The diffusion of opposite opinions in a randomly biased environment

Manuel González-Navarrete*, Rodrigo Lambert †

Abstract

We propose a model for diffusion of two opposite opinions. Here, the decision to be taken by each individual is a random variable which depends on the tendency of the population, as well on its own trend characteristic. The influence of the population trend can be positive, negative or non-existent in a random form. We prove a phase transition in the behaviour of the proportion of each opinion. Namely, the mean square proportions are linear functions of time in the diffusive case, but are given by a power law in the superdiffusive regime.

Keywords: stochastic model, diffusion, phase transition, Bernoulli sequences, Pólya urn, memory lapses

1 Introduction

The spreading of a disease, rumour, political position or music preference has been a widely studied theme along the years and can be found in the literature of several scientific research areas. Most of the times these studies involve some mathematical model. Since it is reasonable to assume that these processes are inherently probabilistic, then it is suitable to adopt a stochastic approach. In this line, statistical mechanics principles can be considered. As illustrations, we quote a voter-model approach [22] and a random-field Ising interaction [6]. Essentially, the main feature is the existence of a phase transition in such models.

Following another direction, Bendor et al. [3] introduced a model for the diffusion of opinions as a probabilistic choice process. This model has its roots on Festinger’s hypothesis [12] and is formulated as follows. Two alternatives, \( \mathcal{A} \) and \( \mathcal{B} \), diffuse through a population (or decision makers, following the author’s terminology). It is assumed that there is a positive initial number of “converted” decision makers. In every period, one decision maker makes up his (her) mind about whether to adopt \( \mathcal{A} \) or \( \mathcal{B} \). The decision taken at the \( n \)-step is denoted by a Bernoulli variable \( X_n \). In this sense, \( X_n = 1 \) means \( n \)-decision was the innovation \( \mathcal{A} \) (and \( X_n = 0 \) means \( \mathcal{B} \)-decision). Let us denote by \( N_n = N_0 + X_1 + \ldots + X_n \) the number of decision makers which adopted \( \mathcal{A} \) and \( M_n \) those that adopted \( \mathcal{B} \), before the \( (n+1) \)-th decision. Therefore, for \( n \geq 1 \) the BHW (Bendor-Huberman-Wu) model is defined by the conditional probabilities

\[
P(X_{n+1} = 1| (N_n, M_n)) = \theta p + (1 - \theta) \frac{N_n}{N_n + M_n},
\]  

where \( 0 < \theta \leq 1 \) and \( P(X_1 = 1) = N_0/(N_0 + M_0) \). When \( \theta = 0 \), we recover the well-known classical Pólya-urn process (for details, see [21]). In another case, if \( \theta = 1 \), \( \{X_i\}_{i \in \mathbb{N}} \) is an independent and

*Departamento de Estadística, Universidad del Bío-Bío, Chile. e-mail: magonzalez@ubiobio.cl
†Faculdade de Matemática, Universidade Federal de Uberlândia and Departamento de Matemática, Universidade Federal da Bahia, Brazil. e-mail: rodrigolambert@ufu.br
identically distributed (i.i.d.) Bernoulli process and then \( N_n \) follows a binomial distribution. We remark, as exposed in [3], that each decision is taken by a linear combination of own and social influences, being represented by parameter \( p \) and proportion \( \frac{N_n}{N_n + M_n} \), respectively.

The role of parameters \( p \) and \( \theta \) in the mean proportion of opinions \( A \) (namely \( \mathbb{E} \left[ \frac{N_n}{N_n + M_n} \right] \)), is discussed in some details in [3]. Particular cases from (1.1) include behaviours such as herding (see [1, 5]) and the so-called soft technologies when \( p \sim 1/2 \) and \( \theta \) is close to zero (see [3]).

In a different framework, Drezner and Farnum [11] introduced a generalized binomial distribution. In that model, the independence between Bernoulli trials is replaced by an assumption similar to (1.1). By using another parametrization and \( P(X_1 = 1) = p \), the authors showed that, as with the case of the binomial distribution, \( \mathbb{E}(X_n) = p \) and \( \text{Var}(X_n) = p(1 - p) \), for all \( n \geq 1 \). However, \( \text{Var}(N_n) \) is not the same as in the binomial distribution.

We remark that this family of models finds applications in different areas, such as, epidemiology [25] or train accidents prevention [19, 20], among others. Moreover, an important attention has been shown in a class of random walks with long-range correlation [17, 23], expressed by conditional probabilities as in (1.1). The relation of such processes with Pólya-type urns was exposed in [2] (we refer the reader to Ref. [13] for further details).

Regarding the asymptotic behaviour of \( N_n \), Heyde [16] has obtained a law of large numbers that provides a limiting proportion for \( N_n \) (and so for \( M_n \)). Formally, the author proved that, for all \( \theta > 0 \), it holds

\[
\lim_{n \to \infty} \frac{N_n}{N_n + M_n} = p \quad \text{a.s.} \tag{1.2}
\]

In the same paper it was stated a central limit theorem whenever \( \theta \geq 1/2 \), and the existence of a phase transition at \( \theta = 1/2 \). These results are presented in a summarized form in the next theorem. We state this results in terms of the BHW model and present it here without proof.

**Theorem 1.** (Heyde, 2004 [16]) Let \( \{X_i\}_{i \geq 1} \) defined by conditional probabilities (1.1) and \( P(X_1 = 1) = p \). The following assertions hold.

(i) If \( \theta > \frac{1}{2} \),

\[
\frac{N_n - np}{\sqrt{n}} \overset{d}{\to} \mathcal{N} \left( 0, \frac{p(1 - p)}{2\theta - 1} \right), \quad \text{as } n \to \infty. \tag{1.3}
\]

(ii) If \( \theta = \frac{1}{2} \),

\[
\frac{N_n - np}{\sqrt{n \log n}} \overset{d}{\to} \mathcal{N}(0, p(1 - p)), \quad \text{as } n \to \infty. \tag{1.4}
\]

(iii) If \( 0 < \theta < \frac{1}{2} \),

\[
\frac{N_n - np}{n^{1-\theta}} \overset{d}{\to} W, \quad \text{as } n \to \infty, \tag{1.5}
\]

where \( W \) is a proper random variable such that

\[
\mathbb{E}(W) = 0, \quad \mathbb{E}(W^2) = \frac{p(1 - p)}{(1 - 2\theta)\Gamma(2(1 - \theta))} \quad \text{and} \quad \mathbb{E}(W^3) \neq 0 \quad \text{if} \quad p \neq 1/2. \tag{1.6}
\]

Here the notation \( \overset{d}{\to} \) denotes convergence in distribution, and \( Z \sim \mathcal{N}(\mu, \sigma^2) \) means that \( Z \) is a normal random variable with mean \( \mu \) and variance \( \sigma^2 \).
We remark some features about the results in this theorem. The diffusive case (i) is \(\sqrt{n}\)-scaled and the critical one (ii) has \(\sqrt{n \log n}\) as scale factor. The superdiffusive regime (iii) converges under the \(n^{1-\theta}\) scaling. We recall that in the diffusive and critical cases the limiting distribution is a normal random variable, instead of the superdiffusive regime, which presents a quite different limiting behaviour. Similar results than (i) – (ii) of this theorem was discussed in [3]. The rigorous proof of these results can be viewed in [16] and [10].

In the present paper we propose a generalized approach for the spreading of two opposite opinions or preferences. The main contribution is the fact that each decision maker presents a random bias, which can be positive, negative or neutral. In words, it means that any decision maker has a tendency factor, which is inherent to each individual, and works as a latent process that imposes a random ambiance for this diffusion. This particular feature characterizes the so-called random environment in which the diffusion occurs.

For this model, we prove a law of large numbers for the proportion of each decision on the population (Theorem 2). Moreover, in the diffusive and critical cases the limiting distributions are showed to be normal random variables and a functional limit theorem can be also obtained (Theorems 3 and 4). Instead, the superdiffusive region exhibits non-normal limiting distribution which depends on the initial proportion of opinions (Theorem 3 (iii)), then a particular case is studied (Example 3).

2 The diffusion model and main results

Imagine a population in which two opposite opinions \(A\) and \(B\) spread. Suppose that this diffusion process starts with \(N_0 + M_0\) seeders, where \(N_0\) is the initial number of individuals which has opinion \(A\), and \(M_0\) those that have opinion \(B\). Each individual in the population has one of the following three intrinsic characteristics: trend-follower, against-trend or indifferent. This intrinsic characteristic is independently imputed to each individual and follows a ternary random variable, defined as follows.

**Definition 1.** For \(\alpha, \beta \in (0, 1)\) such that \(0 \leq \alpha + \beta \leq 1\), the trend process \(\{Y_n\}_{n \in \mathbb{N}}\) is an i.i.d. sequence of ternary random variables (which is also independent of the current proportions of \(A\) and \(B\) decisions on the population) such that, for each \(n \geq 1\)

\[
P(Y_n = y) = \begin{cases} 
  \alpha & \text{if } y = +1 \quad \text{(Trend-follower)}, \\
  \beta & \text{if } y = -1 \quad \text{(Against-trend)}, \\
  1 - \alpha - \beta & \text{if } y = 0 \quad \text{(Indifferent)}.
\end{cases}
\]

We are then ready to introduce the diffusion model with randomly biased environment.

**Definition 2.** The random-trend diffusion model is a stochastic process \(\{X_n\}_{n \in \mathbb{N}}\) taking values on \(\{0, 1\}^{\mathbb{N}}\) defined by the following conditional probabilities

\[
P(X_{n+1} = 1|\mathcal{F}_n, Y_n) = a + bY_n \frac{N_n}{N_n + M_n},
\]

where \(N_n\) (resp. \(M_n\)) is given by the number of \(A\)-decisions (resp. \(B\)-decisions) taken until step \(n\). \(\mathcal{F}_n = \mathcal{F}_n(N_n, M_n)\) is the filtration related to random variables \(N_n, M_n\). In view of Definition 1 we see that \(Y_n\) is independent of \(\mathcal{F}_n\), for all \(n \geq 1\). Parameters \(a\) and \(b\) satisfy \(0 \leq a + b \leq 1\), and also the following condition: if \(\beta \neq 0\), then \(b \leq a\).
In this sense, the trend process \( \{Y_n\}_{n \in \mathbb{N}} \) can be seen as a latent process. In words, it is a process that is not possible to see its input, but that influences the output of each realization of \( \{X_n\}_{n \in \mathbb{N}} \). In what follows, we present an example of how this diffusion can be seen. It is a version of an application presented in [13] and illustrates a potential application for the randomly biased diffusion model.

**Example 1.** At any given time \( n \), a customer can buy a product \((X_n = 1)\) by necessity or by social pressure (history of selling). In the present diffusion model, we consider the conditional probabilities (2.1). Let be parameter \( a \) the quantity related to necessity and parameter \( b \) the quantity related to the social pressure of the product as defined in [3, 26], associated to \((n + 1)\)-th customer. Then, if \( Y_n = 1 \), the individual is a trend-follower, and the social pressure will raise the probability of the customer buying the product. If \( Y_n = -1 \), the customer is against the trend, then the social pressure will decrease the probability of he/she buying the product. Finally, if \( Y_n = 0 \), the customer behaves without taking the social pressure into account.

As a particular case, when \( \beta = 0 \) we recover the dependent random walks with memory lapses introduced in [13]. We say that a \( \ell \)-sized memory lapse is given by a string \( Y_{i+\ell} = 0 \), in which \( Y_{i+\ell} \) is the short notation for \( Y_i, Y_{i+1}, \ldots, Y_{i+\ell} \), and \( 0^\ell \) means \( \ell \) consecutive zeros. In a memory lapse situation, the process is occurring independently of the past. When the lapse stops, the process recovers its whole memory. This includes the piece of memory built at the lapse period too (we refer the reader to [13, 14] for more details, and to [4, 7, 24] to similar ideas).

In what follows we state our first result. It is a strong law of large numbers which provides the limiting proportion of individuals which has taken decisions \( A \) or \( B \), as \( n \) diverges.

**Theorem 2.** Consider the random-trend diffusion model as in Definition 2. If \( 0 \leq a + b \leq 1 \) and \( b \leq a \) we get

\[
\lim_{n \to \infty} \left( \frac{N_n}{N_n + M_n}, \frac{M_n}{N_n + M_n} \right) = \left( \frac{a}{1 - b(\alpha - \beta)}, \frac{1 - a - b(\alpha - \beta)}{1 - b(\alpha - \beta)} \right) \quad a.s. \tag{2.2}
\]

In the case \( \beta = 0 \), this theorem holds for all \( a, b \) such that \( 0 \leq a + b \leq 1 \). In the following results this also happens for \( \beta = 0 \).

Now, we state a phase transition from diffusive to superdiffusive behaviours. In this sense, the parameter space is essentially divided by a critical line which depends on the values of \( b \) and the difference \( \alpha - \beta \). As will be shown, each case demands a particular time-scaling.

**Theorem 3.** Suppose that the hypothesis of Theorem 2 are satisfied

(i) If \( b(\alpha - \beta) < \frac{1}{2} \), then

\[
\lim_{n \to \infty} \frac{1}{\sqrt{n}} \left( N_n, M_n \right) - n \left( \frac{a}{1 - b(\alpha - \beta)}, \frac{1 - a - b(\alpha - \beta)}{1 - b(\alpha - \beta)} \right) \overset{d}{\to} \mathcal{N}(0, \Sigma_1), \tag{2.3}
\]

with covariance matrix \( \Sigma_1 \) given by

\[
\Sigma_1 = \frac{a(1 - a - b(\alpha - \beta))}{(1 - b(\alpha - \beta))^2(1 - 2b(\alpha - \beta))} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}. \tag{2.4}
\]

(ii) If \( b(\alpha - \beta) = \frac{1}{2} \), then


\[
\frac{1}{\sqrt{n \log(n)}} \left[ (N_n, M_n) - n \left( \frac{a}{1 - b(\alpha - \beta)}, \frac{1 - a - b(\alpha - \beta)}{1 - b(\alpha - \beta)} \right) \right] \overset{d}{\rightarrow} \mathcal{N}(0, \Sigma_2), \tag{2.5}
\]

where the covariance matrix \( \Sigma_2 \) is given by
\[
\Sigma_2 = 2a(1 - 2a) \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}. \tag{2.6}
\]

(iii) Suppose that \( \frac{1}{2} < b(\alpha - \beta) < 1 \). Then we have tightness for the sequence of two-dimensional random variables \( \{Z_n\}_{n \geq 2} \) given by
\[
Z_n := \frac{1}{n^{b(\alpha - \beta)}} \left[ (N_n, M_n) - n \left( \frac{a}{1 - b(\alpha - \beta)}, \frac{1 - a - b(\alpha - \beta)}{1 - b(\alpha - \beta)} \right) \right].
\]

Furthermore, there exists a random bi-dimensional vector \( \hat{W} = (\hat{W}_1, \hat{W}_2) \) such that, as \( n \) diverges
\[
Z_n \rightarrow \hat{W} \quad \text{a.s.}
\]

Note that (i) and (ii) are versions of the central limit theorem. In the case of superdiffusive regime in (iii), we recall that condition \( b(\alpha - \beta) > 1/2 \) implies that \( \alpha \) is a bit larger than \( \beta \). In view of the trend process \( \{Y_n\}_{n \in \mathbb{N}} \), one can foretell that the positive trend is expected to be kept along the diffusion. In other words: the statement of this Theorem exhibits an explicit dependence on the initial composition of the model. This result is related to a class of models called large urns, which are usually difficult to obtain explicit limit quantities (for details, see [8]).

In what follows, we state particular cases to exemplify the previous theorem.

**Example 2.** Consider a random-trend model as in Definition 2 with \( \beta = 1 \) and \( a = \frac{1 + b}{2} \). Then, we obtain the so-called critical Polya urn introduced in [24]. For this model, there is no phase transition and the asymptotic behaviour is characterized by a normal random variable given in item (i).

**Example 3.** Let \( \beta = 0 \) and \( \alpha = \frac{1 - 2a}{6} \). Then, there is a phase transition at \( a = \frac{1}{4} \). In particular, if we start with a single A opinion, the random vector \( \hat{W} \) in item (iii), satisfies, for \( a < \frac{1}{4} \)
\[
- \mathbb{E}(\hat{W}) = \frac{1}{2 \Gamma(2(1 - a))} (1, -1);
- \mathbb{E}(\hat{W}^2) = \frac{1}{2 \Gamma(3 - 4a)} \left( 1 + \frac{(1 - 2a)^2}{1 - 4a} \right) (1, -1) \text{ and }
- \mathbb{E}(\hat{W}^3) = \frac{1}{2 \Gamma(2(2 - 3a))} \left( 1 + \frac{(5 - 2a)(1 - 2a)^2}{1 - 4a} \right) (1, -1).
\]

This example is related to the elephant random walk introduced in [23] and the results hold from the cumulants obtained in Example 3.27 in [18]. In particular, the asymptotic position of the elephant can be obtained by computing \( \hat{W}_1 - \hat{W}_2 \).

Finally, we present a functional central limit theorem for our model. In fact, it is a continuous-time version of Theorem 3 (i)-(ii) which deals with a convergence to a two dimensional Brownian motion and provides an explicit limiting covariance matrix. This result is also known in the literature.
as Donsker’s invariance principle with applications arising in estimation, testing, model selection, goodness of fit, regression, and many other common statistical contexts (for details see [9, 15]). We recall that this convergence holds on the function space $D[0,\infty)$ of right-continuous and left-bounded (càdlàg) functions, also known as Skorokhod space.

**Theorem 4.** Suppose that the hypothesis of Theorem 2 are satisfied.

(i) If $b(\alpha - \beta) < \frac{1}{2}$ then, for $n \to \infty$, in $D[0,\infty)$

$$\frac{1}{\sqrt{n}} \left[ (N_{[tn]}, M_{[tn]}) - tn \left( \frac{a}{1 - b(\alpha - \beta)}, \frac{1 - a - b(\alpha - \beta)}{1 - b(\alpha - \beta)} \right) \right] \overset{d}{\to} W_t,$$

where $W_t$ is a continuous bivariate Gaussian process with $W_0 = (0,0)$, $\mathbb{E}(W_t) = (0,0)$ and, for $0 < s \leq t$,

$$\mathbb{E}(W_sW_t^T) = s \left( \frac{t}{s} \right)^{b(\alpha - \beta)} \frac{a(1 - a - b(\alpha - \beta))}{(1 - b(\alpha - \beta))^2(1 - 2b(\alpha - \beta))} \left( \begin{array}{cc} 1 & -1 \\ -1 & 1 \end{array} \right).$$

(ii) If $b(\alpha - \beta) = \frac{1}{2}$ then, for $n \to \infty$, in $D[0,\infty)$

$$\frac{1}{\sqrt{n \log(n)}} \left[ (N_{[nt]}, M_{[nt]}) - nt(2a, 1 - 2a) \right] \overset{d}{\to} W_t,$$

where $W_t$ is a continuous bivariate Gaussian process with $W_0 = (0,0)$, $\mathbb{E}(W_t) = (0,0)$ and, for $0 < s \leq t$,

$$\mathbb{E}(W_sW_t^T) = 2sa(1 - 2a) \left( \begin{array}{cc} 1 & -1 \\ -1 & 1 \end{array} \right).$$

3 Proofs

This section is devoted to provide the paths and explicit proofs for the statements given in the paper. We start by constructing the so-called random replacement matrix, which was defined in [18] as a generalisation of the replacement matrix for general Pólya urn models (see [21], for instance). The eigendecomposition of the random replacement matrix will then provide the main ingredients for the proofs.

### 3.1 Building the random reinforcement matrix

Now we expose the relation between the random-trend diffusion model and a generalized Pólya urn with random reinforcement matrix. This construction will be done in order to bring the random-trend problem to an urn model and therefore apply the well-known results for generalized Pólya urns provided by [18].

Firstly we provide an explicit relationship between the distribution of $(N_n, M_n)_{n \in \mathbb{N}}$ and the distribution of the balls in a two-color urn. Then, we construct the random replacement matrix for the related generalized Pólya urn. We recall that here and hereafter we will follow the notation given in [18]. In this sense, consider the following two column replacement vectors $\xi_1 = (\xi_{11}, \xi_{12})^T$ (red) and
ξ_2 = (ξ_{11}, ξ_{12})^T \text{ (blue)}, \text{ with } ξ_i \in \{(0, 1)^T, (1, 0)^T\} \text{ (a single ball is replaced at each time)}, \text{ and the random replacement matrix given by } M = (ξ_1; ξ_2). \text{ The urn dynamics evolves as follows. If we choose replacement vector } ξ_1 \text{ to reinforce the urn, it means that we will replace } ξ_{11} \text{ red balls and } ξ_{12} \text{ blue balls. Otherwise we choose vector } ξ_2, \text{ and then replace } ξ_{21} \text{ red, and } ξ_{22} \text{ blue balls. At this time, we will be interested in the so-called mean replacement matrix, which is given by (with some abuse of notation)}

\[
A = \mathbb{E}(M) = \begin{pmatrix} \mathbb{E}(ξ_{11}) & \mathbb{E}(ξ_{21}) \\ \mathbb{E}(ξ_{12}) & \mathbb{E}(ξ_{22}) \end{pmatrix}.
\]

Note that \( P(ξ_{ij} = 1) = \mathbb{E}(ξ_{ij}) \), for all \( i, j \). \text{ If at time } n \text{ we get } r \text{ red balls } (R_n = r), \text{ then the probabilities of replacing a red ball (success) or a blue ball (failure) at time } n+1 \text{ (conditioned on a proportion } r/T_n \text{ of red balls, with } T_n \text{ the total number of balls at time } n) \text{ are, respectively, given by}

\[
P(R_{n+1} = r+1|R_n = r) = \mathbb{P}(ξ_{11} = 1) \frac{r}{T_n} + \mathbb{P}(ξ_{21} = 1) \left(1 - \frac{r}{T_n}\right) + \mathbb{E}[ξ_{21}] + (\mathbb{E}[ξ_{11}] - \mathbb{E}[ξ_{21}]) \frac{r}{T_n},
\]

and

\[
P(R_{n+1} = r|R_n = r) = 1 - P(R_{n+1} = r+1|R_n = r) = \mathbb{E}[ξ_{22}] + (\mathbb{E}[ξ_{12}] - \mathbb{E}[ξ_{22}]) \frac{r}{T_n}.
\]

In the case of the random-trend diffusion model in Definition 2, note that, since \( Y_n \) and \((N_n, M_n)\) are independent we get that the \textit{mean probability of a success} is given by

\[
P(X_{n+1} = 1|F_n) = \sum_{y \in \{-1, 0, 1\}} P(X_{n+1} = 1|F_n, Y_n = y)P(Y_n = y) = \frac{a + b(α - β)}{N_n + M_n}.
\]

We also get that the \textit{mean probability of a failure} is given by

\[
P(X_{n+1} = 0|F_n) = 1 - P(X_{n+1} = 1|F_n) = 1 - a - b(α - β) \frac{N_n}{N_n + M_n}.
\]

Now we remark the following: in (3.3) and (3.4), the term \( N_n/(N_n + M_n) \) plays the role of the term \( r/T_n \) in (3.1) and (3.2). This correspondence allows us to obtain the entries of the matrix \( A \) by simply solving the following system:

\[
\begin{align*}
\mathbb{E}[ξ_{21}] &= a, \\
\mathbb{E}[ξ_{22}] &= 1 - a, \\
\mathbb{E}[ξ_{11}] - \mathbb{E}[ξ_{21}] &= b(α - β), \\
\mathbb{E}[ξ_{12}] - \mathbb{E}[ξ_{22}] &= -b(α - β).
\end{align*}
\]

Then we obtain the following matrix

\[
A = \begin{pmatrix} a + b(α - β) & a \\ 1 - a - b(α - β) & 1 - a \end{pmatrix}.
\]

The eigenvalues of \( A \) are given (in decreasing order) by

\[
7
\]
λ₁ = 1 and λ₂ = b(α − β).

Moreover, the correspondent right and left eigenvectors (which normalizations provided by equations (2.2) and (2.3) of [18]) are given by

\[ v₁ = \frac{1}{1 - b(α - β)} \begin{pmatrix} a \\ 1 - a - b(α - β) \end{pmatrix} ; \quad v₂ = \frac{1}{1 - b(α - β)} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \]

\[ u₁ = (1, 1) \quad \text{and} \quad u₂ = (1 - a - b(α - β), -a). \]

Now let’s check conditions (A1)-(A6) in Section 2 of [18]. Since each entry of A is the expected value of a Bernoulli random variable in \{0, 1\}, we get (A1)-(A2). The facts that λ₁ > 0 and λ₂ < λ₁ provide (A3)-(A4). Since both colors belong to a dominating class and the urn starts with a positive number of balls, (A5)-(A6) are then satisfied. Moreover, since the urn is tenable, and since it is impossible to exclude balls in the present dynamics, we get that the essentially non-extinction condition is also satisfied.

3.2 Proof of Theorem 2

Proof. Since λ₂ = b(α − β) < 1 = λ₁ we apply Theorem 3.21 of [18]. It says that if \( R_n \) (respect. \( B_n \)) is the number of red (respect. blue) balls at time \( n \), then for \( U_n := (R_n, B_n) \) we get

\[ \lim_{n \to \infty} \frac{U_n}{n} = λ₁v₁ \text{ a.s.} \]

Since \( M_n + N_n = N_0 + M_0 + n \), we get that \( M_n + N_n \approx n \) when \( n \) is big. This completes the proof.

3.3 Proof of Theorem 3

Proof. (i)

Theorem 3.22 of [18] to our case says that the limiting covariance matrix is given by

\[ Σ₁ = \int_0^∞ ψ_A(s)Bψ_A(s)ᵀe^{-s}ds - v₁v₁ᵀ, \quad (3.5) \]

which in turn depends on the following quantities

(I) \( ψ_A(s) = e^{sA} - v₁(1 1) \int_0^s e^{tA}dt \).

(II) \( B = \sum_{i=1}^2 v₁iB_i \), with \( B_i = \mathbb{E}[ξ_iξ_iᵀ] \).

We start by computing term (II). By definition, \( B = v₁₁B₁ + v₁₂B₂ \), where (again with some abuse of notation)

\[ B_i = \mathbb{E}[ξ_iξ_iᵀ] = \begin{pmatrix} \mathbb{E}[ξ_i²] & \mathbb{E}[ξ_i₁ξ_i₂] \\ \mathbb{E}[ξ_i₂ξ_i₁] & \mathbb{E}[ξ_i₂²] \end{pmatrix} = \begin{pmatrix} \mathbb{E}[ξ_i₁] & 0 \\ 0 & \mathbb{E}[ξ_i₂] \end{pmatrix}, \]

where \( ξ_iᵀ \) is the transpose of \( ξ_i \), and the last equality comes from the following facts: for all \( i \in 1,2 \) \( ξ_{ij} ∈ \{0,1\} \) and the replacement rule does not allow \( ξ_i₁ \) and \( ξ_i₂ \) being 1 at the same time. By a direct computation we conclude that
\[
B = \begin{pmatrix}
\frac{a}{1-b(\alpha-\beta)} & 0 \\
0 & \frac{1-a-b(\alpha-\beta)}{1-b(\alpha-\beta)}
\end{pmatrix}.
\] (3.6)

For term (I), we use diagonalization \( A = V\Lambda V^{-1} \), where \( V \) is a diagonal matrix and then \( e^A = Ve^{\Lambda}V^{-1} \). In this case,
\[
V = \begin{pmatrix} 1 & \frac{a}{1-b(\alpha-\beta)} \\ 1 & \frac{1-a-b(\alpha-\beta)}{1-b(\alpha-\beta)} \end{pmatrix} \quad \text{and} \quad \Lambda = \begin{pmatrix} 1 & 0 \\ 0 & b(\alpha-\beta) \end{pmatrix},
\]
allowing us to obtain \( \psi_A(s) \) directly.

Since the condition \( b(\alpha-\beta) < 1/2 \) provides convergence for the integral in (3.5), we use a mathematical software to compute this and then the proof is finished.

(ii) Since \( A \) is diagonalizable, Theorem 3.23 of [18] to our case says that the limiting covariance matrix is given by
\[
\Sigma_2 = (I - T_1) P_T P_T^* (I - T_1)^T,
\] (3.7)
where \( P_T^* = P_T^T \) is the Hermitian conjugate of \( P_T \) (and the equality holds since it is a real-entries matrix). Moreover \( I \) is the two-dimensional identity matrix and \( T_1 \) is given by
\[
T_1 = v_1.(1 1) = \frac{1}{1-b(\alpha-\beta)} \begin{pmatrix} a & a \\ 1-a-b(\alpha-\beta) & 1-a-b(\alpha-\beta) \end{pmatrix}.
\] (3.8)

We note that the above equality was obtained by recalling that, in view of [18], since \( \det(A^T) = 1/2 \neq 0 \), \( A^T \) has full rank, and then the action vector (which in our case is given by \( v' = (1 1) \)) satisfies \( v' \in \text{Im}(A^T) \). Moreover \( \vec{b}' = (1 1) \) satisfies the condition \( v' = A^T \vec{b}' \). Furthermore the fact that \( \Lambda_{II} = \{\lambda_2\} = \{b(\alpha-\beta)\} = \{1/2\} \) leads to
\[
P_T = v_2 u_2 = 2 \begin{pmatrix} \frac{1}{2} - a & a \\ a - \frac{1}{2} & -a \end{pmatrix}.
\] (3.9)

Since (3.6), (3.8) and (3.9) provide all the ingredients, we finally compute (3.7).

Finally, for item (iii) we just need to notice that the eigenspace \( E_{\lambda_2} \) is the set of vectors on the direction of \((1, -1)\) and then apply Theorem 3.24 of [18]. Theorem 3.26 of [18] provides the moment generator function for the Example 3.

\[\] 3.4 Proof of Theorem 4

Proof. (i) (The diffusive case)

The convergence statement is directly provided by Theorem 3.31 – (i) of [18], with limiting covariance matrix given by
\[
\mathbb{E}(W_sW_t^T) = \int_{-\log s}^{\infty} \psi_A(x + \log s) B \psi_A(x + \log t) e^{-x} dx - sv_1v_1^T,
\]
which is equivalent to obtain (see Remark 5.7 from [18])
\[
\mathbb{E}(W_sW_t^T) = s \Sigma_1 e^{(\log(t/s)A^T).}
\] (3.10)
Now, as done in the proof of Theorem 3, we obtain $A^T = V\Lambda V^{-1}$, and then $e^{A^T} = Ve^{A}V^{-1}$. In this case, they are given by

$$V = \begin{pmatrix} 1 & a+b(\alpha-\beta) \frac{-1}{1} \\ 1 & b(\alpha-\beta) \end{pmatrix}, \quad \Lambda = \begin{pmatrix} 1 & 0 \\ 0 & b(\alpha-\beta) \end{pmatrix} \quad \text{and} \quad V^{-1} = \begin{pmatrix} \frac{a}{1-b(\alpha-\beta)} & \frac{a+b(\alpha-\beta)-1}{b(\alpha-\beta)-1} \\ \frac{a}{b(\alpha-\beta)-1} & \frac{a}{1-b(\alpha-\beta)} \end{pmatrix}.$$ 

A direct calculation completes the proof.

(ii) (The critical case)

Once again the convergence statement is directly provided by Theorem 3.31 – (ii) of [18], with $d := d_{\lambda_2} = 0$, since $A$ is diagonalizable. Now let us compute the limiting covariance matrix. To do this, notice that $\tilde{c}(d, s, t) = s$ and then

$$\mathbb{E}(W_sW^T_t) = s(I - 2v_1P_{\lambda_2})\Sigma_{II}(I - 2P_{\lambda_2}^T1^Tv_1).$$

Actually we use (3.9). Note that,

$$1P_{\frac{1}{2}} = (0 \ 0) ; \ P_{\frac{1}{2}}^T1^T = (0 \ 0)^T.$$ 

Then $\mathbb{E}(W_sW^T_t) = P_{\frac{1}{2}}BP_{\frac{1}{2}}^T$. Concluding the proof, we use (3.6) and (3.9) to obtain the following limiting covariance matrix

$$\mathbb{E}(W_sW^T_t) = 2sa(1 - 2a) \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.\quad \square$$

Acknowledgements

The authors thank Durval Tonon and Alejandra Rada-Mora for helpful comments and discussions. MGN is supported by Fondo Especial DIUBB 1901083-RS and FAPEI from Universidad del Bío-Bío. RL is partially supported by FAPESP (grant 2014/19805-1), and CNPQ (grants 406324/2017-4, 439422/2018-3, and 435470/2018-3). This work is part of the Universal FAPEMIG project “Dinâmica de recorrência para shifts aleatórios e processos com lapsos de memória” (process APQ-00987-18).

References

[1] Banerjee, A.V. (1992) A simple model of herd behavior. Quarterly Journal of Economics 107, 797-818.

[2] Baur, E. and Bertoin, J. (2016) Elephant Random Walks and their connection to Pólya-type urns. Phys. Rev. E 94, 052134.

[3] Bendor, J., Huberman, B. and Wu, F. (2009) Management fads, pedagogies, and other soft technologies. J. Econ. Behav. Organ. 72, 290-304.

[4] Bertoin, J. Noise reinforcement for Lévy processes. Preprint [arXiv:1810.08364], 2018.
[5] Bikhchandani, S., Hirshleifer, D. and Welch, I. (1992) A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy* **100**, 992-1026.

[6] Bouchaud, J.P. (2013) Crises and Collective Socio-Economic Phenomena: Simple Models and Challenges. *J. Stat. Phys.* **151**(3-4), 567–606.

[7] Businger, S. (2018) The Shark Random Swim. *J. Stat. Phys.* **172**(3), 701–717.

[8] Chauvin, B., Pouyanne, N. and Sahnoun, R. (2011) Limit distributions for large Pólya urns. *Ann. Appl. Probab.* **21**, 1–32.

[9] DasGupta, A. (2008) Asymptotic Theory of Statistics and Probability. Springer, New York.

[10] Drezner, Z. (2006) On the Limit of the Generalized Binomial Distribution. *Comm. Statist. Theory Methods* **35**(2), 209–221.

[11] Drezner, Z. and Farnum, N. (1993) A generalized binomial distribution. *Comm. Statist. Theory Methods* **22**, 3051–3063.

[12] Festinger, L. (1954) A theory of social comparison processes. *Hum. Relat.* **7**(2), 117–140.

[13] González-Navarrete, M. and Lambert, R. (2018) Non-Markovian random walks with memory lapses. *J. Math. Phys.* **59**, 113301.

[14] González-Navarrete, M. and Lambert, R. Urn models with two types of strategies. Preprint arXiv:1708.06430, 2019.

[15] Heyde, C.C. (1981) Invariance principles in statistics. *Int. Stat. Rev.* **49**(2), 143–152.

[16] Heyde, C.C. (2004) Asymptotics and criticality for a correlated Bernoulli process. *Aust. N. Z. J. Statist.* **46**, 53–57.

[17] Hod, S. and Keshet, U. (2004) Phase transition in random walks with long-range correlations. *Phys. Rev. E* **70**, 015104(R).

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

[19] Liu, X. and Hong, Y. (2015) Analysis of railroad tank car releases using a generalized binomial model. *Accid. Anal. Prev.* **84**, 20–26.

[20] Liu, X., Liu, C. and Hong, Y. (2017) Analysis of multiple tank car releases in train accidents. *Accid. Anal. Prev.* **107**, 164–172.

[21] Mahmoud, H. (2008) *Pólya urn models*. CRC Press.

[22] Nyczka, P. and Sznajd-Weron, K. (2013) Anticonformity or Independence? Insights from Statistical Physics. *J. Stat. Phys.* **151**(1-2), 174–202.

[23] Schütz, G. and Trimper, S. (2004) Elephants can always remember: Exact long-range memory effects in a non-Markovian random walk. *Phys. Rev. E* **70**, 045101.

[24] Yadav, A. (2018) Critical Pólya urn. *Phys. Rev. E* **98**, 022119.
[25] ZELTERMAN, D. (2004) Discrete Distributions. Applications in the Health Sciences. John Wiley & Sons, New York.

[26] ZHU, H. AND HUBERMAN, B. A. (2014) To switch or not to switch understanding social influence in online choices. Am. Behav. Sci. 58(10), 1329–1344.