Università degli Studi Roma Tre

Facoltà di Scienze Matematiche, Fisiche e Naturali

Large deviations for Generalized Polya Urns with general urn functions

PhD Thesis in Physics

Candidate:

Simone Franchini

Supervisor:

Pietro Caputo

Dept. Mathematics, Università Roma 3

A.A. 2014/2015

Abstract

We consider a generalized two-color Polya urn (black and withe balls) first introduced by Hill, Lane, Sudderth [HLS1980], where the urn composition evolves as follows: let $\pi : [0,1] \to [0,1]$, and denote by x_n the fraction of black balls at step n, then at step $n+1$ a black ball is added with probability $\pi(x_n)$ and a white ball is added with probability $1-\pi(x_n)$. We discuss large deviations for a wide class of continuous urn functions π. In particular, we prove that this process satisfies a Sample-Path Large Deviations principle (SPLDP), also providing a variational representation for the rate function. Then, we derive a variational representation for the limit

$$\phi(s) = \lim_{n\to\infty} \tfrac{1}{n} \log \mathbb{P}\left(\{nx_n = \lfloor sn \rfloor\}\right), \; s \in [0,1], \tag{0.1}$$

where nx_n is the number of black balls at time n, and use it to give some insight on the shape of $\phi(s)$. Under suitable assumptions on π we are able to identify the optimal trajectory. We also find a non-linear Cauchy problem for the cumulant generating function and provide an explicit analysis for some selected examples. In particular, we discuss the linear case, which is strictly related to the so-called Bagchi-Pal urn, giving the exact implicit expression for ϕ in therms of the Cumulant Generating Function.

Contents

1 Introduction

Urns are simple probabilistic models that had a broad theoretical development and applications for several decades, gaining a prominent position within the framework of adaptive stochastic processes. In general, single-urn schemes are Markov chains that consider a set (urn) containing two or more elements of different types: at each step a number of elements is added or removed with some probabilities, dependent on the composition of the urn. Since their introduction, those models where intended to describe phenomena where an underlying tree growth is present [Pem2007, Mam2003, JK1977, Mam2008].

Given the general definition above, an impressive number of variants has been introduced, depending on the number of colors, extraction and replacement rules, etc. This work focuses on Large Deviations Principles (LDP) for a generalization of the classical Polya-Eggenberger two-colors urn scheme, first introduced by Hill, Lane and Sudderth [HLS1980]. Let us consider an infinite capacity urn which contains two kinds of elements, say black and withe balls, and denote by $X_n := \{X_{n,k} : 0 \leq k \leq n\}$ the number of black balls during the urn evolution from time 0 to n: at time k there are k balls in the urn, $X_{n,k}$ of which are black. Given a map $\pi : [0,1] \to [0,1]$ (usually referred to as urn function) the urn evolves as follows: let $x_{n,k} := k^{-1}X_{n,k}$, $1 \leq k \leq n$ be the fraction of black balls in the urn at step k, then a new ball is added at step $k+1$, whose color is black with probability $\pi(x_{n,k})$ and white with probability $1 - \pi(x_{n,k}) = \bar{\pi}(x_{n,k})$ (hereafter we denote the complementary probability by an upper bar),

$$X_{n,k+1} = \begin{cases} X_{n,k} + 1 & \textit{with probability} \quad \pi\left(x_{n,k}\right) \\ X_{n,k} & \textit{with probability} \quad \bar{\pi}\left(x_{n,k}\right) \end{cases} . \tag{1.1}$$

The above model has also been generalized to multicolor urns, whose strong convergence properties have been investigated by Ermoliev et Al. in a series of papers [AEK1986, AEK1986b, AEK1987]. In the present work we restrict our attention to the two-colors case (ie, unidimensional Markov process). Apart form the wide range of behaviors depending on the choice of the urn function, which makes this generalized urn model challenging and rich by itself, attention arises form its relevance to branching phenomena, stochastic approximation and reinforced random walks [HLS1980, Pem2007, Mam2003], as well as some macroeconomic (for example, see [DEK1994]) and biological models ([Pem2007]). We will explore some paradigmatic examples later.

While the a. s. convergence properties of such urns are quite well understood (even in multicolor generalizations, see [AEK1986, AEK1986b, AEK1987]), Large Deviations properties are not. Anyway, the main reason which led us to deal with this model is that it represents a key mathematical tool for a recently introduced semi-phenomenological approach to the random walk range problem [FB2014], where the fraction of occupied sites near the origin for a random walk of given *range* is interpreted as urn function, and used to obtain some insights on the range distribution. We will discuss this arguments in Subsection 2.2, but we remark that the search for Large

Deviations properties of the random walks range embeds topics of critical relevance in physics and applied mathematics, such as the Self-Avoiding Walk problem, the Coil to Globule transition and other long standing issues of polymer theory, applied probability and combinatorics.

We organized this work as follows: in Section 2 we review the main known results about the Generalized Polya Urn (GPU) of Hill, Lane and Sudderth, discussing the classes of urn functions we will consider and introducing some notation. In the same section we also discuss the applications anticipated above, paying particular attention to the Random Walk range problem and the analogies of linear urn function with the Bagchi-Pal urn scheme, a classical urn model with important applications to random trees theory and other related topics (see Subsection 2.2 for a deeper discussion). Then we will expose our main results on large deviations (Section 3): in particular, we will present the theorem statements concerning the Sample Path Large deviations Principle, the large deviations for the event $\{X_{n,n} = \lfloor sn \rfloor\}$, $s \in [0, 1]$ and the Cumulant generating function. In order to keep this section as readable as possible, we will present the theorem statements only, also discussing their applications to the examples of Subsection 2.2. We collected all proof in a dedicated section (Section 4) which contains almost all the technical features of this work. At the end of Section 3 (Subsection 3.4) we also discuss some conjectures and future developments of this work.

2 Urn of Hill, Lane and Sudderth

Here we present the GPU of Hill, Lane and Sudderth, also introducing some non-standard notation which will be useful when dealing with LDPs: we tried to reduce new notation to minimum, keeping the common urn terminology everywhere this was possible.

As we shall see, the initial conditions do not affect the LDPs for the class of urn functions we will consider, unless the urn starts in a monochromatic configurations (all balls are black/white). Then, if not specified otherwise, in this work we set $X_{n,0} := 0$ (the urn starts empty) and $x_{n,0} := 1/2$ by convention: we will discuss the effect of initial conditions on the LDPs in Section 3 . That said, our process $X_n := \{X_{n,k} : 0 \leq k \leq n\}$ is the Markov Chain with transition matrix:

$$\mathbb{P}\left(X_{n,k+1} = X_{n,k} + i \middle| X_{n,k} = j\right) := \pi\left(j/k\right)\mathbb{I}_{\{i=1\}} + \bar{\pi}\left(j/k\right)\mathbb{I}_{\{i=0\}}. \qquad (2.1)$$

We denote by δX_n the associated sequence $\delta X_{n,k} := X_{n,k+1} - X_{n,k} \in \{0,1\}$ for $0 \leq k \leq n-1$. For notational convenience, the dependence on π is not specified.

2.0.1 The urn function π

Throughout this work we will consider a sub-class $\mathcal{U}$ of continuous functions π: $[0,1] \to [0,1]$ such that $\pi(s) \in (0,1)$ for $s \in (0,1)$ and $\pi(0) \in [0,1)$, $\pi(1) \in (0,1]$. We also require that both $\pi(s)$ and $\bar{\pi}(\bar{s})$ admit a factorization that separate eventual zeros near $s \in \{0,1\}$ (ie, if $\pi(0) = 0$ or $\pi(1) = 1$) and such that $1/\pi(s)$ and $1/\bar{\pi}(\bar{s})$ diverges at most algebraically near $s \in \{0,1\}$. Formally, we give the following definition of $\mathcal{U}$:

Definition. *We say that* π: $[0,1] \to [0,1]$ continuous belongs to $\mathcal{U}$ *if we can decompose* $|\log \pi|$ *and* $|\log \bar{\pi}|$ *as* $|\log \pi(s)| = g_0(s) + \nu_0 |\log(s)|$, $|\log \bar{\pi}(\bar{s})| = g_1(\bar{s}) + \nu_1 |\log(\bar{s})|$ *for some* $g_i > 0$ *uniformly continuous and bounded on* $[0,1]$ *and such that some* $f > 0$ *with* $\lim_{\epsilon \to 0} \epsilon \int_\epsilon^1 dz\, |f(z)|/z^2 = 0$ *exists such that* $|g_i(x+\delta) - g_i(x)| \le f(|\delta|)$ *as* $\delta \to 0$.

This is a slightly smaller class of functions than those considered in [HLS1980, Pem2007, Mam2003, Pem1991], where most results are obtained for continuous functions. Even though it may seem quite unnatural, this class has been constructed to include most of the interesting cases that can be described by urn functions, while keeping properties that directly allow to meet the requirements of Varadhan lemma. We will discuss this in Section 3. Also, notice that $\mathcal{U}$ includes some classical urn schemes, such as the Polya-Eggenberger urn (see Section 2.2).

From the above definition it follows that $\pi(0) < 1$, $\pi(1) > 0$ and $\pi(s) \in (0,1)$ for $s \in (0,1)$. Notice that $\nu_0 \log(s)$ and $\nu_1 \log(s)$ are needed in the above decomposition only if $\pi(0) = 0$ and $\pi(1) = 1$. If $0 < \pi(s) < 1$ for all $s \in [0,1]$ we have $\nu_0 = \nu_1 = 0$. It is clear that urn functions with $\pi(0) = 0$ and $\pi(1) = 1$ can be affected by monochromatic initial conditions, since, as example, if $\pi(0) = 0$ and $X_{n,m} = 0$ at some m, then also $X_{n,k} = 0$ for all $k > m$ (the same deterministic evolution would happen for $\pi(1) = 1$ and

$X_{n,m} = m$). We anticipate that these are the only cases in which LDPs can be affected by initial conditions for the considered class of urn functions $\mathcal{U}$.

In the following we introduce some new notation which is intended to ease the description of our results, as well as the limit properties of X_n. Define the following sets:

$$C_\pi := \{s \in (0,1) : \pi(s) = s\}, \, \partial C_\pi := C_\pi \setminus \mathrm{int}(C_\pi), \tag{2.2}$$

where $\mathrm{int}(\mathrm{C}_\pi)$ is the interior of C_π. We will refer to the elements of C_π as *contacts*. Note that for the considered urn functions ∂C_π is a finite set of isolated points: we indicate by $|\partial C_\pi|$ the number of such points for given π.

We can further distinguish the elements of ∂C_π by considering the behavior of $\pi(s)$ in their neighborhood: to do so, we will introduce a partition of the interval $[0,1]$. We remark that the notation we are going to define is not a standard of urn literature, but it will be crucial for the description of our results when dealing with optimal trajectories. First, let us organize the elements of ∂C_π by increasing order, labeling them as

$$\partial C_\pi =: \{s_i, 1 \le i \le |\partial C_\pi| : s_i < s_{i+1}\}. \tag{2.3}$$

Then, we can define the following sequence of intervals

$$\begin{aligned} K_\pi := \big\{K_{\pi,i},\, 0 \le i \le |\partial C_\pi| : K_{\pi,0} := (0, s_1), K_{\pi,|\partial C_\pi|} := \big(s_{|\partial C_\pi|}, 1\big), \\ K_{\pi,j} := (s_j, s_{j+1}), \, 1 \le j \le |\partial C_\pi| - 1\big\}. \end{aligned} \tag{2.4}$$

By definition of ∂C_π, the above intervals are such that $\pi(s) - s$ does not change sign for $s \in K_{\pi,i}$. Then we can associate a variable $a_{\pi,i} \in \{-1, 0, 1\}$ to each interval $K_{\pi,i}$ which expresses the sign of $\pi(s) - s$. We denote such sequence by

$$A_\pi := \{a_{\pi,i},\, 0 \leq i \leq |\partial C_\pi| : \, a_{\pi,i} = \tfrac{\pi(s)-s}{|\pi(s)-s|} \mathbb{I}_{\{\pi(s) \neq s\}},\, s \in K_{\pi,i}\}. \tag{2.5}$$

Some words should be spent on the correct use of this notation when the urn function has $\pi(0) = 0$ or $\pi(1) = 1$ (ie, when $\nu_0 \neq 0$ or $\nu_1 \neq 0$), or both. Consider the first case: if $\pi(0) = 0$ then the smallest element of ∂C_π is $s_1 = 0$. Following our definition of $K_{\pi,0}$ as open interval we would have that $K_{\pi,0} = \emptyset$ and $a_{\pi,0}$ is not well defined. To patch this, we set by convention that $a_{\pi,0} = 1$ if $K_{\pi,0} = \emptyset$ and $a_{\pi,|\partial C_\pi|} = -1$ if $K_{\pi,|\partial C_\pi|} = \emptyset$.

Using the above notation we can now define the subsets $C_\pi(\alpha, \beta)$ of those $s \in \partial C_\pi$ such that $\alpha \in \{+, 0, -\}$ is the sign of $\pi(s') - s'$ for $s' - s \to 0^-$ and $\beta \in \{+, 0, -\}$ is the sign of $\pi(s') - s'$ for $s' - s \to 0^+$.

$$C_\pi(\alpha, \beta) := \{s_i \in \partial C_\pi : \, \mathrm{sign}(a_{\pi,i-1}) = \alpha,\, \mathrm{sign}(a_{\pi,i}) = \beta\} \tag{2.6}$$

References [HLS1980, Pem2007, Mam2003, Pem1991] call $C_\pi(+, -)$ and $C_\pi(-, +)$ respectively *downcrossings* and *upcrossings*, while $C_\pi(+, +)$, $C_\pi(-, -)$ are *touchpoints.* Note that our classification also allows contacts of the kind $C_\pi(\alpha, 0)$ and $C_\pi(0, \beta)$, which are the boundaries of those intervals $K_{\pi,i}$ for which $\pi(s) = s$ ($a_{\pi,i} = 0$).

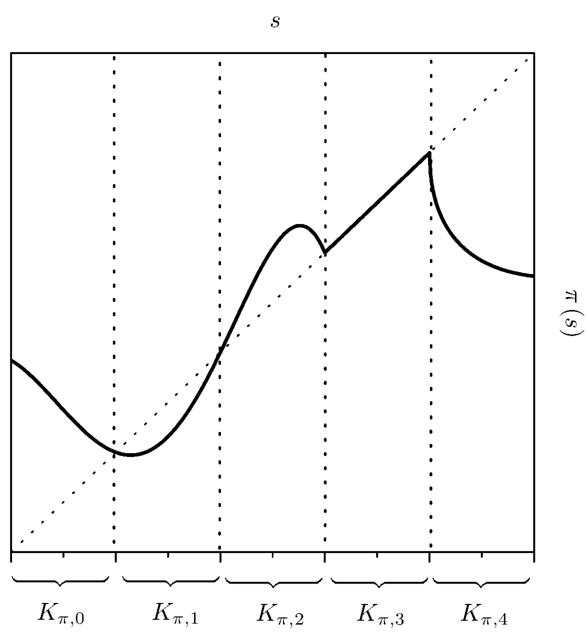


Figure 2.1: Example of urn function $\pi \in \mathcal{U}$ to illustrate the notation introduced in Eq.s (2.4), (2.3). For the function above we have $C_\pi = \{1/5,\, 2/5,\, 3/5,\, 4/5\} \cup (3/5,\, 4/5)$, then $\partial C_\pi = \{1/5,\, 2/5,\, 3/5,\, 4/5\}$, $K_{\pi,0} = [0, 1/5)$, $K_{\pi,4} = (4/5,\, 1]$, $K_{\pi,i} := (1/k,\, 1/\,(k+1))$, $i \in \{1, 2, 3\}$ and $a_\pi = \{1,\, -1,\, 1,\, 0,\, -1\}$.

2.1 Strong convergence

Here we review some of the main known results on *strong convergence*, ie, on the almost sure convergence of $x_{n,n}$). This is a topic widely investigated in [HLS1980, Pem2007, Mam2003, Pem1991]). As example, consider the simplest non trivial urn model, the so called Polya-Eggenberger urn [EP1923], which evolves as follows: at each step draw a ball, if it is black then add a black ball, and add a white one otherwise. This urn is represented in our context by the urn function $\pi(s) = s$. In this case $\mathbb{E}(x_{n,k+1}|x_{n,k}) = x_{n,k}$, so that $x_{n,k}$ is a martingale and $\lim_n x_{n,n}$ exists almost surely. Moreover, let $X_{n,m} = X^*_m \notin \{0, m\}$ at some fixed starting time $m > 0$, then $x_{n,n}$ can reach any state $s \in [X^*_m/n, 1 - m/n + X^*_m/n]$ with positive probability, and therefore $\mathbb{P}(\lim_n x_{n,n} \in [s_1, s_2]) > 0$ for any $0 < s_1 < s_2 < 1$.

The existence of $\lim_n x_{n,n}$ has been shown in [HLS1980] for a much wider class of urn functions (including even non-continuous π). In [HLS1980] it has been shown that if π is a continuous function then $\lim_n x_{n,n}$ exists almost surely, and $\lim_n x_{n,n} \in C_\pi$. The same result holds if π is non-continuous, provided the points s where $\pi(s) - s$ oscillates in sign are not dense in an interval.

Clearly, not all the points of C_π can be the limit of $x_{n,n}$ and several efforts were made to determine if a point belongs to the support of $\lim_n x_{n,n}$ for a given π [HLS1980, Pem1991]. We say that $s \in [0, 1]$ belongs to the support of $\lim_n x_{n,n}$ if we have $\mathbb{P}(|\lim_n x_{n,n} - s| \leq \delta) > 0$ for arbitrary small $\delta > 0$. The following theorem summarizes what is known about the support of $\lim_n x_{n,n}$ in our setting ($\pi \in \mathcal{U}$).

Theorem. *(Hill, Lane, Sudderth, Pemantle [HLS1980, Pem1991]) Let* X_n *be the urn process generated by the urn function* $\pi \in \mathcal{U}$*, and define*

$$\Delta_{\pi,\epsilon}(s) := \epsilon^{-1}\left[\pi(s+\epsilon) - \pi(s)\right]. \tag{2.7}$$

Then the limit $\lim_n x_{n,n}$ *exists almost surely and*

1) Downcrossings $C_\pi(+,-)$ *always belong to the support of* $\lim_n x_{n,n}$ *while upcrossings* $C_\pi(-,+)$ *never do.*

2) If $s \in C_\pi(+,+)$*, then it belongs to the support of* $\lim_n x_{n,n}$ *if and only if some* $\delta > 0$ *exists such that* $\Delta_{\pi,\epsilon}(s) \in (1/2, 1)$ *for* $\epsilon \in (-\delta, 0)$*.*

3) If $s \in C_\pi(-,-)$*, then it belongs to the support of* $\lim_n x_{n,n}$ *if and only if some* $\delta > 0$ *exists such that* $\Delta_{\pi,\epsilon}(s) \in (1/2, 1)$ *for* $\epsilon \in (0, \delta)$*.*

The proof that downcrossings belong to the support of $\lim_n x_{n,n}$, while upcrossings don't, can be found in reference [HLS1980]: it involves Markov chain coupling together with martingale analysis. The statement that touchpoints $C_\pi(+,+)$ with $1/2 < \Delta_{\pi,\epsilon}(s) < 1$ from the left ($\epsilon < 0$) and $C_\pi(-,-)$ with $1/2 < \Delta_{\pi,\epsilon}(s) < 1$ from the right ($\epsilon > 0$) belong to the support of $\lim_n x_{n,n}$ has been proved in [Pem1991] by Pemantle. This seemingly paradoxical statement is actually a deep observation about the dynamics of the process: if the condition on $\Delta_{\pi,\epsilon}(s)$ is fulfilled, then $x_{n,n}$ converges so slowly to $s \in C_\pi(+,+)$ from the left (to $s \in C_\pi(-,-)$ from the right) that it almost surely never crosses this point, accumulating in its left (right) neighborhood. If not, then $x_{n,n}$ crosses s in finite time almost surely, and gets pushed away from the other side toward the closest stable equilibrium (ie, the closest point that belongs to the support of $\lim_n x_{n,n}$).

Even if we left out the cases $C_\pi(\alpha, 0)$, $C_\pi(0, \beta)$ and $s \in K_{\pi,i}$ with $a_{\pi,i} = 0$ from the above theorem, it is believable that they always belong to the support of $\lim_n x_{n,n}$ since in some neighborhood of these points the process behaves like a Polya-Eggenberger urn.

We remark that almost sure convergence is strongly affected by initial conditions: since a detailed discussion of this topic would be far from the scopes of this work, we refer to the reviews [HLS1980, Pem2007, Mam2003, Pem1991].

2.2 Some applications

Below we collect some situations to which the results presented in the next section can be applied. Given the wide variety of models that can be constructed with appropriate urn functions, we will discuss only just a few that we think are particularly relevant, or exemplary in their scope. First, we discuss the connection with the problem of the Random Walk range. Then we will show how to use the urn functions to deal with other urn models, in particular, the Bagchi-Pal model. After that, we will see a classical model of Market Share and, finally, we will discuss a model for neuron growth presented by Khanin and Khanin in [KK2001].

2.2.1 Range of a random walk

Large deviations for the range of a Random Walk (number of different lattice sites visited by the walk) represent a classical topic in pure and applied mathematics, as well as polymer theory and other important physics settings.

Let $\omega_n = \{S_i : 0 \leq i \leq n\}$ be a Simple Random Walk (SRW) of n steps on the $\mathbb{Z}^d$ lattice, with $S_i \in \mathbb{Z}^d$, $|S_{i+1} - S_i| = 1$, $S_0 = 0$. Then, let denote by

$$R[\omega_n] := \sum_{x \in \mathbb{Z}^d} \mathbb{I}_{\{x \in \omega_n\}} \tag{2.8}$$

the cardinality of the range of ω_n. The problem of finding some characterizations for events such $\{R[\omega_n] \leq rn\}$ and $\{R[\omega_n] \leq rn\}$ has been the subject of extensive studies, which led to some important consequences in the theory of large deviations.

For the event $\{R[\omega_n] \leq rn\}$ some very precise results have been obtained. For example, by sub-additivity arguments it is possible to prove that the rate function satisfies

$$\phi_-(r) = -\lim_n n^{-1} \log \mathbb{P}(R[\omega_n] \leq rn) = 0. \tag{2.9}$$

We point out that very precise results concerning so-called *moderate deviations* have been obtained for the volume of the Wiener sausage. The Wiener sausage is the neighborhood of the trace of a standard Brownian motion up to a time t: given a standard Brownian motion $\eta(t)$, the Wiener sausage W_t of unitary radius is the process defined by

$$W_t = \bigcup_{0 \leq \tau \leq t} \left\{ x \in \mathbb{R}^d : |\eta(\tau) - x| < 1 \right\}. \tag{2.10}$$

This is essentially a continuous version of the range for a lattice random walk. In a remarkable work by Van den Berg, Bolthausen, and den Hollander [VBD2001] it is shown that the following limit

$$I\left(r\right)=-\lim_{t\to\infty}t^{-(d+2)/d}\log\mathbb{P}\left(\left|W_t\right|\leq rt\right) \tag{2.11}$$

exists and is finite. They also proved some very precise statements about the behavior of $I\left(r\right)$ (see [VBD2001]) and we expect that the same limit with similar properties also exists for the event $\left\{R\left[\omega_n\right]\leq rn\right\}$.

Despite many efforts, relatively few results have been obtained about the event $\left\{R\left[\omega_n\right]\geq rn\right\}$. In [HK2001] Hamana and Kesten showed that the limit

$$\phi_+\left(r\right):=-\lim_{n\to\infty}t^{-1}\log\mathbb{P}\left(R\left[\omega_n\right]\geq rn\right) \tag{2.12}$$

exist. In the same work it is shown that if $d\geq 2$ then $\phi_+\left(r\right)$ is also convex for $r\in[0,1]$ and strictly increasing for $r\in[1-C_d,1]$, where we denote by $1-C_d$ the probability that a Simple Random Walk never visits again its starting site.

The connection between the above problem and the present work lies on the possibility of dealing with the limit

$$\phi\left(m\right):=\lim_{n}n^{-1}\log\mathbb{P}\left(R\left[\omega_n\right]=\lfloor\left(1-m\right)n\rfloor\right),\ m\in\left(0,1\right), \tag{2.13}$$

by looking at the expectation of the *atmosphere* near $S_0=0$

$$\pi\left[\omega_n\right]:=\left(2d\right)^{-1}\sum_{|x|=1}\mathbb{I}_{\{x\in\omega_n\}} \tag{2.14}$$

conditioned to the event $\{R[\omega_n] = \lfloor rn \rfloor\}$. The concept of *endpoint atmosphere* was introduced some years ago within the study of the Self-Avoiding Walk, as the number of ways in which an $n-$steps SAW can be continued by adding the $(n+1)-$th step. It is natural to interpret $\pi[\omega_n]$ as the probability that a random continuation of ω_n from S_0 will lead to a self-intersection. Then, denote by $\pi_n(m)$ the expectation of $\pi[\omega_n]$ for $R[\omega_n] = \lfloor (1-m)n \rfloor$ and some $m \in [0,1]$:

$$\pi_n(m) = \mathbb{E}\left(\pi[\omega_n] \mid R[\omega_n] = \lfloor (1-m)n \rfloor\right). \tag{2.15}$$

If we take $X_n = n - 1 - R[\omega_n]$ and $x_n = n^{-1}X_n$ it's easy to show that the stochastic process $\{X_n : n \geq 0\}$ is a Markov Chain defined by the following transition matrix

$$\mathbb{P}\left(X_{n,k+1} = X_{n,k} + i \middle| X_{n,k} = j\right) := \pi_n(j/k)\,\mathbb{I}_{\{i=1\}} + \bar{\pi}_n(j/k)\,\mathbb{I}_{\{i=0\}}. \tag{2.16}$$

Therefore, $\{X_n : n \geq 0\}$ is an urn of Hill, Lane, Sudderth with non-homogeneous urn function $\pi_n = \{\pi(m) : m \in [0,1]\}$. In this work we show that the large deviations properties of an urn with urn function π_n are the same of that with urn function $\pi = \lim_n \pi_n$, provided that such π exists, that $\pi \in \mathcal{U}$ and $|\pi_n(m) - \pi(m)| \to 0$ uniformly on $m \in [0,1]$.

In [Fra2011] we studied the event $\{R[\omega_n] = \lfloor (1-m)n \rfloor\}$, $m \in [0,1]$ by numerical simulations, finding a *coil globule transition* (the chain collapses from an extended random coil to a liquid-like cluster) for some critical value

$m = m_c \in (0,1)$. In particular, we studied the limit

$$2\nu_d(m) = \lim_n \left\{ \log(n)^{-1} \log \mathbb{E}\left(S_n^2 | R[\omega_n] = \lfloor (1-m)n \rfloor \right) \right\}, \tag{2.17}$$

concluding that it exists and that $\nu_d(m < m_c) = \nu_d$ (where ν_d is the critical exponent governing the end-to-end distance of the Self-Avoiding Walk), $\nu_d(m = m_c) = \nu_c$, $\nu_d(m > m_c) = 1/d$. We observed that for $d = 3, 4$ $\nu_c = 1/2$ and $m_c = C_d$, while for $d \geq 5$ our numerical simulations suggested that $m_c > C_d$ (see [Fra2011] for further details about this topics).

Concerning the atmosphere π_n, in [Fra2011Th, FB2014] we performed numerical simulations for $3 \leq d \leq 6$ in the region $m < C_d$, where the typical configuration of ω_n is supposed to be in the universality class of the self-avoiding walk (see Figures 2.2, 2.3, 2.4 and 2.5). We observe that π_n approaches to some continuous π uniformly on the considered range. Quite surprisingly, we also observe that for $d \geq 5$ the function π is almost linear (see Figures 2.4). Some other properties of $\pi(m)$ can be deduced from known results of Random and Self-Avoiding Walk theory. For example, it can be shown that $\pi(0) = 1 - \eta_d/2d$, where η_d is the connective constant of the Self-Avoiding Walk. Also, it is possible to compute the first derivative of π for $m = C_d$. Since the aim of this subsection is to provide reasons for studying our urn problem, a detailed treatment of these arguments is to be found in [FB2014] (see also [Fra2011Th] for an early discussion).

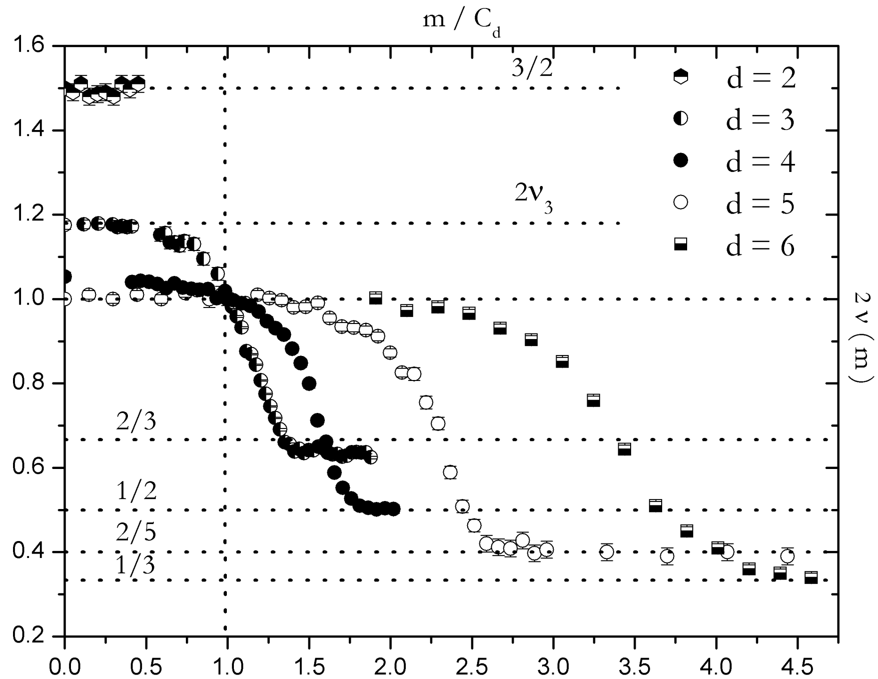


Figure 2.2: Exponents $\nu_d(m)$ vs m/C_d o SRW from simulations at various lattice dimensions (Fig. 1 in [Fra2011]). The graph shows the behavior of $\nu_d(m)$ from fits of $\log\left(\langle S_n^2\rangle_m\right)$ vs $\log(n)$. The range for all fits is $n \in [0.8 \cdot n_M, n_M]$. Maximal length for simulated walks was: $n_M = 2 \cdot 10^3$ for $d = 2$, $n_M = 10^3$ for $d = 3$, $n_M = 10^3$ for $d = 4$, $n_M = 0.5 \cdot 10^3$ for $d = 5$, and $n_M = 0.3 \cdot 10^3$ for $d = 6$. For $d = 3$, the $\log\left(\langle S_n^2\rangle_m\right)$ has been achieved by using $m = C_3 M/\langle M_n\rangle$ to take into account the finite size effects (see Fig. 2 in [Fra2011]). The picture for $d = 4$, $m < m_c$ shows an exponent consistent with the expected logarithmic correction (for chains of length 10^3 an exponent of $2\nu \simeq 1.070$ would be predicted). In the $n \to \infty$ limit, $\nu_d(m)$ is expected to be a step function (see Fig. 3 in [Fra2011]). This graph shows the step at m_c: for $d = 3, 4$ we expect $m_c = C_d$, for $d \geq 5$ simulations suggest $m_c > C_d$.

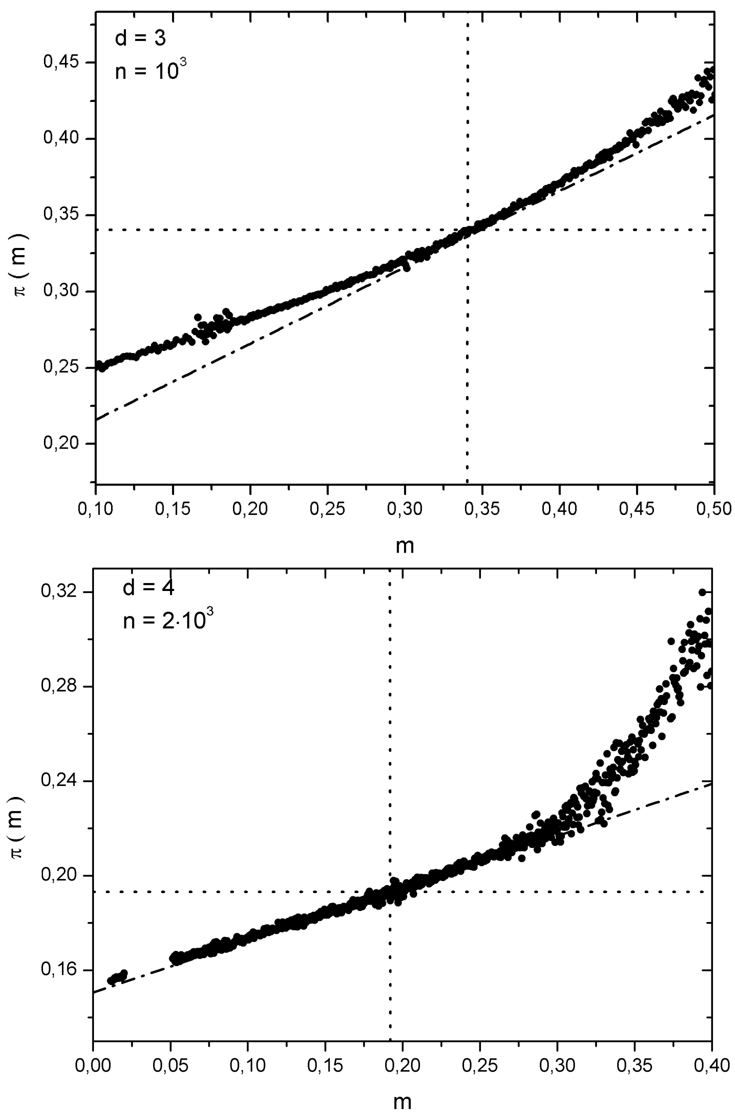


Figure 2.3: The graphs show $\pi_n(m)$ for SRW at $d = 3$, $n = 10^3$ and $d = 4$, $n = 2 \cdot 10^3$ (Fig. 2.9 and 2.10 of [Fra2011Th]). Dotted lines are $m = C_d$, $\pi(m) = C_d$ for $d = 3, 4$ while the dash-dotted lines are $\pi(m) = (1 - B_d)\, C_d + B_d m$, $d = 3, 4$, with $C_3 = 0.3405373$, $C_4 = 0.1932017$, $B_3 = 1/2$ and $B_4 = 0.22081$. See [Fra2011Th, FB2014] for a detailed explanation.

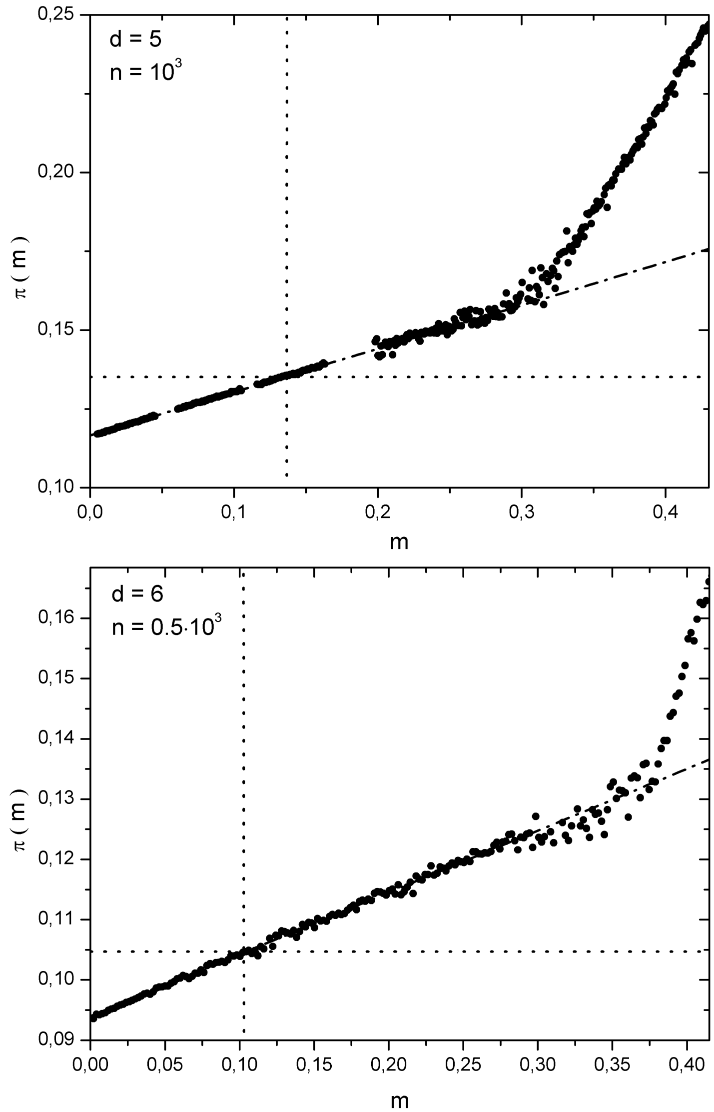


Figure 2.4: The graphs show $\pi_n(m)$ for SRW at $d = 5$, $n = 10^3$ and $d = 6$, $n = 0.5 \cdot 10^3$ (Fig. 2.12 and 2.13 of [Fra2011Th]). Dotted lines are $m = C_d$, $\pi(m) = C_d$ for $d = 3, 4$ while the dash-dotted lines are $\pi(m) = (1 - B_d) C_d + B_d m$, $d = 5, 6$, with $C_5 = 0.1351786$, $C_6 = 0.1047155$, $B_5 = 0.137672$ and $B_6 = 0.10266$. See [Fra2011Th, FB2014] for a detailed explanation.

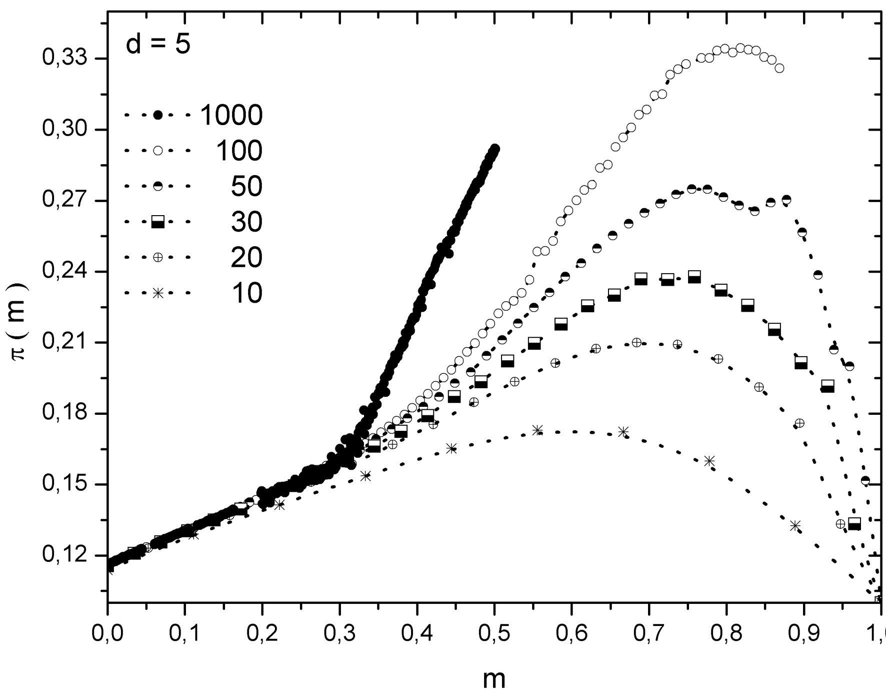


Figure 2.5: Behavior of $\pi_n(m)$ for a SRW at $d = 5$ for some lengths n (Fig. 2.12 of [Fra2011Th]). The graph shows the convergence of $\pi_n(m)$ below the critical point m_c (supposed to be larger than $C_5 = 0.1351786$ for $d = 5$). A slower convergence of $\pi_n(m)$ is also expected also for $m_c > m$ from theoretical arguments (see). Notice the drop for $m \to 1$ due to the finite size effects. See [Fra2011Th, FB2014] for a detailed explanation.

2.2.2 Linear urns and the Bachi-Pal Model

Another interesting topic is the equivalence between linear urns and the Baghi-Pal model, a widely investigated model due to its relevance in studying branching phenomena and random trees [Pem2007, Mam2003]. Consider an urn with black and white balls: at each step a ball is extracted uniformly from the urn, and some new balls are added or discarded according to the square matrix

$$A := \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{2.18}$$

with $a_{ij} \in \mathbb{Z}$, such that if the extraction resulted in a black ball we add a_{11} black balls and a_{12} white balls, otherwise we add a_{21} black balls and a_{22} white balls. If $a_{11} + a_{12} = a_{21} + a_{22} = M$, then the number of balls increases by some deterministic rate M and the urn is said to be *balanced*, if $M > 0$ the urn is said to be also *tenable*.

Our interest raises, inter alia, from the fact that this is the first nontrivial model for which some large deviations results have been obtained. In [FGP2005, FDP2006] the so-called *subtractive* case (negative diagonal entries) is fully analyzed by purely analytic methods, obtaining an explicit characterization of the rate function and other important results: we will discuss this argument again in Section 3.3.

Here we show that the above model is equivalent to a linear urn function $\pi(s) = s_0 + b(s - s_0)$ provided that A fulfills some special conditions. Let

B_k and W_k be the number of black and withe balls of a Bagchi-Pal urn at time k, let $T_k = B_k + W_k$ be the total number of balls and

$$A = \begin{pmatrix} a_{11} & M - a_{11} \\ M - a_{22} & a_{22} \end{pmatrix} \tag{2.19}$$

the reinforcement matrix, where we used the balancing constraint $a_{11}+a_{12} = a_{21} + a_{22} = M$. Since the balancing ensures that $T_k = (B_0 + W_0) + Mk$, we can rescale $k \to k - M^{-1}(B_0 + W_0)$ and consider $k \geq m = M^{-1}(B_0 + W_0)$, $T_k = Mk$. Then, define the variable

$$X_{n,k} = \frac{B_k - (M - a_{22})\,k}{a_{11} + a_{22} - M}, \tag{2.20}$$

with $a_{11} + a_{22} - M \neq 0$: we can show that the process $\{X_{n,k} : m \leq k \leq n\}$ defined by the urn function $\pi(s) = s_0 + b(s - s_0)$, with

$$s_0 = \frac{a_{22} - M}{2M - a_{11} - a_{22}},\ b = \frac{a_{11} + a_{22}}{M} - 1,\ X_{n,m} = \frac{B_0 - (M - a_{22})\,m}{a_{11} + a_{22} - M}. \tag{2.21}$$

is equivalent to a Bagchi-Pal model with reinforcement matrix

$$A = M \begin{pmatrix} b + s_0(1-b) & (1-s_0)(1-b) \\ s_0(1-b) & 1 - s_0(1-b) \end{pmatrix}. \tag{2.22}$$

Since the Bagchi-Pal model usually considers an integer reinforcement matrix, we need M, s_0, b, m such that both $B_0 + W_0$ and the elements of A are integers. If $a_{12} = a_{21} = 0$ we recover the Polya Urn ($b = 1$), while we obviously have to discard the case $a_{11} = a_{21}$ (deterministic evolution of the urn:

$a_{11} + a_{22} - M = 0$). We assume some *tenability* conditions which ensures that the process can't be stopped, ie, that the total number of balls is deterministic and always growing ($M > 0$), that $a_{12} \geq 0$, $a_{21} \geq 0$ and if $a_{11} < 0$ then $(W_0/a_{11}), (a_{21}/a_{11}) \in \mathbb{Z}$, if $a_{22} < 0$ then $(B_0/a_{22}), (a_{12}/a_{22}) \in \mathbb{Z}$. The last two conditions ensure that only balls of the same color of that drawn can be removed from the urn: this prevent from stopping the process by impossible removals.

According to the above discussion, and considering that $B_0/m \in [0, 1]$, it is possible to show that the general urn function describing the balanced Baghi-Pal urns (also tenable if $b < 1$) is the unique solution of

$$\pi(s) = \mathbb{I}_{\{s_0 + b(s - s_0) \geq 1\}} + (s_0 + b(s - s_0))\, \mathbb{I}_{\{0 < s_0 + b(s - s_0) < 1\}}. \tag{2.23}$$

Notice that the above function is not included in $\mathcal{U}$ for general values of s_0 and b, since it can violate the condition $\pi(s) \in (0, 1)$ for $s \in (0, 1)$. As example, the subtractive urn $a_{11} = a_{22} = -1$, $a_{12} = a_{21} = 2$ is described by the urn function $\pi(s) = \mathbb{I}_{\{s \in [0, 1/3)\}} + (2 - 3s)\, \mathbb{I}_{\{s \in [1/3, 2/3]\}}$ that does not belong to $\mathcal{U}$. The conditions $s_0 \in [0, 1]$, $b \in [-s_0/(1 - s_0), 1]$ if $s_0 \in [0, 1/2]$ and $b \in [-(1 - s_0)/s_0, 1]$ if $s_0 \in [1/2, 1]$ ensure that $\pi \in \mathcal{U}$.

From Eq. (2.22) we can see that tenability conditions $a_{12} \geq 0$, $a_{21} \geq 0$ are naturally implemented in our bounds for the coefficients ($s_0 \in [0, 1]$, $b \leq 1$), provided they can produce an integer valued reinforcement matrix and starting values, but only allow to describe tenable urns with positive diagonal entries. In Section 3 we will provide an explicit expression for the Cumulant Generating Function of the case above. We anticipate that in this work we will not deal with the subtractive case considered in [FGP2005], since this would need an extension of our results outside the $\mathcal{U}$ class (even if we believe that such extension would require only minor modifications: this fact will be addressed in section 3.1 shortly after stating the Sample-Path Large Deviation Principle).

2.2.3 Market Share

An influent model of Market Share that con be modeled using urn function has been introduced in the '80s by the economist W. Brian Arthur, and subsequently developed in the urn context by other authors (see Ermoliev, Dosi).

Consider two competing companies that launch onto the market a similar product roughly at same time (as practical example we could think about Samsung, Apple and their smartphones). Suppose that the two products are roughly equivalents, such that there is no practical reason for choosing one over the other. Then, we can imagine that a buyer will base his decision in part on personal opinions (aesthetic tastes, ideologies, advertising, etc.) and in part on those of people that already purchased one of these products.

Then, suppose a simplified situation in which the new buyers are only imperfectly informed about the products, so that they will make their choices by asking to an odd number $z > 1$ of adopters who are already using the products. An alternative hypothesis to the same effect is that there are positive (or negative) externalities in adoption which change the returns to the user along the diffusion process, but adopters in order to estimate them can only sample a fixed number of users. In both cases, we assume that

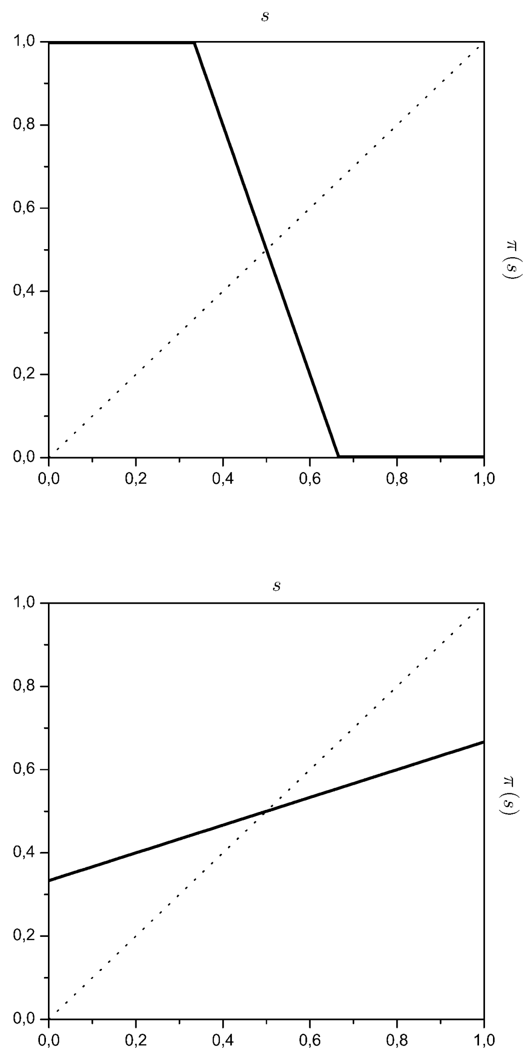


Figure 2.6: Urn functions from Eq.s (2.22) and (2.23) of Bagchi-Pal models for $a_{11} = a_{22} = -1$, $a_{21} = a_{12} = 2$ (upper figure) and $a_{11} = a_{22} = 2$, $a_{21} = a_{12} = 1$ (lower figure). The first one is a subtractive urn of the kind considered in [FGP2005], while the second is an additive and tenable urn. In this work we will give a full solution for the second case only.

any new adopter will choose with probability p the technology used by the majority of the sample z and with probability $\bar{p} = 1 - p$ the technology of the minority of them. The above situation has been considered by Dosi et al. in [DEK1994]. Modeling the owners as black and white balls, they found that the probability of increasing by a one the owners of the first product is

$$\pi(s) = (1-p) + (2p-1)\,\sigma(z,s) + o(1), \tag{2.24}$$

where $o(1)$ goes to zero as $O\left(n^{-1}\right)$ uniformly on $s \in [0,1]$ and $\sigma(z,s)$ is the probability of getting a majority of black balls by extracting z balls from a urn in which the fraction of black balls is s. If $p = 0$ the above models describes a situation in which the customer always buys the product owned by the majority of the sample: this early model has been introduced by Arthur in [AEK1983].

Notice that all the urn functions of the kind in Eq. (2.24) belongs to $\mathcal{U}$ if $p < 1$ and that the case $p \in (0,1)$, $z = 1$ is described by the linear urn $\pi(s) = (1-p) + (2p-1)\,s$. In this work we will also use the case $p = 0$, $z = 3$, described by the urn function $\pi(s) = 3s^2 - 2s^3$, as applicative example for our results on optimal trajectories (we will call it *majority urn*).

2.2.4 Neuron growth

We conclude this section with a model for axons growth introduced by Khanin and Khanin in [KK2001]. From experimental observation it is known that groups of neurites exhibit periods of growth and retraction until one rapidly elongates to become an axon. Since experimental data suggests that any neurite has the potential to be either an axon or a dendrite, they argued that a neuron's length at various stages of growth relative to nearby neurons may influence its development, and hypothesized the following mechanism to simulate axon growth.

Consider a group of q neurites, and denote by $\ell_i(t)$, $1 \leq i \leq q$ the lengths of the neurites at some time t (we set $\ell_i(0) = 1$). They propose an urn model where at each discrete time one of the existing neurites grows by a constant length $l = 1$ with probability $\ell_i(t)^\alpha / \sum_i \ell_i(t)^\alpha$ for some parameter $\alpha > 0$ while the other remains unchanged. If we consider $q = 2$, then $\ell_i(t)$ is an urn process with urn function

$$\pi(s) = \frac{s^\alpha}{s^\alpha + (1-s)^\alpha}. \tag{2.25}$$

The motivating biological question concerns the mechanisms by which apparently identical cells develop into different types. This mechanisms are poorly understood in many important developmental processes. Finally, we remark that the urn function in Eq. (2.25) has also been considered in [Oliv2008], [DFM2002] and [CL2009] in different contexts.

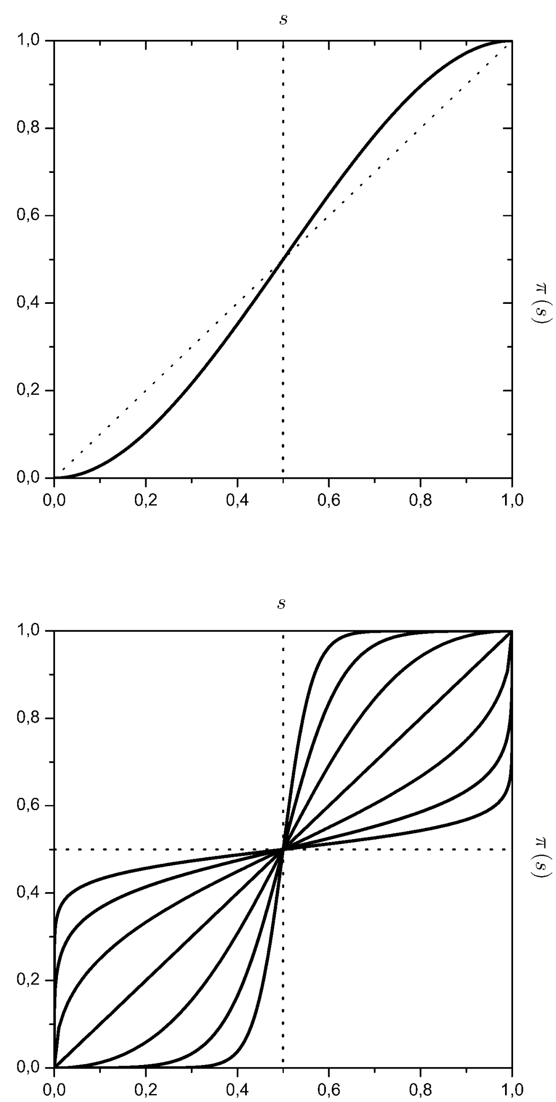


Figure 2.7: Eq. (2.24) for $p = 0$, $z = 3$ (*majority urn*, upper figure) and Eq. (2.25) for some values of α ($\alpha \in \{1/8,\ 1/4,\ 1/2,\ 1,\ 2,\ 3\,4\}$, lower figure).

3 Main results

As stated in the introduction, the aim of this work is to provide a Large Deviation analysis for the Hill, Lane and Sudderth urn model, with urn functions $\pi \in \mathcal{U}$. Apart from the Polya-Eggenberger urn, for which we can explicitly compute the exact urn composition at each time, to the best of our knowledge, large deviations results in urn models have been pioneered by Fajolet et Al. [FGP2005, FDP2006, HKP2007], which provided a detailed analysis of the Bagchi-Pal urn using generating function methods. Since then other authors extended this approach to some related models (of particular interest is [MM2012], a Bagchi-Pal urn with stochastic reinforcement matrix).

In this section we will expose and comment our results on large deviations, also discussing some applications to the examples in Section 2.2. Almost all the proofs of our statements are grouped in Section 4: we will specify where to find them.

3.1 Sample-Path Large Deviation Principle

Our first result is a Sample Path Large Deviation principle (SPLDP) which holds for any $\pi \in \mathcal{U}$. This is a fundamental result in our work, since all other statements will be based on this in one way or another. Then, define the function $\chi_n : [0,1] \to [0,1]$ as follows:

$$\chi_n := \left\{\chi_{n,\tau} = n^{-1}\left[X_{n,\lfloor n\tau\rfloor} + (n\tau - \lfloor n\tau\rfloor)\,\delta X_{n,\lfloor n\tau\rfloor}\right] : \tau \in [0,1]\right\}, \qquad (3.1)$$

where $\lfloor\cdot\rfloor$ denotes the lower integer part, and introduce the subspace of Lipschitz-continuous functions

$$\mathcal{Q} := \{\varphi \in C([0,1]) : \varphi_0 = 0, \, \varphi_{\tau+\delta} - \varphi_\tau \in [0,\delta], \, \delta > 0, \, \tau \in [0,1]\}, \quad (3.2)$$

where $C([0,1])$ is the set of continuous functions on $[0,1]$. Denote by $\|\varphi\| := \sup_{\tau\in[0,1]} |\varphi_\tau|$ the usual supremum norm, and consider the normed metric space $(\mathcal{Q}, \|\cdot\|)$. We show that a good rate function $I_\pi : \mathcal{Q} \to [0,\infty)$ exists such that for every Borel subset $\mathcal{B} \subseteq \mathcal{Q}$:

$$\liminf_{n\to\infty} n^{-1} \log \mathbb{P}(\chi_n \in \operatorname{int}(\mathcal{B})) \geq - \inf_{\varphi\in\operatorname{int}(\mathcal{B})} I_\pi[\varphi], \quad (3.3)$$

$$\limsup_{n\to\infty} n^{-1} \log \mathbb{P}(\chi_n \in \operatorname{cl}(\mathcal{B})) \leq - \inf_{\varphi\in\operatorname{cl}(\mathcal{B})} I_\pi[\varphi]. \quad (3.4)$$

To describe the rate function we introduce a functional $S_\pi : \mathcal{Q} \to (-\infty, 0]$, defined as follows:

$$S_\pi[\varphi] := \int_{\tau\in[0,1]} \left[d\varphi_\tau \, \log \pi(\varphi_\tau/\tau) + d\tilde{\varphi}_\tau \, \log \bar{\pi}(\varphi_\tau/\tau) \right], \quad (3.5)$$

where we denoted $\bar{\pi}(s) = 1 - \pi(s)$ and $\tilde{\varphi}_\tau = \tau - \varphi_\tau$. Then, the following theorem gives the Sample-Path LDP for χ_n:

Theorem 1. *Let $\pi \in \mathcal{U}$, $\varphi \in \mathcal{Q}$, define the function $H(s) := s \log s + \bar{s} \log \bar{s}$, and the functional $J : \mathcal{Q} \to [-\log 2, \infty)$ as follows:*

$$J[\varphi] = \begin{cases} \int_0^1 d\tau \, H(\dot{\varphi}_\tau) & if \ \varphi \in \mathcal{AC} \\ \infty & otherwise, \end{cases} \tag{3.6}$$

where $\mathcal{AC}$ is the class of absolutely continuous functions (we assume the same definition given in Theorem 5.1.2 of [DZ1998]). Then, the law of χ_n satisfies a Sample-Path LDP, with good rate function

$$I_\pi[\varphi] = J[\varphi] - S_\pi[\varphi]. \tag{3.7}$$

Moreover, the LDPs are independent from the initial conditions unless the urn starts in a monochromatic configuration (ie, the balls are all black or all withe).

The proof lies on a change of measure and a straight application of the Varadhan Integral Lemma. We prove this theorem in Section 4.1.

Tanks to the properties of the $\mathcal{U}$ set our proof does need almost nothing more than standard arguments of Large Deviations theory: as we shall see, the requirements of the Varadhan Integral Lemma are met straightforward by the $\mathcal{U}$ set, as well as our approximation arguments. Anyway, we conjecture that both Theorem 1 and its related results can be generalized to urn functions $\pi : [0,1] \to [0,1]$ such that $|\pi(s+\delta) - \pi(s)| \leq f(\delta)$ for some some f with $\lim_{\epsilon \to 0} \epsilon \int_\epsilon^1 dz \, f(z)/z^2 = 0$. This can be done by performing some "surgery" on the set $\mathcal{Q}$ to a priori exclude the trajectories for which technical issues arises in proving the continuity of $S_\pi[\varphi]$ on $(\mathcal{Q}, \|\cdot\|)$ and the approximation argument of Lemma 12 (as examples, those trajectories for which $S_\pi[\varphi]$ diverges). We don't extend Theorem 1 in the present work, but we anticipate that this enhancement will be present in [Fra2014].

Before going ahead, we should spend some words on non homogeneous urn functions. Then, let some $\pi \in \mathcal{U}$ with $\pi(0) > 0$, $\pi(1) < 1$ and consider an urn function π_n such that $\pi_n \in \mathcal{U}$ for every $n \geq 0$ and $\pi_n \to \pi$ uniformly for every $s \in [0,1]$. In section 4.1.1 we show that

Corollary 2. *Let $\pi \in \mathcal{U}$ with $\pi(0) > 0$, $\pi(1) < 1$ and let π_n such that such that $|\pi_n(s) - \pi(s)| \leq \delta_n$, $\lim_n \delta_n = 0$ for all $s \in [0,1]$. Then, the urn process defined by π_n has the same Sample-Path LDP of that defined by π.*

We restricted our statement to urns with $\pi(0) > 0$, $\pi(1) < 1$ to avoid some minor technical issues which would arise if we consider the whole set $\mathcal{U}$, but we are convinced that it is possible to generalize this result on the basis of the same considerations made for Theorem 1. Also this extension will be developed in [Fra2014].

3.2 Entropy of the event $X_{n,n} = \lfloor sn \rfloor$

Our main interest in Theorem 2 comes from the fact that SPLDPs allow to approach some important Large Deviation questions about the urn evolution from the point of view of functional analysis. In this work our attention will mainly focus on the entropy of the event $X_{n,n} = \lfloor sn \rfloor$, $s \in [0,1]$. First we show that the limit

$$\phi(s) := \lim_{n \to \infty} n^{-1} \log \mathbb{P}(X_{n,n} = \lfloor sn \rfloor), \tag{3.8}$$

exists for every $\pi \in \mathcal{U}$, and has the following variational representation:

Theorem 3. *The limit $\phi(s)$ defined in Eq. (3.8) exists for any $\pi \in \mathcal{U}$ and is given by the variational problem*

$$\phi(s) = -\inf_{\varphi \in \mathcal{Q}_s} I_\pi[\varphi], \tag{3.9}$$

where $\mathcal{Q}_s := \{\varphi \in \mathcal{Q} : \varphi_1 = s\}$ and I_π is the rate function of Theorem 1.

Notice that Theorem 1 can not be directly applied to the Eq. 3.8 in order to obtain Theorem 3, since this is a much stronger statement than those usually considered in Large Deviations theory (where we more frequently would meet events of the kind $\{X_{n,n} \leq sn\}$ or $\{X_{n,n} \geq sn\}$). To prove Theorem 3 we combined Theorem 1 with a combinatorial argument: the proof can be found in Section 4.2.1.

3.2.1 Optimal trajectories

Since the variational problem in Theorem 3 heavily depends on the choice of π, a general characterization of $\phi(s)$ would be a quite hard nut to crack.

Anyway, we still can prove some interesting facts on the shape of $\phi(s)$. For example, we can prove that $\phi(s) = 0$ for $s \in [\inf C_\pi, \sup C_\pi]$ and $\phi(s) < 0$ otherwise.

Corollary 4. *For any $\pi \in \mathcal{U}$, $\phi(s) = 0$ for $s \in [\inf C_\pi, \sup C_\pi]$ and $\phi(s) < 0$ otherwise, where C_π is the contact set of π defined by Eq. (2.2).*

An obvious consequence of this is that urn functions with ν_0, $\nu_1 > 0$ have $\phi(s) = 0$ for all $s \in [0, 1]$, since $\inf C_\pi = 0$ and $\sup C_\pi = 1$. For example, this is the case for the *majority urn* $\pi(s) = 3s^2 - 2s^3$ of Section 2.2.

The above corollary is obtained by proving that for any $s \in [\inf C_\pi, \sup C_\pi]$ we can find a trajectory $\varphi^* \in \mathcal{Q}_s$ such that $I_\pi[\varphi^*] = 0$, while this is not possible if $s \in K_{\pi,0} \cup K_{\pi,|\partial C_\pi|}$. Also, are able to give an explicit characterization for the optimal trajectories φ^*. We enunciate this result in two separate corollaries: the first deals with trajectories that end in $s \in K_{\pi,i}$, $1 \leq i \leq |\partial C_\pi| - 1$, while the second deals with trajectories that end in $s \in \partial C_\pi$ (as we shall see, Corollary 4 is an almost direct consequence of the following two) .

Corollary 5. *Let K_π, A_π as in Eq.s (2.4), (2.5), then for any $s \in K_{\pi,i}$ a zero-cost trajectory $\varphi^* \in \mathcal{Q}_s$ with $\tau^{-1}\varphi^*_\tau \in K_{\pi,i} \cup \partial K_{\pi,i}$, $\tau \in [0, 1]$ exists such that $I_\pi[\varphi^*] = 0$, and it can be constructed as follows. If $a_{\pi,i} = 0$ then we can* *take $\varphi^* = s\tau$ as in the Polya-Eggenberger urn. If $a_{\pi,i} \neq 0$ let*

$$F_\pi(s,u) := \int_u^s \frac{dz}{\pi(z) - z}. \tag{3.10}$$

Also, for $s \in K_{\pi,i}$ define

$$s_i^* := \mathbb{I}_{\{a_{\pi,i}=1\}} \inf K_{\pi,i} + \mathbb{I}_{\{a_{\pi,i}=-1\}} \sup K_{\pi,i}, \tag{3.11}$$

$$\tau_{s,i}^* := \exp\left(-\lim{}_{a_{\pi,i}\left(u-s_i^*\right)\to 0^+} \left|F_\pi(s,u)\right|\right). \tag{3.12}$$

and denote by $F_{\pi,s}^{-1}$ the inverse function of $F_\pi(s,u)$ for $u \in K_{\pi,i} \cup \partial K_{\pi,i}$:

$$F_{\pi,s}^{-1} := \left\{F_{\pi,s}^{-1}(q),\, q \in \left[0, \log\left(1/\tau_{s,i}^*\right)\right) : F_\pi\left(s, F_{\pi,s}^{-1}(q)\right) = q\right\}. \tag{3.13}$$

Then, if $a_{\pi,i} \neq 0$ the zero-cost trajectory is given by $\varphi_\tau^* = \tau u_\tau^*$, with

$$u_\tau^* := F_{\pi,s}^{-1}(\log(1/\tau))\, \mathbb{I}_{\left\{\tau\in\left(\tau_{s,i}^*,1\right]\right\}} + s_i^*\, \mathbb{I}_{\left\{\tau\in\left[0,\tau_{s,i}^*\right]\right\}}. \tag{3.14}$$

The proof relies on the fact that any φ^* for which $I_\pi[\varphi^*] = 0$ must satisfy the homogeneous equation $\dot{\varphi}_\tau^* = \pi(\varphi_\tau^*/\tau)$. This is shown in Section 4.2.2.

The above corollary states that the optimal strategy to achieve the event $\{X_{n,n} = \lfloor sn \rfloor\}$, $s \in [\inf C_\pi, \sup C_\pi]$ emanates from the closest unstable equilibrium point, which is on the left if $\pi(s) < s$ and on the right if $\pi(s) > s$.

Notice that u_τ^* is always invertible on $\left(\tau_{s,i}^*, 1\right]$, since it is strictly decreasing from $\sup K_{\pi,i}$ to s if $a_{\pi,i} = -1$, and strictly increasing from $\inf K_{\pi,i}$ to

s if $a_{\pi,i} = 1$.

We can provide an explicit example by the *majority urn* $\pi(s) = 3s^2 - 2s^3$. Since $3s^2 - 2s^3 = s$ has three solutions at 0, $1/2$ and 1 we have $K_{\pi,1} = (1, 1/2)$ and $K_{\pi,2} = (1/2, 1)$, with $a_{\pi,1} = -1$ and $a_{\pi,2} = 1$. Then, applying the above corollary we find that in both cases $s \in K_{\pi,1}$ and $s \in K_{\pi,2}$ we have $\tau^*_{s,1} = 0$, $\tau^*_{s,2} = 0$, and the optimal trajectory is

$$2\,\tau^{-1}\varphi^*_\tau = 1 - \left[1 + \left(1 - 2\,\mathbb{I}_{\left\{s \le \frac{1}{2}\right\}}\right)\frac{\varrho(s)}{\tau}\right]^{-1/2},\ \ \varrho(s) = \tfrac{4s(1-s)}{(2s-1)^2}. \qquad (3.15)$$

A graphical picture of this is in Figure 3.1, where the above trajectories are shown for some values of s.

A curious fact is that the optimal trajectory can be time-inhomogeneous, depending on integrability of $1/(\pi(s) - s)$ as $s \to s^*_i$. If the singularity is integrable, then the equilibrium s^*_i is so unstable that the processes will leave its neighborhood at some $\tau^*_{s,i} > 0$ to end in s. We discuss this interpretation after stating our results for trajectories that end in $s \in \partial C_\pi$.

Corollary 6. *Let K_π, A_π as in Eq.s (2.4), (2.5), and consider $K_{\pi,i}$ for some $1 \le i \le |\partial C_\pi| - 1$. Let $F_\pi(s,u)$ and s^*_i as in Corollary 5 and define*

$$s^\dagger_i := \mathbb{I}_{\{a_{\pi,i}=-1\}} \inf K_{\pi,i} + \mathbb{I}_{\{a_{\pi,i}=1\}} \sup K_{\pi,i}. \qquad (3.16)$$

If $a_{\pi,i} = 0$ the trajectory $\varphi^* = s^\dagger_i \tau$ is the unique zero-cost trajectory in $K_{\pi,i} \cup \partial K_{\pi,i}$. If $a_{\pi,i} \neq 0$ then a family of zero-cost trajectories $\varphi^* \in \mathcal{Q}_{s^\dagger_i}$ with $\varphi^* \in \mathcal{Q}_s$ with $\tau^{-1}\varphi^*_\tau \in K_{\pi,i} \cup \partial K_{\pi,i}$, $\tau \in [0,1]$ exist such that $I_\pi[\varphi^*] = 0$.

If $\lim_{a_{\pi,i}\left(s_i^\dagger - s\right)\to 0^+} |F_\pi(s,\cdot)| = \infty$ *then* $\varphi^* = s_i^\dagger \tau$ *is the unique zero-cost trajectory. If* $\lim_{a_{\pi,i}\left(s_i^\dagger - s\right)\to 0^+} |F_\pi(s,\cdot)| < \infty$ *we define*

$$\theta_i^* := \exp\left(-\lim_{a_{\pi,i}(u-s_i^*)\to 0^+} \lim_{a_{\pi,i}\left(s_i^\dagger - s\right)\to 0^+} |F_\pi(s,u)|\right) \tag{3.17}$$

and the function $F^{-1}_{\pi,s_i^\dagger}$ *as in Corollary 5, with* $s_i^\dagger$, θ_i^* *on behalf of* s, $\tau^*_{s,i}$. *Then* $\varphi_\tau^* = \tau u_\tau^*$ *with*

$$u_\tau^* := s_i^\dagger \mathbb{I}_{\{\tau\in(t,1]\}} + F^{-1}_{\pi,s_i^\dagger}\left(\log(t/\tau)\right)\mathbb{I}_{\left\{\tau\in\left(\theta_i^* t,t\right]\right\}} + s_i^*\,\mathbb{I}_{\left\{\tau\in\left[0,\theta_i^* t\right]\right\}}, \tag{3.18}$$

is a zero-cost trajectory for any $t \in [0,1]$.

As we can can see, the set of zero-cost trajectories that end in a stable equilibrium point is degenerate. Again, this depends only on the integrability of the singularity of $1/(\pi(s)-s)$ for $s \to s_i^\dagger$: if $\lim_{a_{\pi,i}\left(s_i^\dagger - s\right)\to 0^+} |F_\pi(s,\cdot)| = \infty$ the trajectory is simply the martingale $\varphi^* = s_i^\dagger \tau$ and it is unique. If instead $\lim_{a_{\pi,i}\left(s_i^\dagger - s\right)\to 0^+} |F_\pi(s,\cdot)| < \infty$ then we have a family of time-inhomogeneous trajectories, parametrized by the time t at which they hit $s_i^\dagger$, that emanates from the unstable equilibrium s_i^* on the other side of $K_{\pi,i}$. Moreover, if $s_i^\dagger$ is a downcrossing then $s_i^\dagger = \inf K_{\pi,i} = \sup K_{\pi,i-1}$ with $a_{\pi,i} = -1$, $a_{\pi,i-1} = 1$, so that optimal trajectories ending in $s_i^\dagger$ can emanate also from $\inf K_{\pi,i-1}$. Notice that if $1/(\pi(s)-s)$ is integrable also for $s \to s_i^*$ then the $\theta_i^* > 0$ and our optimal trajectories would be doubly time-inhomogeneous, emanating from s_i^* at some $\tau = \theta_i^* t$ and hitting $s_i^\dagger$ at

$\tau = t$. In Figure we give an example of such situation by the urn function

$$\pi(s) = x + \left(\tfrac{1}{4} - x\right)^{\frac{1}{2}} \mathbb{I}_{\left\{x \in \left[0, \frac{1}{4}\right]\right\}} - \left(\tfrac{1}{4} - x\right)^{\frac{1}{2}} \left(x - \tfrac{1}{2}\right)^{\frac{1}{2}} \mathbb{I}_{\left\{x \in \left(\frac{1}{4}, \frac{1}{2}\right]\right\}} + \\ + \left(\tfrac{1}{2} - x\right)^{\frac{1}{2}} \left(x - \tfrac{3}{4}\right)^{\frac{1}{2}} \mathbb{I}_{\left\{x \in \left(\frac{1}{2}, \frac{3}{4}\right]\right\}} - \left(x - \tfrac{3}{4}\right)^{\frac{1}{2}} \mathbb{I}_{\left\{x \in \left(\frac{3}{4}, 1\right]\right\}} \tag{3.19}$$

in the interval $s \in \left[\frac{1}{4}, \frac{1}{2}\right]$, where we have $F_\pi(s,u) = \left[2 \arcsin\left(\sqrt{4z-1}\right)\right]_u^s$ and $\theta_i^* = e^{-\pi}$. By applying Corollary 6 we find that the family of optimal trajectories ending in $s_i^\dagger = 1/2$ is

$$\tau^{-1}\varphi_\tau^* = \tfrac{1}{4}\, \mathbb{I}_{\{x \in [1,t]\}} + \tfrac{1}{2}\, \mathbb{I}_{\{x \in [0, e^{-\pi} t]\}} + \\ + \tfrac{1}{4}\left[1 + \sin^2\left(\tfrac{1}{2}\log\left(t/\tau\right)\right)\right] \mathbb{I}_{\{x \in [e^{-\pi} t, t]\}},\ t \in [0,1], \tag{3.20}$$

while for each $s \in \left(\frac{1}{4}, \frac{1}{2}\right]$ we have $\tau_{s,i}^* = e^{2\arcsin\left(\left(\sqrt{4s-1}\right)\right) - \pi}$ and

$$\tau^{-1}\varphi_\tau^* = \tfrac{1}{4}\left[1 + \sin^2\left(\tfrac{1}{2}\log\left(\tau_{s,i}^*/\tau\right)\right)\right] \mathbb{I}_{\left\{\tau \in \left(\tau_{s,i}^*, 1\right]\right\}} + \tfrac{1}{2}\, \mathbb{I}_{\left\{\tau \in \left[0, \tau_{s,i}^*\right]\right\}}, \tag{3.21}$$

with $\lim_{s \to s_i^\dagger} \tau_{s,i}^* = \theta_i^*$ as expected. The above example has been constructed to explicitly show the effect of integrability on stable and unstable points: as said before, integrability in the neighborhood of an unstable point (like an integrable upcrossing) make it so unstable that the probability mass is in some sense expelled form its neighborhood even on a time scale $O(n)$, and makes it convenient to use a time-inhomogeneous strategy. The inverse picture arises for integrable stable points (like integrable downcrossings), where the process is so attracted that it becomes entropically convenient to hit the equilibrium point in a finite fraction $t \in [0,1)$ of the whole time span (of order $O(n)$), instead of approaching it asymptotically.

It is a quite interesting result that if $a_{\pi,i} \neq 0$ then no trajectory with $\lim_{\tau\to 0}(\varphi_\tau/\tau) \notin \partial C_\pi$ can be optimal, not even if we chose φ_1 to be in a set of stable equilibrium like downcrossings (ie, $\varphi_1 \in C_\pi(+,-)$). We can interpret this result in therms of time spent in a given state: it seems that a process starting with initial conditions $m^{-1}X_{n,m} \notin \partial C_\pi$ concentrates in the neighborhood of ∂C_π in times that are of order $o(n)$, and only those that are in the neighborhood of unstable points can remain there for times $O(n)$, eventually reaching the stable points according to the mechanism suggested by Corollaries 5 and 6.

We believe that the results in this subsection are of particular relevance, since they may eventually open a way to deal with the much harder problem of moderate deviations, ie, to compute limits of the kind

$$\phi_\sigma(s) = \lim_{n\to\infty} \sigma_n^{-1} \log \mathbb{P}(X_{n,n} = \lfloor sn \rfloor) \tag{3.22}$$

for some $\sigma_n = o(n)$, $s \in [\inf C_\pi, \sup C_\pi]$. This can be done by putting $\chi_n = \varphi^* + \delta_{n,\tau}$, where $\delta_{n,\tau}$ is a small perturbation of order $o(n)$, in the expression of $\log \mathbb{P}(X_n = \lfloor sn \rfloor)$ given in Lemma 12. Then, one could proceed with a perturbative analysis to eventually find a variational problem for moderate deviations: this argument will be developed in [Fra2014].

As final remark, we point out that even if our statements fully characterizes the trajectories contained in $K_{\pi,i} \cup \partial K_{\pi,i}$, it is possible to build other trajectories by assembling that of adjacent $K_{\pi,i}$ intervals. As example, this may be the case of some $s_i \in \partial C_\pi$ such that $\pi(s-s_i) - (s-s_i) \sim |s-s_i|^{1/2}$

(s_i is a cuspid touchpoint of the kind $C_\pi(+,+)$). Since $s_i = \inf K_{\pi,i+1}$, $s_i = \sup K_{\pi,i-1}$ and the $|s-s_i|^{-1/2}$, then we can combine a trajectory winch starts in $\inf K_{\pi,i-1}$ and reach s_i at some $t_1 \leq \tau^*_{s,i}$ with that starting from s_i at $\tau^*_{s,i}$ and hitting $s \in K_{\pi,i}$ at $\tau = 1$, to eventually obtain a trajectory that starts $\inf K_{\pi,i-1}$, crosses s_i and ends in some $s \in K_{\pi,i}$.

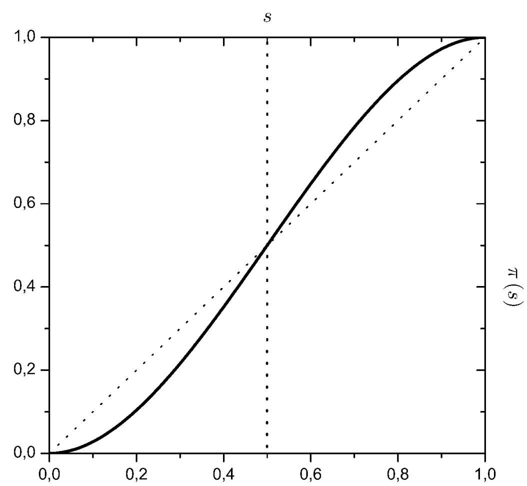
3.3 Cumulant Generating Function

Except that $\phi(s) < 0$, for $s \in [0, \inf C_\pi)$ or $s \in (\sup C_\pi, 1]$ we couldn't extract more informations on the shape of $\phi(s)$ from its variational representation because in such cases the variational problem can't be simplified. We guess that such problem may be approached using the Hamilton-Jacobi-Bellman equation and other Optimal Control tools, but we still haven't obtained any concrete result in this direction.

Anyway, the existence of ϕ proved by probabilistic methods introduces some critical simplifications in treating our problem by slightly easier analytic techniques, provided that π obeys to some additional regularity conditions.

As example we are able to prove the convexity of $-\phi(s)$, $s \in [0, \inf C_\pi)$, or $s \in (\sup C_\pi, 1]$ in case π is invertible on the same intervals and the inverse functions $\pi_-^{-1} : [\pi(0), \pi(\inf C_\pi)) \to [0, \inf C_\pi)$, $\pi_+^{-1} : (\pi(\sup C_\pi), \pi(1)] \to (\sup C_\pi, 1]$ are absolutely continuous Lipschitz functions. Such result is obtained by analyzing the scaling of the Cumulant Generating Function (CGF)

$$\psi(\lambda) := \lim_{n\to\infty} n^{-1} \log \mathbb{E}\left(e^{\lambda X_{n,n}}\right), \ \lambda \in (-\infty, \infty). \tag{3.23}$$



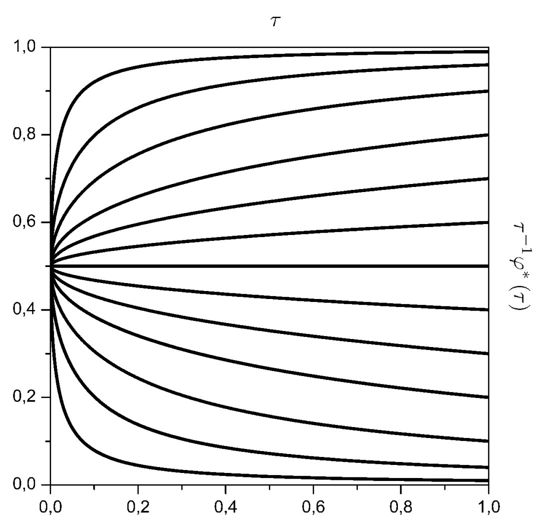
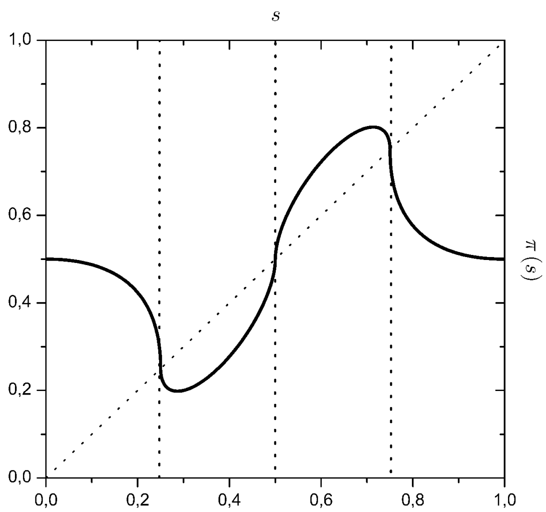


Figure 3.1: *Majority urn* $\pi(s) = 3s^2 - 2s^3$ (upper figure) and its zero-cost trajectories from Eq. (3.15) for some values of s. Here we used $s \in \{0.99, 0.96, 0.9, 0.8, 0.7, 0.5, 0.4, 0.3, 0.2, 0.1, 0.04, 0.01\}$ (lower figure).



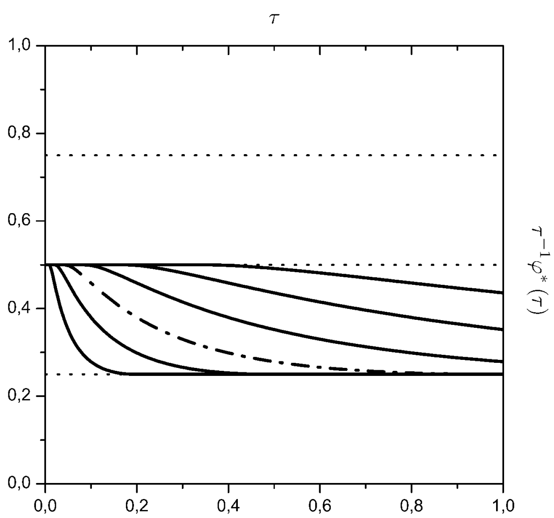


Figure 3.2: Urn function in Eq. (3.19) (upper figure) and some zero-cost trajectories in $K_{\pi,3} \cup \partial K_{\pi,3} = [1/4,\, 1/2]$ from Eq. (3.20), with $s = \frac{1}{4}\left[1 + \sin^2\left(\frac{1}{2}\log(k)\right)\right]$, $k \in \{2,\, 4,\, 8\}$, and Eq. (3.21) with $t = \{1/8,\, 1/2,\, 1\}$. The dash-dotted line is the critical trajectory from Eq. (3.21) with $t = 1$.

First, notice that Theorem 3 implies that the above ψ is well defined [DZ1998]. Then, let $-\hat{\phi}(s) = \operatorname{conv}(-\phi(\mathrm{s}))$ be the convex envelope of $-\phi(s)$ for $s \in [0,1]$. By Theorem 3 and Corollary 4 it follows that $\hat{\phi}(s) = 0$ when $s \in [\inf C_\pi, \sup C_\pi]$ and $\hat{\phi}(s) < 0$ otherwise. Then, define $\hat{\phi}_- : [0, \inf C_\pi) \to (-\infty, 0]$, $\hat{\phi}_+ : (\sup C_\pi, 1] \to (-\infty, 0]$ such that $\hat{\phi}(s) = \hat{\phi}_-(s)$ when $s \in [0, \inf C_\pi)$ and $\hat{\phi}(s) = \hat{\phi}_+(s)$ when $s \in (\sup C_\pi, 1]$, and also define $\psi_- : (-\infty, 0] \to (-\infty, 0]$, $\psi_+ : [0, \infty) \to [0, \infty)$ such that $\psi(\lambda) = \psi_-(\lambda)$ when $\lambda \in (-\infty, 0]$ and $\psi(\lambda) = \psi_+(\lambda)$ when $\lambda \in [0, \infty)$. One can show that $-\hat{\phi}_-$ and $-\hat{\phi}_+$ are the Frenchel-Legendre transforms of $-\psi_-$ and $-\psi_+$ respectively:

$$\hat{\phi}_-(s) = \inf_{\lambda \in (-\infty, 0]} \{\lambda s + \psi_-(\lambda)\}, \; \hat{\phi}_+(s) = \inf_{\lambda \in [0, \infty)} \{\lambda s + \psi_+(\lambda)\}. \quad (3.24)$$

Since the existence of $-\hat{\phi}$ implies the existence of ψ for every $\pi \in \mathcal{U}$, while its convexity ensures that $\psi \in \mathcal{AC}$, we have enough informations to approach ψ by analytic methods.

We remark that a purely analytic treatment via bivariate generating functions has been successfully applied in the special case of the Bagchi-Pal urn by Fajolet et al. [FGP2005], which is strictly related to the linear urn function problem (see Section 2.2.2). In addition to the CGF scaling of the Bagchi-Pal urn, their remarkable paper gives an implicit expression for urn composition at each step. Even if we won't give the composition at each step, the existence of ψ proved by probabilistic methods introduces a critical simplification in treating our problem, since we don't need to face the complex analytic structure which arises from considering a non-linear π at finite time. We will discuss again the Bagchi-Pal model and the case of linear urns shortly after stating the following Cauchy problem for ψ. Concerning the CGF, we will show that it satisfies a non-linear implicit ODE,

$$\pi\left(\partial_\lambda \psi\left(\lambda\right)\right) = \tfrac{e^{\psi(\lambda)}-1}{e^\lambda - 1}, \tag{3.25}$$

for any λ. We stress that the CGF satisfies the above equation for all $\pi \in \mathcal{U}$, but any information would be hard to be extracted if π is not invertible at least on $[0, \inf C_\pi)$ and $(\sup\, C_\pi, 1]$. If this is the case, then the following theorem provides the Cauchy problems for ψ_- and ψ_+:

Theorem 7. *Let $\pi \in \mathcal{U}$ be an invertible function on $[0, \inf C_\pi)$, and denote by $\pi_-^{-1} : [\pi\left(0\right), \pi(\inf C_\pi)) \to [0, \inf C_\pi)$ its inverse. Also, let $\psi_-^* := \psi_-\left(\lambda_-^*\right)$ for some $\lambda_-^* \in (-\infty, 0)$. If π_-^{-1} is $\mathcal{AC}$ and Lipschitz, then for $\lambda \in (-\infty, 0)$ we have $\psi\left(\lambda\right) = \psi_-\left(\lambda\right)$, with $\psi_-\left(\lambda\right)$ solution to the Cauchy problem*

$$\partial_\lambda \psi_-\left(\lambda\right) = \pi_-^{-1}\left(\tfrac{e^{\psi_-(\lambda)}-1}{e^\lambda - 1}\right),\ \psi_-\left(\lambda_-^*\right) = \psi_-^*, \tag{3.26}$$

Let π be invertible on $(\sup\, C_\pi, 1]$, with $\pi_+^{-1} : (\pi(\sup\, C_\pi), \pi\left(1\right)] \to (\sup\, C_\pi, 1]$ its inverse function. If π_+^{-1} is $\mathcal{AC}$ and Lipschitz, then for $\lambda \in (0, \infty)$ we have $\psi\left(\lambda\right) = \psi_+\left(\lambda\right)$, with solution to the Cauchy problem

$$\partial_\lambda \psi_+\left(\lambda\right) = \pi_+^{-1}\left(\tfrac{e^{\psi_+(\lambda)}-1}{e^\lambda - 1}\right),\ \psi_+\left(\lambda_+^*\right) = \psi_+^*, \tag{3.27}$$

where $\psi_+^* := \psi_+\left(\lambda_+^*\right)$ for some $\lambda_+^* \in (0, \infty)$. Since λ_-^* and λ_+^* are finite quantities, a unique global solution exists for both Cauchy problems (3.26), *(3.27), it is $\mathcal{AC}$ and has continuous first derivative.*

This last theorem quite immediately implies the convexity of $-\phi$ since if the cumulants $-\psi_-$ and $-\psi_+$ have continuous first derivatives their Frenchel-Legendre transforms $-\hat{\phi}_-$, $-\hat{\phi}_+$ must be strictly convex, with $\hat{\phi}_- = \phi_-$ and $\hat{\phi}_+ = \phi_+$. Then we can state the following corollary on the shape of ϕ:

Corollary 8. *Let $\pi \in \mathcal{U}$ be an invertible function on $[0, \inf C_\pi)$, and denote by $\pi_-^{-1} : [\pi(0), \pi(\inf C_\pi)) \to [0, \inf C_\pi)$ its inverse function. If π_-^{-1} is $\mathcal{AC}$ and Lipschitz, then ϕ_- is in $\mathcal{AC}$, is strictly concave on $[0, \inf C_\pi)$, and strictly increasing from $\log \bar{\pi}(0)$ to 0. Let π be invertible on $(\sup C_\pi, 1]$, with inverse function $\pi_+^{-1} : (\pi(\sup C_\pi), \pi(1)] \to (\sup C_\pi, 1]$. If π_+^{-1} is $\mathcal{AC}$ and Lipschitz, then ϕ_+ is in $\mathcal{AC}$, it is strictly concave on $(\sup C_\pi, 1]$, and strictly decreasing from 0 to $\log \pi(1)$.*

Another potentially useful application is the inverse problem of obtaining an urn function such that the urn process realizes a given ϕ. Since $-\hat{\phi}(s)$ is convex by definition, then $\psi(\lambda) = \hat{\phi}(\partial_\lambda \psi) - \lambda \partial_\lambda \psi$, from which it follows that $\lambda(s) = -\partial_s \hat{\phi}(s)$ and $\psi(\lambda(s)) = \hat{\phi}(s) - s\partial_s \hat{\phi}(s)$. If $-\phi$ is convex, then obviously $\phi = \hat{\phi}$ and we can state the following corollary:

Corollary 9. *Let $f : [0, 1] \to (-\infty, 0]$ be a bounded and concave $\mathcal{AC}$ func-*

tion, and define the function π_f as follows:

$$\pi_f(s) = \frac{e^{f(s)-s\partial_s f(s)} - 1}{e^{-\partial_s f(s)} - 1},\ s \in [0,1]. \tag{3.28}$$

If the function f is such that $\pi_f \in \mathcal{U}$, then the limit ϕ defined in Eq. (3.8) for an urn process with urn function π_f is $\phi = f$.

As example, we may ask for an urn process whose limit ϕ is $f(s) = -\frac{1}{2}\left(s - \frac{1}{2}\right)^2$, and from the above statement we would find that the urn function π_f of such process is $\pi_f(s) = [\exp(\frac{1}{2}s^2 - \frac{1}{8}) - 1]/[\exp\left(s - \frac{1}{2}\right) - 1]$. We believe that such result could find useful applications in those stochastic approximation algorithms for which the process is required to satisfy some given LDP.

3.3.1 Linear urn functions

An important example on which we can apply Theorem 7 is that of linear maps $\pi(s) = s_0 + b(s - s_0)$. As mentioned earlier, this linear urn function may describe the *tenable* Bagchi-Pal model with positive diagonal entries, provided we take $s_0 \in [0,1]$, $b \in [-s_0/(1-s_0), 1]$ when $s_0 \in [0, 1/2]$, and $b \in [-(1-s_0)/s_0, 1]$ when $s_0 \in [1/2, 1]$ (those bounds on b ensure that both $\pi(0), \pi(1) \in [0,1]$).

Large deviations for the final state of a *subtractive* Bagchi-Pal urn (urn with negative diagonal entries: $a_{11} < 0$, $a_{22} < 0$) have been explicitly computed in [FGP2005] by extracting the limiting properties from the expression of the urn composition at each time. Since we believe that their technique would apply also to the case of positive entries (ie, that embedded in the linear urn functions) we can't consider the following findings as completely innovative in the context of LDPs for the Bagchi-Pal urns. Anyway, we remark that the results we are about expose are new and show some surprising properties that, to the best of our knowledge, are observed for the first time.

Corollary 10. *Let π be such that $\pi(s) = a + bs$ with $a \in (0,1)$, $b \in (-\infty, 1-a)$ for $s \in [0, a/(1-b)]$, and $\pi(s) = c + ds$, $c + d \in (0,1)$, $c \in (0,\infty)$ for $s \in [c/(1-d), 1]$ (these conditions on a, b, c, d ensure that $\pi(s) \in (0,1)$ for both intervals $s \in [0, \inf C_\pi] = [0, a/(1-b)]$ and $s \in [\sup C_\pi, 1] = [c/(1-d), 1]$). Let ψ be as in Eq. (3.23) and define the function*

$$B(\alpha, \beta; x_1, x_2) = \int_{x_1}^{x_2} dt\ (1-t)^{\alpha-1} t^{\beta-1}. \tag{3.29}$$

Then, for $\lambda > 0$ we have $\psi = \psi_+$, with

$$\psi_+(\lambda) = \psi_+(\lambda; d<0)\, \mathbb{I}_{\{d<0\}} + \psi_+(\lambda; d>0)\, \mathbb{I}_{\{d>0\}}, \tag{3.30}$$

where $\psi_+(\lambda; d>0)$, $\psi_+(\lambda; d<0)$ are given by the expressions

$$e^{-\psi_+(\lambda; d>0)} = \tfrac{c}{d} e^{-\frac{c}{d}\lambda} \left(1 - e^{-\lambda}\right)^{\frac{1}{d}} B\left(\tfrac{c}{d}, \tfrac{d-1}{d}; 1 - e^{-\lambda}, 1\right), \tag{3.31}$$

$$e^{-\psi_+(\lambda; d<0)} = \tfrac{c}{d} e^{-\frac{c}{d}\lambda} \left(1 - e^{-\lambda}\right)^{\frac{1}{d}} B\left(\tfrac{c}{d}, \tfrac{d-1}{d}; 0, 1 - e^{-\lambda}\right). \tag{3.32}$$

If $\lambda < 0$ we have instead $\psi = \psi_-$, with

$$\psi_-(\lambda) = \psi_-(\lambda; b<0)\,\mathbb{I}_{\{b<0\}} + \psi_-(\lambda; b>0)\,\mathbb{I}_{\{b>0\}}, \tag{3.33}$$

where $\psi_-(\lambda; b>0)$, $\psi_-(\lambda; b<0)$ are given by

$$e^{-\psi_-(\lambda;b>0)} - 1 = \tfrac{a}{b} e^{-\frac{1-a+b}{b}\lambda} \left(1-e^{\lambda}\right)^{\frac{1}{b}} B\left(\tfrac{1-a}{b}, \tfrac{b-1}{b}; 1-e^{\lambda}, 1\right), \tag{3.34}$$

$$e^{-\psi_-(\lambda;b<0)} - 1 = \tfrac{a}{b} e^{-\frac{1-a+b}{b}\lambda} \left(1-e^{\lambda}\right)^{\frac{1}{b}} B\left(\tfrac{1-a}{b}, \tfrac{b-1}{b}; 0, 1-e^{\lambda}\right). \tag{3.35}$$

The first intriguing property of the above solution is that if $b > 0$ ($c > 0$) then ψ is non-analytic at $\lambda \to 0^-(\lambda \to 0^+)$. We can see this, as example, from the expression of $\psi_-(\lambda; b>0)$: expanding for small λ we find a non vanishing therm $O\left(\lambda^{1/b}\log(\lambda)\right)$ if $1/b \in \mathbb{N}$ and $O\left(\lambda^{1/b}\right)$ if $1/b \notin \mathbb{N}$, which implies that the derivatives of order $\lfloor 1/b \rfloor + 1$ and higher are singular in $\lambda = 0$. We find this singularity even taking $a = c$, $b = d$, while it disappears for any $b < 0$ ($c < 0$). This behavior is not observed in case of *subtractive* urns, for which the rate function is always analytic in $\lambda = 0$ (see [FGP2005]): this is not surprising, since these urns are quite affine to the case $b < 0$ ($d < 0$), for which we also observe a regular solution. Notice that a non-analytic point in $\lambda = 0$ implies divergent cumulants from a certain order onwards. Moreover, if $b > 1/2$ the shape of $\phi(s)$ around its maximum would even be no more Gaussian, since we would have a divergent second cumulant $\partial^2_\lambda \psi(\lambda) = O(\lambda^{-\gamma})$, with $\gamma = 2 - 1/b > 0$. If $b = 1/2$ we observe a logarithmic divergence. A comparative analysis of this solution with respect to that of

[FGP2005, FDP2006] (and to other aspects of the Bagchi-Pal urn) would be an interesting matter, but we believe this would be far from the scopes of the present work, which we try to keep as general as possible.

4 Proofs

In this section we collected most of the proofs and technical features of the present work. The proofs are presented in the order they appeared in the previous section. We will first deal with the Sample-Path Large Deviation Principle, then the entropy of the event $\{X_n = \lfloor sn \rfloor\}$ and, finally, with the Cumulant generating function. We assume that all random variables and processes are defined in a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

4.1 Sample-path Large Deviation Principles

Here we prove the existence of Sample-Path LDPs for χ_n using some standard Large Deviation tools, such as Mogulskii Theorem and the Varadhan Integral Lemma.

Before we get into the core of this topic, we recall that $\|\varphi\| := \sup_{\tau\in[0,1]} |\varphi_\tau|$ is the usual supremum norm, and we consider the metric space $(\mathcal{Q}, \|\cdot\|)$, with $\mathcal{Q}$ defined in Eq. (3.2). Note that $\mathcal{Q}$ is compact with respect to the supremum norm topology. Moreover, since by definition $\|\varphi\| \leq 1$ for any $\varphi \in \mathcal{Q}$ we trivially find that $\mathcal{Q} \subset L_\infty([0,1])$.

Let $S_\pi : \mathcal{Q} \to (-\infty, 0]$ be as in Eq. (3.5). The following lemma shows the continuity of S_π with respect to the supremum norm for any compact subset of $\mathcal{Q}$ and any $\pi \in \mathcal{U}$.

Lemma 11. *Assume $\pi \in \mathcal{U}$. The functional $S_\pi : \mathcal{Q} \to (-\infty, 0]$ is continuous on the metric space $(\mathcal{Q}, \|\cdot\|)$. Moreover, a function $W_\pi : [0,1] \to [0,\infty)$ exists such that $\lim_{s\to 0} W_\pi(s) = 0$ and $|S_\pi[\varphi] - S_\pi[\eta]| \leq W_\pi(\|\varphi - \eta\|)$,*

$\forall \varphi, \eta \in \mathcal{Q}$.

Proof. Take any $\varphi, \eta \in \mathcal{Q}$. By definition of S_π, we can rearrange the terms as follows

$$\begin{aligned} S_\pi[\varphi] - S_\pi[\eta] = & \int_{\tau\in[0,1]} d\varphi_\tau \, \log\pi(\varphi_\tau/\tau) - \int_{\tau\in[0,1]} d\eta_\tau \, \log\pi(\eta_\tau/\tau) + \\ & + \int_{\tau\in[0,1]} d\tilde{\varphi}_\tau \, \log\bar{\pi}(\varphi_\tau/\tau) - \int_{\tau\in[0,1]} d\tilde{\eta}_\tau \, \log\bar{\pi}(\eta_\tau/\tau), \end{aligned} \tag{4.1}$$

where used the notation $\tilde{\varphi} = \tau - \varphi$, $\tilde{\eta} = \tau - \eta$. Let us first consider $\log\pi(s)$: by definition of the set $\mathcal{U}$ we can decompose $|\log\pi(s)| = g_0(s) + \nu_0 |\log(s)|$, with $\|g_0\| < \infty$, $|g_0(x+\delta) - g_0(x)| \leq f(|\delta|)$, $\lim_{\epsilon\to 0} \epsilon \int_\epsilon^1 dz \, |f(z)|/z^2 = 0$. Then, for $g_0(s)$ we have

$$\begin{aligned} & \int_{\tau\in[0,1]} d\varphi_\tau \, g_0(\varphi_\tau/\tau) - \int_{\tau\in[0,1]} d\eta_\tau \, g_0(\eta_\tau/\tau) = \\ & \quad = \int_{\tau\in[0,1]} d\varphi_\tau \, [g_0(\varphi_\tau/\tau) - g_0(\eta_\tau/\tau)] + \int_{\tau\in[0,1]} d(\varphi_\tau - \eta_\tau) \, g_0(\varphi_\tau/\tau). \end{aligned} \tag{4.2}$$

By uniform continuity one has $|g_0(\varphi_\tau/\tau) - g_0(\eta_\tau/\tau)| \leq f(|\varphi_\tau - \eta_\tau|/\tau)$. Moreover, since $\varphi_\tau \leq \tau$ and $\eta_\tau \leq \tau$, we have

$$|\varphi_\tau - \eta_\tau| \leq \min\{\tau, \|\varphi - \eta\|\}, \tag{4.3}$$

and $d\varphi_\tau \leq d\tau$. Then, if we define $s^{-1}H_f(s) := \int_s^1 dz \, |f(z)|/z^2$ the first integral can be bounded as follows

$$\int_{\tau\in[0,1]} d\varphi_\tau \ \left|g_0\left(\varphi_\tau/\tau\right)-g_0\left(\eta_\tau/\tau\right)\right| \le H_f\left(\|\varphi-\eta\|\right), \tag{4.4}$$

while for the second we get

$$\int_{\tau\in[0,1]} d\left(\varphi_\tau-\eta_\tau\right) \ \left|g_1\left(\varphi_\tau/\tau\right)\right| \le \|g_1\| \, \|\varphi-\eta\|\,. \tag{4.5}$$

Let us introduce $H_1(s) := s - s\log(s)$. For the singular part $\nu_0 \log(s)$ we can integrate by parts

$$\begin{aligned}
&\int_{\tau\in[0,1]} d\varphi_\tau \ \log\left(\varphi_\tau/\tau\right) - \int_{\tau\in[0,1]} d\eta_\tau \ \log\left(\eta_\tau/\tau\right) = \\
&= \int_{\tau\in[0,1]} d\varphi_\tau \ \log\left(\varphi_\tau\right) - \int_{\tau\in[0,1]} d\eta_\tau \ \log\left(\eta_\tau\right) - \int_{\tau\in[0,1]} d\left(\varphi_\tau-\eta_\tau\right) \ \log\left(\tau\right) = \\
&\qquad = \left[H_1\left(\eta_\tau\right)-H_1\left(\varphi_\tau\right)\right]_0^1 + \int_{\tau\in[0,1]} d\tau \ \left(\varphi_\tau-\eta_\tau\right)/\tau. \qquad (4.6)
\end{aligned}$$

where we used $\left[\left(\varphi_\tau-\eta_\tau\right)\log\left(\tau\right)\right]_0^1 = 0$. Then, the bound for the singular part is

$$\left|\left[H_1\left(\varphi_\tau\right)-H_1\left(\eta_\tau\right)\right]_0^1\right| + \int_{\tau\in[0,1]} d\tau \left(\left|\varphi_\tau-\eta_\tau\right|\right)/\tau \le 3H_1\left(\|\varphi-\eta\|\right). \tag{4.7}$$

Since by definition H_f, H_1 are positive for $s\in(0,1]$, and $\lim_{s\to 0} H_1(s) = \lim_{s\to 0} H_f(s) = 0$, we can take the limit $\|\varphi-\eta\| \to 0$ in both bounds. Repeating the same steps for the second part, with g_1, ν_1 on behalf of g_0, ν_0 and $\tilde{\varphi}$, $\tilde{\eta}$ on behalf of φ, η will complete the proof.

$\square$

4.1.1 Change of measure

We need a variational representation for the rate function of χ_n in terms of sample paths. Let $\varphi := \{\varphi_\tau : \tau \in [0,1]\}$, and define

$$\mathcal{Q}_n := \{\varphi :\, \varphi_\tau = n^{-1}\left[\, \textstyle\sum_{i=1}^{\lfloor n\tau \rfloor} \theta_i + (n\tau - \lfloor n\tau \rfloor)\, \theta_{\lfloor n\tau \rfloor} \,\right],\, \theta_i \in \{0,1\}\}. \tag{4.8}$$

The above set is the support of χ_n for $n < \infty$: note that $\mathcal{Q}_n \subset \mathcal{Q}$ for all n. We also introduce the following notation:

$$Y_{n,k}(\varphi) := n\varphi_{k/n},\; \delta Y_{n,k}(\varphi) := n\left(\varphi_{(k+1)/n} - \varphi_{k/n}\right), \tag{4.9}$$

Then, let $\varphi \in \mathcal{Q}_n$: by Eq. (2.1) we can write the sample-path probability $\mathbb{P}(\chi_n = \varphi)$ in therms of φ as follows:

$$\mathbb{P}(\chi_n = \varphi) = \prod_{k=1}^{n-1} \pi\left(Y_{n,k}(\varphi)/k\right)^{\delta Y_{n,k}(\varphi)}\, \bar{\pi}\left(Y_{n,k}(\varphi)/k\right)^{1-\delta Y_{n,k}(\varphi)}. \tag{4.10}$$

Here we state a lemma which relates the functional S_π to the entropy of the event $\{\chi_n = \varphi\}$, when $\varphi \in \mathcal{Q}_n$.

Lemma 12. *Take some $\varphi \in \mathcal{Q}_n$, and let $S_\pi : \mathcal{Q} \to (-\infty, 0]$ as in Eq. (3.5): then, $n^{-1} \log \mathbb{P}(\chi_n = \varphi) = S_\pi[\varphi] + O(W_\pi(1/n))$, with W_π as in Lemma 11.*

Proof. Let $\varphi \in \mathcal{Q}_n$. To estimate the difference between $n^{-1} \log \mathbb{P}(\chi_n = \varphi)$

and $S_\pi[\varphi]$ we can proceed as follows. First, we define

$$\epsilon_n := \left\{\epsilon_{n,\tau} = (n\tau/\lfloor n\tau\rfloor)\,\varphi_{\lfloor n\tau\rfloor/n} - \varphi_\tau \,:\, \tau\in[0,1]\right\}, \tag{4.11}$$

such that the difference between $n^{-1}\log\mathbb{P}(\chi_n=\varphi)$ and $S_\pi[\varphi]$ can be written as follows

$$\begin{aligned} n^{-1}\log\mathbb{P}(\chi_n=\varphi) - S_\pi[\varphi] = & \\ = \int_{\tau\in[0,1]} d\varphi_\tau\, \left[\log\pi\left(\left(\varphi_\tau+\epsilon_{n,\tau}\right)/\tau\right) - \log\pi\left(\varphi_\tau/\tau\right)\right] + & \\ + \int_{\tau\in[0,1]} d\tilde{\varphi}_\tau\, \left[\log\bar{\pi}\left(\left(\varphi_\tau+\epsilon_{n,\tau}\right)/\tau\right) - \log\bar{\pi}\left(\varphi_\tau/\tau\right)\right], & \end{aligned} \tag{4.12}$$

Even if ϵ_n is discontinuous at each $\tau=\lfloor n\tau\rfloor/n$, it still satisfies the condition $\epsilon_{n,\tau}\le\min\{\tau,\,\|\epsilon_{n,\tau}\|\}$. Then, we can proceed as in Lemma 11. First consider the $\log\pi$ dependent integral and decompose $|\log\pi(s)| = g_0(s)+\nu_0|\log(s)|$: for the non singular part

$$\int_{\tau\in[0,1]} d\varphi_\tau\, \left|g_0\left(\left(\varphi_\tau+\epsilon_{n,\tau}\right)/\tau\right) - g_0\left(\varphi_\tau/\tau\right)\right| \le H_f\left(\|\epsilon_{n,\tau}\|\right). \tag{4.13}$$

while for the log part we have

$$\begin{aligned} \int_{\tau\in[0,1]} d\varphi_\tau\, \left[\log\left(\left(\varphi_\tau+\epsilon_{n,\tau}\right)/\tau\right) - \log\left(\varphi_\tau/\tau\right)\right] = & \\ = \int_{\tau\in[0,1]} d\varphi_\tau\, \left[\log\left(\varphi_\tau+\epsilon_{n,\tau}\right) - \log\left(\varphi_\tau\right)\right] = & \\ = \left[H_1\left(\varphi_\tau+\epsilon_{n,\tau}\right) - H_1\left(\varphi_\tau\right)\right]_0^1 - \int_{\tau\in[0,1]} d\epsilon_{n,\tau}\, \log\left(\varphi_\tau+\epsilon_{n,\tau}\right). & \end{aligned} \tag{4.14}$$

As before, $\left|\left[H_1\left(\varphi_\tau+\epsilon_{n,\tau}\right)-H_1\left(\varphi_\tau\right)\right]_0^1\right| \leq 2H_1\left(\|\epsilon_{n,\tau}\|\right)$. Then, define $\xi_\tau = \min\left\{\varphi_\tau+\epsilon_{n,\tau},\, \varphi_\tau\right\}$: since $\xi_\tau \geq 0$ we can write

$$\int_{\tau\in[0,1]} d\epsilon_{n,\tau}\ \left|\log\left(\varphi_\tau+\epsilon_{n,\tau}\right)\right| = \int_{\tau\in[0,1]} d\epsilon_{n,\tau}\ \left|\log\left(\xi_\tau+\left|\epsilon_{n,\tau}\right|\right)\right| \leq$$
$$\leq \int_{\tau\in[0,1]} d\epsilon_{n,\tau}\ \left|\log\left(\left|\epsilon_{n,\tau}\right|\right)\right| = \left|\left[H_1\left(\epsilon_{n,\tau}\right)\right]_0^1\right| \leq H_1\left(\|\epsilon_{n,\tau}\|\right). \quad (4.15)$$

Since $\|\epsilon_{n,\tau}\| \leq 1/n$ we conclude that $H_1\left(\|\epsilon_{n,\tau}\|\right) \leq H_1\left(1/n\right)$. Repeating the same steps for the $\log\bar{\pi}$ integral of Eq. (4.10) completes the proof .

□

Let us now introduce the binomial urn process $B_n := \left\{B_{n,k} : 0 \leq k \leq n\right\}$, with constant urn function $\pi(s) = 1/2$ (set $B_0 = 0$). We define $\delta B_{n,k} := B_{n,k+1} - B_{n,k}$. The process δB_n is a sequence of binary i.i.d. random variables with $\mathbb{P}\left(\delta B_{n,k} = 1\right) = \mathbb{P}\left(\delta B_{n,k} = 0\right) = 1/2$, so that each $Y_n(\varphi)$, $\varphi \in \mathcal{Q}_n$ realization of B_n up to time n has constant measure $\mathbb{P}\left(B_n = Y_n(\varphi)\right) = 2^{-n}$. We denote by $\varphi_n : [0,1] \to [0,1]$ the linear interpolation of the $n^{-1}B_k$ sequence for $0 \leq k \leq n$:

$$\beta_n := \left\{\beta_{n,\tau} = n^{-1}\left[B_{n,\lfloor n\tau\rfloor} + \left(n\tau - \lfloor n\tau\rfloor\right)\delta B_{n,\lfloor n\tau\rfloor}\right] : \tau \in [0,1]\right\}. \quad (4.16)$$

Note that $\beta_n \in \mathcal{Q}_n \subset \mathcal{Q}$ for all n. A sample-path LDP for the sequence of functions $\left\{\beta_n : n \in \mathbb{N}\right\}$ is provided by the Mogulskii Theorem [DZ1998].

Lemma 13. *The sequence $\left\{\beta_n : n \in \mathbb{N}\right\}$ defined by Eq.(4.16) with support*

$\mathcal{Q}$ satisfies a LDP in $(\mathcal{Q}, \|\cdot\|)$, with the good rate function

$$I_{1/2}[\varphi] = \begin{cases} \log 2 + \int_0^1 d\tau\, H(\dot{\varphi}_\tau) & \text{if } \varphi \in \mathcal{AC} \\ \infty & \text{otherwise}, \end{cases} \tag{4.17}$$

where $\mathcal{AC}$ is the class of absolutely continuous functions, and $H(s) = s\log s + \bar{s}\log\bar{s}$ as in Theorem 1.

Proof. Since $\beta_n \in \mathcal{Q} \subset L_\infty([0,1])$, Mogulskii Theorem [DZ1998] predicts a LDP for the sequence $\{\beta_n : n \in \mathbb{N}\}$, with good rate function $I_{1/2}[\varphi] = -\int_0^1 d\tau \hat{\Lambda}(\dot{\varphi}_\tau)$ if $\varphi \in \mathcal{AC}$ and $I_{1/2}[\varphi] = \infty$ otherwise, and where $\hat{\Lambda}(s)$ is the Frenchel-Legendre transform of the moment generating function $\Lambda(\lambda) := \mathbb{E}[\exp(\lambda\, \delta Y_{n,1})]$. In our case we have $\Lambda(\lambda) = \left(e^\lambda + 1\right)/2$, then $\hat{\Lambda}(s) = -\log 2 - H(s)$.

□

4.1.2 Proof of Theorem 1

We will use a corollary of the Varadhan Integral Lemma (Lemmas 4.3.2 and 4.3.4 of [DZ1998]) to prove the sample-path LDP for the χ_n sequence, stated in Theorem 1.

Proof. Let $I_\pi[\varphi] := J[\varphi] - S_\pi[\varphi]$ and let $\mathcal{B}$ any subset of $\mathcal{Q}$: we define the following $\mathcal{B}-$dependent functional:

$$S_{\pi,\mathcal{B}}[\varphi] := \begin{cases} S_\pi[\varphi] = J[\varphi] - I_\pi[\varphi] & if\ \varphi \in \mathcal{B} \\ -\infty & otherwise. \end{cases} \tag{4.18}$$

and denote by $\mathbb{E}_0$ the expectation over the possible realizations of the binomial process β_n. By equation (4.10) and Lemma 12 we find that

$$\lim_{n\to\infty} n^{-1} \log \mathbb{P}(\chi_n \in \mathcal{B}) = \log 2 + \lim_{n\to\infty} n^{-1} \log \mathbb{E}_0 \left(e^{nS_\pi[\beta_n]} \mathbb{I}_{\{\beta_n \in \mathcal{B}\}}\right) =$$
$$= \log 2 + \lim_{n\to\infty} n^{-1} \log \mathbb{E}_0 \left(e^{nS_{\pi,\mathcal{B}}[\beta_n]}\right). \tag{4.19}$$

Then, consider $S_{\pi,\mathrm{cl}(\mathcal{B})}$: since $\mathrm{cl}(\mathcal{B})$ is a closed set, and Lemma 11 states that S_π is a continuous functional on $(\mathcal{Q}, \|\cdot\|)$, it follows that $S_{\pi,\mathrm{cl}(\mathcal{B})}$ is upper semicontinuous on $(\mathcal{Q}, \|\cdot\|)$, and Lemma 4.3.2 of [DZ1998] gives the upper bound

$$\log 2 + \limsup_{n\to\infty} n^{-1} \log \mathbb{E}_0 \left(e^{nS_{\pi,\mathrm{cl}(\mathcal{B})}[\beta_n]}\right) \leq$$
$$\leq \log 2 + \sup_{\varphi\in\mathcal{Q}} \left\{S_{\pi,\mathrm{cl}(\mathcal{B})}[\varphi] - I_{1/2}[\varphi]\right\} =$$
$$= \log 2 + \sup_{\varphi\in\mathrm{cl}(\mathcal{B})} \{S_\pi[\varphi] - \log 2 - J[\varphi]\} = - \inf_{\varphi\in\mathrm{cl}(\mathcal{B})} I_\pi[\varphi]. \tag{4.20}$$

Now consider $S_{\pi,\mathrm{int}(\mathcal{B})}$: $\mathrm{int}(\mathcal{B})$ is open and this time we have a lower semicontinuous functional on $(\mathcal{Q}, \|\cdot\|)$, then by Lemma 4.3.3 of [DZ1998] we can write

$$\log 2 + \liminf_{n\to\infty} n^{-1} \log \mathbb{E}_0 \left(e^{nS_{\pi,\mathrm{int}(\mathcal{B})}[\beta_n]} \right) \geq - \inf_{\varphi\in\mathrm{int}(\mathcal{B})} I_\pi [\varphi] . \tag{4.21}$$

which completes the main statement of Theorem 1.

□

4.1.3 Initial conditions and time-inhomogeneous functions

First we deal with the influence of initial conditions on the large deviation properties of our urn process. Until now we considered processes with initial conditions $X_{n,0} = 0$ and $x_{n,0} = 1/2$ by convention: the following lemma shows that fixing $X_{n,m}$ for some $m > 0$ will not affect the rate function for the considered class $\mathcal{U}$ of urn function, provided that m is finite and $0 < X_{n,m} < 1$ for those urn functions with $\pi(0) = 0$ and $\pi(1) = 1$.

Lemma 14. *Let X_n be a urn process with urn function $\pi \in \mathcal{U}$ and initial conditions $0 < X_{n,m} < m < \infty$. Then, the rate function is independent from these initial conditions (if $0 < \pi < 1$ then the same result holds for $0 \leq X_{n,m} \leq m < \infty$).*

Proof. Let $\varphi \in \mathcal{Q}_n$ $x_{m,n} = m^{-1} X_{m,n}$ and $\epsilon_{n,\tau}$ as in Lemma 12. Then we can use the estimates of Lemma 12 to obtain

$$n^{-1}\left|\log\mathbb{P}\left(\chi_n=\varphi\,|\,\varphi_{m/n}=x_{m,n}\right)-\log\mathbb{P}\left(\chi_n=\varphi\right)\right|\leq$$
$$\leq\int_{\tau\in[0,m/n]}d\varphi_\tau\,\left|\log\pi\left(\tfrac{\varphi_\tau+\epsilon_{n,\tau}}{\tau}\right)\right|+\int_{\tau\in[0,m/n]}d\tilde{\varphi}_\tau\,\left|\log\bar{\pi}\left(\tfrac{\varphi_\tau+\epsilon_{n,\tau}}{\tau}\right)\right|\leq$$
$$\leq\left(\|g_0\|+\|g_1\|\right)m/n+3\left(\nu_0+\nu_1\right)H_1\left(m/n\right). \quad (4.22)$$

This difference vanishes as $n\to\infty$ for any $\varphi\in\mathcal{Q}_n$. This obviously implies that the LDPs governing the two processes share the same rate function. Notice that the above inequality would also hold for $\nu_0>0$, $\nu_1>0$ and monochromatic initial conditions $X_{n,m}\in\{0,m\}$ since the process would be deterministic with $X_{n,k}=X_{n,m}$ for all $0\leq k\leq n$. Anyway, we have to discard this case, whose evolution is obviously dependent on weather $X_{n,m}=0$ or $X_{n,m}=m$.

□

Finally, we prove Corollary 2. As remarked, we prove the result only for urn function with $0<\pi<1$, for which the proof is straightforward.

Proof. Let $\pi\in\mathcal{U}$ with $0<\pi<1$ and let π_n such that such that $|\pi_n(s)-\pi(s)|\leq\delta_n$, $\lim_n\delta_n=0$ for all $s\in[0,1]$. By lemma 12 it suffices to show that $|S_{\pi_n}[\varphi]-S_\pi[\varphi]|\to 0$ as $n\to 0$. We can bound $|S_{\pi_n}[\varphi]-S_\pi[\varphi]|$ as follows

$$
\begin{aligned}
|S_{\pi_n}[\varphi] - S_\pi[\varphi]| \le & \int_{\tau\in[0,1]} d\varphi_\tau \; |\log \pi_n(\varphi_\tau/\tau) - \log \pi(\varphi_\tau/\tau)| + \\
& + \int_{\tau\in[0,1]} d\tilde{\varphi}_\tau \; |\log \bar{\pi}_n(\varphi_\tau/\tau) - \log \bar{\pi}(\varphi_\tau/\tau)| \le \\
& \le \int_{\tau\in[0,1]} d\varphi_\tau \, \delta_n / |\pi(\varphi_\tau/\tau)| + \int_{\tau\in[0,1]} d\tilde{\varphi}_\tau \, \delta_n / |\bar{\pi}(\varphi_\tau/\tau)| \le \\
& \le [1/(1 - \|\bar{\pi}\|) + 1/(1 - \|\pi\|)] \, \delta_n .
\end{aligned}
$$

Since for $0 < \pi < 1$ we have $\|\bar{\pi}\| < \infty$, $\|\pi\| < \infty$, the above bound vanishes as $\delta_n \to 0$ and the proof is completed.

□

4.2 Large deviations for the event $X_{n,n} = \lfloor sn \rfloor$

In this section we use the variational representation of Sample-Path LDPs to show Theorem 3 and Corollary 4. Since the event $\{X_{n,n} = \lfloor sn \rfloor\}$ is slightly finer than that usually considered in large deviations theory, its analysis requires some additional estimates. Moreover, note that $\mathcal{Q}_s$ is not an I_π−continuity set because of the fixed endpoint condition $\varphi_1 = s$, which implies $\mathrm{cl}(\mathcal{Q}_s) = \varnothing$. We circumvent this problem as follows

Lemma 15. *Let $s \in [0,1]$, $\delta > 0$ and define $\mathcal{Q}_{s,\delta} := \bigcup_{u-s\in[0,\delta]} \mathcal{Q}_u$, where $\mathcal{Q}_s := \{\varphi \in \mathcal{Q} : \varphi_1 = s\}$, then*

$$
\lim_{n\to\infty} n^{-1} \log \mathbb{P}(\lfloor sn \rfloor \le X_{n,n} \le \lfloor (s+\delta) n \rfloor) = - \inf_{\varphi\in\mathcal{Q}_{s,\delta}} I_{\pi,\epsilon}[\varphi] . \tag{4.23}
$$

Proof. Since $\mathcal{Q}_{s,\delta} := \bigcup_{u-s\in[0,\delta]} \mathcal{Q}_u$ is an $I_\pi-$continuity set when $s \in [0,1]$ and $\delta > 0$, by Theorem 1 we have

$$\lim_{n\to\infty} n^{-1} \log \mathbb{P}\left(\chi_n \in \mathcal{Q}_{s,\delta}\right) = -\inf_{\varphi\in\mathcal{Q}_{s,\delta}} I_\pi\left[\varphi\right]. \tag{4.24}$$

Then, let $0 < \nu < \delta$ so that we can write

$$\begin{aligned}
-\inf_{\varphi\in\mathcal{Q}_{s,\delta-\nu}} I_\pi\left[\varphi\right] = \lim_{n\to\infty} n^{-1} \log \mathbb{P}\left(\chi_n \in \mathcal{Q}_{s,\delta-\nu}\right) \leq \\
\leq \lim_{n\to\infty} n^{-1} \log \mathbb{P}\left(\lfloor sn \rfloor \leq X_{n,n} \leq \lfloor (s+\delta)\, n \rfloor\right) \leq \\
\leq \lim_{n\to\infty} n^{-1} \log \mathbb{P}\left(\chi_n \in \mathcal{Q}_{s,\delta+\nu}\right) = -\inf_{\varphi\in\mathcal{Q}_{s,\delta+\nu}} I_\pi\left[\varphi\right].
\end{aligned} \tag{4.25}$$

Since I_π is continuous on $(\mathcal{Q}, \|\cdot\|)$ and $\mathcal{Q}_{s,\delta'} \subset \mathcal{Q}_{s,\delta} \subset \mathcal{Q}$ for every $\delta' < \delta$, we can take the limit $\nu \to 0$ and the proof is completed.

□

4.2.1 Proof of Theorem 3

Before starting, we remind some notation. Let $\varphi := \{\varphi_\tau : \tau \in [0,1]\}$ and let $Y_{n,k}(\varphi) := n\varphi_k$, $\delta Y_{n,k}(\varphi) := n\left(\varphi_{k+1} - \varphi_k\right)$ as in Eq. (4.9). We also define the set of trajectories

$$\mathcal{Q}_{n,k} := \left\{\varphi \in \mathcal{Q}_n : Y_{n,n}(\varphi) = k\right\}, \tag{4.26}$$

where $\mathcal{Q}_n$ is the support of χ_n as defined in Eq. (4.8).

Proof. We start from the variational representation of $\mathbb{P}\left(\chi_n = \varphi\right)$ in Eq. (4.10): by Lemma 12 we can rewrite $\mathbb{P}\left(X_{n,n} = k\right)$ as

$$\mathbb{P}\left(X_{n,n} = k\right) = \sum_{\varphi \in \mathcal{Q}_{n,k}} \mathbb{P}\left(\chi_n = \varphi\right) = \sum_{\varphi \in \mathcal{Q}_{n,k}} e^{nS_\pi[\varphi] + O(n \cdot W_\pi(1/n))}. \tag{4.27}$$

First, we observe that the following inequalities holds:

$$\mathbb{P}\left(X_{n,n} = k\right) \leq \mathbb{P}\left(k \leq X_{n,n} \leq k'\right) \leq \left(k' - k\right) \sup_{k \leq i \leq k'} \mathbb{P}\left(X_{n,n} = i\right) : \tag{4.28}$$

by defining $k^* : \mathbb{P}\left(X_{n,n} = k^*\right) = \sup_{k \leq i \leq k'} \mathbb{P}\left(X_{n,n} = i\right)$ we can rewrite them as

$$\begin{aligned} &\left|\log \mathbb{P}\left(k \leq X_{n,n} \leq k'\right) - \log \mathbb{P}\left(X_{n,n} = k\right)\right| \leq \\ &\qquad \leq \log\left(k' - k\right) + \left|\log \mathbb{P}\left(X_{n,n} = k^*\right) - \log \mathbb{P}\left(X_{n,n} = k\right)\right|. \end{aligned} \tag{4.29}$$

Let $T^0\left(\varphi\right) := \{i \in \mathbb{N} : \delta Y_{n,i}\left(\varphi\right) = 0\}$, $T^1\left(\varphi\right) := \{i \in \mathbb{N} : \delta Y_{n,i}\left(\varphi\right) = 1\}$ and define the operator $\hat{u}_h$ such that $\hat{u}_h \varphi := \{(\hat{u}_h\varphi)_\tau : \tau \in [0,1]\}$,

$$(\hat{u}_h\varphi)_\tau := \varphi_\tau + \left(\tau - \tfrac{1}{n}\left\lfloor n\tau \right\rfloor\right) \mathbb{I}_{\{n\tau \in [h-1,h]\}} + \tfrac{1}{n}\mathbb{I}_{\{n\tau \in [h,n]\}}. \tag{4.30}$$

If we apply m times this operator to $\varphi \in \mathcal{Q}_{n,k}$ with a suitable sequence of h_i, $1 \leq i \leq m$ we can get a $\hat{u}_{h_m} \ldots \hat{u}_{h_1}\varphi \in \mathcal{Q}_{n,k+m}$. By simple combinatorial arguments it's easy to convince ourselves that the following relation holds

$$\sum_{\varphi\in\mathcal{Q}_{n,k+m}} e^{nS_\pi[\varphi]} = \prod_{j=1}^{m}(k+j)^{-1} \sum_{\varphi\in\mathcal{Q}_{n,k}} \sum_{h_1\in T^0(\varphi)} \sum_{h_2\in T^0\left(\hat{u}_{h_1}\varphi\right)} \cdots$$
$$\cdots \sum_{h_{m-1}\in T^0\left(\hat{u}_{h_{m-2}}\dots\hat{u}_{h_1}\varphi\right)} \sum_{h_m\in T^0\left(\hat{u}_{h_{m-1}}\dots\hat{u}_{h_1}\varphi\right)} e^{nS_\pi\left[\hat{u}_{h_m}\dots\hat{u}_{h_1}\varphi\right]}; \quad (4.31)$$

the product comes from noticing that $\left|T^1(\varphi)\right| = k+j$ when $\varphi\in\mathcal{Q}_{n,k+j}$: it corrects for the exceeding copies of the same path which arise from summing over the $T^0(\dots\hat{u}_{h_2}\hat{u}_{h_1}\varphi)$ sets. Now, since by definition $\|\hat{u}_{h_m}\dots\hat{u}_{h_1}\varphi-\varphi\| = m/n$, from Lemma 11 we have

$$n\left|S_\pi\left[\hat{u}_{h_m}\dots\hat{u}_{h_1}\varphi\right]-S_\pi[\varphi]\right| \le n\,W_\pi(m/n)\,, \quad (4.32)$$

and, given that$\left|T^0(\hat{u}_{h_i}\dots\hat{u}_{h_1}\varphi)\right| = n-k+i-1$ when $\varphi_n\in\mathcal{Q}_{n,k}$, from Eqs. (4.27), (4.31) and (4.32) we can conclude that

$$\begin{aligned}&\left|\log\mathbb{P}(X_{n,n}=k+m)-\log\mathbb{P}(X_{n,n}=k)\right| \le \\ &\qquad \le \left|\textstyle\sum_{i=1}^{m}\log\left((n-k+i-1)/(k+i)\right)\right| + \\ &\qquad\qquad + n\,W_\pi(m/n) + O\left(W_\pi(1/n)\right). \quad (4.33)\end{aligned}$$

Then, we can put together Eqs. (4.27), (4.29), (4.33) and the inequality $k \leq k^* \leq k'$ to get the bound

$$
\begin{aligned}
&\left|\log \mathbb{P}\left(k \leq X_{n,n} \leq k'\right) - \log \mathbb{P}\left(X_{n,n} = k\right)\right| \leq \\
&\qquad \leq \left|\textstyle\sum_{i=1}^{k'-k} \log\left(\left(n-k+i-1\right)/\left(k+i\right)\right)\right| + \\
&\qquad\quad + n\, W_\pi\left(m/n\right) + + O\left(W_\pi\left(1/n\right)\right) + \log\left(k'-k\right), \quad (4.34)
\end{aligned}
$$

By taking $k = \lfloor sn \rfloor$, $k' = \lfloor (s+\delta)\, n \rfloor$, then the limit $n \to \infty$, we find that the sum in the above inequality has the following limiting behavior

$$
\begin{aligned}
&\lim_{n\to\infty} n^{-1} \textstyle\sum_{i=1}^{k'-k} \log\left(\frac{n-k+i-1}{k+i}\right) = \displaystyle\int_{u\in[0,\delta]} du\, \log\left(\left(\bar{s}+u\right)/\left(s+u\right)\right) = \\
&\qquad = H_1\left(s+\delta\right) - H_1\left(s\right) - H_1\left(\bar{s}+\delta\right) + H_1\left(\bar{s}\right) =: H_2\left(s,\delta\right), \quad (4.35)
\end{aligned}
$$

where $H_1(s) = s - s\log s$ as in Lemma 11. Then, applying Lemma 15 and the above relation to Eq. (4.34) we finally obtain the bound

$$
\left|\phi\left(s\right) + \inf_{\varphi\in\mathcal{Q}_{s,\delta}} I_\pi\left[\varphi\right]\right| \leq \left|H_2\left(s,\delta\right)\right| + W_\pi\left(\delta\right) \qquad (4.36)
$$

In the end, since I_π is continuous on $(\mathcal{Q}, \|\cdot\|)$ and $\mathcal{Q}_s \subset \mathcal{Q}_{s,\delta} \subset \mathcal{Q}$, taking $\delta \to 0$ in the above equation will complete our proof. Notice that our bound diverges for $s \in \{0,1\}$, but in such cases the theorem's statement is trivially verified by a direct computation, hence we can assume $s \in (0,1)$.

□

Before dealing with Corollaries 4, 5 and 6 we still need an additional result. We start by finding conditions on φ such $I_\pi[\varphi]=0$. From Theorem 3 we found that $\phi(s)=-\inf_{\varphi\in\mathcal{Q}_s} I_\pi[\varphi]$, and since $I_\pi[\varphi]\geq 0$ our thesis would follow if we can find a trajectory $\varphi\in\mathcal{Q}_s\cap\mathcal{AC}$ such that $I_\pi[\varphi]=0$. The following lemma provides the desired condition on φ

Lemma 16. *Let $\varphi^*:=\{\varphi^*_\tau:\tau\in[0,1]\}$ such that $I_\pi[\varphi^*]=0$. Then, any of such φ^* must satisfy the homogeneous differential equation $\dot{\varphi}^*_\tau=\pi(\varphi^*_\tau/\tau)$ with $\varphi^*\in\mathcal{Q}\cap\mathcal{AC}$.*

Proof. Let $(x,y)\in[0,1]^2$ and $\bar{x}=1-x$, $\bar{y}=1-y$ as usual. Then, define the function $L:[0,1]^2\to(-\infty,0]$ as follows:

$$L(x,y):=x\log(y/x)+\bar{x}\log(\bar{y}/\bar{x}). \tag{4.37}$$

Since by Theorem 3 and Lemma 13 we have $I_\pi[\varphi]=\infty$ when $\varphi\notin\mathcal{AC}$, we can restrict the search for minimizing strategies to the set $\mathcal{Q}\cap\mathcal{AC}$, for which $\dot{\varphi}$ exists almost everywhere. Then, for every $\varphi\in\mathcal{Q}\cap\mathcal{AC}$ we can write $I_\pi[\varphi]$ as

$$I_\pi[\varphi]=-\int_{\tau\in[0,1]} d\tau\, L(\dot{\varphi}_\tau,\pi(\varphi_\tau/\tau)). \tag{4.38}$$

L is a negative concave function for every pair $(x,y)\in[0,1]^2$, with $L(x,y)=0$ if and only if $x=y$. Hence, any choice of φ for which $I_\pi[\varphi]=0$ must satisfy the condition $\dot{\varphi}_\tau=\pi(\varphi_\tau/\tau)$ for every $\tau\in[0,1]$.

□

4.2.2 Proof of Corollaries 4, 5 and 6

We can now prove the corollaries of Theorem 3 concerning optimal trajectories. Since Corollary 4 is an almost obvious consequence of 5 and 6, we first concentrate on the last two, and prove Corollary 4in the end of this subsection.

Proof. Lemma 16 states that every trajectory for which $I_\pi[\varphi^*] = 0$ is in $\mathcal{AC}$ and must satisfy the homogeneous differential equation $\dot{\varphi}^*_\tau = \pi(\varphi^*_\tau/\tau)$. Then our zero-cost trajectory, if existent, must be a solution to the homogeneous Cauchy Problem

$$\dot{\varphi}^*_\tau = \pi(\varphi^*_\tau/\tau), \; \varphi^*_1 = s. \tag{4.39}$$

To characterize the solution we first define $u^* : [0,1] \to [0,1]$ as

$$u^* := \{u^*_\tau, \tau \in [0,1] : u^*_\tau = \varphi^*_\tau/\tau\}, \tag{4.40}$$

such that we can rewrite the Cauchy problem (4.39) as

$$\dot{u}^*_\tau = \tfrac{1}{\tau}\left[\pi(u^*_\tau) - u^*_\tau\right], \; u^*_1 = s. \tag{4.41}$$

If $a_{\pi,i} = 0$ then $\pi(s) - s = 0$ for $s \in K_{\pi,i}$, and the solution is trivially $u^* = s$, then we concentrate on $a_{\pi,i} \neq 0$. We recall that for $a_{\pi,i} \neq 0$ the boundary $\partial K_{\pi,i}$ of $K_{\pi,i}$ is a set of two isolated points. Then, let $\partial K_{\pi,i} = \left\{s^*_i, s^\dagger_i\right\}$ with

$$s^*_i := \mathbb{I}_{\{a_{\pi,i}=1\}} \inf K_{\pi,i} + \mathbb{I}_{\{a_{\pi,i}=-1\}} \sup K_{\pi,i}, \tag{4.42}$$

$$s_i^{\dagger} := \mathbb{I}_{\{a_{\pi,i}=-1\}} \inf K_{\pi,i} + \mathbb{I}_{\{a_{\pi,i}=1\}} \sup K_{\pi,i}, \tag{4.43}$$

such that $\pi(s) - s,\ s \in K_{\pi,i}$ is always decreasing in the neighborhood of $s_i^*$and increasing in that of $s_i^{\dagger}$ at least if $1 \leq i \leq |\partial C_\pi| - 1$.

First, we notice that both constant trajectories $u_\tau^* = s_i^{\dagger}$ and $u_\tau^* = s_i^*$ satisfy the Cauchy problem in Eq. (4.41). To simplify the exposition, we consider $a_{\pi,i} = -1$, such that $s_i^{\dagger} < s_i^*$ and, by Eq. (4.41), u_τ^* must be a decreasing function of $\tau \in [0,1]$ with $u_\tau^* \in [u_1^*,\, u_0^*] \subseteq K_{\pi,i} \cup \partial K_{\pi,i}$.

Given that, we have only two possible kinds of optimal trajectory u_τ^* for the variational problem with $s \in K_{\pi,i} \cup \partial K_{\pi,i}$. The first is that u_τ^* decreases from some $u_0^* < s_i^*$ to $u_1^* = s$, while the second is such that $u_\tau^* = s_i^*$ constant from $\tau = 0$ to some $\tau_{s,i}^* \in [0,1)$, and then it decreases from $s_i^*$to eventually reach s at $\tau = 1$. Then, define

$$F_\pi(s,u) := \int_u^s \frac{dz}{\pi(z) - z} \tag{4.44}$$

for some $s \in K_{\pi,i}$, so that the solution to the Cauchy problem can be written in implicit form as $F_\pi(s, u_\tau^*) = -\log(\tau)$. We can easily see that $\tau(u) = e^{-F_\pi(s,u)}$ is a decreasing function with $\tau(u) = 0$ only if $F_\pi(s,u) = \infty$. Since by definition $F_\pi(s,u)$ can diverge only for $u \to s_i^*$ we conclude that only trajectories of the second kind, with $u_\tau^* = s_i^*$ until some $\tau_{s,i}^* \in [0,1)$, can meet our requirements for being optimal. Moreover, we can compute $\tau_{s,i}^*$ by integrating backward in time the solution from $\tau = 1$. We find that

$$\tau_{s,i}^* := \exp(-\lim\nolimits_{a_{\pi,i}(u - s_i^*) \to 0^+} |F_\pi(s,u)|), \tag{4.45}$$

where the above expression holds for both $a_{\pi,i} = 1$ and $a_{\pi,i} = -1$. Define the inverse function $F_{\pi,s}^{-1} : (\tau_{s,i}^{*}, 1] \to (s, s_i]$ of π on $(s, s_i]$:

$$F_{\pi,s}^{-1} := \left\{ F_{\pi,s}^{-1}(q), \, q \in \left[0, \log\left(1/\tau_{s,i}^{*}\right)\right) : \, F_{\pi,s}\left(F_{\pi,s}^{-1}(q)\right) = q \right\} \tag{4.46}$$

Then we can write the global solution to our Cauchy problem as

$$u_{\tau}^{*} := F_{\pi,s}^{-1}\left(\log\left(1/\tau\right)\right) \mathbb{I}_{\{\tau \in (\tau_{s,i}^{*}, 1]\}} + s_i^{*} \, \mathbb{I}_{\{\tau \in [0, \tau_{s,i}^{*}]\}}, \tag{4.47}$$

The same reasoning can be obviously applied to the case $a_{\pi,i} = 1$, with $\dot{u}_{\tau}^{*} > 0$ and u_{τ}^{*} increasing in τ. We remark that the homogeneity of the above solution depends critically on the integrability of $1/\left|\pi(u) - u\right|$ when $\left|u - s_i^{*}\right| \to 0$: if $\lim_{a_{\pi,i}(u - s_i^{*}) \to 0^-} \left|F_{\pi}(s, u)\right| = \infty$, then obviously $\tau_{s,i}^{*} = 0$, while $0 < \tau_{s,i}^{*} < 1$ otherwise.

A similar reasoning can be applied to the case $u_1^{*} = s_i^{\dagger}$. Let us again consider $u_{\tau}^{*} \in K_{\pi,i} \cup \partial K_{\pi,i}$, $a_{\pi,i} = -1$ and take $s = s_i^{\dagger}$ in Eq. (4.41). Here the picture is slightly more complex, since it also depends on the behavior of $\left|F_{\pi}(s, u)\right|$, $s < u$, as $s - s_i^{\dagger} \to 0^{+}$.

In general, if $\left|F_{\pi}(s, u)\right|$, $s < u$, diverges as $s - s_i^{\dagger} \to 0^{+}$ then it is clear that the only possible trajectory $u_{\tau}^{*} \in K_{\pi,i} \cup \partial K_{\pi,i}$ that ends in $s_i^{\dagger}$ is $u_{\tau}^{*} = s_i^{\dagger}$. Anyway, if $\left|F_{\pi}(s, u)\right|$ remains finite then we can have optimal trajectories that hit $s_i^{\dagger}$ at some time $\tau = t < 1$ and stay in $s_i^{\dagger}$ for the remaining $\tau \in [t, 1]$. This is equivalent to set $u_t^{*} = s_i^{\dagger}$ as boundary condition of the Cauchy Problem in Eq. (4.41), so that the implicit expression of the optimal trajectory is $F_{\pi}\left(s_i^{\dagger}, u_{\tau}^{*}\right) = \log(t) - \log(\tau)$, where $t \in [0, 1]$ is free parameter. Since the above expression is simply a shifted version of that for $u_1^* \in K_{\pi,i}$, with $s_i^\dagger$ on behalf of s, t/τ on behalf of τ and $\theta_i^* t$,

$$\theta_i^* := \exp\left(-\lim\nolimits_{a_{\pi,i}\left(u-s_i^*\right)\to 0^+} \lim\nolimits_{a_{\pi,i}\left(s_i^\dagger - s\right)\to 0^+} \left|F_\pi\left(s,u\right)\right|\right), \tag{4.48}$$

on behalf of $\tau_{s,i}^*$, we can proceed as in the case $u_1^* \in K_{\pi,i}$ to find that

$$u_\tau^* := s_i^\dagger \mathbb{I}_{\{\tau\in(t,1]\}} + F_{\pi,s}^{-1}\left(\log\left(t/\tau\right)\right)\mathbb{I}_{\left\{\tau\in\left(\theta_i^* t,t\right]\right\}} + s_i^* \,\mathbb{I}_{\left\{\tau\in\left[0,\theta_i^* t\right]\right\}}. \tag{4.49}$$

It only remains to show that there is no solution to the Cauchy Problem in Eq. (4.41) for boundary conditions $u_1^* \in K_{\pi,0} \cup K_{\pi,|\partial C_\pi|}$. Let consider $K_{\pi,0}$, for which always we have $a_{\pi,0} = 1$ (the same result for $K_{\pi,|\partial C_\pi|}$ can be obtained by a similar reasoning). Since if $K_{\pi,0} \neq \varnothing$, then $\pi\left(0\right) > 0$ and in this case $s_0^\dagger = 0$ is not a zero-cost trajectory. Then, u_τ^* should increase from some $u_0^* < u_1^*$ to some $u_1^* < s_0^*$, but the general form of the Cauchy Problem in Eq. (4.41) rules out this possibility. We conclude that no trajectory $\varphi_\tau^* = \tau u_\tau^*$, $u_1^* \in K_{\pi,0}$ such that $I\left[\varphi^*\right] = 0$ exists, and by Lemma 16 this implies that $I\left[\varphi\right] > 0$ for every $\varphi_\tau = \tau u_\tau$ with $u_1 \in K_{\pi,0}$ as stated in Corollary 4.

□

4.3 Cumulant Generating Function

In this section we use conditional expectations and Picard-Lindelof theorem to get a non-linear Cauchy problem for $\psi\left(\lambda\right)$. Since the arguments are quite standard, we won't indulge in details except this is necessary. Then, let define the CGF scaling up to time n

$$\psi_n(\lambda) := \log \mathbb{E}\left(e^{\lambda X_{n,n}}\right), \ \lambda \in (-\infty, \infty), \tag{4.50}$$

so that $\psi(\lambda) := \lim_n \psi_n(\lambda)$. Hereafter we denote by $\mathbb{P}_\lambda(X_n = X)$ the tilted measure $\mathbb{P}_\lambda(X_{n,n} = X) := \exp[\lambda X - n\psi_n(\lambda)]\,\mathbb{P}(X_{n,n} = X)$, and by $\mathbb{E}_\lambda$ the tilted expectation. First we prove some trivial properties for $\psi_n(\lambda)$.

Lemma 17. *Let $\psi_n(\lambda)$ in Eq. (4.50), and define*

$$\gamma_n(\lambda) := (n+1)\left[\psi_{n+1}(\lambda) - \psi_n(\lambda)\right], \tag{4.51}$$

then $|\psi_n(\lambda)| \leq \lambda$, $\partial_\lambda \psi_n(\lambda) \in [0,1]$ and $|\gamma_n(\lambda)| \leq 2|\lambda|$ for all $n \in \mathbb{N}$, $\lambda \in \mathbb{R}$.

Proof. That $|\psi_n(\lambda)| \leq \lambda$ follows directly from definitions: since $0 \leq X_{n,n} \leq n$, then obviously $n^{-1}\left|\log \mathbb{E}\left(e^{\lambda X_{n,n}}\right)\right| \leq |\lambda|$. Similarly, $\partial_\lambda \psi_n(\lambda) = n^{-1}\mathbb{E}_\lambda(X_{n,n})$, hence $\partial_\lambda \psi_n(\lambda) \in [0,1]$. We shall now find a recursive relation for the Moment Generating Function $\mathbb{E}\left(e^{\lambda X_{n,n}}\right)$. Consider the conditional expectation $\mathbb{E}\left(e^{\lambda X_{n+1,n+1}}|\mathcal{F}_n\right)$: from Eq. (2.1) it's quite easy to check the Moment Generating Function obeys the following recursion rule:

$$\mathbb{E}\left(e^{\lambda X_{n,k+1}}\right) - \mathbb{E}\left(e^{\lambda X_{n,k}}\right) = \left(e^\lambda - 1\right)\mathbb{E}\left[\pi(x_{n,k})\,e^{\lambda X_{n,k}}\right]. \tag{4.52}$$

After few manipulations we can write the above relation as

$$\gamma_n(\lambda) = -\psi_n(\lambda) + \log\left\{1 + \left(e^\lambda - 1\right)\mathbb{E}_\lambda\left[\pi(x_{n,n})\right]\right\}, \tag{4.53}$$

Since by definition $\pi(x) \in [0,1]$, then $\mathbb{E}\left[\pi(x_{n,n}) e^{\lambda X_{n,n}}\right] \leq \mathbb{E}\left(e^{\lambda X_{n,n}}\right)$ and $\mathbb{E}_\lambda\left[\pi(x_{n,n})\right] \in [0,1]$, so that $|\gamma_n(\lambda)|$ can be bounded as

$$|\gamma_n(\lambda)| \leq |\psi_n(\lambda)| + \left|\log\left(1 + \left|e^\lambda - 1\right|\right)\right| \leq 2|\lambda|, \tag{4.54}$$

which completes the proof.

□

From last relation we found that $\lim_n |\psi_{n+1}(\lambda) - \psi_n(\lambda)| = 0$, but this is not enough to state whether $\lim_n \gamma_n(\lambda) = 0$ for every $\lambda \in \mathbb{R}$. Before presenting our proof we still need the following lemma

Lemma 18. *Let $\{f_n,\, n \in \mathbb{N}\}$ be a bounded real sequence. Then, the sequence $g_n := (n+1)(f_{n+1} - f_n)$ either converges to 0 or does not converge.*

Proof. Let suppose that g_n converges to some $g > 0$. Then $h > 0$ and $\epsilon > 0$ exist such that $0 < \epsilon \leq g_n$ for $n \geq h$. Follows that $f_n \geq \epsilon \sum_{k=h}^{n-1}(k+1)^{-1} + f_h$ would diverge for $n \to \infty$, which contradicts that f_n is bounded. A similar reasoning taking $g < 0$ will lead to the conclusion that g can be neither strictly positive nor strictly negative, hence we must have $g = 0$.

□

4.3.1 Proof of Theorem 7

Before starting we remark that even if the the statement of Theorem 7 asks for some additional properties for $\pi \in \mathcal{U}$, the first part of this proof, devoted to obtain the implicit ODE (3.25), does not.

Proof. Lemma 18 implies that if both $\lim_n \psi(\lambda)$ and $\lim_n \mathbb{E}_\lambda[\pi(x_{n,n})]$ exist, then we would have $\lim_n \gamma_n(\lambda) = 0$. The existence of $\psi(\lambda)$ follows from Theorem 3, while, since π is continuous and bounded, that of $\lim_n \mathbb{E}_\lambda[\pi(x_{n,n})]$ follows from weak convergence. Moreover, since $\psi \in \mathcal{AC}$ by definition of CGF, weak convergence also imply that

$$\lim_{n\to\infty} \mathbb{E}_\lambda[\pi(x_{n,n})] = \pi(\lim_{n\to\infty} \mathbb{E}_\lambda(x_{n,n})) = \pi(\partial_\lambda\psi(\lambda)). \tag{4.55}$$

Hence, from the above relations and by Lemma 18 we obtain the following non linear implicit ODE for ψ:

$$\psi(\lambda) = \log\left[1 + \left(e^\lambda - 1\right)\pi(\partial_\lambda\psi(\lambda))\right]. \tag{4.56}$$

The above ODE holds for every $\pi \in \mathcal{U}$, but its explicitation obviously require that π is invertible at least in the co-domain of $\partial_\lambda\psi(\lambda)$. By Corollary 4 we know that $\partial_\lambda\psi(\lambda) \in [0, \inf C_\pi)$ for $\lambda \in (-\infty, 0]$ and $\partial_\lambda\psi(\lambda) \in (\sup C_\pi, 1]$ for $\lambda \in [0, \infty)$, then we can restrict our invertibility requirements to those domains. Notice that since for $\lambda \in [0, \infty)$

$$\begin{aligned}\inf\nolimits_\lambda\{\partial_\lambda\psi(\lambda)\} \leq \frac{e^{\lambda\cdot\inf_\lambda\{\partial_\lambda\psi(\lambda)\}} - 1}{e^\lambda - 1} \leq \frac{e^{\psi(\lambda)} - 1}{e^\lambda - 1} \leq \\ \leq \frac{e^{\lambda\cdot\sup_\lambda\{\partial_\lambda\psi(\lambda)\}} - 1}{e^\lambda - 1} \leq \sup\nolimits_\lambda\{\partial_\lambda\psi(\lambda)\},\end{aligned} \tag{4.57}$$

then also $\left(e^{\psi(\lambda)}-1\right)/\left(e^{\lambda}-1\right)$ has co-domain $(\sup C_\pi, 1]$. Similarly, for $\lambda \in (-\infty, 0]$, we find a co-domain $[0, \inf C_\pi)$ as for $\partial_\lambda \psi(\lambda)$.

Let $\pi \in \mathcal{U}$ be an invertible function on $[0, \inf C_\pi)$, as required by the statement of Theorem 7, and denote by $\pi_-^{-1} : [\pi(0), \pi(\inf C_\pi)) \to [0, \inf C_\pi)$ its inverse. Moreover, let $\psi_-\left(\lambda_-^*\right) = \psi_-^*$ for some $\lambda_-^* \in (-\infty, 0)$. Then, $\psi(\lambda) = \psi_-(\lambda)$, with $\psi_-(\lambda)$ solution to the Cauchy problem

$$\partial_\lambda \psi_-(\lambda) = \pi_-^{-1}\left(\tfrac{e^{\psi_-(\lambda)}-1}{e^{\lambda}-1}\right), \; \psi_-\left(\lambda_-^*\right) = \psi_-^*, \tag{4.58}$$

If $\pi_-^{-1} \in \mathcal{AC}$ and Lipschitz, then we can apply the Picard-Lindelof theorem, which ensure the existence and uniqueness of ψ_- for any $\lambda \in (-\infty, 0)$. Notice that we have to discard $\lambda = 0$ and $\lambda = \infty$ since for those point the Lipschitz continuity in ψ required by the Picard-Lindelof theorem is not fulfilled. The same proceeding can be applied to the case $\lambda \in (0, \infty)$: let $\pi_+^{-1} : (\pi(\sup C_\pi), \pi(1)] \to (\sup C_\pi, 1]$ the inverse of π on $(\sup C_\pi, 1]$, let $\pi_+^{-1} \in \mathcal{AC}$ and Lipschitz, then for $\lambda \in (0, \infty)$ we have $\psi(\lambda) = \psi_+(\lambda)$, with $\psi_+(\lambda)$ solution to the Cauchy problem

$$\partial_\lambda \psi_+(\lambda) = \pi_+^{-1}\left(\tfrac{e^{\psi_+(\lambda)}-1}{e^{\lambda}-1}\right), \; \psi_+\left(\lambda_+^*\right) = \psi_+^*, \tag{4.59}$$

and this completes our proof. Finally, that $\partial_\lambda \psi(\lambda)$ is continuous comes from the fact that both $\pi_\pm^{-1}$ and $\left(e^{\psi_\pm(\lambda)}-1\right)/\left(e^{\lambda}-1\right)$ are continuous functions by definitions.

$\square$

4.3.2 Linear urn functions

The last goal of this section is the proof of Corollary 10, which gives the shape of ψ in case for $s \in [0, \inf C_\pi] \cup [\sup C_\pi, 1]$ the urn function π is a linear function.

Proof. Let $\pi(s) = a + bs$ for $s \in [0, a/(1-b)]$, and let $\pi(s) = c + ds$ for $s \in [c/(1-d), 1]$. To ensure that $\pi(s) \in [0,1]$ in both intervals, we need that $a \in [0,1]$, $b \in (-\infty, 1-a]$, $c + d \in [0,1]$, $a \in [0, \infty)$. Given these conditions, let first consider the case $s \in [c/(1-d), 1]$, which gives the ψ for $\lambda > 0$. Then, let $\pi(s) = c + ds$, so that the ODE to solve is

$$c + d\,\partial_\lambda \psi(\lambda) = \frac{e^{\psi(\lambda)} - 1}{e^\lambda - 1}. \tag{4.60}$$

We use the transformations $y(z(\lambda)) = e^{-\psi(\lambda)} - 1$, $z(\lambda) = 1 - e^{-\lambda}$, so that for $\lambda \in [0, \infty)$ we have $\psi(\lambda(z)) = -\log(1 + y(z))$, $\lambda(z) = -\log(1-z)$ and

$$\partial_z y(z) = \left[\tfrac{c}{d(1-z)} + \tfrac{1}{dz}\right] y(z) + \left[\tfrac{c}{d(1-z)}\right], \tag{4.61}$$

with $z \in [0,1]$. By Laplace method, we can rewrite the above equation as

$$\partial_z \left[y(-z)(1-z)^{\frac{c}{d}} z^{-\frac{1}{d}}\right] = \tfrac{c}{d}(1-z)^{\frac{c}{d}-1} z^{-\frac{1}{d}}. \tag{4.62}$$

Then, we define the function

$$B(\alpha, \beta; x_1, x_2) = \int_{x_1}^{x_2} dt\ (1-t)^{\alpha-1} t^{\beta-1}. \tag{4.63}$$

If $d > 0$, since $a \geq 0$ we have that $(1-z)^{\frac{a}{b}} z^{-\frac{1}{b}}$ is regular at $z = 1$, then

$$y\left(z; d>0\right) = \left(1-z\right)^{-\frac{c}{d}} z^{\frac{1}{d}} \left[K_1^* - \tfrac{c}{d} B\left(\tfrac{c}{d}, \tfrac{d-1}{d}; z, 1\right)\right], \tag{4.64}$$

where K_1^* defends on the initial conditions. Since when $\lambda \to \infty$ we must have $\partial_\lambda \psi\left(\lambda\right) \to 1$, from Eq. (4.60) we can write $\lim_{z\to 1} y\left(z; d>0\right) = -1$. Then, it can be shown that

$$\lim_{z\to 1} \left(1-z\right)^{-\frac{c}{d}} z^{\frac{1}{d}} B\left(\tfrac{c}{d}, \tfrac{d-1}{d}; z, 1\right) = \frac{d}{c}. \tag{4.65}$$

It follows that $K_1^* = 0$, and substituting $y\left(z\left(\lambda\right)\right) = e^{-\psi(\lambda)} - 1$, $z\left(\lambda\right) = 1 - e^{-\lambda}$ we find the following expression for $\lambda > 0$, $d > 0$

$$e^{-\psi_+(\lambda; d>0)} = \tfrac{c}{d} e^{-\frac{c}{d}\lambda} \left(1 - e^{-\lambda}\right)^{\frac{1}{d}} \left|B\left(\tfrac{c}{d}, \tfrac{d-1}{d}; 1-e^{-\lambda}, 1\right)\right| \tag{4.66}$$

If $d < 0$, we have instead that $(1-z)^{\frac{a}{d}} z^{-\frac{1}{d}}$ is regular at $z = 0$ and we take

$$y\left(z; d<0\right) = \left(1-z\right)^{-\frac{c}{d}} z^{\frac{1}{d}} \left[K_2^* + \tfrac{c}{d} B\left(\tfrac{c}{d}, \tfrac{d-1}{d}; 0, z\right)\right]. \tag{4.67}$$

This time we use $\lim_{z\to 0} y\left(z; d<0\right)/z = -\pi\left(c/\left(1-d\right)\right) = -c/\left(1-d\right)$ and

$$\lim_{z\to 0} \left(1-z\right)^{-\frac{c}{d}} z^{\frac{1}{d}-1} B\left(\tfrac{c}{d}, \tfrac{d-1}{d}; z, 1\right) = -\frac{d}{1-d} \tag{4.68}$$

to find that $K_2^* = 0$. Substituting as before we get the ψ for $\lambda > 0$ and $d > 0$:

$$e^{-\psi_+(\lambda; d<0)} = \tfrac{c}{d} e^{-\frac{c}{d}\lambda} \left(1 - e^{-\lambda}\right)^{\frac{1}{d}} \left|B\left(\tfrac{c}{d}, \tfrac{d-1}{d}; 0, 1-e^{-\lambda}\right)\right| \tag{4.69}$$

Then, let consider the case $s \in [0, a/(1-b)]$, which means $\lambda < 0$: this time we use take $y'(z'(\lambda)) = e^{\psi(\lambda)} - 1$ and $z'(\lambda) = 1 - e^{\lambda}$, so that again $z' \in [0,1]$. We can directly use the previous results for $\lambda > 0$ by applying the transformations $y(z) = -y'(z')/[1 + y'(z')]$ and $z = -z'/1 - z'$. Substituting in Eq. (4.61) and using Laplace method we find

$$\partial_z \left[\tfrac{y'(z')}{1+y'(z')} \left(1 - z'\right)^{\frac{1-a}{b}} \left(z'\right)^{-\frac{1}{b}} \right] = \tfrac{a}{b} \left(1 - z'\right)^{\frac{1-a}{b} - 1} \left(z'\right)^{-\frac{1}{b}} . \tag{4.70}$$

Again, since $a \in [0,1]$, for $b > 0$ the therm $(1 - z')^{\frac{1-a}{b}} (z')^{-\frac{1}{b}}$ is regular at $z' = 1$, then we take

$$\frac{y'(z'; b > 0)}{1 + y'(z'; b > 0)} = \left(1 - z'\right)^{-\frac{1-a}{b} - 1} \left(z'\right)^{\frac{1}{b}} \left[K_3^* - \tfrac{a}{b} B\left(\tfrac{1-a}{b}, \tfrac{b-1}{b}; z', 1\right)\right] \tag{4.71}$$

and use $\lim_{z' \to 1} y'(z'; b > 0) = -\pi(0) = -a$ and

$$\lim_{z' \to 1} \left(1 - z'\right)^{-\frac{1-a}{b} - 1} \left(z'\right)^{\frac{1}{b}} B\left(\tfrac{1-a}{b}, \tfrac{b-1}{b}; z', 1\right) = -\frac{b}{1-a} \tag{4.72}$$

to find that, again, $K_3^* = 0$. Substituting $y'(z'(\lambda)) = e^{\psi(\lambda)} - 1$ and $z'(\lambda) = 1 - e^{\lambda}$, for $\lambda < 0$, $b > 0$ we find

$$e^{-\psi_-(\lambda; b>0)} - 1 = \tfrac{a}{b} e^{-\frac{1-a+b}{b}\lambda} \left(1 - e^{\lambda}\right)^{\frac{1}{b}} \left|B\left(\tfrac{1-a}{b}, \tfrac{b-1}{b}; 1 - e^{\lambda}, 1\right)\right| \tag{4.73}$$

Finally, if $b < 0$ we can write down our solution as

$$\frac{y'(z'; b < 0)}{1 + y'(z'; b < 0)} = \left(1 - z'\right)^{-\frac{1-a}{b} - 1} \left(z'\right)^{\frac{1}{b}} \left[K_4^* + \tfrac{a}{b} B\left(\tfrac{1-a}{b}, \tfrac{b-1}{b}; 0, z'\right)\right] . \tag{4.74}$$

Then, from $\lim_{z \to 0} y\left(z'; b<0\right)/z' = -a/\left(1-b\right)$ and

$$\lim_{z' \to 0} \left(1-z'\right)^{-\frac{1-a}{b}-1} \left(z'\right)^{\frac{1}{b}-1} B\left(\tfrac{1-a}{b}, \tfrac{b-1}{b}; z', 1\right) = -\frac{b}{1-b} \tag{4.75}$$

we find that also the last constant is $K_4^* = 0$, and that

$$e^{-\psi_-(\lambda; b<0)} - 1 = \tfrac{a}{b} e^{-\frac{1-a+b}{b}\lambda} \left(1-e^{\lambda}\right)^{\frac{1}{b}} \left|B\left(\tfrac{1-a}{b}, \tfrac{b-1}{b}; 0, 1-e^{\lambda}\right)\right|. \tag{4.76}$$

This completes the proof. Notice that the boundary conditions we used to compute ψ fall outside the statement of Theorem 7, which requires the knowledge of $\psi\left(\lambda_\pm^*\right)$ for some finite $\lambda_\pm^* \neq 0$. The fact that our solution can be determined by boundary conditions at $\lambda \to 0$ and $\lambda \to \pm\infty$ is a special property of the linear urn, and can't be generalized to arbitrary urn functions.

$\square$

We remark that in the above proof the cases $b = 0$ and $d = 0$ are not considered, since we would get a Bernoulli process whose ϕ can be trivially computed by elementary techniques. Anyway, taking the limit $b \to 0$ in the above expressions should return the desired result.

Acknowledgments

I am grateful to Pietro Caputo (Università degli Studi Roma 3) for his critical help in preparing this work, and for teaching me what I know about Probability Theory. I would also thank Giorgio Parisi (Sapienza Università di Roma) and Riccardo Balzan (Université Paris-Descartes) for interesting discussions and suggestions.

References

[AEK1983] W.B. Arthur, Yu.M. Ermoliev and Yu.M. Kaniovski, *The Generalized Urn Problem and Its Application*, Kibernetika No.1 (1983), 49-56.

[AEK1986] W. B. Arthur, Y. Ermoliev, M. Kaniovsky, *Strong laws for a class of path-dependent stochastic processes with applications*, Stochastic Optimization Lecture Notes in Control and Information Sciences 81 (1986), 287-300.

[AEK1986b] W. B. Arthur, Y. Ermoliev, M. Kaniovsky, *Limit Theorems for Proportions of Balls in a Generalized Urn Scheme*, IIASA Working Paper WP-87-111 (1987).

[AEK1987] W. B. Arthur, Y. Ermoliev, M. Kaniovsky, *Non-Linear Urn Processes: Asymptotic Behavior and Applications*, IIASA Working Paper WP-87-085 (1987).

[AEK1987b] B. Arthur, Y. Ermoliev, and Y. Kaniovskii, *Path dependent processes and the emergence of macro-structure.* Eur. J. Oper. Res., 30 (1987), 294–303.

[BP1983] A. Bagchi, A. K. Pal, *Asymptotic Normality in the Generalized Polya–Eggenberger Urn Model, with an Application to Computer Data Structures*, SIAM. J. on Algebraic and Discrete Methods, 6 (1983), 394–405.

[CL2009] C. Cotar and V. Limic *Attraction time for strongly reinforced random walks*, Ann. Appl. Probab. Volume 19, Number 5 (2009), 1972-2007.

[DEK1994] G. Dosi, Y. Ermoliev, Y. Kaniovski, *Generalized urn schemes and technological dynamics*, Journal of Mathematical Economics 23 (1994), 1-19.

[DZ1998] A. Dembo, O. Zeitouni, *Large deviation techniques and applications* (Springer, New York, 1998).

[DFM2002] E. Drinea, A. Frieze, and M. Mitzenmacher, *Balls in bins processes with feedback* In Proceedings of the 11th Annual ACM-SIAM Symposium on Discrete Algorithms, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, (2002), 308–315.

[EP1923] F. Eggenberger, G. Pólya, *Uber die statistik verketetter vorgage* Zeitschrift fur Angewandte Mathematik und Mechanik, 1, (1923), 279–289.

[FGP2005] P. Fajolet, J. Gabarro, H. Pekari, *Analytic Urns,* The Annals of Probability 33 (2005), 1200-1233.

[FP2005] P. Flajolet and V. Puyhaubert, *Analytic combinatorics at OK Corral*, Technical memorandum, unpublished (2005).

[FDP2006] P. Flajolet, P. Dumas, and V. Puyhaubert, *Some exactly solvable models of urn process theory*, Fourth Colloquium on Mathematics and Computer Science, DMTCS proc. AG (2006), 59-118.

[Fra2011] S. Franchini, *Ideal Chains with fixed self intersections*, Phys. Rev. E 84, 051104 (2011).

[Fra2011Th] Simone Franchini, *Catene Ideali con numero fissato di intersezioni*, MS thesis, Sapienza Università di Roma, (2011).

[Fra2014] S. Franchini, *Large deviations for Generalized Polya Urns with general Urn Functions*, in preparation (2014).

[FB2014] S. Franchini, R. Balzan, *Self-intesection distribution in Simple Random Walks by atmospheric methods*, in preparation (2014).

[HK2001] Y. Hamana and H. Kesten, *A large deviation result for the range of random walk and for the Wiener sausage*, Probab. Th. Rel. F., 120 (2001), 183-208.

[HKP2007] H.-K. Hwang, M. Kuba, A. Panholzer, *Analysis of some exactly solvable diminishing urn models*, Formal Power Series and Algebraic Cominatorics, Nakai University, Tianjin, China (2007).

[HLS1980] B. M. Hill, D. Lane, W. Sudderth, *A Strong Law for some Generalized Urn Processes*, The Annals of Probability 8 (1980), 214-226.

[JK1977] L. Johnson, S. Kotz, *Urn models and their application*, Wiley (1977).

[KB1997] S. Kotz and N. Balakrishnan, *Advances in urn models during the past two decades*, Advances in combinatorial methods and applications to probability and statistics, Stat. Ind. Technol., Birkhauser Boston, Boston (1997), 203-257.

[KK2001] K. Khanin and R. Khanin. *A probabilistic model for establishment of neuron polarity*, J. Math. Biol., 42 (2001), 26–40.

[KMR2000] S. Kotz, H. Mahmoud, and P. Robert, *On generalized Polya urn models*, Statistics & Probability Letters 49, 2 (2000), 163-173.

[Mam2003] H. M. Mahmoud, *Polya Urn Models and Connections to Random Trees: A Review*, Journal of the Iranian Statistical Society 2 (2003), 53-114.

[Mam2008] H .M . Mahmoud, *Polya Urn Models*, Taylor & Francis (2008).

[MM2012] B. Morcrette, H. Mahmoud, *Exactly solvable balanced tenable urns with random entries via the analytic methodology*, Discrete Mathematics and Theoretical Computer Science, proc. AQ (2012), 219-232.

[Oliv2008] R. Oliveira, *Balls-in-bins processes with feedback and brownian motion*, Journal Combinatorics, Probability and Computing archive, Volume 17 Issue 1, (2008), 87-110.

[Pem1991] R. Pemantle, *When are Touchpoints Limits for Generalized Polya Urns?*, Proceedings of the American Mathematical Society 113 (1991), 235-243.

[Pem2007] R. Pemantle, *A survey of random processes with reinforcement*, Probability Surveys 4 (2007), 1-79.

[VBD2001] M. van den Berg, E Bolthausen, F den Hollander, *Moderate deviations for the volume of the Wiener sausage* , Annals of Mathematics, 153, (2001), 355-406.