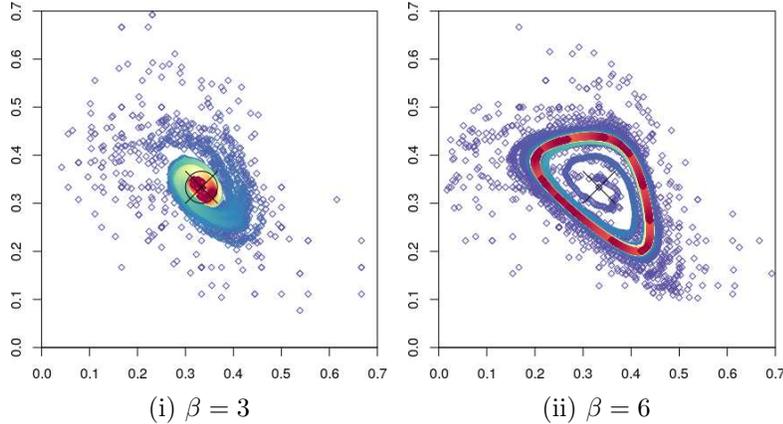
$$\mathcal{P}_{a,3}(z) = (z - 1)(az^3 + (a - 1)z^2 + (a - 1)z + 2a),$$

where the cubic factor has one real root (which is negative) when  $a > a_c = \frac{1}{166}(3 \cdot (2)^{1/2} \cdot (3)^{1/4} + 24 \cdot (3)^{1/2} + 13 \cdot (2)^{1/2} \cdot (3)^{3/4} - 20) = 0.4160306$  and three real roots (one of which is negative) when  $a < a_c$ . (In the notation of Theorem 1.1,  $\beta_1(a_c) = 3$ .) Hence for  $a > a_c$  we get almost sure convergence to  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ , for  $a < \frac{2}{5}$  we get almost sure convergence to one of three asymmetric stationary points, and for  $\frac{2}{5} < a < a_c$  the process may converge either to  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  or to an asymmetric stationary point, each with positive probability.

## 5.3 The cyclic model

Figure 2i shows 20 simulations of the cyclic model when  $\beta = 3$ , showing convergence to  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ . Figure 2ii shows 20 simulations with  $\beta = 6$ , showing apparent convergence to a single limit cycle.

Figure 2: 20 simulations of the cyclic model



## 5.4 The symmetric case with more than three types

It is natural to extend Section 3 to  $d > 3$  types, letting  $A$  be the  $d \times d$  matrix with  $a_{ii} = 1$  for each  $i$  and  $a_{ij} = a$  for  $i \neq j$ . It is straightforward to extend the Lyapunov function (25)

to this case, meaning that Lemma 3.2 applies. However, the later calculations, starting with Proposition 3.1, do not apply. It thus may be interesting to investigate whether more complex patterns of phases can occur in this case than when  $d = 3$ ; however, numerical solution of the stationary point equations for particular values of  $a$  when  $d = 4, 5, 6$  suggests that the behaviour is in fact very similar to the  $d = 3$  case, with three phases which parallel those found in Theorem 1.1.

## 5.5 A more general cyclic case

It is natural to extend Section 4 by allowing the matrix  $A$  to take the form

$$\begin{pmatrix} 1 & a & 0 \\ 0 & 1 & a \\ a & 0 & 1 \end{pmatrix},$$

allowing for different strengths of the cyclic reinforcement. It is straightforward to extend Lemma 4.1 to this case, showing that the stationary point at  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  is linearly stable if  $a \geq 2$  or if  $a < 2$  and  $\beta < \frac{2(1+a)}{2-a}$ , and linearly unstable if  $a < 2$  and  $\beta > \frac{2(1+a)}{2-a}$ . However Lemma 4.2 does not apply for general  $a$  and there may be other stationary points.

Numerical investigation when  $\beta = 2$  suggests that there are three phases in  $a$ : in addition to a phase with almost sure convergence to  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  and a phase with convergence to a limit cycle, there is a phase up to  $a \approx 0.25057$  where there are stable stationary points other than  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  and that the process usually converges to one of these.

## Acknowledgements

M.C.'s research was supported by CONICET [grant number 10520170300561CO]. The authors are grateful to Andrew Wade for suggestions that improved the presentation.

## References

- [1] M. Benaïm. Dynamics of stochastic approximation algorithms. *Séminaire de Probabilités, XXXIII*, 1709:1–68, 1999.

- [2] M. Benaïm, I. Benjamini, J. Chen, and Y. Lima. A generalized Pólya’s urn with graph based interactions. *Random Structures & Algorithms*, 46(4):614–634, 2015.
- [3] M. Costa. *Cooperative Models of Stochastic Growth*. PhD thesis, Durham University, 2018.
- [4] B. Davis. Reinforced random walk. *Probability Theory and Related Fields*, 84:203–229, 1990.
- [5] K. Khanin and R. Khanin. A probabilistic model for the establishment of neuron polarity. *Journal of Mathematical Biology*, 42:26–40, 2001.
- [6] Sophie Laruelle, Gilles Pagès, et al. Nonlinear randomized urn models: a stochastic approximation viewpoint. *Electronic Journal of Probability*, 24, 2019.
- [7] M. Launay and V. Limic. Generalized interacting urn models. <https://arxiv.org/abs/1207.5635>, 2012.
- [8] R. I. Oliveira. The onset of dominance in balls-in-bins processes with feedback. *Random Structures & Algorithms*, 34(4):454–477, 2009.
- [9] R. Pemantle. Nonconvergence to unstable points in urn models and stochastic approximations. *Ann. Probab.*, 18(2):698–712, 04 1990.
- [10] R. Pemantle. A survey of random processes with reinforcement. *Probability Surveys*, 4:1–79, 2007.
- [11] R. van der Hofstad, M. Holmes, A. Kuznetsov, and W. Ruzel. Strongly reinforced Pólya urns with graph-based competition. *The Annals of Applied Probability*, 26(4):2494–2539, 2016.