Gaussian Phases Toward Statistical Equilibrium in Some Urn Models

by Tong Zhang

B.S., July 2006, University of Science and Technology of China, China

A dissertation submitted to

The Faculty of
Columbian College of Arts and Sciences
of the George Washington University
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

August 31, 2011

Dissertation directed by

Hosam Mahmoud
Professor of Statistics

The Columbian College of Arts and Sciences of The George Washington University certifies that Tong Zhang has passed the Final Examination for the degree of Doctor of Philosophy as of June 8th, 2011. This is the final and approved form of the dissertation.

Gaussian Phases Toward Statistical Equilibrium
in Some Urn Models

Tong Zhang

Hosam Mahmoud, Professor of Statistics, Dissertation Director

Joseph Gastwirth, Professor of Statistics, Committee Member

Subrata Kundu, Associate Professor of Statistics, Committee Member

Dedication

To my wife, Jiayu, for her love and support.

Acknowledgments

I would like to thank, with deepest gratitude, my advisor and mentor, Dr. Hosam Mahmoud. The completion of this dissertation is the most rewarding adventure in my life, and his encouragement two years ago started it. Since then, I have gained countless help from him. His knowledge, and his fearlessness in pursuing the knowledge lead us through number of difficulties. His humorous made this journey a fun memory. I owe him a lot.

I also want to thank my committee members. Dr. Gastwirth is nice and caring for my study all the years. Dr. Kundu is friendly and helpful. Dr. Ullman read the thesis thoroughly and gave many advices and suggestions. He is a true scholar. Special thanks to Dr. Hitczenko, who came a long way from Drexel. He is an expert in this field, and his advices become lines of future studies. They are all very nice and kind persons. I could not get a better committee.

Words can not describe how much I want to thank my parents, who have always been my final resort to console. Their support is without reservation. I feel so lucky to have them as parents and have been missing them a lot. At the the same time, I have had the Department of Statistics at GWU like a family here in the US. Dr. Modarres gave me lot of advices both on study and on life. Dr. Balaji is the kindest professor you can work with....The list is very long and I am filled with thankfulness. I appreciate all the professors and my fellow graduate students for sharing precious years of time with me, and I thank my friends around the world. I wish I could tell you all how your simple encouragement inspired me.

Abstract

Gaussian Phases Toward Statistical Equilibrium in Some Urn Models

We are interested in a group of urn models with mean reverting property, that is, they reach statistical equilibrium, usually at the mean, when scale space and/or time reach infinity. Both Ehrenfest and Bernoulli-Laplace urns are within this category. We analyze these urns in different phases (*sublinear*, *linear* and *superlinear*), along which the models approach equilibrium, and try to identify and explain the cutoff phenomenon (Diaconis, 1996) and the changes at the "seamlines" between phases (Balaji, Mahmoud and Zhang, 2010). Essentially, we want to investigate the probability laws of all the embedded phases. Analyzing the phases will help us further understand the nature of these models.

We introduce a generalization of both the Ehrenfest and Bernoulli-Laplace urns, in which samples of independent identically distributed random sizes with a general generating discrete distribution (the generator) are taken out of an urn (Ehrenfest) or urns (Bernoulli-Laplace) at each time epoch. Generators with the same mean and variance exert the same influence on the phases. The influence of the generators gets attenuated from the sublinear phase to the linear phase and disappears at the superlinear phase.

This methodology can be generalized to study other similar random structures and is left for future developments.

Contents

List of Tables

Chapter 1

Introduction

1.1 Background I—Urn models

Urns, like coins, dice, card packs and chessboards, are stochastic devices producing random results, and are tools to understand, model and develop probability theory. As one of the most popular topics in probability theory, urn models are found in various applications in fields like Computer Science (data structures (Bagchi and Pal, 1985)), Economics (market shares (Arthur, Ermoliev and Kaniovski, 1987)), Political Sciences (voting strategies (Berg, 1985)), Law (business schemes (Gastwirth and Bhattacharya, 1983)), Public Health (clinical trials (Wei, 1978)), and Biology (evolution (Ewens, 1972)), etc. For more background and applications of urns, we refer the reader to two comprehensive books on this topic (Johnson and Kotz, 1977 and Mahmoud, 2008).

A typical urn model has an urn containing balls of different colors, the number of colors can be finite or countably infinite. At every time epoch $j = 1, 2, 3, \ldots$, a ball is randomly picked from the urn, the color is noted and the content of the urn is then changed depending on the color. The probability of choosing a specific color of ball is equal to the proportion of that color in the urn. Variations on the basic urn model can be made to accommodate for more than one urn, and/or more than one ball to be picked at each time, etc.

We are interested in a group of mixing models which can be constructed using urn models and variations therein, including the Ehrenfest urn and the Bernoulli-Laplace urn.

1.1.1 The Ehrenfest urn

In the Ehrenfest urn (Ehrenfest and Ehrenfest, 1907 and Feller, 1968), we have one urn that contains two different colors, say white and red, of balls of total number n. At each discrete point in time, we pick a ball at random from the urn. We paint that ball with the opposite color and put it back in the urn. The process goes on with the next time point. Earlier studies of the number of white balls in the Ehrenfest urn consider a Markov model where the transition matrix $P = \{p_{ij}\}_{i,j=0}^{n}$ has all entries 0 except

$$p_{i,i-1} = \frac{i}{n}, \qquad p_{i,i+1} = \frac{n-i}{n}.$$

Thus, the underlying Markov chain has period 2 with no limiting distribution.

1.1.2 The Bernoulli-Laplace urn

In the parallel model of the Bernoulli-Laplace, we have two urns, urn A starts with n white balls and urn B starts with n red balls. At each time epoch, a single ball is drawn randomly from each urn and is transferred to the other urn. The process is repeated at the next time epoch, and so on. The earliest models of this type can be traced back to Bernoulli (1768 and 1769), Laplace (1812) and are discussed by Feller (1968).

Earlier studies often consider the Markov model of the number of white balls in urn A, which has transition matrix $P = \{p_{ij}\}_{i,j=0}^{n}$ with all entries 0 except

$$p_{i,i-1} = \left(\frac{i}{n}\right)^2, \qquad p_{i,i} = \frac{2i}{n}\left(\frac{n-i}{n}\right), \qquad p_{i,i+1} = \left(\frac{n-i}{n}\right)^2.$$

This underlying Markov chain is irreducible and aperiodic.

1.2 Background II—Statistical mechanics

Urn models are found in the literature of statistical mechanics, especially the Ehrenfest urn and the Bernoulli-Laplace urn, they serve as fundamental tools to explain the theory of thermodynamics (Ehrenfest and Ehrenfest, 1907).

The Ehrenfest urn is first proposed as a model for the diffusion of nonreacting gases (Ehrenfest and Ehrenfest, 1907). The model consists of two chambers (say A and B) containing gases (possibly the same). The two chambers are connected through a pipe controlled by a valve. The valve is opened at time 0 and the diffusion proceeds over epochs of time, which we can take as the unity. In each time unit (diffusion step) one molecule of gas randomly chosen from the population of molecules in both chambers jumps from its chamber to the other one. This continual switching of sides affects a gradual diffusion; inducing change in the amount of gas in each chamber. It is of interest to know the amount of gas (number of molecules) in chamber A after a certain period of time.

This physical model of gas diffusion can be visualized in terms of a scheme of drawing balls from an urn. We can think of the molecules in chamber A as balls of a certain color (say white) and those in chamber B as balls of an antithetical color (say red). The gas model with n molecules can then be viewed as n balls of two colors all residing in one urn, which evolves in the following manner. At each discrete point in time, we pick a ball at random from the urn. We paint that ball with the opposite color and put it back in the urn. In this equivalent model, the interest is to know the number of white balls (the amount of gas in chamber A) after a certain period of time.

In the context of Boltzmann's statistical theory of gases, the Ehrenfest urn is also regarded as a simple physics model for heat exchange between two isolated bodies, with the number of balls representing temperature (Kac, 1947). In the Ehrenfest model, the expected number of white balls in the urn after the kth step of mixing, W_k, is given by

$$\mathbf{E}[W_k] = \frac{n}{2} + \Big(1 - \frac{2}{n}\Big)^k \Big(W_0 - \frac{n}{2}\Big).$$

Let τ be the unit of time, $k\tau = t =$ constant time, the factor $(1 - \frac{2}{n})^k$ in the limit, as $k \to \infty$, $n \to \infty$, and $\tau \to \infty$, in such a way that $\frac{n\tau}{2} \to \gamma^{-1}$($\gamma$ being a constant), approaches the limit $e^{-\gamma t}$, which is consistent with Newton's law of cooling (Kac, 1947 and Seneta, 1982). In Seneta (1982), the author describes the connections between the Ehrenfest urn and the Bernoulli-Laplace urn and the notion of entropy. For an overview see Bingham (1991).

Chapter 2

Terminology

2.1 Asymptotic phases

The asymptotic behavior of urn models is extensively studied, especially for the Pólya urn. Take the Pólya-Eggenberger urn (Eggenberger and Pólya, 1923) and Friedman's urn for example, see Freedman (1965). Smythe (1996) defines the extended Pólya urn model and considers its asymptotic normality. For an overview, see Bai, Hu and Zhang (2002) and Mahmoud (2008). One common feature of these urn models is that the size of their population changes as the number of draws grows, and the relation between the number of draws k and the population in the urn n is deterministic by the mechanisms of the urns.

There is another type of Pólya urn including "Coupon collector's urn" (Feller 1968) and the "Ehrenfest urn" (Feller 1968), etc. These urns share one common character that the total number of balls in the urn, the population, stays constant despite that the number of draws increases. We call this type of urn "zero-balanced," for a recent study on the "zero-balanced" urn, see Kholfi and Mahmoud (2010). Previous asymptotic analysis of the zero-balanced urn focuses on the asymptotics of the urn as the number of draws $k \to \infty$.

There have been very few asymptotic studies on the zero-balanced urn that consider the cases when the population n goes to ∞, until Diaconis and Shahshahani (1987) and Diaconis (1988). Diaconis (1988) considers the Ehrenfest urn process, as $n \to \infty$, when the number of draws k is $\lfloor \frac{1}{4} n \ln n + cn \rfloor$, for any positive constant c. The study shows that asymptotically the Ehrenfest urn process approaches a stationary distribution with a Gaussian error function $\mathrm{Erf}(e^{-c/2}/\sqrt{8}\,)$, where $\mathrm{Erf}(z) = (2/\sqrt{\pi}\,) \int_0^z e^{-\frac{t^2}{2}} dt$, also see Voit (1996).

Recently, in a study of coupon collector's urn, Mahmoud (2010) introduced the concept of "phases" which specifies the relation between the number of draws k_n and the population of urn n, and conducts an asymptotic analysis, as $n \to \infty$, within each phase. This author shows that Gaussian distributions are obtained in properly defined "sublinear", "linear", and "superlinear" phases of k_n via a martingale approach.

We will show that the idea of phases, as a well defined concept of asymptotic analysis in certain zero-balanced urns, provides an explanation to the cut-off phenomenon in the Markovian approach of the Ehrenfest urn and the Bernoulli-Laplace urn process (Diaconis and Shahshahani, 1987 and Diaconis, 1988). Also, in the field of statistical mechanics, "Newton's law of cooling" is a concrete example of the Ehrenfest urn in the linear phase, we hope analyzing the probability laws in phases can provide more insightful understanding of these phenomenons.

2.2 Randomization

Randomness in Pólya urn models and their variations are well studied too. Athreya and Karlin (1968) first considered the asymptotic properties of the Generalized Friedman urn model, where the reinforcements are random.

For the zero-balanced urn, since there is no reinforcement into the urn after each draw, we consider the randomness in the size of the sample of balls for each draw and call it random sampling. Selke (1995) investigates the number of independent and identically distributed (i.i.d.) random samples needed to see all the balls in the coupon collector's urn. See Kobza (2007) for a survey of random sample sizes in coupon collector's problem.

We introduce the idea of random sampling to the Ehrenfest urn and the Bernoulli-Laplace urn. The sizes of the draws are i.i.d. random variables instead of being a constant. We call the sampling "*subcritical*" when the range of the random sample size is relatively small compared to the population size, and "*critical*," when that is significant in comparison to the population size; we shall make these notions precise when we need them.

2.3 A martingale approach

Because of the dependent structure of urn models, martingales have been one of the most powerful and popular tools to analyze urn problems (other customary methods include the method of moments (Hwang, Kuba, and Panholzer, 2007), analytical methods (Flajolet, Dumas, and Puyhaubert, 2006), Stein's method (Chatterjee, Diaconis, and Meches, 2005). Ultimately, we want to study the probability laws of the phases of the urn models. Often, we rely on the martingale central limit theorem. Sufficient conditions for the central limit theorem for a zero-mean martingale $X_{j,n}$ are conditional Lindeberg's condition and the conditional variance condition on the martingale differences $\nabla X_{j,k_n} = X_{j,k_n} - X_{j-1,k_n}$; (see Theorem 3.2 and Corollary 3.1, P. 58 in Hall and Heyde, 1980).

Conditional Lindeberg's condition requires that, for all $\varepsilon > 0$,

$$U_n := \sum_{j=1}^{k_n} \mathbf{E}\Big[\Big(\nabla X_{j,k_n}\Big)^2 \mathbf{1}_{\Big\{\Big|\nabla X_{j,k_n}\Big| > \varepsilon\Big\}} \Big| \mathcal{F}_{j-1}\Big] \xrightarrow{\mathcal{P}} 0,$$

where the indicator $\mathbf{1}_{\mathcal{E}}$ is a function of a sample space that assumes the value 1 if $\mathcal{E}$ occurs, and otherwise it assumes the value 0, and, for a constant c, a c–conditional variance condition requires that

$$V_n := \sum_{j=1}^{k_n} \mathbf{E}\Big[\Big(\nabla X_{j,k_n}\Big)^2 \Big| \mathcal{F}_{j-1}\Big] \xrightarrow{\mathcal{P}} c. \tag{2.1}$$

When both conditions hold, the sum $\sum_{j=1}^{k_n} \nabla X_{j,k_n} = (X_{k_n} - X_0) = (X_{k_n} - X_0(n))$ converges to the normally distributed random variable $\mathcal{N}(0, c^2)$.

Chapter 3

Organization and notation

In this chapter we go over the remaining sections and over the terminology and various mathematical symbols used in this dissertation.

3.1 Organization

The main focus of this dissertation is the Gaussian phases of the Ehrenfest urn and the Bernoulli-Laplace urn. The rest of this dissertation has the following organization. In Chapter 4 we start the discussion with the Ehrenfest urn. In Chapter 5, we continue the discussion on the Generalized Ehrenfest urn. In Chapter 6, we move on to the Generalized Bernoulli-Laplace urn. In the last chapter, Chapter 7, we discuss future developments along this line of research.

Within each chapter, we will first give an introduction to the urn model focusing on the stochastic recurrence of the random structure. Then, we will derive the first few exact moments. The main theorem will be stated and the analysis of the Gaussian phases will be presented next. In the end, the critical region at the "seam-lines" between two conjoining phases will be discussed and examples will be shown to illustrate the main theorem.

3.2 Notation

The following notation will be used frequently in all the chapters. The notation $\mathrm{Bin}(n, p)$ stands for a binomial random variable on n trials with rate of success p per trial. Let $\mathrm{Hypergeo}(n, m, w)$ be a hypergeometric random variable that is the number of white balls in a sample of size m balls taken at random (all subsets of size m being equally likely) from an urn containing a total of n white and red balls, of which w are white. The mean and variance for this standard distribution are:

$$\begin{aligned} \mathbf{E}[\mathrm{Hypergeo}(n, m, w)] &= \frac{mw}{n}, \\ \mathbf{Var}[\mathrm{Hypergeo}(n, m, w)] &= \frac{mw(n-m)(n-w)}{n^2(n-1)}, \end{aligned}$$

see Stuart, Ord and Arnold (2009).

The normally distributed random variate with mean 0 and variance ν^2 is denoted by $\mathcal{N}(0, \nu^2)$, and the notation $\xrightarrow{\mathcal{D}}$ stands for convergence in distribution. We shall use the symbol $\xrightarrow{\mathcal{P}}$ for convergence in probability. The notation $o_{\mathcal{L}_1}(g(n))$ will stand for a sequence of random variables that is $O(g(n))$ in the $\mathcal{L}_1$ norm, that is there exist a positive constant C and a positive integer n_0, such that $\mathbf{E}[|X_n|] \le C|g(n)|$, for all $n \ge n_0$. We let $\mathcal{F}_j$ be the sigma field generated by $W_0, \ldots, W_j$. Note that this sequence of sigma fields is increasing. Thus, $\{\mathcal{F}_j\}_{j=0}^{\infty}$ can be the filtration of a martingale sequence. Unless otherwise stated, all asymptotic equivalents and bounds are taken as $n \to \infty$.

We shall also use the backward difference operator ∇, which when applied to a function $h(i)$, with integer argument i, gives the difference between two successive steps; that is, $\nabla h(i) = h(i) - h(i-1)$.

Further notation used in each chapter are restricted to that chapter.

Chapter 4

The Ehrenfest urn model

The Ehrenfest urn is a model for the diffusion of gases between two chambers. Classic research deals with this system as a Markovian model with a fixed number of balls, and derives the steady-state behavior as a binomial distribution (which can be approximated by a normal distribution). We study the gradual change for an urn containing n (a very large number) balls from the initial condition to the steady state. We look at the status of the urn after k_n draws. We identify three phases of k_n: The growing sublinear, the linear, and the superlinear. In the growing sublinear phase the amount of gas in either chamber is normally distributed, with parameters that are influenced by the initial conditions. In the linear phase a different normal distribution applies, in which the influence of the initial conditions is attenuated. The steady state is not a good approximation until a certain superlinear amount of time has elapsed. At the superlinear stage the mix is nearly perfect, with a nearly perfect symmetrical normal distribution in which the effect of the initial conditions is completely washed away. We give interpretations for how the results in different phases conjoin at the seam lines. In fact, these Gaussian phases are all manifestations of one master theorem. The results are obtained via martingale theory.

4.1 Introduction

The steady-state behavior of the classic Ehrenfest model has been studied by Bellman and Harris (1951), Blom (1989) and Karlin and McGregor (1965); for an overview see Mahmoud (2008).

Antognini (2005) looks at the speed of diffusion for a fixed number of particles. In a physical system the number of gas molecules is very large, we shall take it to be n, and is apportioned as $\lfloor \alpha n \rfloor \sim \alpha n$ in chamber A and $n - \lfloor \alpha n \rfloor \sim (1-\alpha)n$ in Chamber B, for some $\alpha \in (0, 1)$. One is interested in knowing the behavior of the gases after a certain finite interval of time. So, the question is *How many white balls are in the urn after $k = k_n$ draws, for functions k_n of various growth rates?* The study of the evolution of urns through various stages prior to the steady state is a topic of recent interest, see for example, Mikhailov (1977, 1980), Vatutin and Mikhailov (1982), Mahmoud (2010) , Smythe (2011) and Mahmoud and Smythe (2011+).

4.2 Scope

We shall identify three phases of k_n:

(a) The sublinear phase, when $k_n = o(n)$;

(b) The linear phase, when $k_n = \lambda_n n$, for some $\lambda_n > 0$ of a magnitude separated from zero and infinity;

(c) The superlinear phase, when $n = o(k_n)$.

We shall prove the following general trends. Trivially, at the very low end of the sublinear phase, when $k_n = O(1)$, as $n \to \infty$, there is not much change in the content of the two chambers, only a finite perturbation on the initial conditions can be felt. Changes begin to happen, when k_n grows to infinity.

Theorem 1 *Let W_{k_n} be the number of white balls in the Ehrenfest urn (molecules in Chamber A) after k_n draws (gas diffusion steps) from an urn with n balls, of which initially the number of white balls is $W_0(n) = \lfloor \alpha n \rfloor$, where $k_n \to \infty$, in a sublinear, linear or superlinear fashion. Then,*

$$\frac{W_{k_n} - \Big(\frac{n}{2} + \Big(W_0(n) - \frac{n}{2}\Big)\Big(\frac{n-2}{n}\Big)^{k_n}\Big)}{\sqrt{\frac{1}{4}n + \Big(\Big(\frac{n}{2} - W_0(n)\Big)^2 - \frac{n}{4}\Big)\Big(\frac{n-4}{n}\Big)^{k_n} - \Big(\frac{n}{2} - W_0(n)\Big)^2\Big(\frac{n-2}{n}\Big)^{2k_n}}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

The shift and scale are the mean and standard deviation of W_{k_n}. This theorem has the following manifestations in various phases. (We use three different approximation techniques in each of the three phases.) When k_n grows sublinearly to ∞, one sees normal behavior in the amount of gas in each chamber, even for a fairly slowly growing function k_n. We call the phase when k_n grows sublinearly to ∞, the *growing sublinear phase*. Functions that are asymptotically as small as $\frac{1}{20}\ln\ln n$, for example, are sufficient to give a normally distributed mix in each chamber. For the sublinear phase, the initial conditions persist, and the asymptotic normal result in this case contains the initial condition α. The Gaussian law in Theorem 1 takes the form

$$\frac{W_{k_n} - n\Big(\frac{1}{2} + \Big(\alpha - \frac{1}{2}\Big)\Big(\frac{n-2}{n}\Big)^{k_n}\Big)}{\sqrt{k_n}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 4\alpha(1-\alpha)).$$

Normality continues to hold in the linear and superlinear phases. However, in each phase we get a different normal distribution; the mean and scale factors are essentially different. In the linear phase a different normal distribution (in the usual style of central limit theorems) is in effect, and the parameters of the distribution depend on both the initial condition α and the coefficient of linearity.

A typical instance of the linear phase is when $k_n = cn + o(\sqrt{n}\,)$, for a positive constant c, in which case the Gaussian law in Theorem 1 takes the form

$$\frac{W_{k_n} - \Bigl(\Bigl(\alpha - \dfrac{1}{2}\Bigr)e^{-2c} + \dfrac{1}{2}\Bigr)\, n}{\sqrt{\dfrac{e^{4c} - 1 - 4c(2\alpha - 1)^2}{4e^{4c}}\, n}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

Note how the influence of the initial conditions is attenuated as we get deeper (large c) in the linear phase.

As one might expect, after a very long period of time, as in the superlinear case, the diffusion is nearly complete. In the superlinear phase, and if additionally $k_n = \frac{1}{4} n \ln n + g_n$, for any function g_n such that $g_n/n \to \infty$, Theorem 1 takes the symmetric form

$$\frac{W_{k_n} - \dfrac{1}{2} n}{\sqrt{n}} \xrightarrow{\mathcal{D}} \mathcal{N}\Bigl(0, \frac{1}{4}\Bigr),$$

which is the usual approximation of the binomial distribution by the normal. Note also how the effect of any initial conditions is washed away.

Diaconis (1996) takes a different view and discusses the "cutoff phenomenon" in the Ehrenfest urn model, where the total variation distance to the stationary distribution experiences a sharp decline after a large number of draws (while keeping the size of the urn fixed).

The problem can also be viewed as an allocation scheme of balls in urns, where k_n balls are dropped randomly in n urns. A ball in the Ehrenfest urn is represented by an urn in the allocation scheme. As a ball changes color during the history of the Ehrenfest process, the number of balls in the corresponding urn in the allocation model changes parity. More specifically, suppose the balls in the Ehrenfest urn are labeled $1, \ldots, n$. We label the urns in the allocation scheme with $1, \ldots, n$, say from left to right, and the ith urn represents the ith ball in the Ehrenfest urn. The urns labeled $1, \ldots, W_0(n)$ are the initial white balls in the Ehrenfest model, and the urns labeled $W_0(n) + 1, \ldots, n$ are the initial red balls in the Ehrenfest model. After k_n ball drops, an urn among the $W_0(n)$ leftmost urns containing an even number of balls indicates that the corresponding ball in the Ehrenfest urn (initially white) has been drawn an even number of times, and it is now white; let the number of such urns be L_{k_n}. Likewise, an urn among the $n - W_0(n)$ rightmost urns containing an odd number of balls indicates that the corresponding ball in the Ehrenfest urn (initially red) has been drawn an odd number of times, and it is now white; let the number of such urns be R_{k_n}. Then,

$$W_{k_n} = L_{k_n} + R_{k_n}.$$

Allocation scheme formulations, such as this one, received quite a bit of attention, see for example Kolchin, Sevastyanov, and Chistyakov (1976), where the schemes where handled by the method of moments.

In this chapter, unless otherwise stated, all asymptotics will mean asymptotic equivalents and bounds as $n \to \infty$. The number $n/(n-2)$ will appear often, and we shall give it the designation ρ_n. We shall repeatedly use well-known facts about ρ_n^{yn}, for $y > 0$, such as the fact that ρ_n^{yn} is asymptotically $e^{2y} + O(1/n)$.

4.3 Exact moments

Let $W_j = W_j(n)$ be the number of white balls (molecules in chamber A) after j such draws (diffusion steps). Let I_n^W and I_n^R respectively be the indicators of picking a white or a red ball in the nth step. Because of their mutual exclusion, we have $I_n^R = 1 - I_n^W$. There is stochastic dependence between W_{j-1} and W_j. After $j - 1$ draws, the number of white balls in the urn is W_{j-1}, and the number of white balls will increase by 1 after one drawing, if a red ball is picked, but will decrease by 1, if a white ball is picked. And so,

$$W_j = W_{j-1} + I_n^R - I_n^W = W_{j-1} + 1 - 2I_n^W. \tag{4.1}$$

A recurrence for the mean follows from the expectation of the stochastic recurrence (4.1) and a recurrence for the variance follows from the expectation of its square. Solving these recurrences we obtain

$$\begin{aligned}
\mathbf{E}[W_j] &= \mathbf{E}[\mathbf{E}[W_j|\mathcal{F}_{j-1}]] \\
&= \Big(1-\frac{2}{n}\Big)\mathbf{E}[W_{j-1}] + 1 \\
&= \Big(1-\frac{2}{n}\Big)\Big(\Big(1-\frac{2}{n}\Big)\mathbf{E}[W_{j-1}] + 1\Big) + 1 \\
&\vdots \\
&= \Big(1-\frac{2}{n}\Big)^j \mathbf{E}[W_0] + \sum_{i=0}^{j-1}\Big(1-\frac{2}{n}\Big)^i \\
&= \Big(1-\frac{2}{n}\Big)^j W_0(n) + \frac{1-\Big(1-\frac{2}{n}\Big)^j}{1-\Big(1-\frac{2}{n}\Big)} \\
&= \frac{1}{2}n + \Big(W_0(n) - \frac{n}{2}\Big)\Big(1-\frac{2}{n}\Big)^j,
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{E}[W_j^2] &= \mathbf{E}[\mathbf{E}[W_j^2|\mathcal{F}_{j-1}]] \\
&= \mathbf{E}\big[\mathbf{E}\big[(W_{j-1} + 1 - 2I_n^W)^2|\mathcal{F}_{j-1}\big]\big] \\
&= \mathbf{E}\Big[W_{j-1}^2 + 1 + 4\frac{W_{j-1}}{n} + 2W_{j-1} - 4W_{j-1}\frac{W_{j-1}}{n} - 4\frac{W_{j-1}}{n}\Big] \\
&= \Big(1-\frac{4}{n}\Big)\mathbf{E}[W_{j-1}^2] + \Big(n - (2W_0(n) - n)\Big(1-\frac{2}{n}\Big)^{j-1}\Big) + 1
\end{aligned}$$

$$
\begin{aligned}
&= \Big(1-\frac{4}{n}\Big)\Big(\Big(\Big(1-\frac{4}{n}\Big)\mathbf{E}[W_{j-2}^2]+(n-2(W_0(n)-n))\Big(1-\frac{2}{n})^{j-2}\Big)+1\Big)\\
&\quad +\Big(n-(2W_0(n)-n)\Big(1-\frac{2}{n}\Big)^{j-1}\Big)+1\\
&\vdots\\
&= \Big(1-\frac{4}{n}\Big)^j W_0^2(n)+(2W_0(n)-n)\sum_{i=0}^{j-1}\Big(1-\frac{2}{n}\Big)^i\Big(1-\frac{4}{n}\Big)^{j-1-i}\\
&\quad +(n+1)\sum_{j=0}^{j-1}\Big(1-\frac{4}{n}\Big)^i\\
&= \frac{n}{2}+\Big(\Big(\frac{n}{2}-W_0(n)\Big)^2-\frac{n}{4}\Big)\Big(\frac{n-4}{n}\Big)^j+n\Big(W_0(n)-\frac{n}{2}\Big)\Big(\frac{n-2}{n}\Big)^j.
\end{aligned}
$$

So, the exact expectation and variance are

$$
\mathbf{E}[W_j] = \frac{1}{2}n+\Big(W_0(n)-\frac{n}{2}\Big)\Big(\frac{n-2}{n}\Big)^j; \tag{4.2}
$$

$$
\begin{aligned}
\mathbf{Var}\big[W_j\big] &= \mathbf{E}[W_j^2]-(\mathbf{E}[W_j])^2\\
&= \frac{1}{4}n+\Big(\Big(\frac{n}{2}-W_0(n)\Big)^2-\frac{n}{4}\Big)\Big(\frac{n-4}{n}\Big)^j\\
&\qquad -\Big(\frac{n}{2}-W_0(n)\Big)^2\Big(\frac{n-2}{n}\Big)^{2j}.
\end{aligned} \tag{4.3}
$$

(The special case of $W_0(2n)=n$ is developed in Antognini, 2005.)

Note that under the assumption that $W_0(n)=\lfloor\alpha n\rfloor\sim\alpha n$, the mean after k_n diffusion steps experiences phases according to how fast k_n grows. For the growing sublinear, linear and superlinear phases we have the mean asymptotics

$$
\mathbf{E}[W_{k_n}]\sim\begin{cases}\alpha\, n, & \text{for } k_n=o(n),\\ \Big(\Big(\alpha-\dfrac{1}{2}\Big)e^{-2\lambda_n}+\dfrac{1}{2}\Big)\,n, & \text{for } k_n=\lambda_n n,\\ \dfrac{1}{2}\,n, & \text{for } n=o(k_n).\end{cases}
$$

Like the mean, under the assumption that $W_0(n)=\lfloor\alpha n\rfloor\sim\alpha n$, the variance of the amount of gas in chamber A after k_n diffusion steps experiences phases according to how fast k_n grows. For the growing sublinear, linear and superlinear phases we have the variance asymptotics

$$\mathbf{Var}[W_{k_n}] \sim \begin{cases} 4\alpha(1-\alpha)\,k_n, & \text{for } k_n, = o(n), \\ \dfrac{e^{4\lambda_n} - 1 - 4\lambda_n(2\alpha-1)^2}{4e^{4\lambda_n}}\, n, & \text{for } k_n = \lambda_n n, \\ \dfrac{1}{4}\, n, & \text{for } n = o(k_n). \end{cases}$$

Observe how the average and variance of the three phases meet at the seam lines. The linear phase with $\lambda_n = 0$ gives the result in the growing sublinear phase, and with $\lambda_n = \infty$ gives the result of the superlinear phase.

4.4 A martingale underlying gas diffusion

Conditioning the recurrence (4.1) on the content of the sigma field $\mathcal{F}_{j-1}$, one gets

$$\mathbf{E}[W_j \mid \mathcal{F}_{j-1}] = \Big(1 - \frac{2}{n}\Big) W_{j-1} + 1. \tag{4.4}$$

There is an associated martingale as in the following lemma.

Lemma 1 *For* $j = 0, 1, \ldots,$

$$M_j := \rho_n^j W_j - \frac{\rho_n^{j+1} - \rho_n}{\rho_n - 1}$$

is a martingale, where $\rho_n = n/(n-2)$.

Proof. Introduce the transformation

$$M_j = a_j W_j + b_j.$$

We wish to turn M_j into a martingale with suitable choices of deterministic sequences a_j and b_j. So, M_j must satisfy

$$\mathbf{E}[M_j \,|\, \mathcal{F}_{j-1}] = M_{j-1} = a_{j-1} W_{j-1} + b_{j-1}. \tag{4.5}$$

We compute

$$\begin{aligned} \mathbf{E}[M_j \,|\, \mathcal{F}_{j-1}] &= \mathbf{E}[a_j W_j + b_j \,|\, \mathcal{F}_{j-1}] \\ &= a_j \mathbf{E}[W_j \,|\, \mathcal{F}_{j-1}] + b_j. \end{aligned}$$

From (4.4) we proceed with

$$\mathbf{E}[M_j \,|\, \mathcal{F}_{j-1}] = a_j \Big(1 - \frac{2}{n}\Big) W_{j-1} + a_j + b_j.$$

Matching the coefficients of this equality with those in (4.5), we arrive at recurrences for a_j and b_j. We have $a_j = \rho_n a_{j-1}$. This recurrence unfolds easily to give $a_j = \rho_n^j a_0$, for any arbitrary value of a_0; we take $a_0 = 1$.

We also have the recurrence $b_j = b_{j-1} - a_j$, which unwinds into

$$b_j = b_0 - \sum_{k=1}^{j} \rho_n^k,$$

for arbitrary b_0; we take $b_0 = 0$ and simplify the sum to

$$b_j = -\frac{\rho_n^{j+1} - \rho_n}{\rho_n - 1}. \quad \square$$

The fact that M_j is a martingale is key to proving Gaussian limits in all the phases. We shall deal with the centered martingale

$$\tilde{M}_j = M_j - W_0(n)$$

(which has mean 0) to employ the martingale central limit theorem, which requires calculations on a zero-mean martingale. Specifically in our case, conditional Lindeberg's condition requires that, for some positive increasing sequence ξ_n, and for all $\varepsilon > 0$,

$$U_n := \sum_{j=1}^{k_n} \mathbf{E}\Big[\Big(\frac{\nabla \tilde{M}_j}{\xi_n}\Big)^2 \mathbf{1}_{\Big\{\Big|\frac{\nabla \tilde{M}_j}{\xi_n}\Big| > \varepsilon\Big\}} \,\Big|\, \mathcal{F}_{j-1}\Big] \xrightarrow{\mathcal{P}} 0,$$

and, for a constant c, a c–conditional variance condition requires that

$$V_n := \sum_{j=1}^{k_n} \mathbf{E}\Big[\Big(\frac{\nabla \tilde{M}_j}{\xi_n}\Big)^2 \,\Big|\, \mathcal{F}_{j-1}\Big] \xrightarrow{\mathcal{P}} c. \qquad (4.6)$$

When both conditions hold, the sum $\sum_{j=1}^{k_n} \nabla \tilde{M}_j/\xi_n = (M_{k_n} - M_0)/\xi_n = (M_{k_n} - W_0(n))/\xi_n$ converges to the normally distributed random variable $\mathcal{N}(0, c^2)$.

In all the phases we take $\xi_n = \rho_n^{k_n}\sqrt{\mathbf{Var}[W_{k_n}]}$. For calculations involved in conditional Lindeberg's condition the following uniform bound is helpful in all the phases.

Lemma 2

$$\Big|\frac{\nabla \tilde{M}_j}{\rho_n^j}\Big| \le 4.$$

Proof. With the help of (4.1) write the absolute differences as

$$\begin{aligned}
|\nabla \tilde{M}_j| &= \big|(M_j - W_0(n)) - (M_{j-1} - W_0(n))\big| \\
&= \big|(\rho_n^j W_j + b_j) - (\rho_n^{j-1} W_{j-1} + b_{j-1})\big| \\
&= \big|\big(\rho_n^j (W_{j-1} + I_j^R - I_j^W) + b_j\big) - (\rho_n^{j-1} W_{j-1} + b_{j-1})\big| \\
&\le \rho_n^{j-1}\big((\rho_n - 1)W_{j-1} + \rho_n |I_j^R - I_j^W| + \rho_n\big) \\
&\le \rho_n^{j-1}\Big(\frac{2}{n-2}\, W_{j-1} + \frac{2n}{n-2}\Big).
\end{aligned}$$

The number of white balls at any stage is at most n, and the lemma follows. □

Lemma 2 enables us to verify conditional Lindeberg's condition in all the phases.

Lemma 3

$$U_n = \sum_{j=1}^{k_n} \mathbf{E}\Big[\Big(\frac{\nabla \tilde{M}_j}{\rho_n^j \sqrt{\mathbf{Var}[W_{k_n}]}}\Big)^2 \mathbf{1}_{\Big\{\Big|\frac{\nabla \tilde{M}_j}{\rho_n^j \sqrt{\mathbf{Var}[W_{k_n}]}}\Big| > \varepsilon\Big\}} \,\Big|\, \mathcal{F}_{j-1}\Big] \xrightarrow{\mathcal{P}} 0.$$

Proof. In all the growing phases, the variance grows with n. Therefore, for any given $\varepsilon > 0$, the uniform bound in Lemma 2 asserts that the sets $\{|\nabla \tilde{M}_j| > \varepsilon \rho_n^j \sqrt{\mathbf{Var}[W_{k_n}]}\ \}$ are all empty, for all n greater than some positive integer $n_0(\varepsilon)$.

For large n we have

$$
\begin{aligned}
U_n &= \sum_{j=1}^{n_0(\varepsilon)} \mathbf{E}\Big[\Big(\frac{\nabla\tilde{M}_j}{\rho_n^j\sqrt{\mathbf{Var}[W_{k_n}]}}\Big)^2 \mathbf{1}_{\Big\{\Big|\frac{\nabla\tilde{M}_j}{\rho_n^j\sqrt{\mathbf{Var}[W_{k_n}]}}\Big| > \varepsilon\Big\}} \,\Big|\, \mathcal{F}_{j-1}\Big] \\
&\le \frac{1}{\mathbf{Var}[W_{k_n}]} \sum_{j=1}^{n_0(\varepsilon)} \mathbf{E}\Big[\Big(\frac{\nabla\tilde{M}_j}{\rho_n^j}\Big)^2 \,\Big|\, \mathcal{F}_{j-1}\Big] \\
&\le \frac{16 n_0(\varepsilon)}{\mathbf{Var}[W_{k_n}]} \\
&\to 0, \qquad \text{as } n \to \infty. \quad \square
\end{aligned}
$$

For calculations involved in conditional Lindeberg's condition we need to find the expression of $\mathbf{E}\Big[(\nabla\tilde{M}_j)^2 \,|\, \mathcal{F}_{j-1}\Big]$ (see the definition of V_n in (4.6)), which is

$$
\begin{aligned}
\mathbf{E}\big[(\nabla\tilde{M}_j)^2 \,|\, \mathcal{F}_{j-1}\big] &= \mathbf{E}\big[M_j^2 \,|\, \mathcal{F}_{j-1}\big] - M_{j-1}^2 \\
&= \mathbf{E}\big[(\rho_n^j W_j + b_j)^2 \,|\, \mathcal{F}_{j-1}\big] - (\rho_n^{j-1} W_{j-1} + b_{j-1})^2 \\
&= \mathbf{E}\big[(\rho_n^j (W_{j-1} + 1 - 2I_j^W) + b_j)^2 \,|\, \mathcal{F}_{j-1}\big] \\
&\qquad -(\rho_n^{j-1} W_{j-1} + b_{j-1})^2.
\end{aligned}
$$

After some laborious but straightforward calculation involving (4.1) we get

$$
\begin{aligned}
\mathbf{E}\big[(\nabla\tilde{M}_j)^2 \,|\, \mathcal{F}_{j-1}\big] &= \Big(\rho_n^{2j} - \rho_n^{2j-2} - \frac{4}{n}\rho_n^{2j}\Big) W_{j-1}^2 \\
&\quad + \Big(2\rho_n^{2j} + 2b_j\rho_n^j - \frac{4}{n}\, b_j\rho_n^j - 2\rho_n^{j-1} b_{j-1}\Big) W_{j-1} \\
&\quad + 2b_j\rho_n^j + \rho_n^{2j} + b_j^2 - b_{j-1}^2.
\end{aligned}
$$

This further simplifies to

$$
\mathbf{E}\big[(\nabla\tilde{M}_j)^2 \,|\, \mathcal{F}_{j-1}\big] = -\frac{4}{n^2}\,\rho_n^{2j} W_{j-1}^2 + \frac{4}{n}\,\rho_n^{2j} W_{j-1}.
$$

Summing we reconstruct V_n as

$$V_n = \frac{1}{\rho_n^{2k_n}\mathbf{Var}[W_{k_n}]}\Big(-\frac{4}{n^2}\sum_{j=1}^{k_n}\rho_n^{2j}W_{j-1}^2 + \frac{4}{n}\sum_{j=1}^{k_n}\rho_n^{2j}W_{j-1}\Big). \tag{4.7}$$

4.5 Phases during long-term drawing

Suppose the gas diffusion process is perpetuated indefinitely. We shall see that as the ball drawing continues from the Ehrenfest urn the process experiences different phases.

4.5.1 The growing sublinear phase

Let k_n be in the growing sublinear phase (k_n grows to ∞, and $k_n = o(n)$). The number of white balls after $0 \le j \le k_n$ draws has obvious bounds—if all the draws are of red balls an increase by j goes in favor of the number of white balls over their initial number, and if all the draws are of white balls, a deficit of j occurs against the initial number of white balls. We have the inequalities

$$W_0(n) - j \le W_j \le W_0(n) + j.$$

We can ascertain that

$$W_j = \alpha n + O(k_n), \tag{4.8}$$

for all $0 \le j \le k_n$.

Proof of Theorem 1 in the sublinear phase. In Lemma 3, conditional Lindeberg's condition has been verified throughout the growing sublinear phase. The proof will be complete if we show that V_n converges to a constant in probability.

In (4.7) replace W_{j-1} by the asymptotic equivalent in (4.8) to get

$$\begin{aligned} V_n &= \frac{1}{\rho_n^{2k_n}\mathbf{Var}[W_{k_n}]}\Big(-\frac{4}{n^2}\sum_{j=1}^{k_n}\rho_n^{2j}\big(\alpha n + O(k_n)\big)^2 \\ &\qquad +\frac{4}{n}\sum_{j=1}^{k_n}\rho_n^{2j}\big(\alpha n + O(k_n)\big)\Big) \\ &= \frac{1}{4\alpha(1-\alpha)k_n\big(1+o(1)\big)}\Big(4\alpha(1-\alpha)+O\Big(\frac{k_n}{n}\Big)+O\Big(\frac{k_n^2}{n^2}\Big)\Big)\sum_{j=1}^{k_n}\rho_n^{2j}. \end{aligned}$$

Recall that $\rho_n = n/(n-2)$, and we can bound the remaining sum asymptotically:

$$k_n \le \sum_{j=1}^{k_n}\rho_n^{2j} \le k_n\rho_n^{2k_n} = k_n\big(1+o(1)\big).$$

And so, we have

$$\begin{aligned} V_n &= \frac{1}{k_n\big(1+o(1)\big)}\Big(1+O\Big(\frac{k_n}{n}\Big)+O\Big(\frac{k_n^2}{n^2}\Big)\Big)\big(k_n+o(k_n)\big) \\ &\to 1. \end{aligned}$$

The 1–conditional variance condition has been verified in the growing sublinear phase.

With both conditions checked, the martingale central limit theorem gives

$$\sum_{j=1}^{k_n}\frac{\nabla\tilde{M}_j}{\rho_n^{k_n}\sqrt{\mathbf{Var}[W_{k_n}]}} = \frac{M_{k_n}-W_0(n)}{\rho_n^{k_n}\sqrt{\mathbf{Var}[W_{k_n}]}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

Subsequently, we write

$$\frac{\rho_n^{k_n}W_{k_n} - \dfrac{\rho_n^{k_n+1}-\rho_n}{\rho_n-1} - W_0(n)}{\rho_n^{k_n}\sqrt{\mathbf{Var}[W_{k_n}]}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1),$$

which after reorganization is the statement of the theorem. □

4.5.2 The linear phase

In the linear phase $k_n = \lambda_n n$, for some $\lambda_n > 0$ of a magnitude uniformly bounded from above and below, that is, for two positive constants, S_1 and S_2, and for all n,

$$S_1 \le \lambda_n \le S_2.$$

At this phase of the gas diffusion we have the following asymptotic equivalents (as $n \to \infty$), following from (4.2) and (4.3):

$$\mathbf{E}[W_{k_n}] = \mu_n n + o(n), \tag{4.9}$$

and

$$\mathbf{Var}[W_{k_n}] = v_n n + o(n), \tag{4.10}$$

where

$$\mu_n = \Big(\alpha - \frac{1}{2}\Big)e^{-2\lambda_n} + \frac{1}{2},$$

and

$$v_n = \frac{e^{4\lambda_n} - 1 - 4\lambda_n(1-2\alpha)^2}{4e^{4\lambda_n}} = O(1).$$

We start with a first-order result for W_{k_n}.

Lemma 4 *For $k_n = \lambda_n n$, for some $\lambda_n > 0$ of a magnitude separated from zero and infinity,*

$$\frac{W_{k_n}}{\Big(\Big(\alpha - \frac{1}{2}\Big)e^{-2\lambda_n} + \frac{1}{2}\Big)\, n} \xrightarrow{\mathcal{P}} 1.$$

Proof. By Chebyshev's inequality,

$$\begin{aligned}
\mathbf{Prob}\big(\big|W_{k_n} - \mathbf{E}[W_{k_n}]\big| \ge \varepsilon \mathbf{E}[W_{k_n}]\big) &\le \frac{\mathbf{Var}[W_{k_n}]}{\varepsilon^2(\mathbf{E}[W_{k_n}])^2} \\
&\sim \frac{v_n n}{\varepsilon^2 \mu_n^2 n^2} \\
&\to 0, \qquad \text{as} \quad n \to \infty.
\end{aligned}$$

Hence,

$$\frac{W_{k_n}}{\mathbf{E}[W_{k_n}]} \xrightarrow{\mathcal{P}} 1.$$

From the convergence $\mathbf{E}[W_{k_n}]/(\mu_n n) \to 1$, and Slutsky's Theorem in its multiplicative form (see Karr, 1993, P. 147), we obtain

$$\frac{W_{k_n}}{\mu_n n} \xrightarrow{\mathcal{P}} 1. \quad \square$$

Before we dwell on the proof of a central limit theorem for the amount of gas in Chamber A by the end of some linear phase, we need a technical lemma, which shows that W_{k_n} grows linearly with n, like its mean, with correction terms that are $o_{\mathcal{L}_1}(n)$. The purpose of this calculation is for later summation to verify conditional Lindeberg's condition.

Lemma 5 *Let W_{k_n} be the number of white balls in the urn after k_n draws, where $k_n = \lambda_n n$, for some λ_n, such that $0 < S_1 \le \lambda_n \le S_2 < \infty$. Then*

$$W_{k_n} = \mu_n n + o_{\mathcal{L}_1}(n),$$

and

$$W_{k_n}^2 = \mu_n^2 n^2 + o_{\mathcal{L}_1}(n^2),$$

Proof. From the asymptotics of the mean and variance, as given in (4.9) and (4.10), for large n we have

$$\begin{aligned}\mathbf{E}\big[\big(W_{k_n} - \mu_n n\big)^2\big] &= \mathbf{Var}[W_{k_n}] + \big(\mathbf{E}[W_{k_n}] - \mu_n n\big)^2 \\ &= v_n\, n + o(n^2) \\ &= o(n^2).\end{aligned}$$

So, by Jensen's inequality

$$\mathbf{E}\Big[\Big|W_{k_n} - \mu_n n\Big|\Big] \le \sqrt{\mathbf{E}\big[\big(W_{k_n} - \mu_n n\big)^2\big]} = o(n),$$

which implies

$$W_{k_n} = \mu_n n + o_{\mathcal{L}_1}(n),$$

and the second equation in the lemma is valid by taking the square. □

Proof of Theorem 1 in the linear phase. In Lemma 3, conditional Lindeberg's condition has been verified throughout the linear phase. It remains to verify the conditional variance condition.

Recall the expressions for V_n (see (4.7)). In this phase the asymptotic equivalents in Lemma 5 apply only in the linear phase. However, before the linear phase the obvious bound n on W_{j-1} is sufficient for our purpose.

To asymptotically handle the sums in conditional Lindeberg's condition (going over the range of indexes 1 to $k_n = \lambda_n n$) let us break them up at some point near the beginning of the linear phase. Choose a small positive $\epsilon < S_1$ and break up the sums in V_n into sums going from 1 to $\lfloor \varepsilon n \rfloor - 1$ and sums starting at $\lfloor \varepsilon n \rfloor$ and ending at k_n. Applying the asymptotics of Lemma 5, we write (4.7) in the form

$$\begin{aligned} V_n &= \frac{1}{\rho_n^{2k_n}(v_n n + o(n))}\Big(-\frac{4}{n^2}\sum_{j=1}^{\lfloor \varepsilon n\rfloor -1} \rho_n^{2j} W_{j-1}^2 + \frac{4}{n}\sum_{j=1}^{\lfloor \varepsilon n\rfloor -1} \rho_n^{2j} W_{j-1} \\ &\quad -\frac{4}{n^2}\sum_{j=\lfloor \varepsilon n\rfloor}^{k_n} \rho_n^{2j}\Big(\Big(\Big(\alpha-\frac{1}{2}\Big)e^{-2j/n+o(1)}+\frac{1}{2}\Big)^2 n^2 + o_{\mathcal{L}_1}(n^2)\Big) \\ &\quad +\frac{4}{n\rho_n^{2k_n}}\sum_{j=\lfloor \varepsilon n\rfloor}^{k_n} \rho_n^{2j}\Big(\Big(\Big(\alpha-\frac{1}{2}\Big)e^{-2j/n+o(1)}+\frac{1}{2}\Big)n + o_{\mathcal{L}_1}(n)\Big)\Big) \\ &= \frac{1}{1+o(1)}(C_n + C_n' + D_n + H_n), \end{aligned}$$

where

$$\begin{aligned} C_n &:= -\frac{4}{n^3 v_n e^{4\lambda_n}}\sum_{j=1}^{\lfloor \varepsilon n\rfloor -1} \rho_n^{2j} W_{j-1}^2, \\ C_n' &:= \frac{4}{n^2 v_n e^{4\lambda_n}}\sum_{j=1}^{\lfloor \varepsilon n\rfloor -1} \rho_n^{2j} W_{j-1}, \\ D_n &:= -\frac{4}{n^3 v_n e^{4\lambda_n}}\sum_{j=\lfloor \varepsilon n\rfloor}^{k_n} \rho_n^{2j}\Big(\Big(\Big(\alpha-\frac{1}{2}\Big)e^{-2j/n+o(1)}+\frac{1}{2}\Big)^2 n^2 + o_{\mathcal{L}_1}(n^2)\Big), \end{aligned}$$

and

$$H_n := \frac{4}{n^2 v_n e^{4\lambda_n}}\sum_{j=\lfloor \varepsilon n\rfloor}^{k_n} \rho_n^{2j}\Big(\Big(\Big(\alpha-\frac{1}{2}\Big)e^{-2j/n+o(1)}+\frac{1}{2}\Big)n + o_{\mathcal{L}_1}(n)\Big).$$

For large n, we have

$$\begin{aligned} |C_n| &\le \frac{4}{n^3 v_n e^{4\lambda_n}}\sum_{j=1}^{\lfloor \varepsilon n\rfloor -1} \rho_n^{2j} n^2 \\ &\le \frac{4}{n v_n e^{4\lambda_n}}\sum_{j=1}^{\lfloor \varepsilon n\rfloor} 2e^{4S_1} \\ &= O(\varepsilon), \qquad \text{as } \ \varepsilon \to 0. \end{aligned}$$

Likewise, we have

$$|C'_n| = O(\varepsilon), \qquad \text{as } \ \varepsilon \to 0.$$

The formula for D_n reduces to

$$\begin{aligned}
D_n &= -\frac{4}{nv_n e^{4\lambda_n}} \sum_{j=\lfloor \varepsilon n\rfloor}^{k_n} \rho_n^{2j}\Big(\Big(\Big(\alpha-\frac{1}{2}\Big)e^{-2j/n+o(1)}+\frac{1}{2}\Big)^2 + o_{\mathcal{L}_1}(1)\Big) \\
&= -\frac{4}{nv_n e^{4\lambda_n}} \sum_{j=\lfloor \varepsilon n\rfloor}^{k_n} \rho_n^{2j}\Big(\Big(\alpha-\frac{1}{2}\Big)^2 e^{-4j/n} + \Big(\alpha-\frac{1}{2}\Big)e^{-2j/n} + \frac{1}{4} \\
&\qquad + o_{\mathcal{L}_1}(1)\Big) \\
&= -\frac{4}{nv_n e^{4\lambda_n}} \Big(\Big(\alpha-\frac{1}{2}\Big)^2 \Big(\sum_{j=0}^{k_n} \rho_n^{2j} e^{-4j/n} - \sum_{j=0}^{\lfloor \varepsilon n\rfloor -1} \rho_n^{2j} e^{-4j/n}\Big) \\
&\qquad + \Big(\alpha-\frac{1}{2}\Big)\Big(\sum_{j=0}^{k_n} \rho_n^{2j} e^{-2j/n} - \sum_{j=0}^{\lfloor \varepsilon n\rfloor -1} \rho_n^{2j} e^{-2j/n}\Big) \\
&\qquad + \frac{1}{4}\Big(\sum_{j=0}^{k_n} \rho_n^{2j} - \sum_{j=0}^{\lfloor \varepsilon n\rfloor -1} \rho_n^{2j}\Big) \\
&\qquad + o_{\mathcal{L}_1}(1) \sum_{j=\lfloor \varepsilon n\rfloor}^{k_n} \rho_n^{2j}\Big).
\end{aligned}$$

This calculation involves several sums of the form

$$\sum_{j=0}^{b_n-1} \rho_n^{2j} e^{-\frac{\gamma j}{n}} = \frac{\Big(\dfrac{n}{n-2}\Big)^{2b_n} e^{-\frac{\gamma b_n}{n}} - 1}{\Big(\dfrac{n}{n-2}\Big)^2 e^{-\frac{\gamma}{n}} - 1},$$

with $b_n = \beta_n n + r_n$, and the remainder function r_n is $o(n)$. Using the asymptotic relation $(n/(n-2))^{2\beta_n n} = e^{4\beta_n} + 4\beta_n e^{4\beta_n}/n + O(1/n^2)$, and the standard local expansion

$$e^{\frac{c}{n}} = 1 + \frac{c}{n} + \frac{c^2}{2n^2} + O\Big(\frac{1}{n^3}\Big),$$

we get

$$\sum_{j=0}^{b_n-1} \rho_n^{2j} e^{-\frac{\gamma j}{n}} = \frac{\Big(\Big(e^{4\beta_n} + \frac{4\beta_n e^{4\beta_n}}{n} + O\Big(\frac{1}{n^2}\Big)\Big)\Big(\frac{n}{n-2}\Big)^{2r_n} e^{-\frac{\gamma\beta_n n+\gamma r_n}{n}} - 1\Big)}{\Big((4-\gamma)n + \Big(\frac{\gamma^2}{2} - 4\Big) + O\Big(\frac{1}{n}\Big)\Big)} \times (n-2)^2$$

$$= \frac{e^{(4-\gamma)\beta_n}\Big(1 + \frac{4\beta_n}{n} + O\Big(\frac{1}{n^2}\Big)\Big) e^{2r_n(\frac{2}{n}+O(\frac{1}{n^2}))} e^{-\frac{\gamma r_n}{n}} - 1}{\Big((4-\gamma)n + \Big(\frac{\gamma^2}{2} - 4\Big) + O\Big(\frac{1}{n}\Big)\Big)} (n-2)^2.$$

Applying these formulas with $\gamma = 4, 2, 0$, we have

$$D_n = -\frac{1}{v_n e^{4\lambda_n}}\Big[4\Big(\alpha - \frac{1}{2}\Big)^2 \lambda_n + 2\Big(\alpha - \frac{1}{2}\Big)\big(e^{2\lambda_n} - 1\big) + \frac{1}{4}\big(e^{4\lambda_n} - 1\big)\Big] + O(\varepsilon) + o(1) + o_{\mathcal{L}_1}(1).$$

Similarly, we have

$$H_n = \frac{1}{v_n e^{4\lambda_n}}\Big(2\Big(\alpha - \frac{1}{2}\Big)\big(e^{2\lambda_n} - 1\big) + \frac{1}{2}\big(e^{4\lambda_n} - 1\big)\Big) + O(\varepsilon) + o(1) + o_{\mathcal{L}_1}(1).$$

Consequently, we have

$$V_n = \frac{1}{1+o(1)}\big(O(\varepsilon) + 1 + o(1) + o_{\mathcal{L}_1}(1)\big).$$

Taking the limit, as $\varepsilon \to 0$, we get

$$\lim_{\varepsilon\to 0} V_n = V_n = \frac{1}{1+o(1)}\big(1 + o(1) + o_{\mathcal{L}_1}(1)\big).$$

Now, let $n \to \infty$ to get

$$V_n \xrightarrow{\mathcal{P}} 1.$$

According to the martingale central limit theorem

$$\sum_{j=1}^{k_n} \frac{\nabla \tilde{M}_j}{\rho_n^{k_n}\sqrt{\mathbf{Var}[W_{k_n}]}} = \frac{M_{k_n} - M_0}{\rho_n^{k_n}\sqrt{\mathbf{Var}[W_{k_n}]}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

Subsequently, we write

$$\frac{\rho_n^{k_n} W_{k_n} - \dfrac{\rho_n^{k_n+1} - \rho_n}{\rho_n - 1} - W_0(n)}{\rho_n^{k_n}\sqrt{\mathbf{Var}[W_{k_n}]}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1). \quad \square$$

4.5.3 The superlinear phase

Suppose that the gas diffusion continued for a long period of time. As seen from the behavior of the average, the initial conditions are attenuated through the linear phase and the fixed average component $\frac{1}{2}n$ becomes more pronounced and eventually dominates in the superlinear phase. Many of the principles of the proof for the linear phase apply within the superlinear phase, so we shall be a bit brief in presenting an adjustment of these proofs. For instance, via the asymptotic equivalents of the mean and variance in the superlinear phase, we can mimic the proof of Lemma 4, and get a similar result. Namely, when $n = o(k_n)$, we have

$$\frac{W_{k_n}}{n} \xrightarrow{\mathcal{P}} \frac{1}{2}.$$

Also, in view of the mean and variance asymptotics we can replicate the result of Lemma 5. We only need to replace μ_n by $\frac{1}{2}$, and the proof goes through verbatim to obtain

$$W_{k_n} = \frac{1}{2}n + o_{\mathcal{L}_1}(n),$$

and

$$W_{k_n}^2 = \frac{1}{4}n^2 + o_{\mathcal{L}_1}(n^2).$$

Proof of Theorem 1 in the superlinear phase. For the sum in the conditional variance condition we apply the bound $W_{j-1} \le n$ until the superlinear phase. More precisely, to asymptotically handle the sums in conditional Lindeberg's condition (going over the range of indices 1 to k_n) we break up the sums in V_n into sums going from 1 to $k'_n - 1$, which is any superlinear function of order less than k_n (giving ignorable contribution) and sums starting at k'_n and ending at k_n (most of the contribution comes near k_n). We can take $k'_n = \lfloor k_n/\ln(k_n/n)\rfloor$. Then

$$\begin{aligned}
V_n &= \frac{1}{\rho_n^{2k_n}\mathbf{Var}[W_{k_n}]}\Bigg(-\frac{4}{n^2}\sum_{j=1}^{k'_n-1}\rho_n^{2j}W_{j-1}^2 + \frac{4}{n}\sum_{j=1}^{k'_n-1}\rho_n^{2j}W_{j-1} \\
&\qquad -\frac{4}{n^2}\sum_{j=k'_n}^{k_n}\rho_n^{2j}\Big(\frac{n^2}{4} + o_{\mathcal{L}_1}(n^2)\Big) \\
&\qquad +\frac{4}{n}\sum_{j=k'_n}^{k_n}\rho_n^{2j}\Big(\Big(\frac{n}{2} + o_{\mathcal{L}_1}(n)\Big)\Big) \\
&= \frac{1}{1+o(1)}(\tilde{C}_n + \tilde{D}_n + \tilde{H}_n),
\end{aligned}$$

where

$$\begin{aligned}
\tilde{C}_n &:= -\frac{16}{n^3\rho_n^{2k_n}}\sum_{j=1}^{k'_n-1}\rho_n^{2j}W_{j-1}^2 + \frac{16}{n^2\rho_n^{2k_n}}\sum_{j=1}^{k'_n-1}\rho_n^{2j}W_{j-1}, \\
\tilde{D}_n &:= -\frac{16}{n^3\rho_n^{2k_n}}\sum_{j=\lfloor\varepsilon n\rfloor}^{k_n}\rho_n^{2j}\Big(\frac{n^2}{4} + o_{\mathcal{L}_1}(n^2)\Big),
\end{aligned}$$

and

$$\tilde{H}_n := \frac{16}{n^2\rho_n^{2k_n}}\sum_{j=k'_n}^{k_n}\rho_n^{2j}\Big(\frac{n}{2} + o_{\mathcal{L}_1}(n)\Big).$$

We have

$$
\begin{aligned}
|\tilde{C}_n| &\leq \frac{32}{n\rho_n^{2k_n}} \sum_{j=1}^{k_n'-1} \rho_n^{2j} \\
&= O\Big(\rho_n^{2k_n'-2k_n} \Big).
\end{aligned}
$$

We also have

$$
\tilde{D}_n + \tilde{H}_n = \frac{4}{n\rho_n^{2k_n}} \Big(\frac{\rho_n^{2k_n+2} - \rho_n^{k_n'}}{\rho_n^2 - 1} \Big) \big(1 + o_{\mathcal{L}_1}(1) \big) = 1 + o_{\mathcal{L}_1}(1).
$$

Putting the terms together we see that

$$
V_n \xrightarrow{\mathcal{P}} 1.
$$

According to the martingale central limit theorem

$$
\frac{\rho_n^{k_n} W_{k_n} - \dfrac{\rho_n^{k_n+1} - \rho_n}{\rho_n - 1} - W_0(n)}{\rho_n^{k_n} \sqrt{\mathbf{Var}[W_{k_n}]}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1). \quad \square
$$

Chapter 5

A Generalized Ehrenfest urn model

We introduce a generalization of the Ehrenfest urn model in which samples of independent identically distributed random sizes with a general generating discrete distribution (the generator) are taken out of an urn containing white and red balls (n in total). Each ball in the sample is repainted with the opposite color and the sample is replaced in the urn. For generators that are not nearly deterministic, in the growing sublinear phase the number of white balls is normally distributed, with parameters that are influenced by the initial conditions and the generator. A similar result holds in a properly defined sublinear phase for nearly deterministic generators. In the linear phase a different normal distribution applies, in which the influence of the initial conditions and the generator are attenuated. At the superlinear phase the mix is nearly perfect, with a nearly symmetrical normal distribution in which the effect of the initial conditions and the generator is obliterated.

5.1 Generalized Ehrenfest urns

We introduce a model that is a generalization of the Ehrenfest urn scheme. We shall use a language corresponding to the mixing of population through travel and migration between countries, and we shall use the more common phraseology concerning urns. For example, we refer to the objects in the urn as balls.

Consider a population of two equally-affluent countries, n in total, that is initially split into $W_0(n)$ in the first country, and $R_0(n)$ in the second (with $W_0(n) + R_0(n) = n$). On any given day, a random number of residents travel to the other side for business and numerous other purposes. We assume that on day j a random number S_j of travelers cross the borders, and $S_j \stackrel{\mathcal{L}}{=} S$ are discrete independent identically distributed random variables with law like a generic S with support in $\{0, 1, \ldots, s_n\}$, for $0 \le s_n \le n$. We call the distribution of S the *generating distribution*, or simply the *generator*. We assume that $S_j = S_j(n) \stackrel{\mathcal{L}}{=} S = S(n)$ obeys a certain discrete generating distribution with mean $\mu_S(n)$, and variance $\sigma_S^2(n)$. As we shall see, most of the results are governed by these two parameters in the distribution of S. To avoid trivial degeneracy, we assume $\mathbf{Prob}(S > 0) > 0$, and thus $\mu_S(n) > 0$, too.

For the demographic applications we have in mind, where small changes occur daily as travelers arrive in planes, trains and ships, we assume s_n to be small relative to n, that is, $s_n = o(\sqrt{n}\,)$. We shall call this *subcritical sampling*; cases where s_n is of higher order will be called *critical sampling*. We shall impose a few minor additional restrictions on s_n to obtain Gaussian laws.

We are interested in the daily census in the two countries. Let $W_j = W_j(n)$ be the number of residents in the first country and $R_j = R_j(n)$ be the number of residents in the second at the end of day j. Critical sampling is less practical and less interesting than subcritical cases, as it may lead to an unstable status without limit distributions.

For example, if $S \equiv n$, we have

$$W_j = \begin{cases} W_0(n), & \text{if } j \text{ is even;} \\ n - W_0(n), & \text{if } j \text{ is odd.} \end{cases}$$

Here there are no limit distribution or phases. In the sequel we focus on subcritical sampling. We shall discuss one example of critical sampling at the end of the chapter.

Let us represent the residents of the two countries by colored balls in an urn, white for the residents of the first country and red for the residents of the second. Travel on a particular day corresponds to drawing samples from the urn. At each draw, a random number of balls, following a general discrete distribution like that of S, is taken out of the urn. Each ball in the sample is repainted with the opposite color (indicating travel), and the sample is returned to the urn. Thus, the classical Ehrenfest scheme is the very special case of the urn being discussed, where $S_j \overset{\mathcal{L}}{=} S \equiv 1$. In the language of the urn W_j is the number of white balls in the urn and R_j is the number of red balls in the urn after j draws. Note that throughout, the total population in the two countries remains the same. We study the daily census W_j and R_j of the two countries.

5.2 Results and organization

It is our intention to discover phases during the evolution, find interpretation for how they conjoin at the seam lines, and investigate the influence of the underlying generator. We identify three phases of $k = k_n$ draws. In each phase, under the appropriate centering and scaling, a Gaussian distribution is obtained. The three laws are different and the results elicit the critical transition at the seam lines between the different Gaussian phases.

We identify three phases:

(a) The sublinear phase, when $\mu_S(n)k_n = o(n)$. Nothing much of interest takes place in the phase $\mu_S(n)k_n = O(1)$, where not enough time has elapsed to effect a significant change from the starting conditions. Certain appropriate regularity conditions are sufficient for Gaussian laws in upper sublinear phases:

  (i) For the nondegenerate case, where $\sigma_S^2(n) \nrightarrow 0$, we develop a limit Gaussian law in the upper sublinear phase in which

$$s_n^2 = o\Big(\Big(4\frac{W_0(n)}{n}\Big(1 - \frac{W_0(n)}{n}\Big)\mu_S(n) + \Big(1 - 2\frac{W_0(n)}{n}\Big)^2 \sigma_S^2(n)\Big)\Big).$$

  (ii) For the degenerate case, where $\sigma_S^2(n) \to 0$, we develop a limit Gaussian law in the upper sublienar phase in which

$$s_n^2 = o\Big(4\frac{W_0(n)}{n}\Big(1 - \frac{W_0(n)}{n}\Big)\mu_S(n)k_n + 2\frac{\mu_S^2(n)k_n^2}{n}\Big).$$

(b) The linear phase, when $\mu_S(n)k_n = \lambda_n n$, for some $\lambda_n > 0$ of a magnitude bounded from above and below by positive constants. That is, we assume the existence of two constants Q_1 and Q_2, such that $0 < Q_1 \le \lambda_n \le Q_2 < \infty$, for all n. For instance λ_n can be a simple constant like the case $\lambda_n = 3$, may converge to a constant like the case $\lambda_n = 5 + \frac{1}{n}$, or may include oscillations, such as the case $\lambda_n = 7 + 4\cos n$.

(c) The superlinear phase, when $n = o(\mu_S(n)k_n)$.

Note that the notion of linearity is a combination of the mean of the generator and the number of draws. We do not deal with cases where $n^{-1}W_0(n)$ hits the values 0 or 1 infinitely often. We also exclude the cases where $0 = \liminf_{n\to\infty} \sigma_S^2(n) <$

$\limsup_{n\to\infty} \sigma_S^2(n)$. The main result of this chapter is the following.

Theorem 2 *Let W_{k_n} be the number of white balls in a generalized Ehrenfest urn containing n balls after k_n draws of subcritical sampling, of which initially the number of white balls is $W_0(n)$, where $k_n \to \infty$, in a sublinear (under the regularity conditions (i) and (ii)), linear or superlinear fashion. Then*

$$\frac{W_{k_n} - \Big(\Big(1 - \frac{2\mu_S(n)}{n}\Big)^{k_n}\Big(W_0(n) - \frac{n}{2}\Big) + \frac{n}{2}\Big)}{\sqrt{v_n}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1),$$

where

$$\begin{aligned} v_n \;=\; & \Big(\Big(W_0(n) - \frac{n}{2}\Big)^2 - \frac{n}{4}\Big) \\ & \times\Big(\frac{4\mu_S^2(n) + 4\sigma_S^2(n) - 4n\mu_S(n) + n^2 - n}{n^2 - n}\Big)^{k_n} \\ & + n\Big(W_0(n) - \frac{n}{2}\Big)\Big(1 - \frac{2\mu_S(n)}{n}\Big)^{k_n} + \frac{1}{4}n(n+1) \\ & - \Big(\Big(1 - \frac{2\mu_S(n)}{n}\Big)^{k_n}\Big(W_0(n) - \frac{n}{2}\Big) + \frac{n}{2}\Big)^2. \end{aligned}$$

Remark: *Under subcritical sampling, the condition s_n^2/k_n is automatically satisfied in the linear and superlinear case. In the sublinear phase, we have to keep s_n small to get Gaussian laws.*

Note that the result depends on the generator only through its mean and variance; all generating distributions with the same mean and variance exert the same influence on the long term mixing. The theorem has different interpretation in different phases. In the sublinear phase, the initial conditions persist. For instance, if $W_0(n) \sim \alpha n$, for some $\alpha \in (0,1)$, Theorem 2 takes the form

$$\frac{W_{k_n} - \Big(\Big(1 - \dfrac{2\mu_S(n)}{n}\Big)^{k_n}\Big(\alpha - \dfrac{1}{2}\Big)n + \dfrac{n}{2}\Big)}{\sqrt{(4\alpha(1-\alpha)\mu_S(n) + (1-2\alpha)^2\sigma_S^2(n))k_n}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

For the sublinear phase, the initial conditions persist and the influence of the generator is quite pronounced. In the linear phase a different asymptotic normal distribution holds. However, the influence of the generator and initial conditions are attenuated as we get deeper in the linear phase (larger λ_n). Theorem 2 takes the form

$$\frac{W_{k_n} - \Big(\Big(\frac{W_0(n)}{n} - \frac{1}{2}\Big)e^{-2\lambda_n} + \dfrac{1}{2}\Big)\,n}{\sqrt{\dfrac{e^{-4\lambda_n}}{4}\Big(e^{4\lambda_n} - 1 - 4\lambda_n\Big(\dfrac{2W_0(n)}{n} - 1\Big)^2 + 4\lambda_n\dfrac{\sigma_S^2(n)}{\mu_S(n)}\Big(\dfrac{2W_0(n)}{n} - 1\Big)^2\Big)\,n}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

After a very long period of time, as in the superlinear phase, the mixing is nearly complete, and the result is a central limit theorem in which the effect of any initial conditions is obliterated, and the generator has no effect. For instance if k_n is of order higher than $n \ln n$, Theorem 2 takes the form

$$\frac{W_{k_n} - \dfrac{1}{2}n}{\sqrt{n}} \xrightarrow{\mathcal{D}} \mathcal{N}\Big(0, \frac{1}{4}\Big).$$

These results (also reported in Zhang and Mahmoud, 2011+) generalize the results of Chapter 4 (reported in Balaji, Mahmoud and Zhang, 2010), where phases in the classical Ehrenfest model are discussed. Diaconis (1996) takes a different view of these phases and explores the so-called cutoff phenomenon in finite Markov chains. There is general interest in phases in the long-term drawing from urns; see for example Mahmoud (2010), Mikhailov (1977, 1980), Vatutin and Mikhailov (1982), and Smythe (2011) for phases associated with coupon collection.

For the generalized Ehrenfest urn, given W_{j-1} and S_j, the number S_j^W of white balls in the jth sample follows a Hypergeo(n, S_j, W_{j-1}). Accordingly, we have

$$\begin{aligned}
\mathbf{E}[S_j^W | \mathcal{F}_{j-1}, S_j] &= \frac{S_j W_{j-1}}{n}; \\
\mathbf{Var}[S_j^W | \mathcal{F}_{j-1}, S_j] &= \frac{S_j \frac{W_{j-1}}{n}\Big(1 - \frac{W_{j-1}}{n}\Big)(n - S_j)}{n-1}.
\end{aligned}$$

Of the S_j balls that appear in the jth sample, let S_j^W and $S_j^R = S_j - S_j^W$ be respectively the number of white and red balls in that sample. According to the sampling and replacement scheme, we have a basic stochastic recurrence:

$$W_j(n) = W_j = W_{j-1} + S_j^R - S_j^W = W_{j-1} + S_j - 2S_j^W, \tag{5.1}$$

and subsequently we find

$$\begin{aligned}
\mathbf{E}[W_j \,|\, \mathcal{F}_{j-1}, S_j] &= W_{j-1} + S_j - 2\mathbf{E}[S_j^W \,|\, \mathcal{F}_{j-1}, S_j] \\
&= W_{j-1} + S_j - \frac{2S_j W_{j-1}}{n} \\
&= W_{j-1}\Big(1 - \frac{2S_j}{n}\Big) + S_j.
\end{aligned}$$

By the law of iterated expectations and the independence of S_j and $\mathcal{F}_{j-1}$, we have

$$\begin{aligned}
\mathbf{E}[W_j \,|\, \mathcal{F}_{j-1}] &= \mathbf{E}_{S_j \,|\, \mathcal{F}_{j-1}}\big[\mathbf{E}[W_j \,|\, \mathcal{F}_{j-1}, S_j] \,|\, \mathcal{F}_{j-1}\big] \\
&= \mathbf{E}_{S_j \,|\, \mathcal{F}_{j-1}}\Big[W_{j-1}\Big(1 - \frac{2S_j}{n}\Big) + S_j \,|\, \mathcal{F}_{j-1}\Big] \\
&= W_{j-1}\Big(1 - \frac{2\mu_S(n)}{n}\Big) + \mu_S(n).
\end{aligned} \tag{5.2}$$

5.3 Moments and asymptotical phases within

We are interested in finding the first few moments of the number of white balls in the urn as it evolves.

5.3.1 The mean

Taking double expectations in (5.2) we get a recurrence, with solution

$$\mathbf{E}[W_j] = \Big(1 - \frac{2\mu_S(n)}{n}\Big)^j \Big(W_0(n) - \frac{n}{2}\Big) + \frac{n}{2}.$$

Note that all generators with the same mean, give the same average number of white balls after j draws. Observe the strong balancing effect of the very special cases where $\mu_S(n) = \frac{1}{2}n$, or $W_0(n) = \frac{1}{2}n$: in these cases $\mathbf{E}[W_j] = \frac{1}{2}n$, for all $j \geq 1$.

For the cases of interest with subcritical sampling, we have $s_n = o(\sqrt{n}\,)$, and $\mu_S(n) = o(\sqrt{n}\,)$, too. So, we can get the following local expansions from the exact expression for the mean:

$$\begin{aligned}
\mathbf{E}[W_{k_n}] &= \Big(1 - \frac{2\mu_S(n)}{n}\Big)^{k_n} \Big(W_0(n) - \frac{n}{2}\Big) + \frac{n}{2} \\
&= \exp\Big(k_n \log\Big(1 - \frac{2\mu_S(n)}{n}\Big)\Big) \Big(W_0(n) - \frac{n}{2}\Big) + \frac{n}{2} \\
&= \exp\Big(-k_n\Big(\frac{2\mu_S(n)}{n} + O\Big(\frac{\mu_S^2(n)}{n^2}\Big)\Big)\Big) \Big(W_0(n) - \frac{n}{2}\Big) + \frac{n}{2}.
\end{aligned}$$

With subcritical sampling we find

$$\mathbf{E}[W_{k_n}] = \begin{cases} W_0(n) + \Big(1 - \dfrac{2W_0(n)}{n}\Big)\mu_S(n)k_n + O\big(\dfrac{\mu_S^2(n)k_n^2}{n}\big), & k_n \quad \text{sublinear;} \\ e^{-2\lambda_n}\Big(W_0(n) - \dfrac{n}{2}\Big) + \dfrac{n}{2} + o(n), & k_n \quad \text{linear;} \\ \dfrac{n}{2} + o(n), & k_n \quad \text{superlinear.} \end{cases}$$

Observe that in the sublinear phase the average is essentially a perturbation on the initial value $W_0(n)$. If $W_0(n)$ is of order n, we essentially have what we started with throughout the sublinear phase. But if the initial value is of smaller order than the order of $\mu_S(n)k_n$, this perturbation comes in the picture as the bigger influence. In the linear phase the steady-state component $\frac{n}{2}$ appears, and the influence of the generator asymptotically disappears. The effect of the initial conditions get attenuated quickly at an exponential rate with the linearity coefficient: The deeper we are in the linear phase, the weaker is the influence of the initial condition. In the superlinear phase the effect of the generator and the initial condition is asymptotically obliterated. We get a plain $\frac{1}{2}n$ asymptotic mean, reflecting an even split on average.

5.3.2 The variance

We go further with the variance. The basic stochastic recurrence (5.1) gives us a recurrence for the second moment:

$$\begin{aligned}
\mathbf{E}[W_j^2|\mathcal{F}_{j-1}, S_j] &= W_{j-1}^2 + S_j^2 + 4\mathbf{E}\big[(S_j^W)^2 \,|\, \mathcal{F}_{j-1}, S_j\big] + 2W_{j-1}S_j \\
&\quad -4W_{j-1}\mathbf{E}\big[S_j^W \,|\, \mathcal{F}_{j-1}, S_j\big] - 4\mathbf{E}\big[S_j^W \,|\, \mathcal{F}_{j-1}, S_j\big]S_j \\
&= \Big(1 - \frac{4S_j}{n(n-1)} + \frac{4S_j^2}{n^2(n-1)} + \frac{S_j^2}{n^2} - \frac{4S_j}{n^2}\Big)W_{j-1}^2 \\
&\quad +\Big(\frac{4S_j}{n-1} - \frac{4S_j^2}{n(n-1)} + 2S_j - \frac{4S_j^2}{n^2}\Big)W_{j-1} + S_j^2,
\end{aligned}$$

which leads to

$$
\begin{aligned}
\mathbf{E}\big[W_j^2 \,|\, \mathcal{F}_{j-1}\big] &= \mathbf{E}_{S_j \,|\, \mathcal{F}_{j-1}}\big[\mathbf{E}_{W_j \,|\, \mathcal{F}_{j-1}, S_j}\big[W_j^2 \,|\, \mathcal{F}_{j-1}, S_j\big] \,|\, S_j\big] \\
&= \Big(1 - \frac{4\mu_S(n)}{n(n-1)} + \frac{4(\mu_S^2(n) + \sigma_S^2(n))}{n^2(n-1)} + \frac{\mu_S^2(n) + \sigma_S^2(n)}{n^2} \\
&\qquad - \frac{4\mu_S(n)}{n^2}\Big) W_{j-1}^2 \\
&\qquad + \Big(\frac{4\mu_S(n)}{n-1} - \frac{4(\mu_S^2(n) + \sigma_S^2(n))}{n(n-1)} + 2\mu_S(n) \\
&\qquad\quad - \frac{4(\mu_S^2(n) + \sigma_S^2(n))}{n^2}\Big) W_{j-1} \\
&\qquad + \mu_S^2(n) + \sigma_S^2(n).
\end{aligned}
$$

A double expectation operation gives a recurrence for $\mathbf{E}[W_j^2]$, which involves $\mathbf{E}[W_j]$ (obtained in Subsection 5.3.1). After solving the recurrence and simplifying, we get a formula for the second moment, from which we get the exact variance

$$
\begin{aligned}
\mathbf{Var}[W_{k_n}] &= \Big(\Big(W_0(n) - \frac{n}{2}\Big)^2 - \frac{n}{4}\Big) \\
&\qquad \times \Big(\frac{4\mu_S^2(n) + 4\sigma_S^2(n) - 4n\mu_S(n) + n^2 - n}{n^2 - n}\Big)^{k_n} \\
&\qquad + n\Big(W_0(n) - \frac{n}{2}\Big)\Big(1 - \frac{2\mu_S(n)}{n}\Big)^{k_n} + \frac{1}{4}n(n+1) \\
&\qquad - \Big(\Big(1 - \frac{2\mu_S(n)}{n}\Big)^{k_n}\Big(W_0(n) - \frac{n}{2}\Big) + \frac{n}{2}\Big)^2.
\end{aligned}
$$

Note that all generators with the same mean and variance, give the same variance for the number of white balls after j draws.

Conducting an asymptotical analysis of this variance formula, we find the following phases in the variance:

1. In the sublinear phase we have

$$\begin{aligned}\mathbf{Var}[W_{k_n}] &= 4\Big(\frac{W_0(n)}{n}\Big(1-\frac{W_0(n)}{n}\Big)\mu_S(n)+\Big(1-2\frac{W_0(n)}{n}\Big)^2\sigma_S^2(n)\Big)\,k_n\\ &+\frac{2\Big(\Big(2\mu_S^2(n)-2\mu_S(n)\sigma_S^2(n)\Big)\Big(\frac{2W_0(n)}{n}-1\Big)^2-\mu_S^2(n)\Big)}{n}\,k_n^2\\ &+O\Big(\frac{\mu_S^2(n)k_n}{n}\Big)+O\Big(\frac{\mu_S(n)\sigma_S^2(n)k_n}{n}\Big).\end{aligned}$$

Note that, if $n^{-1}W_0(n)$ stays bounded away from 0 and 1, the asymptotic variance is

$$\mathbf{Var}[W_{k_n}] \sim 4\Big(\frac{W_0(n)}{n}\Big(1-\frac{W_0(n)}{n}\Big)\mu_S(n)+\Big(1-2\frac{W_0(n)}{n}\Big)^2\sigma_S^2(n)\Big)\,k_n.$$

In the degenerate case, where $\sigma_S^2(n)\to 0$, the situation depends on the initial conditions. Here, if $n^{-1}W_0(n)$ stays bounded away from 0 and 1, we still have

$$\mathbf{Var}[W_{k_n}] \sim 4\frac{W_0(n)}{n}\Big(1-\frac{W_0(n)}{n}\Big)\mu_S(n)k_n.$$

But, if the conditions are extremal (i.e. $n^{-1}W_0(n)$ converges to 0 or 1), we have the asymptotic equivalent

$$\mathbf{Var}[W_{k_n}] \sim 4\frac{W_0(n)}{n}\Big(1-\frac{W_0(n)}{n}\Big)\mu_S(n)k_n+\frac{2\mu_S^2(n)k_n^2}{n}.$$

2. In the linear phase we have

$$\begin{aligned}\mathbf{Var}[W_{k_n}] &\sim \Big(\Big(e^{4\lambda_n}-1-4\lambda_n\Big(1-2\frac{W_0(n)}{n}\Big)^2\\ &\quad+4\lambda_n\frac{\sigma_S^2(n)}{\mu_S(n)}\Big(1-2\frac{W_0(n)}{n}\Big)^2\Big)e^{-4\lambda_n}\Big)\frac{n}{4}.\end{aligned}$$

3. In the superlinear phase we have

$$\mathbf{Var}[W_{k_n}] \sim \frac{n}{4}.$$

5.4 Asymptotic distributions in various phases

We first find martingales associated with the migration process to be able to appeal to the martingale central limit theorem.

5.4.1 A martingale formulation

Starting with (5.1), we have

$$\mathbf{E}\big[W_j \,|\, \mathcal{F}_{j-1}\big] = \mathbf{E}_{S_j \,|\, \mathcal{F}_{j-1}}\big[\mathbf{E}[W_j \,|\, \mathcal{F}_{j-1}, S_j] \,|\, \mathcal{F}_{j-1}\big] = W_{j-1}\Big(1 - \frac{2\mu_S(n)}{n}\Big) + \mu_S(n).$$

We want to construct a martingale based on $\{W_j : j \geq 0\}$. So, let $M_j = a_j W_j + b_j$, where a_j and b_j are deterministic coefficients, yet to be appropriately determined to render M_j a martingale. We proceed with

$$\begin{aligned}
\mathbf{E}[M_j \,|\, \mathcal{F}_{j-1}] &= a_j\Big(W_{j-1}\Big(1 - \frac{2\mu_S(n)}{n}\Big) + \mu_S(n)\Big) + b_j \\
&= M_{j-1} \\
&= a_{j-1}W_{j-1} + b_{j-1}.
\end{aligned}$$

Equating the coefficient of W_{j-1} on the first and last lines of this display, we get a recurrence for a_j, for arbitrary a_0, which we take to be 1 for simplicity. The solution is

$$a_j = \Big(\frac{n}{n - 2\mu_S(n)}\Big)^{j_n} =: \rho_n^{j_n}.$$

Likewise, we have a recurrence for b_j, with arbitrary b_0, which we take as 0, and get

$$b_j = -\mu_S(n)\frac{\rho_n^{j+1} - \rho_n}{\rho_n - 1}.$$

Hence

$$M_j = M_j(n) = \rho_n^j W_j - \mu_S(n)\frac{\rho_n^{j+1} - \rho_n}{\rho_n - 1}$$

is a martingale.

The martingale central limit theorem applies to centered martingales (with mean 0); let us introduce the martingale

$$\tilde{M}_j(n) = \tilde{M}_j = M_j - W_0(n), \tag{5.3}$$

with martingale differences

$$\nabla\tilde{M}_j = \big(M_j - W_0(n)\big) - \big(M_{j-1} - W_0(n)\big) = M_j - M_{j-1}.$$

The style of this proof parallels that of Subsection 4.4; the two conditions for the martingale central limit theorem will be checked.

Lemma 6

$$\big|\nabla\tilde{M}_j\big| \le 4s_n\rho_n^j.$$

Proof. We have

$$\begin{aligned}\big|\nabla\tilde{M}_j\big| &= |M_j - M_{j-1}| \\ &= \left|\left(\rho_n^j W_j - \mu_S(n)\frac{\rho_n^{j+1} - \rho_n}{\rho_n - 1}\right) - \left(\rho_n^{j-1} W_{j-1} - \mu_S(n)\frac{\rho_n^{j} - \rho_n}{\rho_n - 1}\right)\right|\end{aligned}$$

$$
\begin{aligned}
&= \Big|\Big(\rho_n^j(W_{j-1} + S_j^R - S_j^W) - \mu_S(n)\frac{\rho_n^{j+1} - \rho_n}{\rho_n - 1}\Big) \\
&\qquad -\Big(\rho_n^{j-1}W_{j-1} - \mu_S(n)\frac{\rho_n^j - \rho_n}{\rho_n - 1}\Big)\Big| \\
&= \big|\rho_n^{j-1}(\rho_n - 1)W_{j-1} + \rho_n^j(S_j^R - S_j^W) - \mu_S(n)\rho_n^j\big| \\
&\le \rho_n^{j-1}\Big(\frac{2\mu_S(n)}{n - 2\mu_S(n)}W_{j-1} + \rho_n\big|S_j^R - S_j^W\big| + \rho_n\mu_S(n)\Big) \\
&\le \rho_n^j(3\mu_S(n) + s_n).
\end{aligned}
$$

The last line ensues from the fact that W_{j-1} is bounded by n, and $|S_j^R - S_j^W|$ is bounded by s_n. With $\mu_S(n) \le s_n$, the lemma follows. □

The following two lemmas, holding in all phases, will help us throughout.

Lemma 7

$$
\begin{aligned}
V_n &= \frac{1}{\rho_n^{2k_n}\mathbf{Var}[W_{k_n}]}\Big(\frac{4(\sigma_S^2(n)n + \mu_S^2(n) - \mu_S(n)n)}{n^2(n-1)}\sum_{j=1}^{k_n}\rho_n^{2j}W_{j-1}^2 \\
&\qquad -\frac{4(n\sigma_S^2(n) + \mu_S^2(n) - n\mu_S(n))}{n(n-1)}\sum_{j=1}^{k_n}\rho_n^{2j}W_{j-1} \\
&\qquad +\sigma_S^2(n)\sum_{j=1}^{k_n}\rho_n^{2j}\Big).
\end{aligned}
$$

Proof. Squaring out the centered martingale in (5.3), and taking various respective expectations we arrive at

$$
\begin{aligned}
\mathbf{E}[(\nabla\tilde{M}_j)^2 \,|\, \mathcal{F}_{j-1}] &= \mathbf{E}[M_j^2 - 2M_jM_{j-1} + M_{j-1}^2 \,|\, \mathcal{F}_{j-1}] \\
&= \mathbf{E}[M_j^2 \,|\, \mathcal{F}_{j-1}] - M_{j-1}^2 \\
&= \mathbf{E}_{S_j \,|\, \mathcal{F}_{j-1}}[\mathbf{E}[M_j^2 \,|\, \mathcal{F}_{j-1}, S_j] \,|\, \mathcal{F}_{j-1}] - M_{j-1}^2
\end{aligned}
$$

$$
\begin{aligned}
&= \mathbf{E}_{S_j \,|\, \mathcal{F}_{j-1}}[\mathbf{E}[(\rho_n^j W_j + b_j)^2 \,|\, \mathcal{F}_{j-1}, S_j] \,|\, \mathcal{F}_{j-1}] \\
&\qquad -(\rho_n^{j-1} W_{j-1} + b_{j-1})^2 \\
&= \mathbf{E}_{S_j \,|\, \mathcal{F}_{j-1}}[\mathbf{E}[(\rho_n^j (W_{j-1} + S_j - 2S_j^W) \\
&\qquad + b_j)^2 \,|\, \mathcal{F}_{j-1}, S_j] \,|\, \mathcal{F}_{j-1}] - (\rho_n^{j-1} W_{j-1} + b_{j-1})^2 \\
&= \rho_n^{2j} W_{j-1}^2 + 2\rho^{2j} W_{j-1} \mu_S(n) - 4\frac{\rho_n^{2j} W_{j-1}^2 \mu_S(n)}{n} \\
&\qquad + 2\rho_n^j W_{j-1} b_j + \rho_n^{2j}\big(\sigma_S^2(n) + \mu_S^2(n)\big) \\
&\qquad - \frac{4\rho_n^{2j}\Big(\sigma_S^2(n) + \mu_S^2(n)) W_{j-1}}{n} + 2\rho_n^j \mu_S(n) b_j \\
&\qquad + 4\rho_n^{2j}\Big(\frac{(\sigma_S^2 + \mu_S^2(n)) W_{j-1}^2}{n} \\
&\qquad + \big(n^2 \mu_S(n) W_{j-1} - (\sigma_S^2(n) + \mu_S^2(n)) W_{j-1} n\big) \\
&\qquad\quad - \mu_S(n) W_{j-1}^2 n + \big(\sigma_S^2(n) + \mu_S^2\big) W_{j-1}^2 \\
&\qquad\qquad \times \big(n^{-2}(n-1)^{-1}\big)\Big) \\
&\qquad - \frac{4\rho_n^j \mu_S(n) W_{j-1} b_j}{n} + b_j^2 \\
&\qquad\quad - \rho_n^{2j-2} W_{j-1}^2 - 2\rho^{j-1} W_{j-1} b_{j-1} - b_{j-1}^2.
\end{aligned}
$$

This further simplifies as in the statement of the lemma. □

Lemma 8 *Under regularity conditions (i) and (ii), we have*

$$
U_n = \sum_{j=1}^{k_n} \mathbf{E}\Big[\Big(\frac{\nabla \tilde{M}_j}{\rho_n^j \sqrt{\mathbf{Var}[W_{k_n}]}}\Big)^2 \mathbf{1}_{\Big\{\Big|\frac{\nabla \tilde{M}_j}{\rho_n^j \sqrt{\mathbf{Var}[W_{k_n}]}}\Big| > \varepsilon\Big\}} \,\Big|\, \mathcal{F}_{j-1}\Big] \xrightarrow{\mathcal{P}} 0.
$$

Proof. Consider first the nondegenerate case and the degenerate but nonextremal case. In all the growing phases, the variance grows with n. Therefore, for any given $\varepsilon > 0$, the uniform bound in Lemma 6 asserts that the sets $\{|\nabla \tilde{M}_j| > \varepsilon \rho_n^j \sqrt{\mathbf{Var}[W_{k_n}]}\,\}$ are all empty, for all n greater than some positive integer $n_0(\varepsilon)$.

Recall that in the linear and superlinear phases, the variance is of order n whereas in the sublinear phase it is of order k_n. For large n, under regularity condition (i), we have

$$\begin{aligned} U_n &= \sum_{j=1}^{n_0(\varepsilon)} \mathbf{E}\Big[\Big(\frac{\nabla \tilde{M}_j}{\rho_n^j \sqrt{\mathbf{Var}[W_{k_n}]}}\Big)^2 \mathbf{1}_{\Big\{\Big|\frac{\nabla \tilde{M}_j}{\rho_n^j \sqrt{\mathbf{Var}[W_{k_n}]}}\Big| > \varepsilon\Big\}} \Big| \mathcal{F}_{j-1}\Big] \\ &\le \frac{1}{\mathbf{Var}[W_{k_n}]} \sum_{j=1}^{n_0(\varepsilon)} \mathbf{E}\Big[\Big(\frac{\nabla \tilde{M}_j}{\rho_n^j}\Big)^2 \Big| \mathcal{F}_{j-1}\Big] \\ &\le \frac{16 s_n^2 n_0(\varepsilon)}{\mathbf{Var}[W_{k_n}]} \\ &\to 0, \qquad \text{as } n \to \infty. \end{aligned}$$

Consider next the degenerate extremal case. For this case the asymptotic variance is $2\mu_S^2(n)k_n^2/n$, and so, the sets $\{|\nabla \tilde{M}_j| > \varepsilon \rho_n^j \sqrt{\mathbf{Var}[W_{k_n}]}\ \}$ are all empty, but only in the upper sublinear phase. A similar calculation for U_n goes through, and U_n converges in probability to 0 under regularity condition (ii). □

5.4.2 The growing sublinear phase

In this phase $k_n \to \infty$, and $\mu(n)k_n = o(n)$. Therefore,

$$W_0(n) - js_n \le W_j \le W_0(n) + js_n,$$

for every $0 \le j \le k_n$, and $W_j = W_0(n) + O(k_n s_n)$. This crude bound is good enough to deal with the case of samples with random size. However, when the samples are of nearly deterministic size ($\sigma_S^2(n) \to 0$) and either $W_0(n) = o(n)$, or $B_0(n) = o(n)$, a finer asymptotic approximation is required. We present this approximation next.

Proof of Theorem 2 in the sublinear phase. Assume the regularity conditions (i) and (ii). Conditional Lindeberg's condition holds (see Lemma 8). It remains to verify the conditional variance condition. Note that

$$k_n = \sum_{j=1}^{k_n} 1 \le \sum_{j=1}^{k_n} \rho_n^{2j} \le \sum_{j=1}^{k_n} \rho_n^{2k_n} = k_n + o(k_n).$$

Consider first the nondegenerate case. Replace W_{j-1} by $W_0(n) + O(k_n s_n)$. If $\sigma_S^2(n)$ does not converge to zero, for any starting conditions we have the asymptotic equivalent

$$\begin{aligned} V_n &= \frac{1}{\rho_n^{2k_n}\mathbf{Var}[W_{k_n}]}\Big(\frac{4(\sigma_S^2(n)n + \mu_S^2(n) - \mu_S(n)n)}{n^2(n-1)} \\ &\quad \times\big(k_n + o(k_n)\big)\big(W_0(n) + O(k_n s_n)\big)^2 \\ &\quad -\frac{4(\sigma_S^2(n)n + \mu_S^2(n) - \mu_S(n)n)}{n(n-1)}\big(k_n + o(k_n)\big)\big(W_0(n) + O(k_n s_n)\big) \\ &\quad +(k_n + o(k_n))\sigma_S^2(n)\Big) \\ &\sim \frac{1}{\mathbf{Var}[W_{k_n}]}\Big(4\frac{W_0(n)}{n}\Big(1 - \frac{W_0(n)}{n}\Big)\mu_S(n) + \Big(1 - 2\frac{W_0(n)}{n}\Big)^2\sigma_S^2(n)\Big)k_n \\ &\to 1. \end{aligned}$$

In the degenerate case $\sigma_S^2(n) \to 0$, the situation depends on the initial conditions. Here, if $W_0(n)/n$ stays bounded away from 0 and 1, we still have

$$V_n \sim \frac{1}{\mathbf{Var}[W_{k_n}]}\Big(4\frac{W_0(n)}{n}\Big(1 - \frac{W_0(n)}{n}\Big)\mu_S(n)\Big)k_n \to 1.$$

However, if $n^{-1}W_0(n)$ approaches 0 or 1,

$$V_n \sim 4\frac{W_0(n)}{n}(1 - \frac{W_0(n)}{n})\mu_S(n)k_n + \frac{2\mu_S^2(n)k_n^2}{n}.$$

In the case of $n^{-1}W_0(n)$ approaches 0,

$$\begin{aligned}
V_n &= \frac{1}{\rho_n^{2k_n}\mathbf{Var}[W_{k_n}]}\Big(\frac{4\Big(\sigma_S^2(n)n+\mu_S^2(n)+\mu_S(n)n\Big)}{n^2(n-1)} \\
&\qquad \times\sum_{j=1}^{k_n}\rho_n^{2j}\Big(W_0(n)+\mu_S(n)j+o_P(\mu_S(n)j)\Big)^2 \\
&\qquad -\frac{4\Big(\sigma_S^2(n)n+\mu_S^2(n)+\mu_S(n)n\Big)}{n(n-1)} \\
&\qquad \times\sum_{j=1}^{k_n}\rho_n^{2j}\Big(W_0(n)+\mu_S(n)j+o_P(\mu_S(n)j)\Big) \\
&\qquad +(k_n+o(k_n))\sigma_S^2(n)\Big) \\
&\sim \frac{4\Big(\mu_S^2(n)-\mu_S(n)n\Big)}{\mathbf{Var}[W_{k_n}]n(n-1)} \\
&\qquad \times\Big(\frac{1}{n}\sum_{j=1}^{k_n}\Big(W_0(n)+\mu_S(n)j\Big)^2-k_nW_0(n)-\mu_S(n)\sum_{j=1}^{k_n}j\Big) \\
&\sim \frac{4\Big(\mu_S^2(n)-\mu_S(n)n\Big)}{\mathbf{Var}[W_{k_n}]n(n-1)}\Big(\frac{k_nW_0^2(n)}{n}+\frac{2W_0(n)\mu_S}{n}\frac{k_n(k_n+1)}{2} \\
&\quad +\frac{\mu_S^2(n)}{n}\Big(\frac{k_n(k_n+1)(2k_n+1)}{6}\Big)-k_nW_0(n)-\mu_S(n)\frac{k_n(k_n+1)}{2}\Big)
\end{aligned}$$

$$\begin{aligned}
&\sim \frac{1}{\mathbf{Var}[W_{k_n}]}\Big(4\frac{W_0(n)}{n}(1-\frac{W_0(n)}{n})\mu_S(n)k_n \\
&\qquad +\frac{2\mu_S^2(n)k_n^2}{n}+O(\frac{\mu_S^3(n)k_n^3}{n^2})\Big) \\
&\to 1.
\end{aligned}$$

A similar proof is valid in the case of $n^{-1}W_0(n)$ approaches 1. In all the relevant cases, the variance condition checks out.

By the Martingale Central Limit Theorem, we have

$$\begin{aligned}
\sum_{j=1}^{k_n} \frac{\nabla \tilde{M}_j}{\rho_n^{k_n}\sqrt{\mathbf{Var}[W_{k_n}]}} &= \frac{M_{k_n} - W_0(n)}{\rho_n^{k_n}\sqrt{\mathbf{Var}[W_{k_n}]}} \\
&= \frac{\rho_n^{k_n} W_{k_n} - \mu_S(n)\frac{\rho_n^{k_n+1}-\rho_n}{\rho_n - 1} - W_0(n)}{\rho_n^{k_n}\sqrt{\mathbf{Var}[W_{k_n}]}} \\
&\xrightarrow{\mathcal{L}} \mathcal{N}(0,1).
\end{aligned}$$

In this phase $\rho_n^{k_n} \to 1$; Theorem 2 follows in its stated form upon an application of Slutsky's theorem (see Karr, 1993, P. 146). $\square$

5.4.3 The linear phase

Consider the linear phase, in which $\mu_S(n)k_n = \lambda_n n$, for λ_n uniformly bounded from above and below by two positive constants, that is, there are two real constants Q_1 and Q_2 such that $0 < Q_1 \le \lambda_n \le Q_2 < \infty$, for all $n \ge 1$. From the exact mean and variance relations (developed in Subsections 5.3.1 and 5.3.2) we see that

$$\begin{aligned}
\mathbf{E}[W_{k_n}] &= \mu_n n + o(n); \\
\mathbf{Var}[W_{k_n}] &= \nu_n n + o(n),
\end{aligned}$$

where

$$\begin{aligned}
\mu_n &= \Big(\frac{W_0(n)}{n} - \frac{1}{2}\Big)e^{-2\lambda_n} + \frac{1}{2}, \\
\nu_n &= \frac{e^{4\lambda_n} - 1 - 4\Big(1 - 2\frac{W_0(n)}{n}\Big)^2\lambda_n + 4\Big(1 - 2\frac{W_0(n)}{n}\Big)^2\lambda_n\frac{\sigma_S^2(n)}{\mu_S(n)}}{4e^{4\lambda_n}}.
\end{aligned}$$

Lemma 9 *Under subcritical sampling, in the linear phase*

$$\frac{W_{k_n}}{\Big(\Big(\frac{W_0(n)}{n} - \frac{1}{2}\Big)e^{-2\lambda_n} + \frac{1}{2}\Big)n} \xrightarrow{\mathcal{P}} 1.$$

Proof. By Chebyshev's inequality,

$$\begin{aligned}
\mathbf{Prob}(|W_{k_n} - \mathbf{E}[W_{k_n}]| \ge \varepsilon \mathbf{E}[W_{k_n}]) &\le \frac{\mathbf{Var}[W_{k_n}]}{\varepsilon^2(\mathbf{E}[W_{k_n}])^2} \\
&\sim \frac{\nu_n n}{\varepsilon^2 \mu_n^2 n^2} \\
&\to 0.
\end{aligned}$$

So,

$$\frac{W_{k_n}}{\mathbf{E}[W_{k_n}]} \xrightarrow{\mathcal{P}} 1, \qquad \frac{W_{k_n}}{\mu_n n} \xrightarrow{\mathcal{P}} 1,$$

and an application of Slutsky's theorem (Karr, 1993, P. 146) gives the result. □

Proof of Theorem 2 in the linear phase.

It was shown in Lemma 6 that under subcritical sampling, $U_n \to 0$ in the linear phase. We need only check the conditional variance condition.

In Lemma 7, we have sums over *random* values of W_{j-1}^2 and W_{j-1}. We handle these sums by approximating these random variables by their leading terms in the $\mathcal{L}_1$ sense, as in the following lemma.

Lemma 10 *In the prescribed linear phase, we have the approximation*

$$W_{k_n} = \mu_n n + O_{\mathcal{L}_1}(\sqrt{n}\,).$$

Proof. From the asymptotics of the mean and variance, for large n, we have

$$\begin{aligned}\mathbf{E}[(W_{k_n}-\mu_n n)^2] &= \mathbf{Var}[W_{k_n}] + \big(\mathbf{E}[W_{k_n}-\mu_n n]\big)^2 \\ &= \nu_n n + o_{\mathcal{L}_1}(n^2) \\ &= o_{\mathcal{L}_1}(n^2).\end{aligned}$$

So, by Jensen's inequality,

$$\mathbf{E}[|W_{k_n}-\mu_n n|] \leq \sqrt{\mathbf{E}[(W_{k_n}-\mu_n)^2]} = o_{\mathcal{L}_1}(n),$$

which implies that

$$W_{k_n} = \mu_n n + o_{\mathcal{L}_1}(n). \quad \square$$

Replace W_{j-1} by $(((W_0(n)/n-\frac{1}{2})e^{-\frac{2\mu_S(n)j}{n}+O(1)}+\frac{1}{2})n+O_{L_1}(n^{\frac{1}{2}}))$, choose a small positive $\varepsilon < S_1$, partition the sum into a sum with indexes at most $\lfloor \varepsilon n \rfloor - 1$ and a sum with indexes at least $\lfloor \varepsilon n \rfloor$, then,

$$\begin{aligned}V_n &= \frac{4\big(-\mu_S(n)n+\mu_S^2(n)+\sigma_S^2(n)n\big)}{n^2(n-1)\xi_n^2}\sum_{j=1}^{k_n}\rho_n^{2j}W_{j-1}^2 \\ &\quad -\frac{4(-\mu_S(n)n+\mu_S^2(n)+\sigma_S^2(n)n)}{n(n-1)\xi_n^2}\sum_{j=1}^{k_n}\rho_n^{2j}W_{j-1}+\frac{\sigma_S^2(n)}{\xi_n^2}\sum_{j=1}^{k_n}\rho_n^{2j}\end{aligned}$$

$$
\begin{aligned}
= \quad & \frac{4(-\mu_S(n)n+\mu_S^2(n)+\sigma_S^2(n)n)}{n^2(n-1)\xi_n^2}\sum_{j=1}^{\lfloor \varepsilon n\rfloor -1}\rho_n^{2j}W_{j-1}^2 \\
& -\frac{4(-\mu_S(n)n+\mu_S^2(n)+\sigma_S^2(n)n)}{n(n-1)\xi_n^2}\sum_{j=1}^{\lfloor \varepsilon n\rfloor -1}\rho_n^{2j}W_{j-1} \\
& +\frac{4(-\mu_S(n)n+\mu_S^2(n)+\sigma_S^2(n)n)}{n^2(n-1)\xi_n^2}\sum_{\lfloor \varepsilon n\rfloor}^{k_n}\rho_n^{2j}\Big(\Big(\Big(\frac{W_0(n)}{n}-\frac{1}{2}\Big) \\
& e^{-\frac{2\mu_S(n)j}{n}+O(1)}+\frac{1}{2}\Big)^2 n^2+O_{\mathcal{L}_1}(n^{\frac{3}{2}})\Big) \\
& -\frac{4(-\mu_S(n)n+\mu_S^2(n)+\sigma_S^2(n)n)}{n(n-1)\xi_n^2}\sum_{\lfloor \varepsilon n\rfloor}^{k_n}\rho_n^{2j}\Big(\Big(\Big(\frac{W_0(n)}{n}-\frac{1}{2}\Big)e^{-\frac{2\mu_S(n)j}{n}+O(1)} \\
& +\frac{1}{2}\Big)n+O_{\mathcal{L}_1}(n^{\frac{1}{2}})\Big)+\frac{\sigma_S^2(n)}{\xi_n^2}\sum_{j=1}^{k_n}\rho_n^{2j} \\
& = C_n+D_n+C_n'+D_n'+G_n,
\end{aligned}
$$

where

$$
C_n=\frac{4(-\mu_S(n)n+\mu_S^2(n)+\sigma_S^2(n)n)}{n^2(n-1)\xi_n^2}\sum_{j=1}^{\lfloor \varepsilon n\rfloor -1}\rho_n^{2j}W_{j-1}^2.
$$

For large n,

$$
\begin{aligned}
|C_n| \quad & \leq \quad \frac{4(-\mu_S(n)n+\mu_S^2(n)+\sigma_S^2(n)n)}{n^3(n-1)}\sum_{j=1}^{\lfloor \varepsilon n\rfloor -1}\rho_n^{2j}n^2 \\
& \leq \quad \frac{4(-\mu_S(n)n+\mu_S^2(n)+\sigma_S^2(n)n)}{n(n-1)}\sum_{j=1}^{\lfloor \varepsilon n\rfloor -1}4e^{4\mu_S(n)Q_2} \\
& = \quad O(\varepsilon), \qquad \text{as} \;\; n\to\infty.
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
D_n \quad & = \quad \frac{4(-\mu_S(n)n+\mu_S^2(n)+\sigma_S^2(n)n)}{n(n-1)\xi_n^2}\sum_{j=1}^{\lfloor \varepsilon n\rfloor -1}\rho_n^{2j}W_{j-1}; \\
|D_n| \quad & = \quad O(\varepsilon), \qquad \text{as } n\to\infty,
\end{aligned}
$$

Further,

$$
\begin{aligned}
C_n' &= \frac{4(-\mu_S(n)n+\mu_S^2(n)+\sigma_S^2(n)n)}{n^2(n-1)\xi_n^2}\sum_{\lfloor\varepsilon n\rfloor}^{k_n}\rho_n^{2j}\Big(\Big(\Big(\frac{W_0(n)}{n}-\frac{1}{2}\Big) \\
&\qquad e^{-\frac{2\mu_S(n)j}{n}+O(1)}+\frac{1}{2}\Big)^2 n^2+O_{\mathcal{L}_1}(n^{\frac{3}{2}})\Big)\Big) \\
&= \frac{4(-\mu_S(n)n+\mu_S^2(n)+\sigma_S^2(n)n)+o(1)}{(n-1)\mathbf{Var}[W_{k_n}]e^{4\lambda_n}}\sum_{\lfloor\varepsilon n\rfloor}^{k_n}\rho_n^{2j}\Big(\Big(\Big(\frac{W_0(n)}{n}-\frac{1}{2}\Big) \\
&\qquad \times e^{-\frac{2\mu_S(n)j}{n}+O(1)}+\frac{1}{2}\Big)^2+O_{\mathcal{L}_1}(n^{-\frac{1}{2}})\Big)\Big)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{4(-\mu_S(n)n+\mu_S^2(n)+\sigma_S^2(n)n)+o(1)}{(n-1)\mathbf{Var}[W_{k_n}]e^{4\lambda_n}}\Big(\Big(\frac{W_0(n)}{n}-\frac{1}{2}\Big)^2 \\
&\qquad \times\Big(\sum_{j=0}^{k_n}\rho_n^{2j}e^{-\frac{4\mu_S(n)j}{n}}-\sum_{j=0}^{\lfloor\varepsilon n\rfloor-1}\rho_n^{2j}e^{-\frac{4\mu_S(n)j}{n}}\Big) \\
&\qquad +\Big(\frac{W_0(n)}{n}-\frac{1}{2}\Big)\Big(\sum_{j=0}^{k_n}\rho_n^{2j}e^{-\frac{4\mu_S(n)j}{n}}-\sum_{j=0}^{\lfloor\varepsilon n\rfloor-1}\rho_n^{2j}e^{-\frac{4\mu_S(n)j}{n}}\Big) \\
&\qquad +\frac{1}{4}\Big(\sum_{j=0}^{k_n}\rho_n^{2j}-\sum_{j=0}^{\lfloor\varepsilon n\rfloor-1}\rho_n^{2j}\Big) \\
&\qquad +O_{\mathcal{L}_1}(n^{-\frac{1}{2}})\sum_{j=\lfloor\varepsilon n\rfloor}^{k_n}\rho_n^{2j}\Big).
\end{aligned}
$$

We can show that

$$
\sum_{j=0}^{\beta_n-1}\rho_n^{2j}e^{-\frac{\gamma\mu_S(n)j}{n}} = \begin{cases} \dfrac{e^{(4-\gamma)\mu_S(n)\beta_n}-1}{(4-\gamma)\mu_S(n)}\,n+o(n), & \gamma\neq 4; \\[2ex] \beta_n n, & \gamma = 4. \end{cases}
$$

Applying these formulas, we have

$$
\begin{aligned}
C_n' &= \frac{1}{\mathbf{Var}[W_{k_n}]e^{4\lambda_n}}\Big(4\Big(\frac{W_0(n)}{n}-\frac{1}{2}\Big)^2\big(\sigma_S^2(n)-\mu_S(n)\big)\frac{\lambda_n}{\mu_S(n)}n \\
&\quad +4\Big(\frac{W_0(n)}{n}-\frac{1}{2}\Big)\big(\sigma_S^2(n)-\mu_S(n)\big) \\
&\quad \times\frac{e^{2\lambda_n}-1}{2\mu_S(n)}n+\big(\sigma_S^2(n)-\mu_S(n)\big)\frac{e^{4\lambda_n}-1}{4\mu_S(n)}n+O(\varepsilon)+o(1) \\
&\quad +O_{\mathcal{L}_1}\Big(\frac{1}{\sqrt{n}}\Big)\Big).
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
D_n' &= -\frac{1}{\mathbf{Var}[W_{k_n}]e^{4\lambda_n}}\Big(4\Big(\frac{W_0(n)}{n}-\frac{1}{2}\Big)(\sigma_S^2(n)-\mu_S(n))\frac{e^{2\lambda_n}-1}{2\mu_S(n)}n \\
&\quad +2(\sigma_S^2(n)-\mu_S(n))\frac{\sigma_S^2(n)\big(e^{4\lambda_n}-1\big)}{4\mu_S(n)}n+O(\varepsilon)+o(1)+O_{\mathcal{L}_1}\Big(\frac{1}{\sqrt{n}}\Big)\Big); \\
G_n &= \frac{1}{\mathbf{Var}[W_{k_n}]e^{4\lambda_n}}\frac{\sigma_S^2(n)(e^{4\lambda_n}-1)}{4\mu_S(n)}n
\end{aligned}
$$

Thus,

$$
\begin{aligned}
C_n+D_n &= O(\varepsilon)\to 0, \qquad \text{as } \varepsilon\to 0; \\
C_n'+D_n'+G_n &= \frac{1}{\mathbf{Var}[W_{k_n}]e^{4\lambda_n}} \\
&\quad \times\Big(\frac{e^{4\lambda_n}-1-4(1-2\frac{W_0(n)}{n})^2\lambda_n+4(1-2\frac{W_0(n)}{n})^2\lambda_n\frac{\sigma_S^2(n)}{\mu_S(n)}}{4}n \\
&\quad +O(\varepsilon)+o(1)+O_{\mathcal{L}_1}\Big(\frac{1}{\sqrt{n}}\Big)\Big) \\
&\to 1 \qquad \text{as } \ \varepsilon\to 0, n\to\infty.
\end{aligned}
$$

By the Martingale Central Limit Theorem,

$$\begin{aligned}\sum_{j=1}^{k_n} \frac{\nabla \tilde{M}_j}{\rho_n^{k_n}\sqrt{\mathbf{Var}[W_{k_n}]}} &= \frac{M_{k_n} - W_0(n)}{\rho_n^{k_n}\sqrt{\mathbf{Var}[W_{k_n}]}} \\ &= \frac{\rho_n^{k_n} W_{k_n} - \mu_S(n)\frac{\rho_n^{k_n+1}-\rho_n}{\rho_n-1} - W_0(n)}{\rho_n^{k_n}\sqrt{\mathbf{Var}[W_{k_n}]}} \\ &\xrightarrow{\mathcal{L}} \mathcal{N}(0,1).\end{aligned}$$

In the linear phase, we have the asymptotic relation

$$\frac{\mu_S(n)\dfrac{\rho_n^{k_n+1}-\rho_n}{\rho_n - 1} + W_0(n) - \left(e^{-2\lambda_n\mu_S(n)}\left(W_0(n) - \dfrac{n}{2}\right) + \dfrac{n}{2}\right)}{\rho_n^{k_n}\sqrt{\mathbf{Var}[W_{k_n}]}} \to 0.$$

An application of Slutsky's theorem (Karr, 1993, P. 146) completes the proof. □

5.4.4 The superlinear phase

In the superlinear phase, $n = o(\mu_S(n)k_n)$.

Proof of Theorem 2 in the superlinear phase.

It was shown in Lemma 6 that under subcritical sampling, $U_n \to 0$ in the superlinear phase. In this phase we can show that

$$\begin{aligned}W_{k_n} &= \frac{1}{2}n + O_{\mathcal{L}_1}(\sqrt{n}\,); \\ W_{k_n}^2 &= \frac{1}{4}n^2 + O_{\mathcal{L}_1}(n^{3/2}).\end{aligned}$$

To asymptotically handle the sums in the conditional Lindeberg's condition, apply the bound $W_{j-1} \leq n$ until the superlinear phase. Beyond that point replace W_{j-1} with $\frac{n}{2} + O_{\mathcal{L}_1}(\sqrt{n}\,)$. More precisely, partition the range $\{1, \ldots, k_n\}$ of indexes into a subrange going from 1 to $k_n' - 1$, which is any superlinear function of order less than

k_n (giving ignorable contribution) and subrange of indexes starting at k'_n and ending at k_n (most of the contribution comes near k_n). We can take $k'_n = \lfloor k_n / \ln(k_n/n) \rfloor$. Then

$$\begin{aligned} V_n &= \frac{4(\mu_S^2(n) + n\sigma_S^2(n) - n\mu_S(n))}{n^2(n-1)\xi_n^2} \sum_{j=1}^{k'_n-1} \rho_n^{2j} W_{j-1}^2 \\ &\quad - \frac{4(\mu_S^2(n) + n\sigma_S^2(n) - n\mu_S(n))}{n(n-1)\xi_n^2} \sum_{j=1}^{k'_n-1} \rho_n^{2j} W_{j-1} \\ &\quad + \frac{4(\mu_S^2(n) + n\sigma_S^2(n) - n\mu_S(n))}{n^2(n-1)\xi_n^2} \sum_{j=k'_n}^{k_n} \rho_n^{2j} \Big(\frac{n^2}{4} + O_{\mathcal{L}_1}\big(n^{3/2}\big) \Big) \\ &\quad - \frac{4(\mu_S^2(n) + n\sigma_S^2(n) - n\mu_S(n))}{n(n-1)\xi_n^2} \sum_{j=k'_n}^{k_n} \rho_n^{2j} \Big(\frac{n}{2} + O_{\mathcal{L}_1}\big(n^{1/2}\big) \Big) \\ &\quad + \frac{\sigma_S^2(n)}{\xi_n^2} \sum_{j=1}^{k'_n-1} \rho_n^{2j} + \frac{\sigma_S^2(n)}{\xi_n^2} \sum_{j=k'_n}^{k_n} \rho_n^{2j}. \end{aligned}$$

The first two sums and the fourth are $O(\rho_n^{2k'_n - 2k_n})$; the remaining three are

$$\begin{aligned} & \frac{\sigma_S^2(n) - \mu_S(n)}{n\rho_n^{2k}} \sum_{j=k'_n}^{k_n} \rho_n^{2j} - \frac{2(\sigma_S^2(n) - \mu_S(n))}{n\rho_n^{2k}} \sum_{j=k'_n}^{k_n} \rho_n^{2j} + \frac{\sigma_S^2}{n\rho_n^{2k}} \sum_{j=k'_n}^{k_n} \rho_n^{2j} + O\Big(\frac{1}{\sqrt{n}}\Big) \\ &= \frac{\mu_S(n)}{n\rho_n^{2k_n}} \Big(\Big(\frac{\rho_n^{2k_n+2} - \rho_n^{k'_n}}{\rho_n^2 - 1} \Big) + O\Big(\frac{1}{\sqrt{n}}\Big) \\ &= \frac{1}{4} + O\Big(\frac{1}{\sqrt{n}}\Big) \\ &\to \frac{1}{4}. \end{aligned}$$

By the Martingale Central Limit Theorem,

$$\begin{aligned}\sum_{j=1}^{k_n} \frac{\nabla \tilde{M}_j}{\rho_n^{k_n}\sqrt{\mathbf{Var}[W_{k_n}]}} &= \frac{M_{k_n} - W_0(n)}{\rho_n^{k_n}\sqrt{\mathbf{Var}[W_{k_n}]}} \\ &= \frac{\rho_n^{k_n} W_{k_n} - \mu_S(n)\frac{\rho_n^{k_n+1}-\rho_n}{\rho_n-1} - W_0(n)}{\rho_n^{k_n}\sqrt{\mathbf{Var}[W_{k_n}]}} \\ &\xrightarrow{\mathcal{L}} \mathcal{N}(0,1).\end{aligned}$$

In this phase,

$$\begin{aligned}\frac{W_0(n)}{\rho_n^{k_n}\sqrt{\mathbf{Var}[W_{k_n}]}} &\to 0; \\ -\frac{\mu_S(n)\rho_n}{(\rho_n-1)\rho_n^{k_n}\sqrt{\mathbf{Var}[W_{k_n}]}} &\to 0; \\ \frac{\frac{n}{2} - \mu_S(n)\rho_n^{k_n+1}}{(\rho_n-1)\rho_n^{k_n}\sqrt{\mathbf{Var}[W_{k_n}]}} &\to 0.\end{aligned}$$

An application of Slutsky's theorem (Karr, 1993, P. 146) gives

$$\frac{M_{k_n} - W_0(n)}{\rho_n^{k_n}\sqrt{\mathbf{Var}[W_{k_n}]}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1). \quad \square$$

5.5 Illustrative examples

We give a few illustrating examples with subcritical and critical sampling. We also explore degenerate cases with extremal and nonextremal initial conditions.

5.5.1 An example with a fixed generator

Let $S_j \equiv s$, a fixed positive integer ($\sigma_S^2(n) = 0$). This is a degenerate case. If $W_0(n) = \lfloor \frac{2n}{3} \rfloor$, the case is nonextremal. According to Theorem 2, after $k_n = \lceil \ln n \rceil$

draws we have

$$\frac{W_{\lceil \ln n \rceil} - \frac{2}{3}n + \frac{1}{3}s\ln n}{\sqrt{\ln n}} \xrightarrow{\mathcal{D}} \mathcal{N}\Big(0, \frac{8}{9}s\Big).$$

For this degenerate case, according to Theorem 2 after $\lfloor \lambda n \rfloor$ draws we have,

$$\frac{W_{\lfloor \lambda n \rfloor} - (\frac{1}{6}e^{-2\lambda} + \frac{1}{2})n}{\sqrt{n}} \xrightarrow{\mathcal{D}} \mathcal{N}\Big(0, \frac{1}{4}(1 - (\frac{4}{9}\lambda + 1)e^{-4\lambda}\Big),$$

and after n^2 draws

$$\frac{W_{n^2} - \frac{1}{2}n}{\sqrt{n}} \xrightarrow{\mathcal{D}} \mathcal{N}\Big(0, \frac{1}{4}\Big).$$

However, in the extremal case of $n^{-1}W_0(n) \to 0$, say if $W_0(n) = \lfloor 2\ln n \rfloor$, and $k_n = \lfloor n^{3/4} \rfloor$, we have

$$\frac{W_{\lfloor n^{3/4} \rfloor} - (2\ln(n) + sn^{\frac{3}{4}} - s^2 n^{\frac{1}{2}} + \frac{2}{3}s^3 n^{\frac{1}{4}})}{\sqrt{n^{\frac{1}{2}}}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 2s^2).$$

5.5.2 An example with a uniformly bounded generator

Let the generator S be $3 + \text{Bin}(8, \frac{1}{3})$. Suppose the starting split is such that $W_0(n) = \lfloor \frac{1}{9}n \rfloor$. According to Theorem 2, after $k_n = \lceil \ln\ln n \rceil$ draws we have

$$\frac{W_{\lceil \ln\ln n \rceil} - (\frac{1}{9}n + \frac{119}{27}\ln\ln n)}{\sqrt{\ln\ln n}} \xrightarrow{\mathcal{D}} \mathcal{N}\Big(0, \frac{2416}{729}\Big);$$

after $2n$ draws we have,

$$\frac{W_{2n} - (\frac{1}{2} - \frac{7}{18}e^{-\frac{68}{3}})n}{\sqrt{n}} \xrightarrow{\mathcal{D}} \mathcal{N}\Big(0, \frac{1}{4} - \frac{289}{36}e^{-\frac{136}{3}}\Big).$$

Note how close the number $\frac{1}{2} - \frac{7}{18}e^{-\frac{68}{3}}$ is to $\frac{1}{2}$, and how close $\frac{1}{4} - \frac{289}{36}e^{-\frac{136}{3}}$ is to $\frac{1}{4}$. While $2n$ is near the lower edge of the linear phase, the case is already close to the superlinear phase.

5.5.3 An example with a generator with a growing range

Suppose the generating distribution is uniformly distributed on the three-point set $\{7, \lfloor n^{0.3} \rfloor, 2\lfloor n^{0.3} \rfloor\}$, $W_0(n) = \lceil \frac{1}{5} n \rceil$, and $k_n = \lceil 2n^{0.35} + 13 \ln n \rceil$. In this case, $s_n = 2\lfloor n^{0.3} \rfloor$, $\mu_S(n) \sim n^{0.3}$, $\sigma_S^2(n) \sim n^{0.3}$. As $\mu_S(n) k_n \sim 2n^{0.65} = o(n)$, the case is sublinear. According to Theorem 2

$$\frac{W_{\lceil 2n^{0.35}+13\ln n \rceil} - \frac{1}{5}n - \frac{6}{5}n^{0.65}}{n^{0.325}} \xrightarrow{\mathcal{D}} \mathcal{N}\Big(0, \frac{104}{25}\Big),$$

and according to Theorem 2, after $k_n = \lceil n^{0.7} \rceil$ draws

$$\frac{W_{\lceil n^{0.7} \rceil} - (\frac{1}{2} - \frac{3}{10}e^{-2})n}{\sqrt{n}} \xrightarrow{\mathcal{D}} \mathcal{N}\Big(0, \frac{1 - e^{-4}}{4}\Big).$$

5.5.4 An example with critical sampling

If the sampling distribution is critical, balance occurs very rapidly. Consider, for example, the critical case where the generating distribution is $\mathrm{Bin}(n, \frac{1}{2})$.

It is clear that the process is a Markov chain, with stationary $\mathrm{Bin}\big(n, \frac{1}{2}\big)$ distribution. We shall show that, regardless of the starting split in the two countries, $W_1(n)$ is distributed like a $\mathrm{Bin}(n, \frac{1}{2})$ random variable. Thus, in one step the process enters its stationary distribution, and of course, stays in it thereafter. In other words $W_j(n) \overset{\mathcal{L}}{=} \mathrm{Bin}(n, \frac{1}{2})$, for all $j \geq 1$. To show this, we calculate

$$\begin{aligned}
\mathbf{Prob}\big(W_1(n) = r\big) &= \sum_{W_0(n)-k+(k-i)=r} \mathbf{Prob}(S_1 = i, S_{1,W} = k) \\
&= \sum_{W_0(n)+i-2k=r} \mathbf{Prob}\big(S_{1,W} = k \,|\, S_1 = i\big)\, \mathbf{Prob}(S_1 = i)
\end{aligned}$$

$$= \frac{1}{2^n} \sum_{W_0(n)+i-2k=r} \frac{\binom{W_0(n)}{k}\binom{n-W_0(n)}{i-k}}{\binom{n}{i}} \times \binom{n}{i}$$

$$= \frac{1}{2^n} \sum_{m=0}^{W_0(n)} \binom{W_0(n)}{m}\binom{n-W_0(n)}{n-r-m)}.$$

Recall that

$$\frac{1}{\binom{n}{n-r}} \sum_{m=0}^{W_0(n)} \binom{W_0(n)}{m}\binom{n-W_0(n)}{n-r-m}$$

is the sum of all the probabilities for Hypergeo$(n, n-r, W_0(n))$, and thus must be 1. Hence,

$$\mathbf{Prob}\big(W_1(n) = r\big) = \frac{1}{2^n}\binom{n}{r},$$

and $W_1(n) \stackrel{\mathcal{L}}{=} \text{Bin}(n, \frac{1}{2})$. It follows that, for all $j \geq 1$, $W_j(n) \stackrel{\mathcal{L}}{=} \text{Bin}(n, \frac{1}{2})$. Note that there are no phases here—after any number of draws we have the same binomial distribution, with

$$\frac{W_j(n) - \frac{1}{2}n}{\sqrt{n}} \xrightarrow{\mathcal{D}} \mathcal{N}\Big(0, \frac{1}{4}\Big),$$

as in the superlinear phase. The sample size is too large, inducing very rapid mixing that readily puts the process in a well mixed position.

Chapter 6

A Generalized Bernoulli-Laplace urn model

Consider a generalized Bernoulli-Laplace urn model in which balls of i.i.d. random sample sizes, $S_j, j = 0, 1, 2, \ldots$, are transferred at each time epoch between two urns A and B, with urn A containing $a(n)$ white balls and urn B containing $b(n)$ red balls at the beginning. We let $a(n)$, $b(n)$ and the number of draws k_n be increasing functions of n and study the composition of the model after k_n steps of mixing as n grows. We derive the exact mean and variance for the difference of the proportions of white balls, D_{k_n}, in each of the two urns. This serves as a vehicle to understand the composition of the urns. For instance, the number of white balls in either urn can be easily extracted from this difference at any step. With a concept of linearity of k_n, based on a combination of the sample size S_j and the total number of balls n, and as n goes to infinity, we show that D_{k_n} follows a Gaussian distribution in properly defined sublinear, linear, and superlinear phases using the martingale central limit theorem. We give interpretations for how the results in different phases conjoin at the seam lines (boundaries between phases).

6.1 Generalized Bernoulli-Laplace model

We first introduce the generalized Bernoulli-Laplace model and show some preliminary results for the model.

6.1.1 Introduction

Consider two urns A and B. Initially we have $a(n)$ white balls in urn A, $b(n)$ red balls in urn B. At the jth time epoch, instead of picking just one ball at a time, which is the classic Bernoulli-Laplace model (Diaconis, 1987, and Donnelly, Lloyd and Sudbury, 1994), a sample of balls of random size S_j is picked randomly from both urn A and a sample of size S_j is picked randomly from urn B. The sample taken out of urn A is deposited in urn B and the sample taken out of urn B is deposited in urn A. The process goes on for $j = 1, 2, 3, \ldots$ steps.

Let $S_j := S_j(n)$ denote the sample size generating random variable or simply the *generator*. We consider the case when $S_j, j = 1, 2, 3, \ldots$ are *i.i.d.* with mean $\mathbf{E}[S(n)] = \mu_S(n)$, and variance $\mathbf{Var}[S(n)] = \sigma_S^2(n)$. The *index of dispersion* (Cox and Lewis, 1966, and Perry and Mead, 1979) of the random variable is then $\delta_S(n) = \frac{\sigma_S^2(n)}{\mu_S(n)}$; let s_n be the range of S_j. We assume s_n to be small relative to the sizes of the urns, i.e. $s_n = o(\sqrt{\min\{a(n), b(n)\}})$. We shall call this *subcritical sampling*. The cases when s_n is of higher order will be called *critical sampling*. Let $S_{A,j}^W$ and $S_{B,j}^W$ be the number of white balls in the jth random samples from urn A and urn B, respectively; $S_{A,j}^R$ and $S_{B,j}^R$ be the number of red balls in the jth random samples from urn A and urn B, respectively. Then,

$$\begin{aligned} S_j &= S^W_{A,j} + S^R_{A,j} \\ &= S^W_{B,j} + S^R_{B,j}, \\ W_{A,j} &= W_{A,j-1} + S^W_{B,j} - S^W_{A,j}. \end{aligned}$$

In the generalized Bernoulli-Laplace model, at any time epoch, the composition of the two urns can be determined by the number of white balls in urn A, $W_{A,j}$, so we will focus on the white balls in the samples picked, i.e. $S^W_{A,j}$ and $S^W_{B,j}$. The conditional distribution of $S^W_{A,j}$ and $S^W_{B,j}$ are

$$\begin{aligned} S^W_{A,j} \,|\, \mathcal{F}_{j-1}, S_j &\overset{\mathcal{L}}{=} \text{Hypergeo}(a(n), S_j, W_{A,j-1}), \\ S^W_{B,j} \,|\, \mathcal{F}_{j-1}, S_j &\overset{\mathcal{L}}{=} \text{Hypergeo}(b(n), S_j, a(n) - W_{A,j-1}). \end{aligned}$$

The conditional moments are

$$\begin{aligned} \mathbf{E}[S^W_{A,j} \,|\, \mathcal{F}_{j-1}, S_j] &= \frac{S_j W_{A,j-1}}{a(n)}, \\ \mathbf{E}[S^W_{B,j} \,|\, \mathcal{F}_{j-1}, S_j] &= \frac{S_j (a(n) - W_{A,j-1})}{b(n)}, \end{aligned}$$

$$\begin{aligned} \mathbf{E}[(S^W_{A,j})^2 \,|\, \mathcal{F}_{j-1}, S_j] &= \frac{S_j W_{A,j-1}(a(n) - S_j)(a(n) - W_{A,j-1})}{a^2(n)(a(n)-1)} \\ &\quad + \Big(\frac{S_j W_{A,j-1}}{a(n)}\Big)^2, \end{aligned}$$

$$\begin{aligned}\mathbf{E}[(S_{B,j}^W)^2 \,|\, \mathcal{F}_{j-1}, S_j] &= \frac{1}{b^2(n)} S_j(a(n) - W_{A,j-1})(b(n) - S_j) \\ &\quad \times (b(n) - a(n) + W_{A,j-1})(a(n) - W_{A,j-1}) \\ &\quad + \Big(\frac{S_j(a(n) - W_{A,j-1})}{b(n)}\Big)^2,\end{aligned}$$

$$\begin{aligned}\mathbf{E}[S_{A,j}^W S_{B,j}^W \,|\, \mathcal{F}_{j-1}, S_j] &= \mathbf{E}[S_{A,j}^W \,|\, \mathcal{F}_{j-1}, S_j]\mathbf{E}[S_{B,j}^W \,|\, \mathcal{F}_{j-1}, S_j] \\ &= \frac{S_j^2 W_{A,j-1}(a(n) - W_{A,j-1})}{a(n)b(n)}.\end{aligned}$$

In the classic Bernoulli-Laplace model, when $S \equiv 1$, let $I_{A,j}^W$ and $I_{B,j}^W$ be the indicators of picking a white ball in the jth step from urn A and urn B respectively. In this case, $S_{A,j}^W = I_{A,j}^W, S_{B,j}^W = I_{B,j}^W$, and we have

$$\begin{aligned}I_{A,j}^W \,|\, \mathcal{F}_{j-1} &\overset{\mathcal{L}}{=} \text{Ber}\Big(\frac{W_{A,j-1}}{a(n)}\Big), \\ I_{B,j}^W \,|\, \mathcal{F}_{j-1} &\overset{\mathcal{L}}{=} \text{Ber}\Big(\frac{a(n) - W_{A,j-1}}{b(n)}\Big).\end{aligned}$$

Let n grow and let $a(n)$, $b(n)$ and k_n be increasing functions of n. We are interested in understanding the asymptotic distributions of the composition of the urns as k_n grows to infinity in a sublinear, linear and superlinear manner, which we define as:

1. The sublinear phase, when $\mu_S(n)k_n = o(\min\{a(n), b(n)\})$, certain appropriate regularity conditions are sufficient for Gaussian laws in upper sublinear phases:

   (a) For the nondegenerate case, where $\sigma_S^2(n) \nrightarrow 0$, we develop a limit Gaussian law under the condition that

   $$\frac{s_n^2}{\sigma_S^2(n)} = o(k_n).$$

   (b) For the degenerate case, where $\sigma_S^2(n) \to 0$, we develop a limit Gaussian law under the condition that

$$\frac{s_n\sqrt{\min\{a(n), b(n)\}}}{\mu_S(n)} = o(k_n).$$

2. The linear phase, when $\mu_S(n)k_n \sim \lambda_n \min\{a(n), b(n)\}$, for some $\lambda_n > 0$ of a magnitude bounded by two constant $0 < M_1 \le \lambda_n \le M_2 < \infty$ from below and above.

3. The superlinear phase, when $\min\{a(n), b(n)\} = o(\mu_S(n)k_n)$.

We rely on martingale central limit theorem to find the limiting distributions within each phase.

6.1.2 Moments

In order to utilize the martingale central limit theorem, we first find the exact mean and variance of $W_{A,j}$:

$$\begin{aligned}
\mathbf{E}[W_{A,j} \,|\, \mathcal{F}_{j-1}, S_j] &= \mathbf{E}[W_{A,j-1} + S^W_{B,j} - S^W_{A,j} \,|\, \mathcal{F}_{j-1}, S_j] \\
&= W_{A,j-1} + \frac{a(n) - W_{A,j-1}}{b(n)} S_j - \frac{W_{A,j-1}}{a(n)} S_j \\
&= W_{A,j-1}\Big(1 - \frac{S_j}{a(n)} - \frac{S_j}{b(n)}\Big) + \frac{a(n)}{b(n)} S_j.
\end{aligned}$$

It follows that

$$\begin{aligned}
\mathbf{E}[W_{A,j} \,|\, \mathcal{F}_{j-1}] &= \mathbf{E}[\mathbf{E}[W_{A,j} \,|\, \mathcal{F}_{j-1}, S_j]] \\
&= W_{A,j-1}\Big(1 - \frac{\mu_S(n)}{a(n)} - \frac{\mu_S(n)}{b(n)}\Big) + \frac{a(n)}{b(n)} \mu_S(n).
\end{aligned}$$

Take expectation and by induction, we have

$$\mathbf{E}[W_{A,j}] \;=\; \mathbf{E}[W_j] = \frac{a(n)b(n)}{a(n)+b(n)}\Big(1-\frac{\mu_S(n)}{a(n)}-\frac{\mu_S(n)}{b(n)}\Big)^j + \frac{a^2(n)}{a(n)+b(n)}.$$

Similarly, for the second moment,

$$\begin{aligned}\mathbf{E}[W_{A,j}^2 \,|\, \mathcal{F}_{j-1}, S_j] &= \mathbf{E}[(W_{A,j-1}+S_{B,j}^W - S_{A,j}^W)^2 \,|\, \mathcal{F}_{j-1}, S_j] \\ &= W_{A,j-1}^2 + \mathbf{E}[(S_{A,j}^W)^2 \,|\, \mathcal{F}_{j-1}, S_j] + \mathbf{E}[(S_{B,j}^W)^2 \,|\, \mathcal{F}_{j-1}, S_j] \\ &\quad -2W_{A,j-1}\mathbf{E}[S_{B,j}^W \,|\, \mathcal{F}_{j-1}, S_j] + 2W_{A,j-1}\mathbf{E}[S_{A,j}^W \,|\, \mathcal{F}_{j-1}, S_j] \\ &\quad -2\mathbf{E}[S_{A,j}^W \,|\, \mathcal{F}_{j-1}, S_j]\mathbf{E}[S_{B,j}^W \,|\, \mathcal{F}_{j-1}, S_j].\end{aligned}$$

Take double expectation and by induction, we can get $\mathbf{E}[W_{A,j-1}^2]$ and we just show here the formula for the variance, which is

$$\begin{aligned}\mathbf{Var}[W_{A,j}] &= \frac{a(n)b(n)(a(n)-1)(b(n)-1)}{(a(n)+b(n)-1)(a(n)+b(n)-2)} \\ &\quad \times\Big(1-\frac{a(n)+b(n)-1}{a(n)b(n)(a(n)-1)(b(n)-1)} \\ &\quad\quad \times\Big((2a(n)b(n)-a(n)-b(n))\mu_S(n) \\ &\quad\quad\quad -(a(n)+b(n)-2)(\mu_S^2(n)+\sigma_S^2(n))\Big)\Big)^j \\ &\quad -\frac{a^2(n)b^2(n)}{(a(n)+b(n))^2}\Big(1-\frac{\mu_S(n)}{a(n)}-\frac{\mu_S(n)}{b(n)}\Big)^{2j} \\ &\quad +\frac{a(n)b(n)(a(n)-b(n))^2}{(a(n)+b(n)-2)(a(n)+b(n))^2}\Big(1-\frac{\mu_S(n)}{a(n)}-\frac{\mu_S(n)}{b(n)}\Big)^j \\ &\quad +\frac{a^2(n)b^2(n)}{(a(n)+b(n))^2(a(n)+b(n)-1)},\end{aligned}$$

where $\mathbf{Var}[W_{A,1}] = \sigma_S^2(n)$.

For the classic Bernoulli-Laplace model, take $a(n) = b(n) = n, \mu_S(n) = 1$, and

$\sigma_S^2(n) = 0$, the exact mean and variance become

$$\begin{aligned}
\mathbf{E}[W_{A,j}] &= \frac{n}{2}\Big(1 - \frac{2}{n}\Big)^j + \frac{n}{2}, \\
\mathbf{Var}[W_{A,j}] &= \frac{n^2}{4}\Big(\frac{2n-2}{2n-1}\Big(1 - \frac{2(2n-1)}{n^2}\Big)^j - \Big(1 - \frac{2}{n}\Big)^{2j} + \frac{1}{2n-1}\Big),
\end{aligned}$$

which has been shown in Diaconis and Shahshahani (1987). Notice that when $j = 1$, $\mathbf{Var}[W_{A,1}] = 0$, there is no variability in the first step.

6.1.3 Asymptotical limits under phases

Define

$$D_j := \frac{W_{A,j}}{a(n)} - \frac{W_{B,j}}{b(n)} = W_{A,j}\Big(\frac{1}{a(n)} + \frac{1}{b(n)}\Big) - \frac{a(n)}{b(n)}.$$

This random variable represents the difference of the proportions of the white balls in the two urns, which is expected to go from 1 at the start to 0 after long-term mixing. The inverse relation is between D_j and $W_{A,j}$ is

$$W_{A,j} = \frac{a(n)b(n)}{a(n)+b(n)}D_j + \frac{a^2(n)}{a(n)+b(n)}.$$

We can obtain the exact and asymptotic expectation and variance of D_j from $W_{A,j}$. Because $a(n)$ is of the same order as $b(n)$, $\frac{a(n)}{b(n)} \to 0$ or $\frac{b(n)}{a(n)} \to 0$, the expressions of the asymptotics are essentially the same. Without loss of generality, we will focus on the situation when $a(n)$ is of the same order of $b(n)$. We obtain the exact first moment of D_j from $W_{A,j}$,

$$\mathbf{E}[D_j] = \mathbf{E}\Big[\frac{W_{A,j}}{a(n)} - \frac{W_{B,j}}{b(n)}\Big]$$

$$
\begin{aligned}
&= \mathbf{E}\Big[W_{A,j}\Big(\frac{1}{a(n)} + \frac{1}{b(n)}\Big) - \frac{a(n)}{b(n)}\Big] \\
&= \Big(1 - \frac{\mu_S(n)}{a(n)} - \frac{\mu_S(n)}{b(n)}\Big)^j \\
&= \rho_n^j.
\end{aligned}
$$

We then give the specific asymptotic expression of $\mathbf{E}[D_j]$ in the three phases:

1. In the growing sublinear phase we have

$$
\mathbf{E}[D_{k_n}] = 1 + o(1).
$$

2. In the linear phase we have

$$
\mathbf{E}[D_{k_n}] = \exp\Big(-\frac{\lambda_n n \mu_S(n)}{a(n)} + \frac{\lambda_n n \mu_S(n)}{b(n)}\Big) + o(1).
$$

3. In the superlinear phase we have

$$
\mathbf{E}[D_{k_n}] = o(1).
$$

In a similar way, we can obtain the exact variance of D_j,

$$
\begin{aligned}
\mathbf{Var}[D_j] &= \mathbf{Var}[\frac{W_{A,j}}{a(n)} - \frac{W_{B,j}}{b(n)}] \\
&= \mathbf{Var}[W_j\Big(\frac{1}{a(n)} + \frac{1}{b(n)}\Big) - \frac{a(n)}{b(n)}] \\
&= \Big(\frac{1}{a(n)} + \frac{1}{b(n)}\Big)^2 \mathbf{Var}[W_{A,j}]
\end{aligned}
$$

$$
\begin{aligned}
= \;& \Big(1-\frac{1}{a(n)}-\frac{1}{b(n)}-\frac{1}{a(n)b(n)}\Big)\Big(1-\frac{1}{a(n)+b(n)}\Big)\Big(1-\frac{2}{a(n)+b(n)}\Big)\\
&\times\Big(1-\Big(\frac{1}{a(n)}+\frac{1}{b(n)}+\frac{1}{a(n)-1}+\frac{1}{b(n)-1}\Big)\mu_S(n)\\
&\quad+\Big(\frac{1}{a(n)}+\frac{1}{b(n)}-\frac{1}{a(n)b(n)}\Big)\\
&\quad\times\Big(\frac{1}{a(n)-1}+\frac{1}{b(n)-1}\Big)\big(\mu_S^2(n)+\sigma_S^2(n)\big)\Big)^j\\
&-\Big(1-\frac{\mu_S}{a(n)}-\frac{\mu_S}{b(n)}\Big)^{2j}\\
&+\frac{(a(n)-b(n))^2}{a(n)b(n)(a(n)+b(n)-2)}\Big(1-\frac{\mu_S}{a(n)}-\frac{\mu_S}{b(n)}\Big)^j\\
&+\frac{1}{a(n)+b(n)-1}.
\end{aligned}
$$

Conducting an asymptotical analysis of the variance formula, we find the following phases in the variance:

1. In the growing sublinear phase we have, for $\sigma_S^2(n) \nrightarrow 0$,

$$
\mathbf{Var}[D_{k_n}] \sim k_n\sigma_S^2(n)\Big(\frac{1}{a(n)}+\frac{1}{b(n)}\Big)^2,
$$

and for $\sigma_S^2(n) \to 0$,

$$
\mathbf{Var}[D_{k_n}] \sim \frac{\mu_S^2(n)k_n^2}{2}\Big(\frac{1}{a(n)}+\frac{1}{b(n)}\Big)^3.
$$

2. In the linear phase we have

$$
\begin{aligned}
\mathbf{Var}[D_{k_n}] \;\sim\; & \frac{1}{a(n)+b(n)} \frac{1}{\exp\Big(2k_n\mu_S(n)\Big(\frac{1}{a(n)}+\frac{1}{b(n)}\Big)\Big)} \\
& \times\Big(\exp\Big(2k_n\mu_S(n)\Big(\frac{1}{a(n)}+\frac{1}{b(n)}\Big)\Big) \\
& +\Big(\frac{a(n)}{b(n)}+\frac{b(n)}{a(n)}-2\Big)\exp\Big(k_n\mu_S(n)\Big(\frac{1}{a(n)}+\frac{1}{b(n)}\Big)\Big) \\
& +\Big(\frac{3k_n}{a(n)}+\frac{a(n)k_n}{b^2(n)}+\frac{3k_n}{b(n)}+\frac{b(n)k_n}{a^2(n)}\Big)\sigma_S^2(n) \\
& +\Big(-\frac{k_n}{a(n)}-\frac{k_n}{b(n)}-\frac{a(n)k_n}{b^2(n)}-\frac{b(n)k_n}{a^2(n)}\Big)\mu_S(n) \\
& -\frac{a(n)}{b(n)}-\frac{b(n)}{a(n)}+1\Big).
\end{aligned}
$$

3. In the superlinear phase we have

$$\mathbf{Var}[D_{k_n}] \sim \frac{1}{a(n)+b(n)}.$$

6.2 The main result

In this section, we first state the main theorem, then construct a martingale formulation. In the rest of the section we prove the main theorem within three specific phases via the martingale central limit theorem. (These results are reported in Zhang, 2011+.)

6.2.1 The main theorem

For simplicity, we assume $a(n) = b(n) = n$, so $\rho_n^j = \Big(1 - \frac{2\mu_S(n)}{n}\Big)^j$.

Theorem 3 *Let D_{k_n} be the difference of proportions of the white balls in a generalized Bernoulli-Laplace urn model under subcritical sampling, where $k_n \to \infty$, in an upper* *sublinear, linear or superlinear manner. Then,*

$$\frac{D_{k_n} - \mathbf{E}[D_{k_n}]}{\sqrt{\mathbf{Var}[D_{k_n}]}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1),$$

where

$$\begin{aligned} \mathbf{E}[D_{k_n}] &= \Big(1 - \frac{2\mu_S(n)}{n}\Big)^{k_n}, && (6.1) \\ \mathbf{Var}[D_{k_n}] &= \frac{2(n-1)}{2n-1}\Big(1 + \frac{1}{n^2(n-1)^2}\Big((2n-1)(2n-2)(\mu_S^2(n) + \sigma_S^2(n)) \\ &\quad -2n(2n-1)(n-1)\mu_S(n)\Big)\Big)^{k_n} - \Big(1 - \frac{2\mu_S(n)}{n}\Big)^{2k_n} + \frac{1}{2n-1}. && (6.2) \end{aligned}$$

The result depends on the generator only through its mean and variance. All generators with same mean and variance exert the same influence on the long-term mixing.

6.2.2 A martingale formulation

There is stochastic dependence between D_j and D_{j-1},

$$\begin{aligned} D_j \,|\, \mathcal{F}_{j-1}, S_j &= \frac{W_{A,j}}{n} - \frac{W_{B,j}}{n} \\ &= \frac{W_{A,j-1} + S^W_{B,j} - S^W_{A,j}}{n} - \frac{W_{B,j-1} + S^W_{A,j} - S^W_{B,j}}{n} \\ &= \frac{W_{A,j-1}}{n} - \frac{W_{B,j-1}}{n} + \frac{S^W_{B,j} - S^W_{A,j}}{n} - \frac{S^W_{A,j} - S^W_{B,j}}{n} \\ &= D_{j-1} - \frac{2}{n}\big(S^W_{A,j} - S^W_{B,j}\big). \end{aligned}$$

First, we construct a martingale based on $\{D_j, j \geq 0\}$.

Lemma 11 *For* $j = 0, 1, \ldots,$

$$M_j = \frac{D_j}{\rho_n^j}$$

is a martingale.

Proof.

$$\begin{aligned}
\mathbf{E}[D_j|\mathcal{F}_{j-1}] &= \mathbf{E}_{S_j|\mathcal{F}_{j-1}}\Big[\mathbf{E}\Big[D_{j-1} - \frac{2(S^W_{A,j} - S^W_{B,j})}{n}\,\Big|\,\mathcal{F}_{j-1}, S_j\Big]\,\Big|\,\mathcal{F}_{j-1}\Big] \\
&= D_{j-1} - \frac{2}{n}\Big(\frac{\mu_S(n)W_{A,j-1}}{n} - \frac{\mu_S(n)(n - W_{A,j-1})}{n}\Big) \\
&= \rho_n D_{j-1},
\end{aligned}$$

so that,

$$\mathbf{E}\Big[\frac{D_j}{\rho_n^j}\,\Big|\,\mathcal{F}_{j-1}\Big] = \frac{D_{j-1}}{\rho_n^{j-1}}.$$

So, $\{M_j\}_{j=0}^{\infty}$ is a martingale with $M_0 = 1$. $\square$

The central limit theorem applies to the centered martingales (with mean 0); let us introduce the centered martingale

$$\tilde{M}_j(n) = \tilde{M}_j = M_j - M_0 = M_j - 1,$$

with martingale differences

$$\nabla\tilde{M}_j = (M_j - 1) - (M_{j-1} - 1) = M_j - M_{j-1}.$$

The following lemmas will help us check conditional Lindeberg's condition in all the phases.

Lemma 12

$$\big|\nabla\tilde{M}_j\big| \le \frac{6s_n}{\rho_n^j n}.$$

Proof. We have

$$\begin{aligned}
\left|\nabla\tilde{M}_j\right| &= |M_j - M_{j-1}| \\
&= \Big|\frac{D_j}{\rho_n^j} - \frac{D_{j-1}}{\rho_n^{j-1}}\Big| \\
&= \frac{1}{\rho_n^j}\Big|\Big(\frac{W_{A,j}}{n} - \frac{W_{B,j}}{n}\Big) - \rho_n\Big(\frac{W_{A,j-1}}{n} - \frac{W_{B,j-1}}{n}\Big)\Big| \\
&= \frac{1}{\rho_n^j}\Big|\frac{W_{A,j} - \rho_n W_{A,j-1}}{n} - \frac{W_{B,j} - \rho_n W_{B,j-1}}{n}\Big| \\
&= \frac{1}{\rho_n^j}\Big|\frac{S_{B,j}^W - S_{A,j}^W + \Big(\frac{2\mu_S(n)}{n}\Big)W_{A,j-1}}{n} - \frac{S_{A,j}^W - S_{B,j}^W + \Big(\frac{2\mu_S(n)}{n}\Big)W_{B,j-1}}{n}\Big| \\
&= \frac{1}{\rho_n^j}\Big|(S_{B,j}^W - S_{A,j}^W)\Big(\frac{2}{n}\Big) + \Big(\frac{2\mu_S(n)}{n}\Big)\Big(\frac{W_{A,j-1}}{n} + \frac{W_{B,j-1}}{n}\Big)\Big| \\
&\leq \frac{1}{\rho_n^j}\Big|s_n \times \frac{2}{n} + \frac{2s_n}{n} \times 1\Big| \\
&\leq \frac{6s_n}{\rho_n^j n}. \quad \square
\end{aligned}$$

The conditional Lindeberg's condition for all three phases is checked as follows.

Lemma 13 *We have*

$$U_n = \sum_{j=1}^{k_n} \mathbf{E}\Big[\Big(\frac{\nabla\tilde{M}_j}{\sqrt{\mathbf{Var}[M_{k_n}]}}\Big)^2 \mathbf{1}_{\Big\{\Big|\frac{\nabla\tilde{M}_j}{\sqrt{\mathbf{Var}[M_{k_n}]}}\Big| > \varepsilon\Big\}} \Big| \mathcal{F}_{j-1}\Big] \xrightarrow{\mathcal{P}} 0.$$

Proof. Choose any $\varepsilon > 0$. Under the assumption of subcritical sampling we have $s_n = o(\sqrt{n})$, and thus the uniform bound in Lemma 12 asserts that the sets $\{|\nabla\tilde{M}_j| > \varepsilon\sqrt{\mathbf{Var}[M_{k_n}]}\ \}$ are all empty, for all n greater than some positive integer $n_0(\varepsilon)$. In the linear and super-linear phases, we can show that $\mathbf{Var}[D_{k_n}] \geq \frac{1}{n} + o(\frac{1}{n})$, and under the conditions in the upper sublinear phase, we have

$$
\begin{aligned}
U_n &= \sum_{j=1}^{n_0(\varepsilon)} \mathbf{E}\Big[\Big(\frac{\nabla \tilde{M}_j}{\sqrt{\mathbf{Var}[M_{k_n}]}}\Big)^2 \mathbf{1}_{\Big\{\Big|\frac{\nabla \tilde{M}_j}{\sqrt{\mathbf{Var}[M_{k_n}]}}\Big| > \varepsilon\Big\}} \Big| \mathcal{F}_{j-1}\Big] \\
&\leq \frac{1}{\mathbf{Var}[M_{k_n}]} \sum_{j=1}^{n_0(\varepsilon)} \mathbf{E}\big[\big(\nabla \tilde{M}_j\big)^2 \,\big|\, \mathcal{F}_{j-1}\big] \\
&\leq \frac{\frac{36 n_0(\varepsilon) s_n^2}{\rho_n^{2k_n} n^2}}{\frac{\mathbf{Var}[D_{k_n}]}{\rho_n^{2k_n}}} \\
&\to 0, \qquad \text{as } n \to \infty.
\end{aligned}
$$

In order to prove the conditional variance condition, we want to find the expression for the V_n.

Lemma 14

$$
\begin{aligned}
V_n &= \frac{\rho_n^{2k_n}}{\mathbf{Var}[D_{k_n}]}\Big(\frac{2(-n\mu_S(n) + 2n\sigma_S^2(n) - \sigma_S^2(n) + \mu_S^2(n))}{n^2(n-1)} \sum_{j=0}^{k_n} \frac{1}{\rho_n^{2j}} D_{j-1}^2 \\
&\quad - \frac{2(-n\mu_S(n)) + \sigma_S^2(n) + \mu_S^2(n))}{n^2(n-1)} \sum_{j=0}^{k_n} \frac{1}{\rho_n^{2j}}\Big).
\end{aligned} \tag{6.3}
$$

Proof.

Taking square of the D_{k_n} after noticing that

$$
\begin{aligned}
\mathbf{E}[S_{A,j}^W - S_{B,j}^W | \mathcal{F}_{j-1}, S_j] &= S_j\Big(\frac{W_{A,j-1}}{n} - \frac{n - W_{A,j-1}}{n}\Big) \\
&= S_j D_{j-1},
\end{aligned}
$$

we arrive at

$$
\begin{aligned}
\mathbf{E}[D_j^2 \,|\, \mathcal{F}_{j-1}, S_j] &= \mathbf{E}\Big[\Big(D_{j-1} - \frac{2}{n}(S_{A,j}^W - S_{B,j}^W)\Big)^2 \,\Big|\, \mathcal{F}_{j-1}, S_j\Big] \\
&= D_{j-1}^2 - \frac{4}{n} D_{j-1} \mathbf{E}[S_{A,j}^W - S_{B,j}^W \,|\, \mathcal{F}_{j-1}, S_j] \\
&\quad + \mathbf{E}\Big[\Big(\frac{2}{n}(S_{B,j}^W - S_{A,j}^W)\Big)^2 \,\Big|\, \mathcal{F}_{j-1}, S_j\Big] \\
&= D_{j-1}^2 - \frac{4}{n} S_j D_{j-1}^2 + \Big(\frac{2}{n}\Big)^2 \Big(\mathbf{E}[(S_{A,j}^W)^2 \,|\, \mathcal{F}_{j-1}, S_j] \\
&\quad + \mathbf{E}[(S_{B,j}^W)^2 \,|\, \mathcal{F}_{j-1}, S_j] - 2\mathbf{E}[S_{A,j}^W S_{B,j}^W \,|\, \mathcal{F}_{j-1}, S_j]\Big) \\
&= \Big(1 - \frac{4S_j}{n}\Big) D_{j-1}^2 + \frac{2(2n+1)S_j^2 - 2nS_j}{n^2(n-1)} D_{j-1}^2 \\
&\quad + \frac{1}{n^2(n-1)}(2nS_j^2 - 2S_j),
\end{aligned}
$$

which implies

$$
\begin{aligned}
V_n &= \sum_{j=0}^{k_n} \mathbf{E}\Big[\Big(\frac{\nabla \tilde{M}}{\sqrt{\mathbf{Var}[M_{k_n}]}}\Big)^2 \,\Big|\, \mathcal{F}_{j-1}\Big] \\
&= \frac{\rho_n^{2k_n}}{\mathbf{Var}[D_{k_n}]} \sum_{j=0}^{k_n} \mathbf{E}[\big(\nabla \tilde{M}\big)^2 \,|\, \mathcal{F}_{j-1}] \\
&= \frac{\rho_n^{2k_n}}{\mathbf{Var}[D_{k_n}]} \sum_{j=0}^{k_n} \Big(\mathbf{E}_{S_j \,|\, \mathcal{F}_{j-1}}[\mathbf{E}[M_j^2 \,|\, \mathcal{F}_{j-1}, S_j] \,|\, \mathcal{F}_{j-1}] - M_{j-1}^2\Big) \\
&= \frac{\rho_n^{2k_n}}{\mathbf{Var}[D_{k_n}]} \sum_{j=0}^{k_n} \Big(\mathbf{E}_{S_j \,|\, \mathcal{F}_{j-1}}\Big[\mathbf{E}\Big[\Big(\frac{D_j}{\rho_n^j}\Big)^2 \,\Big|\, \mathcal{F}_{j-1}, S_j\Big] \,\Big|\, \mathcal{F}_{j-1}\Big] - \Big(\frac{D_{j-1}}{\rho_n^{j-1}}\Big)^2\Big) \\
&= \frac{\rho_n^{2k_n}}{\mathbf{Var}[D_{k_n}]} \Big(\frac{2(-n\mu_S(n) + 2n\sigma_S^2(n) - \sigma_S^2(n) + \mu_S^2(n))}{n^2(n-1)} \sum_{j=0}^{k_n} \frac{1}{\rho_n^{2j}} D_{j-1}^2 \\
&\quad - \frac{2(-n\mu_S(n)) + \sigma_S^2(n) + \mu_S^2(n))}{n^2(n-1)} \sum_{j=0}^{k_n} \frac{1}{\rho_n^{2j}}\Big). \quad \square
\end{aligned}
$$

6.2.3 The growing sublinear phase

Recall that $a(n) = b(n)$, in this phase $k_n \to \infty$, and $\mu_S(n)k_n = o(n)$. In the case that $\sigma_S^2(n) \nrightarrow 0$, $\mathbf{Var}[D_{k_n}] \to \frac{4k_n\sigma_S^2(n)}{n^2}$; in the case that $\sigma_S^2(n) \to 0$, $\mathbf{Var}[D_{k_n}] \to \frac{4k_n^2\mu_S^2(n)}{n^3}$. A crude bound for D_{k_n} in the sublinear phase is

$$D_{k_n} = 1 - O_{\mathcal{L}_1}\Big(\frac{k_n s_n}{n}\Big).$$

Proof of Theorem 3 in the sublinear phase.

It is sufficient to verify the conditional variance condition (6.3), for this we compute

$$\begin{aligned} V_n &= \sum_{j=0}^{k_n} \mathbf{E}\Big[\Big(\frac{\nabla \tilde{M}_j}{\sqrt{\mathbf{Var}[M_{k_n}]}}\Big)^2 \Big| \mathcal{F}_{j-1}\Big] \\ &= \frac{\rho_n^{2k_n}}{\mathbf{Var}[D_{k_n}]}\Big(\frac{2(-n\mu_S(n) + 2n\sigma_S^2(n) - \sigma_S^2(n) + \mu_S^2(n))}{n^2(n-1)} \\ &\quad \times \sum_{j=0}^{k_n} \frac{1}{\rho_n^{2j}}\Big(1 - O_{\mathcal{L}_1}\Big(\frac{k_n s_n}{n}\Big)\Big)^2 \\ &\quad - \frac{2(-n\mu_S(n) + \sigma_S^2(n) + \mu_S^2(n))}{n^2(n-1)} \sum_{j=0}^{k_n} \frac{1}{\rho_n^{2j}}\Big). \end{aligned}$$

In this phase, we have $\rho_n^{2k} \to 1$ and $\displaystyle\sum_{j=0}^{k_n} \frac{1}{\rho_n^{2j}} = k_n + o(k_n)$.

Let

$$\begin{aligned} A_n &= \frac{2(-n\mu_S(n) + 2n\sigma_S^2(n) - \sigma_S^2(n) + \mu_S^2(n))}{n^2(n-1)} \sum_{j=0}^{k_n} \frac{1}{\rho_n^{2j}}\Big(1 - O_{\mathcal{L}_1}\Big(\frac{k_n s_n}{n}\Big)\Big)^2 \\ &\sim \frac{2(-n\mu_S(n) + 2n\sigma_S^2(n) - \sigma_S^2(n) + \mu_S^2(n))}{n^2(n-1)} k_n, \end{aligned}$$

$$\begin{aligned} B_n &= -\frac{2(-n\mu_S(n) + \sigma_S^2(n) + \mu_S^2(n))}{n^2(n-1)} \sum_{j=0}^{k_n} \frac{1}{\rho_n^{2j}} \\ &\sim -\frac{2(-n\mu_S(n) + \sigma_S^2(n) + \mu_S^2(n))}{n^2(n-1)} k_n. \end{aligned}$$

Wether $\sigma_S^2(n)$ approaches 0 or not,

$$V_n = \frac{\rho_n^{2k_n}}{\mathbf{Var}[D_{k_n}]}(A_n + B_n) \to 1.$$

This gives us

$$\frac{D_{k_n} - \mathbf{E}[D_{k_n}]}{\sqrt{\mathbf{Var}[D_{k_n}]}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1).$$

6.2.4 The linear phase

Recall that $a(n) = b(n) = n$, consider the linear phase, in which $\mu_S(n)k_n = \lambda_n n$. From the exact mean and variance in (6.1) and (6.2), we see that,

$$\begin{aligned} \mathbf{E}[D_{k_n}] &= \Big(1 - \frac{2\mu_S(n)}{n}\Big)^{k_n} \\ &= \exp(-2\lambda_n) + o(1). \end{aligned}$$

In the case of $\delta_S(n) \nrightarrow \infty$, we have

$$\mathbf{Var}[D_{k_n}] = \frac{\nu_n}{2n} + o\Big(\frac{1}{n}\Big),$$

where

$$\nu_n = \frac{e^{4\lambda_n} - 4\lambda_n + 8\delta_S(n)\lambda_n - 1}{e^{4\lambda_n}}; \tag{6.4}$$

in the case of $\delta_S(n) \to \infty$, we have

$$\mathbf{Var}[D_{k_n}] \sim \frac{4\delta_S(n)\lambda_n}{e^{4\lambda_n} n}.$$

In both cases, $\delta_S(n) = \frac{\sigma_S^2(n)}{\mu_S(n)} < \frac{s_n^2}{1}$ is $o(n)$, so that we have the following two lemmas

Lemma 15

$$\frac{D_{k_n}}{e^{-2\lambda_n}} \xrightarrow{\mathcal{L}_1} 1.$$

Proof. By Chebyshev's inequality,

$$\begin{aligned}\mathbf{Prob}(|D_{k_n} - \mathbf{E}[D_{k_n}]| \geq \varepsilon \mathbf{E}[D_{k_n}]) &\leq \frac{\mathbf{Var}[D_{k_n}]}{\varepsilon^2 \mathbf{E}[D_{k_n}]^2} \\ &\to 0.\end{aligned}$$

Hence,

$$\frac{D_{k_n}}{\mathbf{E}[D_{k_n}]} \xrightarrow{\mathcal{P}} 1.$$

With $\frac{\mathbf{E}[D_{k_n}]}{e^{-2\lambda_n}} \to 1$, and Slutsky's Theorem, we conclude that $\frac{D_{k_n}}{e^{-2\lambda_n}} \xrightarrow{\mathcal{P}} 1$, and since $e^{-2\lambda_n}$ is bounded, we have $\frac{D_{k_n}}{e^{-2\lambda_n}} \xrightarrow{\mathcal{L}_1} 1$. □

Proof of Theorem 3 in the linear phase.

The expression of the conditional variance condition V_n (6.3) in the linear phase is

$$V_n = \sum_{j=0}^{k_n} \mathbf{E}[\Big(\frac{\nabla \tilde{M}_j}{\sqrt{\mathbf{Var}[M_{k_n}]}}\Big)^2 \,|\, \mathcal{F}_{j-1}]$$

$$
\begin{aligned}
&= \frac{\rho_n^{2k_n}}{\mathbf{Var}[D_{k_n}]}\Big(\frac{2(-n\mu_S(n)+2n\sigma_S^2(n)-\sigma_S^2(n)+\mu_S^2(n))}{n^2(n-1)}\sum_{j=1}^{\lfloor \varepsilon n\rfloor}\frac{1}{\rho_n^{2j}}D_{j-1}^2 \\
&\quad +\frac{2(-n\mu_S(n)+2n\sigma_S^2(n)-\sigma_S^2(n)+\mu_S^2(n))}{n^2(n-1)}\sum_{j=\lfloor \varepsilon n\rfloor+1}^{k_n}\frac{1}{\rho_n^{2j}}D_{j-1}^2 \\
&\quad -\frac{2(-n\mu_S(n))+\sigma_S^2(n)+\mu_S^2(n))}{n^2(n-1)}\sum_{j=1}^{k_n}\frac{1}{\rho_n^{2j}}\Big). \\
&=: C_n + C_n' + D_n
\end{aligned}
$$

For large n, we have

$$
\begin{aligned}
|C_n| &\leq \frac{\rho_n^{2k_n}}{\mathbf{Var}[D_{k_n}]}\Big(\frac{2(-n\mu_S(n)+2n\sigma_S^2(n)-\sigma_S^2(n)+\mu_S^2(n))}{n^2(n-1)}\sum_{j=1}^{\lfloor \varepsilon n\rfloor}\frac{1}{\rho_n^{2j}}(1)\Big) \\
&\leq \frac{e^{-4\lambda_n}}{\mathbf{Var}[D_{k_n}]}\Big(\frac{2(2\sigma_S^2(n)-\mu_S(n))}{n(n-1)}\sum_{j=1}^{\lfloor \varepsilon n\rfloor}e^{4M_1}\Big) \\
&= O(\varepsilon), \quad \text{as } \varepsilon\to 0,
\end{aligned}
$$

and

$$
\begin{aligned}
C_n' &= \frac{e^{-4\lambda_n}}{\mathbf{Var}[D_{k_n}]}\Big(\frac{2(-n\mu_S(n)+2n\sigma_S^2(n)-\sigma_S^2(n)+\mu_S^2(n))}{n^2(n-1)} \\
&\quad \times\sum_{j=\lfloor \varepsilon n\rfloor+1}^{k_n}\frac{1}{\rho_n^{2j}}\big(e^{-4j/n}+o_{\mathcal{L}_1}(1)\big)\Big)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{e^{-4\lambda_n}}{\mathbf{Var}[D_{k_n}]}\left(\frac{2(-n\mu_S(n)+2n\sigma_S^2(n)-\sigma_S^2(n)+\mu_S^2(n))}{n^2(n-1)}\right. \\
&\quad \left.\times\left(\sum_{j=0}^{k_n}\frac{1}{\rho_n^{2j}}(e^{-4j/n})-\sum_{j=0}^{\lfloor \varepsilon n\rfloor}\frac{1}{\rho_n^{2j}}(e^{-4j/n})+o_{\mathcal{L}_1}(1)\sum_{j=\lfloor \varepsilon n\rfloor+1}^{k_n}\frac{1}{\rho_n^{2j}}\right)\right).
\end{aligned}
$$

For $b_n = \beta_n n + r_n$ with $r_n = o(n)$, we have

$$\sum_{j=0}^{b_n-1} \frac{1}{\rho_n^{2j}} e^{-4j/n} \quad = \quad \frac{(n/(n-2\mu_S(n)))^{2b_n} e^{-4b_n/n} - 1}{(n/(n-2\mu_S(n)))^2 e^{-4/n} - 1} = \beta_n(n) + o(n).$$

So,

$$C'_n \quad = \quad \frac{e^{-4\lambda_n}}{\mathbf{Var}[D_{k_n}]} \frac{2(2\sigma_S^2(n) - \mu_S(n))}{n(n-1)} \Big(\frac{\lambda_n}{\mu_S(n)} - \frac{\varepsilon}{\mu_S(n)} \Big) n + o(1).$$

Likewise, we have

$$\sum_{j=0}^{b_n-1} \frac{1}{\rho_n^{2j}} = \Big(\frac{e^{4\beta_n} - 1}{2\mu_S(n)} \Big) n + o(n),$$

$$D_n = \frac{e^{-4\lambda_n}}{\mathbf{Var}[D_{k_n}]} \Big(\frac{2\mu_S(n)}{n(n-1)} \Big) \Big(\frac{e^{4\lambda_n} - 1}{4\mu_S(n)} - \frac{e^{4\varepsilon} - 1}{4\mu_S(n)} \Big) n + o_{\mathcal{L}_1}(1).$$

Taking the limit as $\varepsilon \to 0$, we obtain

$$\begin{aligned}
\lim_{\varepsilon \to 0} V_n \quad &= \quad \lim_{\varepsilon \to 0} \Big(O(\varepsilon) + \Big(\frac{e^{-4\lambda_n}}{\mathbf{Var}[D_{k_n}]} \frac{2(2\sigma_S^2(n) - \mu_S(n))}{n(n-1)} \Big(\frac{\lambda_n}{\mu_S(n)} - \frac{\varepsilon}{\mu_S(n)} \Big) n + o(1) \Big) \\
&\quad + \Big(\frac{e^{-4\lambda_n}}{\mathbf{Var}[D_{k_n}]} \Big(\frac{2\mu_S(n)}{n(n-1)} \Big) \Big(\frac{e^{4\lambda_n} - 1}{4\mu_S(n)} - \frac{e^{4\varepsilon} - 1}{4\mu_S(n)} \Big) n + o_{\mathcal{L}_1}(1) \Big) \Big) \\
&= \quad \frac{1}{\mathbf{Var}[D_{k_n}]} \Big(\Big(e^{-4\lambda_n} \frac{2(2\sigma_S^2(n) - \mu_S(n))}{n(n-1)} \Big(\frac{\lambda_n}{\mu_S(n)} \Big) n \\
&\quad + \Big(e^{-4\lambda_n} \Big(\frac{2\mu_S(n)}{n(n-1)} \Big) \Big(\frac{e^{4\lambda_n} - 1}{4\mu_S(n)} \Big) n \Big) \Big) + o_{\mathcal{L}_1}(1)
\end{aligned}$$

Recall that, in the case of $\delta_S(n) \nrightarrow \infty$, we have

$$\mathbf{Var}[D_{k_n}] \quad = \quad \frac{\nu_n}{2n} + o\Big(\frac{1}{n} \Big),$$

where

$$\nu_n \quad = \quad \frac{e^{4\lambda_n} - 4\lambda_n + 8\delta_S(n)\lambda_n - 1}{e^{4\lambda_n}};$$

in the case of $\delta_S(n) \to \infty$, we have

$$\mathbf{Var}[D_{k_n}] \quad \sim \quad \frac{4\delta_S(n)\lambda_n}{e^{4\lambda_n} n}.$$

Thus, in both cases, we have

$$\lim_{\varepsilon \to 0} V_n \to 1 + o_{\mathcal{L}_1}(1), \quad as \quad n \to \infty.$$

We arrive at

$$\frac{D_{k_n} - e^{-2\lambda_n}}{\sqrt{\frac{\nu_n}{2n}}} \quad \xrightarrow{\mathcal{D}} \quad \mathcal{N}(0,1), \qquad \text{for } \delta_S(n) \nrightarrow \infty,$$

(see (6.4) for the definition of ν_n), and

$$\frac{D_{k_n} - e^{-2\lambda_n}}{\sqrt{\frac{4\delta_S(n)\lambda_n}{e^{4\lambda_n} n}}} \quad \xrightarrow{\mathcal{D}} \quad \mathcal{N}(0,1), \qquad \text{for } \delta_S(n) \to \infty.$$

6.2.5 The superlinear phase

Recall that $a(n) = b(n) = n$, in the superlinear phase, $n = o(k_n)$, $\mathbf{E}[D_{k_n}] = o(1)$, $\mathbf{Var}[D_{k_n}] = \frac{1}{2n} + o(\frac{1}{n})$.

Lemma 16

$$D_{k_n} = o_{\mathcal{L}_1}(1).$$

Proof. By Chebyshev's inequality,

$$\begin{aligned}
\mathbf{Prob}(|D_{k_n} - \mathbf{E}[D_{k_n}]| \geq \varepsilon) &\leq \frac{\mathbf{Var}[D_{k_n}]}{\varepsilon^2} \\
&\sim \frac{1}{\varepsilon^2(2n)} \\
&\to 0.
\end{aligned}$$

Hence,

$$D_{k_n} - \mathbf{E}[D_{k_n}] \xrightarrow{\mathcal{P}} 0$$

$$\mathbf{E}[D_{k_n}] \to 0$$

An application of Slutsky's theorem gives $D_{k_n} \xrightarrow{\mathcal{P}} 0$, and since $D_{k_n} \in [0,1]$, $D_{k_n} = o_{\mathcal{L}_1}(1)$. □

Proof of Theorem 3 in the superlinear phase.

The expression of the conditional variance condition V_n (6.3) in the superlinear phase is

$$\begin{aligned}
V_n &= \sum_{j=0}^{k_n} \mathbf{E}\Big[\Big(\frac{\nabla \tilde{M}_j}{\sqrt{\mathbf{Var}[M_{k_n}]}}\Big)^2 \,|\, \mathcal{F}_{j-1}\Big] \\
&= \frac{\rho_n^{2k_n}}{\mathbf{Var}[D_{k_n}]}\Big(\frac{2(-n\mu_S(n) + 2n\sigma_S^2(n) - \sigma_S^2(n) + \mu_S^2(n))}{n^2(n-1)} \sum_{j=k_n}^{k_n} \frac{1}{\rho_n^{2j}} D_{j-1}^2 \\
&\quad - \frac{2(-n\mu_S(n)) + \sigma_S^2(n) + \mu_S^2(n))}{n^2(n-1)} \sum_{j=1}^{k_n} \frac{1}{\rho_n^{2j}}\Big) \\
&=: E_n + F_n,
\end{aligned}$$

and

$$
\begin{aligned}
E_n &= \frac{\rho_n^{2k_n}}{\mathbf{Var}[D_{k_n}]}\Big(\frac{2(-n\mu_S(n)+2n\sigma_S^2(n)-\sigma_S^2(n)+\mu_S^2(n))}{n^2(n-1)}\sum_{j=k_n}^{k_n}\frac{1}{\rho_n^{2j}}o_{\mathcal{L}_1}(1)\Big)\\
&= \frac{\rho_n^{2k_n}}{\mathbf{Var}[D_{k_n}]}\Big(\frac{2(-n\mu_S(n)+2n\sigma_S^2(n)-\sigma_S^2(n)+\mu_S^2(n))}{n^2(n-1)}\\
&\quad\times\frac{(\frac{n-2\mu_S(n)}{n})^{-1}(1-(\frac{n-2\mu_S(n)}{n})^{-2k_n-2})}{1-(\frac{n-2\mu_S(n)}{n})^{-2}}\Big)o_{\mathcal{L}_1}(1)\\
&\to 0,\\
F_n &= \frac{\rho_n^{2k_n}}{\mathbf{Var}[D_{k_n}]}\Big(-\frac{2(-n\mu_S(n))+\sigma_S^2(n)+\mu_S^2(n))}{n^2(n-1)}\sum_{j=1}^{k_n}\frac{1}{\rho_n^{2j}}\Big)\\
&= \frac{(\frac{n-2\mu_S(n)}{n})^{2k_n}}{\frac{1}{2n}+o(\frac{1}{n})}\Big(-\frac{2(-n\mu_S(n))+\sigma_S^2(n)+\mu_S^2(n)}{n^2(n-1)}\\
&\quad\times\frac{(\frac{n-2\mu_S(n)}{n})^{-1}(1-(\frac{n-2\mu_S(n)}{n})^{-2k_n-2})}{1-(\frac{n-2\mu_S(n)}{n})^{-2}}\Big)\\
&\to 1.
\end{aligned}
$$

So,

$$V_n \xrightarrow{\mathcal{P}} 1.$$

We arrive at

$$\frac{D_{k_n}}{\sqrt{\frac{1}{2n}}}\xrightarrow{\mathcal{D}}\mathcal{N}(0,1).$$

6.3 Illustrative examples

With a few illustrating examples, we explore the cases with degenerate and non-degerate sublinear phase, and cases in the linear phase with index of dispersion $\delta_S(n)\to\infty$ and $\delta_S(n)\nrightarrow\infty$.

6.3.1 An example with a fixed generator

Let $S_j(n) \equiv s$, a fixed positive integer ($\sigma_S^2(n) = 0$). This is a degenerate case. After $k_n = \lfloor n^{\frac{2}{3}} \rfloor$ draws, regularity condition (b) is met, and according to Theorem 3,

$$\frac{D_{\lfloor n^{\frac{2}{3}} \rfloor} - \left(1 - 2s\left(\frac{1}{n}\right)^{\frac{1}{3}} + 2s^2\left(\frac{1}{n}\right)^{\frac{2}{3}}\right)}{\sqrt{\left(\frac{1}{n}\right)^{\frac{5}{3}}}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 4s^2).$$

6.3.2 An example with a generator with a growing range

Let $S_j(n)$ be uniformly distributed on the two-point set $\{1, \lfloor n^{\frac{1}{3}} \rfloor\}$. This is a nondegenerate case. After $k_n = \lfloor \ln n \rfloor$ draws, $\frac{s_n^2}{\sigma_S^2(n)} = o(k_n)$, regularity condition (a) is met. According to Theorem 3,

$$\frac{D_{\lfloor \ln n \rfloor} - \left(1 - \frac{\ln n}{2n^{\frac{2}{3}}}\right)}{\sqrt{\frac{\ln n}{n^{\frac{4}{3}}}}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1).$$

6.3.3 An example with a degenerate generator

Let $S_j(n)$ be Poisson($\ln n$) distributed random random variables. Then $\delta_S(n) = 1$, after $k_n = \lfloor \frac{2n}{\ln n} \rfloor$ draws, and according to Theorem 3,

$$\frac{D_{\lfloor \frac{2n}{\ln n} \rfloor} - e^{-4}}{\sqrt{\frac{1}{n}}} \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{e^8 - 55}{2e^8}\right).$$

6.3.4 An example with a nondegenerate generator

Let $S_j(n)$ be discrete Uniform$(1, \lfloor\sqrt{n}\rfloor)$ distributed random random variables. Then $\delta_S(n) \sim \frac{\sqrt{n}}{6} \to \infty$, after $k_n = \lfloor 2\sqrt{n}\rfloor$ draws, according to Theorem 3,

$$\frac{D_{\lfloor 2\sqrt{n}\rfloor} - e^{-4}}{\sqrt{\frac{1}{n^{\frac{1}{2}}}}} \xrightarrow{\mathcal{D}} \mathcal{N}\Big(0, \frac{\frac{4}{3}}{e^8}\Big).$$

6.4 Concluding remarks

The three phases are connected in a way that can be shown from the general form. Recall that in the linear phase,

$$\frac{D_{k_n} - e^{-2\lambda_n}}{\sqrt{\frac{1}{2n}\Big(\frac{e^{4\lambda_n} - 4\lambda_n + 8\delta_S(n)\lambda_n - 1}{e^{4\lambda_n}}\Big)}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1).$$

The sublinear and the superlinear phases can be viewed as special cases of this form.
Let $\lambda_n = \frac{k_n}{n} \to 0$, this will become the Gaussian limit of the sublinear phase.
Let $\lambda_n \to \infty$, this will become the Gaussian limit of the superlinear phase.

6.5 Simulation

We have simulations of the classic Bernoulli-Laplace model with $n = 100, 200, 500, 1000,$ and 2000. Let k_n be $\lfloor n^{2/3} + 29\rfloor$, $0.5n$, $\lfloor\frac{1}{4}n\log n + n\rfloor$, and $\lfloor\frac{1}{4}n\ln n + n\ln n\ln n\ln n\rfloor$, which are examples of upper sublinear phase, linear phase, lower superlinear phase, and upper superlinear phase respectively. We calculate the theoretical means and variances of W_{A,k_n} (Avg. The. and Var. The.); and, with 100 repetitions, we get the mean of the means and variances of W_{A,k_n} from the simulation (Avg. Sim. and Var. Sim.) to compare with the theoretical ones.

Table 6.1: Bernoulli-Laplace Simulation Comparison

k_n \ n		100		200		500		1000		2000	
$\lfloor n^{2/3}+29 \rfloor$		50		63		91		129		187	
Avg. The.	Avg. Sim.	68.21	68.17	153.09	153.06	423.60	423.66	886.20	886.24	1829.37	1829.73
Var. The.	Var. Sim.	25.00	7.60	19.85	8.22	16.56	10.92	16.64	9.47	17.48	15.10
$0.5n$		50		100		250		500		1000	
Avg. The.	Avg. Sim.	68.39	68.17	136.79	136.30	341.97	342.19	683.94	683.39	1367.88	1366.54
Var. The.	Var. Sim.	7.42	7.60	14.85	14.96	37.12	38.60	74.25	61.63	148.50	139.39
$\lfloor \frac{1}{4} n \ln n + n \rfloor$		210		464		1276		2726		5800	
Avg. The.	Avg. Sim.	50.68	50.25	100.97	100.87	251.52	251.21	502.14	502.42	1003.03	1002.08
Var. The.	Var. Sim.	12.50	12.56	25.00	30.83	62.50	64.09	125.00	128.39	250.00	253.39
$\lfloor \frac{1}{4} n \ln n$ $+ n \ln \ln n \rfloor$		267		598		1690		3659		7856	
Avg. The.	Avg. Sim.	50	49.4	100.00	100.45	250.00	250.75	500.00	499.66	1000.00	998.82
Var. The.	Var. Sim.	12.50	11.66	25.00	23.83	62.50	61.63	125.00	100.60	250.00	233.27

Chapter 7

Further developments

Further interests include the distribution of the lower sublinear phases, where the asymptotic distributions are possibly not Gaussian, and mixing models that have more than two urns, among others.

7.1 Poisson limit in the sublinear phase

A Poisson convergence limit theorem follows from the results for infinitely divisible laws developed by Brown and Eagleson (1971); see also Freedman (1974). It can be used in our models at the lower edges of the Gaussian phases. Let $A_{n,j}$ be an array of events with $A_{n,j}$ being $\mathcal{F}_{n,j}$-measurable, and let $I(A_{n,j})$ be their indicators.

Theorem 4 (Freedman, 1974)*:*

If, for some constant λ*,*

$$(i) \qquad \sum_{j=1}^{n} \mathbf{Prob}(A_{n,j} \,|\, \mathcal{F}_{n,j-1}) \xrightarrow{\mathcal{P}} \lambda;$$

and

$$(ii) \qquad \max_{j<n} \mathbf{Prob}(A_{n,j} \,|\, \mathcal{F}_{n,j-1}) \xrightarrow{\mathcal{P}} 0,$$

then $T_n = \sum_{j=1}^{n} I(A_{n,j}) \xrightarrow{\mathcal{D}} \text{Poisson}(\lambda)$.

7.1.1 The classic Ehrenfest urn model

In the generalized Ehrenfest urn model, let W_{k_n} be the number of white balls in the urn after the k_nth draw. Let $\mathcal{F}_j, j = 0, 1, 2, \ldots$ be the sigma-field generated by $W_0, W_1, W_2, \ldots$. To instantiate the type of Poisson limits one might get, we first look at an example of the classic Ehrenfest urn in Chapter 4. Let A_j be the event that jth draw is white. Then

$$\mathbf{Prob}(A_j \,|\, \mathcal{F}_{j-1}) = \frac{W_{j-1}}{n}.$$

If we start with an urn containing all the red balls at the beginning, that is $W_0(n) := W_0 = 0$ or $W_0(n)$ is asymptotically small compared to n, say $\frac{k_n W_0}{n} \to 0$, as $n \to \infty$, then, there is a large probability that the first k_n draws are all red, where k_n is in the sublinear phase.

We have the following lemma for the generalized Ehrenfest urn model, which applies to our example of the classic Ehrenfest urn model.

Lemma 17 *In the sublinear phase of the generalized Ehrenfest urn (Zhang and Mah-*

moud, 2011+), let $s_n k_n = o(n)$, in the degenerate extremal case $W_0(n) = o(n)$, we have

$$W_j = W_0(n) + \mu_S(n)j + o_P(\mu_S(n)j), \quad \textit{for} \quad 0 \le j \le k_n.$$

Proof. The degenerate case, where $\sigma_S^2(n) \to 0$ is deterministic or nearly so. In this case $S_j - \mu_S(n) \xrightarrow{\mathcal{P}} 0$; we also have the asymptotic approximation $S_j = \mu_S(n) + o_P(\mu_S(n))$. Observe that this convergence takes place in $\mathcal{L}_1$, too as $S_j \le s_n$, for all j. Thus, we also have the asymptotic approximation $S_j = \mu_S(n) + o_{L_1}(\mu_S(n))$.

The event that $W_j = W_0(n) + \sum_{i=1}^{j} S_i$ occurs when the ith sample is all red, for $1 \le i \le j \le k_n$, and

$$\mathbf{Prob}\Big(W_j = W_0(n) + \sum_{i=1}^{j} S_i \,\Big|\, S_1, \dots S_j\Big) = \frac{\binom{n - W_0(n)}{S_1}}{\binom{n}{S_1}}$$

$$\times \frac{\binom{n - W_0(n) - S_1}{S_2}}{\binom{n}{S_2}} \times \frac{\binom{n - W_0(n) - S_1 - S_2}{S_3}}{\binom{n}{S_3}}$$

$$\times \cdots \times \frac{\binom{n - W_0(n) - \sum\limits_{i=1}^{j-1} S_i}{S_j}}{\binom{n}{S_j}}$$

$$= \prod_{i=1}^{j} \frac{\Gamma(n - W_0(n) - \sum_{r=1}^{i-1} S_r + 1)\,\Gamma(n - S_i + 1)}{\Gamma(n+1)\,\Gamma\Big(n - W_0(n) - \sum\limits_{r=1}^{i} S_r + 1\Big)}.$$

After an application of Stirling's approximation of the ratio of the Gamma functions, the probability takes the asymptotic form

$$\begin{aligned}
&\mathbf{Prob}\Big(W_j = W_0(n) + \sum_{i=1}^{j} S_i \,\Big|\, S_1, \dots S_j\Big) \\
&= \prod_{i=1}^{j} \frac{\big(n - W_0(n)\big)^{S_i}}{n^{S_i}} \Big(1 + O\Big(\frac{1}{n}\Big)\Big) \\
&= \Big(1 - \frac{W_0(n)}{n}\Big)^{\sum_{i=1}^{j} S_i} + \Big(1 + O\Big(\frac{j}{n}\Big)\Big) \\
&= \Big(1 - \frac{W_0(n)}{n}\Big)^{\mu_S(n)j + o_{\mathcal{L}_1}(\mu_S(n)j)} + \Big(1 + O\Big(\frac{k_n}{n}\Big)\Big) \\
&\xrightarrow{\mathcal{L}_1} 1.
\end{aligned}$$

The unconditional probability follows by taking expectations. □

In our example of the classic Ehrenfest urn model, it is a degenerate and extremal case of the generalized Ehrenfest urn model, with sample size generator being deterministically 1, so we have

$$W_j = W_0(n) + j + o_P(j), \quad \text{for} \quad 0 \le j \le k_n.$$

We have the following Poisson limit theorem by checking the two conditions in the theorem in Freedman (1974).

Theorem 5 *If* $k_n = \sqrt{n}$*, then* $\dfrac{W_0(n) + k_n - W_{k_n}}{2}$ *has a* Poisson$(\frac{1}{2})$ *distribution.*

Proof.

$$\begin{aligned}
\sum_{j=1}^{k_n} \mathbf{Prob}(A_j \mid \mathcal{F}_{j-1}) &= \frac{\sum_{j=1}^{k_n} W_{j-1}}{n} \\
&= \frac{1}{n}\Big(\sum_{j=1}^{k_n}(W_0 + (j-1) + o_P(j-1))\Big) \\
&= \frac{1}{n}\Big((k_n-1)W_0 + \frac{(0+(k_n-1))k_n}{2} + o_P(k_n^2)\Big) \\
&\xrightarrow{\mathcal{P}} \lambda = \frac{1}{2},
\end{aligned}$$

and

$$\begin{aligned}
\max_{j<k_n} \mathbf{Prob}(A_j \mid \mathcal{F}_{j-1}) &= P(A_{k_n} | \mathcal{F}_{k_n-1}) \\
&= \frac{W_{k_n-1}}{n} \\
&= \frac{1}{n}\Big(W_0(n) + (k_n-1) + o_P(k_n-1)\Big) \\
&\xrightarrow{\mathcal{P}} 0.
\end{aligned}$$

Then,

$$\begin{aligned}
T_{k_n} &= \sum_{j=1}^{k_n} I(A_j) \\
&= \frac{B_{k_n} - (n - W_0(n) - k_n)}{2} \\
&= \frac{W_0(n) + k_n - W_{k_n}}{2}
\end{aligned}$$

has a Poisson($\frac{1}{2}$) distribution. □

Similar results of Poisson limit from a scenario of disposal of particles into cells have been shown by Mikhailov (1977).

7.1.2 The classic coupon collection urn

In the coupon collection question, we have an urn containing balls of two colors: white and red. White balls represent unobserved coupons, and red balls represent coupons that have already been picked in a previous draw. At each stage a ball is picked randomly, if the ball is red, we put it back to the urn, if the ball is white, we paint it red and put it back in the urn. The process repeats at the next step and so on. In the sublinear case of the classic coupon collection problem (Mahmoud, 2010 and Smythe, 2011), we start with all the coupons uncollected, in the classic coupon collection question, when the sample size generator $S = 1$ deterministically (that is, $\sigma_S^2 = 0$), we have the asymptotic representation

$$W_j = n - j + o_P(k_n).$$

Let A_j be the event that the jth draw is red, then

$$\mathbf{Prob}(A_j \,|\, \mathcal{F}_{j-1}) = \frac{n - W_{j-1}}{n}.$$

The following result is due to Kolchin, Sevastyanov, and Chistyakov (1976) and can be proved by checking the conditions of the theorem in Freedman (1974).

Theorem 6 *If $k_n = \sqrt{n}$, then $W_{k_n} - (n - k_n)$ has a* Poisson$\left(\frac{1}{2}\right)$ *distribution.*

Proof. We check the two conditions in the Theorem in Freedman (1974).

$$\begin{aligned}
\sum_{j=1}^{k_n} \mathbf{Prob}(A_j \,|\, \mathcal{F}_{j-1}) &= \frac{\sum_{j=1}^{k_n}(n - W_{j-1})}{n} \\
&= \frac{1}{n}\Big(\sum_{j=1}^{k_n}((j-1) + o_P(k_n))\Big) \\
&= \frac{1}{n}\Big(\frac{(0 + (k_n - 1))k_n}{2} + o_P(k_n^2)\Big) \\
&\xrightarrow{\mathcal{P}} \lambda = \frac{1}{2},
\end{aligned}$$

and

$$\begin{aligned}
\max_{j<k_n} \mathbf{Prob}(A_j \,|\, \mathcal{F}_{j-1}) &= P(A_{k_n} \,|\, \mathcal{F}_{k_n-1}) \\
&= \frac{n - W_{k_n-1}}{n} \\
&= \frac{1}{n}\Big((k_n - 1) + o_P(k_n)\Big) \xrightarrow{\mathcal{P}} 0.
\end{aligned}$$

Then, $T_{k_n} = \sum_{j=1}^{k_n} I(A_j) = W_{k_n} - (n - k_n)$ has a Poisson$\Big(\frac{1}{2}\Big)$ distribution. □

These results are indicative of a whole class of Poisson limits that can be obtained under different assumptions for the sampling distribution, and the kind of "phase transition" at the lower edge of Gaussian phase. We shall pursue this in the future.

7.2 A general mixing urn

Seneta (1982) sets up a general mixing urn as follows: Let the number of urns be $N \geq 2$ and the total number of balls in all the urns be n. A ball is randomly selected from all the balls, taken from its urn, and randomly placed into urn $t, t = 1, \ldots, N$. Let $W_j, j = 0, 1, 2, \ldots$ denote the number of balls in any specific urn at time j, then $\{W_j\}$

is a Markov chain and we have the following transition matrix $P = \{p_{ij}\}_{i,j=0}^{n}$ with all p_{ij} being 0 except

$$\begin{aligned} p_{i,i-1} &= \frac{i(N-1)}{Nn}, \\ p_{i,i+1} &= \frac{n-i}{Nn}, \\ p_{i,i} &= 1 - \frac{1}{N} - \frac{i}{n} + \frac{2i}{nN}. \end{aligned}$$

In the case N=2, if we consider W_j only at its epochs of change of state, it becomes the Ehrenfest urn model. Results given by Seneta (1982) show that

$$\begin{aligned} \mathbf{E}\Big[W_j - \frac{n}{N}\Big|\mathcal{F}_{j-1}\Big] &= \Big(1-\frac{1}{n}\Big)\Big(W_{j-1} - \frac{n}{N}\Big), \\ \mathbf{E}[W_j] &= \frac{n}{N} + \Big(W_0 - \frac{n}{N}\Big)\Big(1-\frac{1}{n}\Big)^j, \\ \mathbf{E}\Big[\Big(W_j - \frac{n}{N}\Big)^2\Big|W_j\Big] &= \Big(W_{j-1} - \frac{n}{N}\Big)^2\Big(1-\frac{2}{n}\Big) \\ &\quad + \Big(\frac{W_{j-1}}{n} - \frac{1}{N}\Big)\Big(1-\frac{2}{N}\Big) + \frac{2}{N}\Big(1-\frac{1}{N}\Big). \end{aligned}$$

So that we can get the exact first two moments of the general mixing urn:

$$\begin{aligned} \mathbf{E}\left[\frac{W_j - \frac{n}{N}}{(W_0 - \frac{n}{N})}\right] &= \Big(1-\frac{1}{n}\Big)^j, \\ \mathbf{Var}\left[\Big(\frac{W_j - \frac{n}{N}}{W_0 - \frac{n}{N}}\Big)^2\right] &= \Big(1-\frac{2}{n}\Big)^j + \frac{N^2 - n + 2W_0N}{(W_0N-n)^2}\Big(1-\frac{2}{n}\Big)^j \\ &\quad + \frac{W_0N^2 + 2n - 2W_0N - 2Nn}{(W_0N-n)^2}\Big(1-\frac{1}{n}\Big)^j + \frac{(N-1)n}{(W_0N-n)^2}. \end{aligned}$$

We can conduct asymptotic analysis from here and get a more general result on the phases of the mixing urns.

Bibliography

[1] Antognini, B. (2005). On the speed of convergence of some urn designs for the balanced allocation of two treatments. *Metrika*, **62**, 309–322.

[2] Arthur, W. , Ermoliev, Y. and Kaniovskib, Y. (1987). Path-dependent processes and the emergence of macro-structure. *European Journal of Operational Research*. **30, 3**, 294–303.

[3] Athreya, B. and Karlin, S. (1968). Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *Annals of Mathematical Statisitics*. **39, 6**, 1801-1817.

[4] Bagchi, A. and Pal, A. (1985). Asymptotic normality in the generalized Pólya-Eggenberger urn model, with an application to computer data structures. *SIAM Journal on Algebraic and Discrete Methods*, **6**. 394–405.

[5] Balaji, S. , Mahmoud, H. and Zhang, T. (2010). Phases in the diffusion of gases via the Ehrenfest urn model. *Journal of Applied Probability*, **47, 3**. 841–855.

[6] Bai, Z. , Hu, F. and Zhang, L. (2002). Gaussian approximation theorems for urn models and their applications. *Annals of Applied Probability*, **12, 4**, 1149–1173.

[7] Bellman, R. and Harris, T. (1951). Recurrence times for the Ehrenfest model. *Pacific Journal of Mathematics*, **1**, 184–188.

[8] Berg, S. (1985). Paradox of voting under an urn model: The effect of homogeneity. *Public Choice*, **47, 2**, 377–387.

[9] Bernoulli, D. (1768). De usu algorithmi infinitesimalis in arte conjectandi specimen. *Novi Commentarii Academiae Scientiarum Imperialis Petropolitanae*, **12**, 87–98.

[10] Bernoulli, D. (1769). Disquisitiones analyticae de novo problemate conjecturall. *Novi Commentarii Academiae Scientiarum Imperialis Petropolitanae*, **14**, 3–25.

[11] Bingham, N. (1991). Fluctuation Theory for the Ehrenfest Urn. *Advances in Applied Probability*, **23**, 598–611.

[12] Blom, G. (1989). Mean transition times for the Ehrenfest urn model. *Advances in Applied Probability*, **21**, 479–480.

[13] Brown, B. and Eagleson, G. (1971). Martingale convergence to infinitely divisible laws with finite variances. *Transactions of The American Mathematical Society*, **162**, 449–453.

[14] Chatterjee, S. , Diaconis, P. and Meches, E. (2005). Exchangeable pairs and Poisson approximation. *Probability Surveys*, **2**, 64–106.

[15] Cox, D. and Lewis, P. (1966). *The Statistical Analysis of Series of Events.* Methuen, London.

[16] Diaconis, P. (1988). Group Representaions in Probability and Statistics. *Institute of Mathematical Statistics*, Hayward, CA.

[17] Diaconis, P. (1996). The cutoff phenomenon in finite Markov chains. *Proceedings of the National Academy of Science, USA*, **93**, 1659–1664.

[18] Diaconis, P. and Shahshahani, M. (1987). Time to reach stationarity in the Bernoulli-Laplace Diffusion Model. *SIAM Journal on Mathematical Analysis*, **18**, No.1., 208–218.

[19] Donnelly, P. , Lloyd, P. and Sudbury, A. (1994). Approach to stationarity of the Bernoulli-Laplace diffusion model. *Advances in Applied Probability*, **26**, 715–727.

[20] Eggenberger, F. and Pólya, G. (1923). Über die Statistik verketter Vorgänge. *Zeitschrift für Angewandte Mathematik und Mechanik*, **1**, 279–289

[21] Ehrenfest, P. and Ehrenfest, T. (1907). Über zwei bekannte einwände gegen das Boltzmannsche H-theorem. *Physikalische Zeitschrift*, **8**, 311–314.

[22] Ewens, W. (1972). The sampling theory of selectively neutral alleles. *Theoretical Population Biology*, **3**, 87–112.

[23] Feller, W. (1968). *An Introduction to Probability and Its Applications. Vol. I, 3rd Edition.* John Wiley, New York.

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

[25] Freedman, D. (1965). Bernard Friedman's urn. *Annals of Mathematical Statistics*, **36, 3**, 956–970.

[26] Freedman, D. (1974). The Poisson approximation for dependent events. *Annals of Probability*, **2**, 256–269.

[27] Gastwirth, J. and Bhattacharya, P. (1984). Two probability models of pyramids or chain letter schemes demonstrating that their promotional claims are unreliable. *Operations Research*, **32**, 527–536.

[28] Hall, P. and Heyde, C. (1980). *Martingale Limit Theory and Its Applications*. Academic Press, New York.

[29] Hwang, H. , Kuba, M. and Panholzer, A. (2007). Analysis of some exactly solvable diminishing urn models. *Formal Power Series and Algebraic Combinatorics, Nankai University, Tianjin, China, 2007.*

[30] Johnson, N. and Kotz, S. (1977). *Urn Models and their Applications: An Approach to Modern Discrete Probability*. Wiley, New York.

[31] Kac, M. (1947). Random walk and the theory of Brownian motion. *American Mathematical Monthly*, **54**, 369–391.

[32] Karlin, S. and McGregor, J. (1965). Ehrenfest urn models. *Journal of Applied Probability*, **2**, 352–376.

[33] Karr, A. (1993). *Probability*. Springer, New York.

[34] Kholfi, S. and Mahmoud, H. (2010). The class of tenable zero-balanced Polya urn schemes: characterization and gaussian phases. *2011 Proceedings of the Eighth Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*, 117–126.

[35] Kobza, J., and Jacobson, S. and Vaughan, D. (2007). A survey of the coupon collectors problem with random sample sizes. *Methodology and Computing in Applied Probability*, **9, 4**, 573–584.

[36] Kolchin, V. Sevastyanov, B. and Chistyakov, V. (1976). *Random Allocations*, Moscow.

[37] Laplace, P. (1812). Théorie analytique des probabilités. *Oeuvres Complétes*, **7**, 413–414.

[38] Mahmoud, H. (2008). *Pólya Urn Models*. Chapman-Hall, Boca Raton, Florida.

[39] Mahmoud, H. (2010). Gaussian phases in generalized coupon collection. *Advances in Applied Probability*, **42**, 994–1012.

[40] Mahmoud, H. and Smythe. R. (2011+). On the joint behavior of types of coupons in generalized coupon collection. *Submitted manuscript.*

[41] Mikhailov, V. (1977). A Poisson limit theorem in the scheme of group disposal of particles. *Theory of Probability and Its Applications*, **22**, 152–156.

[42] Mikhailov, V. (1980). Asymptotic normality of the number of empty cells for group allocation of particles. *Theory of Probability and Its Applications*, **25**, 82–90.

[43] Perry, J. and Mead, R. (1979). On the power of the index of dispersion test to detect spatial pattern. *Biometrics*, **35**, 613–622.

[44] Sellke, T. (1995). How many i.i.d. samples does it take to see all the balls in a box? *The Annals of Applied Probability*, **5**, 294–309.

[45] Seneta, E. (1982). Entropy and martingales in Markov chain models. *Journal of Applied Probability*, **19**, 367–381.

[46] Smythe, R. (1996). Central limit theorems for urn models. *Stochastic Processes and their Applications*, **65**, 115–137.

[47] Smythe, R. (2011). Phases in generalized coupon collection. *Advances in Applied Probability*, **48**, 189–199.

[48] Stuart, A. Ord, K and Arnold, S. (2009). *The Advanced Theory of Statistics, Volume 2, Sixth edition*. Wiley, New York.

[49] Vatutin, V. and Mikhailov, V. (1982). Limit theorem for the number of empty cells in an equiprobable scheme for group allocation of particles. *Theory of Probability and Its Applications*, **27**, 734–743.

[50] Voit, M. (1996). Asymptotic distributions for the Ehrenfest urn and related random walks. *Journal of Applied Probability*, **33**, 340–356.

[51] Wei, L. (1978). An application of an urn model to the design of sequential controlled clinical trials. *Journal of the American Statistical Association*, **73**, 559–563.

[52] Zhang, T. and Mahmoud, H. (2011+). An urn model for population mixing and the phases within. *Submitted manuscript.*

[53] Zhang, T. (2011+). Gaussian phases in a generalized Bernoulli-Laplace model. *Manuscript.*