

Limits of Pólya urns with innovations

Jean Bertoin*

Abstract

We consider a version of the classical Pólya urn scheme which incorporates innovations. The space S of colors is an arbitrary measurable set. After each sampling of a ball in the urn, one returns C balls of the same color and additional balls of different colors given by some finite point process ξ on S , where the distribution P_s of the pair (C, ξ) depends on the sampled color s . We suppose that the average number of copies $E_s(C)$ is the same for all $s \in S$, and that the intensity measures of innovations have the form $E_s(\xi) = a(s)\mu$ for some finite measure μ and a modulation function a on S that is bounded away from 0 and ∞ . We then show that the empirical distribution of the colors in the urn converges to the normalized intensity $\bar{\mu}$. In turn, different regimes for the fluctuations are observed, depending on whether $E(C)$ is larger or smaller than $\mu(a)$.

Keywords: Pólya urn; innovation; empirical distribution; martingale central limit theorem.

MSC2020 subject classifications: 60F17; 60G44; 60J85; 62G30.

Submitted to EJP on November 17, 2022, final version accepted on October 21, 2023.

1 Introduction

At each step of an original Pólya urn scheme, a ball is drawn uniformly at random from an urn, independently of the preceding steps. One observes its color; then the ball is returned in the urn together with an additional ball of the same color. More generally, one can consider random replacement schemes, for which one rather returns random numbers of balls of different colors according to some fixed distribution that depends only on the color of the sampled ball. The quantities of interest are the proportions of balls with given colors after a large number of steps. Most often, it is assumed that the set of all possible colors is finite, with the notable exception of certain recent contributions which will be discussed at the end of this Introduction. General Pólya urn schemes are used as models in a variety of fields including Computer Sciences, Biology, Social Sciences, etc.; see e.g. [20] and references therein.

We are interested in a variation of the Pólya urn scheme which incorporates innovations, in the sense that at each step, balls with new colors that have never been used

*Institute of Mathematics, University of Zurich, Switzerland. E-mail: jean.bertoin@math.uzh.ch

before can be returned in the urn, and the space of colors is an arbitrary measurable space. Such an urn scheme with innovation has been first introduced¹ in 1955 by Herbert Simon [24] in terms of a simple model for producing a text by random iterations. One starts with a first word, say w_1 , and when one has already written a text of length $k \geq 1$, say $w_1 \dots w_k$, one appends the next word w_{k+1} which is either, with fixed probability p , a repetition of $w_{u(k)}$, where $u(k)$ has the uniform distribution on $\{1, \dots, k\}$, or, with complementary probability $1 - p$, a new word different from all the preceding. Further situations where innovations or novelties are combined with urn schemes have been considered by [11, 25, 12], amongst others. Needless to recall, the association of stochastic reinforcement and innovation is an important aspect of machine learning on which there exists of course a huge literature.

In a recent work [8] that was partly motivated by the study of certain step reinforced random walks, we analyzed the effects of linear reinforcement on the sequence of empirical distributions in Simon model. It was observed there that when one takes for the different words independent uniform random variables on $[0, 1]$, the Glivenko-Cantelli theorem still holds, that is there is the almost-sure convergence of the empirical distribution functions of the reinforced sequence towards the identity function on $[0, 1]$, but the Donsker theorem is affected. Indeed, the sequence of empirical processes converges in law to a Brownian bridge only up to a constant factor when the repetition probability $p < 1/2$; a further rescaling is needed when $p > 1/2$ and the limit is then a bridge with exchangeable increments and discontinuous paths.

Here, we consider more generally a Borel space S as the set of possible colors of balls. We write \mathcal{L}^∞ for the space of bounded measurable functions $f : S \rightarrow \mathbb{R}$, and often use the short notation

$$\nu(f) := \int_S f(s) \nu(ds)$$

for a finite measure ν on S and $f \in \mathcal{L}^\infty$. In particular, the total mass of ν is denoted by $\nu(\mathbf{1})$, where $\mathbf{1}$ stands for the function identical to 1 on S . Whenever $\nu \neq 0$, we write

$$\bar{\nu} = \nu(\mathbf{1})^{-1} \nu$$

for the normalized probability measure associated to ν .

A typical replacement consists of a pair (C, ξ) , where C is a random variable with values in $\{-1, 0, 1, 2, \dots\}$ which represents the number of copies of the sampled ball which are returned in the urn (the case $C = -1$ accounts for the situation where the sampled ball is removed from the urn), and ξ a point process on S which represents the random family of new balls which are simultaneously added. The dynamics are hence fully encoded by the kernel of laws $(P_s)_{s \in S}$ that specify the distributions of the pair (C, ξ) as a function of the color s of the sampled ball. We implicitly assume that the probability kernel $s \mapsto P_s$ is measurable, and to avoid discussing situations where the urn could be emptied, we always assume henceforth that for all $s \in S$,

$$P_s(C + \xi(\mathbf{1}) \geq 0) = 1, \tag{1.1}$$

where, by the preceding notation, $\xi(\mathbf{1})$ accounts for the number of innovative balls.

More precisely, at each step, independently of the previous steps, we sample a ball uniformly at random from the urn and observe its color, say s , as well as a pair (C, ξ) with law P_s . Given $C = k$ and $\xi = \sum_{i=1}^j \delta_{s_i}$, with δ_s denoting the Dirac measure at s , we then return that ball in the urn together with k balls with the same color s and j

¹Simon model bears close connections to the earlier work of Udny Yule on the frequency distribution of the sizes of biological genera, and in turn has been rediscovered and applied in a variety of frameworks, see e.g. the discussion in [23].

further balls with colors s_1, \dots, s_j (it is not required that the s_i are all distinct). We represent the composition of the urn after n steps by a counting measure U_n on S such that the atoms of U_n are the colors of the balls present in the urn after n steps, and their multiplicities the number of balls of these colors. Then for $f \in \mathcal{L}^\infty$, $U_n(f)$ is the sum of the $f(s)$ for the colors $s \in S$ present in the urn after n steps and repeated according to their multiplicities.

We now introduce the three crucial assumptions of our study. First, we suppose that the average number of copies that are returned at a typical step does not depend on the sampled color, viz.

$$\text{the function } s \mapsto E_s(C) := E(C) \text{ is constant on } S, \tag{1.2}$$

where the notation E_s is used for the mathematical expectation under P_s . Second, we suppose that the intensity measures of innovations are all proportional to some fixed (non-zero) measure μ on the space of colors. Specifically, we assume that there exists a measurable function a on S that is bounded away from 0 and from ∞ , viz. with

$$0 < \inf_{s \in S} a(s) \leq \sup_{s \in S} a(s) := \|a\|_\infty < \infty,$$

such that for every $s \in S$, the intensity measure of the point processes ξ under P_s is given by

$$E_s(\xi(f)) = a(s)\mu(f), \quad \text{for all } f \in \mathcal{L}^\infty. \tag{1.3}$$

We should think of a as a factor which modulates the intensity of innovations as a function of the color of the sampled ball. Last, we shall also need an assumption of uniform upper and lower boundedness for the total number of balls which are returned at each step, namely that

$$0 < \inf_{s \in S} P_s(C + \xi(\mathbf{1}) \geq 1) \quad \text{and} \quad \sup_{s \in S} E_s((C + \xi(\mathbf{1}))^2) < \infty. \tag{1.4}$$

We then introduce two fundamental constants,

$$\lambda_1 := E(C) + \mu(a) \quad \text{and} \quad \lambda_2 := E(C). \tag{1.5}$$

As the notation suggests, these two quantities shall be viewed as eigenvalues. More precisely, consider the operator $\mathcal{R} : \mathcal{L}^\infty \rightarrow \mathcal{L}^\infty$ defined by

$$\mathcal{R}f(s) := E_s(Cf(s) + \xi(f)) = \lambda_2 f(s) + a(s)\mu(f), \tag{1.6}$$

which describes averaged replacements as a function of the color of the sampled ball. Then λ_1 and λ_2 arise as the two eigenvalues of \mathcal{R} ; specifically

$$\mathcal{R}a = \lambda_1 a \quad \text{and} \quad \mu(\mathcal{R}f) = \lambda_1 \mu(f) \quad \text{for all } f \in \mathcal{L}^\infty, \tag{1.7}$$

and also

$$\mathcal{R}f = \lambda_2 f \quad \text{whenever } \mu(f) = 0. \tag{1.8}$$

Observe from (1.1) and (1.4) that

$$\lambda_1 = \int_S E_s(C + \xi(\mathbf{1}))\bar{\mu}(ds) > 0,$$

and also that $\lambda_1 > \lambda_2$, since $\mu(a) > 0$.

When the set of colors S is finite, say $S = \{1, \dots, d\}$, the present model with innovation merely rephrases the classical setting of urns with random replacements schemes and

the average replacement operator \mathcal{R} is given by a $d \times d$ matrix. The requirements (1.2) and (1.3) entail that the latter has the form $\mathcal{R} = \lambda_2 \text{Id} + \mathcal{R}_1$, with \mathcal{R}_1 a matrix of rank 1. In this situation, the spectral properties of \mathcal{R} observed above are immediate, and the results in this work should not come much as a surprise. Nonetheless dealing with a general space of colors cannot be reduced (e.g. by approximations) to the case when S is finite. Although our analysis will of course borrow the same guiding lines as for the classical setting, in particular spectral analysis of the mean replacement operator and martingale limit theorems, new ideas are also needed to resolve several issues.

An important result for urn schemes with finitely many colors is that the first order asymptotic of the contents of the urn as the number of steps goes to infinity is determined by the largest eigenvalue of the mean replacement matrix and its eigenvectors. The same feature holds in the present setting; it implies the almost-sure convergence of the empirical distribution of colors to the normalised intensity of innovation.

Theorem 1.1. *Assume (1.2), (1.3), and (1.4). For every $f \in \mathcal{L}^\infty$, we have*

$$\lim_{n \rightarrow \infty} n^{-1} U_n(f) = \lambda_1 \bar{\mu}(f) \quad \text{a.s.},$$

where $\bar{\mu} := \mu(\mathbf{1})^{-1} \mu$ is the normalised intensity of innovation. As a consequence, we have also in terms of the empirical distribution of colors that

$$\lim_{n \rightarrow \infty} \bar{U}_n(f) = \bar{\mu}(f) \quad \text{a.s.} \tag{1.9}$$

We next turn our attention to the fluctuations of the empirical distributions; in order to state our main result, we need to introduce first a few more objects. For any $f \in \mathcal{L}^\infty$, we set

$$\sigma^2(f) := \int_S E_s ((f(s)C + \xi(f))^2) \bar{\mu}(ds), \tag{1.10}$$

which defines a positive semidefinite quadratic form on \mathcal{L}^∞ . We denote by $G = (G(f) : f \in \mathcal{L}^\infty)$ the associated Gaussian process, i.e. G is centered with covariance

$$\mathbb{E}(G(f)G(g)) = \int_S E_s ((f(s)C + \xi(f))(g(s)C + \xi(g))) \bar{\mu}(ds). \tag{1.11}$$

In general, the variables $G(f)$ and $G(g)$ are not always independent when f and g are orthogonal in $\mathcal{L}^2(\mu)$, due the component $\mathbb{E}(\xi(f)\xi(g))$ in the covariance. We also introduce the Gaussian bridge $G^{(\text{br})} = (G^{(\text{br})}(f) : f \in \mathcal{L}^\infty)$, where

$$G^{(\text{br})}(f) := G(f) - \bar{\mu}(f)G(\mathbf{1}).$$

Plainly, the Gaussian bridge is translation invariant in the sense that $G^{(\text{br})}(f) = G^{(\text{br})}(f + c)$ for any $c \in \mathbb{R}$, and coincides with G on the hyperplane of functions $f \in \mathcal{L}^\infty$ with $\mu(f) = 0$.

It is well-known for classical urn schemes with finitely many colors, that the fluctuations of the empirical distributions of colors in the urn depend crucially of whether the largest eigenvalue of the mean replacement matrix is larger or smaller than twice the real part of the second largest eigenvalue. This incites us to introduce

$$\rho := \frac{\lambda_2}{\lambda_1} \in (-\infty, 1). \tag{1.12}$$

In particular, $\rho > 1/2$ if and only if $E(C) > \mu(a)$, which we interpret as reinforcement being stronger than innovation.

Theorem 1.2. Assume (1.2), (1.3), and (1.4). Then the following limits hold:

(i) If $\rho > 1/2$, then we have for every $f \in \mathcal{L}^\infty$:

$$\lim_{n \rightarrow \infty} n^{1-\rho} (\bar{U}_n(f) - \bar{\mu}(f)) = L_f \quad \text{a.s.},$$

where L_f is some random variable which is non degenerate, except when there is a constant $c \in \mathbb{R}$ such that $P_s(Cf(s) + \xi(f) = c) = 1$ for μ almost every $s \in S$.

(ii) If $\rho < 1/2$, and further a slightly stronger version of (1.4),

$$\sup_{s \in S} E_s ((C + \xi(\mathbf{1}))^\alpha) < \infty \quad \text{for some } \alpha > 2, \quad (1.13)$$

holds, then the sequence of processes

$$\left(\lambda_1 \sqrt{(1-2\rho)n} (\bar{U}_n(f) - \bar{\mu}(f)) : f \in \mathcal{L}^\infty \right), \quad n \geq 1$$

converges in the sense of finite dimensional distributions to the Gaussian bridge $G^{(\text{br})}$.

(iii) If $\rho = 1/2$ and (1.13) holds, then the sequence

$$\left(\lambda_1 \sqrt{\frac{n}{\log n}} (\bar{U}_n(f) - \bar{\mu}(f)) : f \in \mathcal{L}^\infty \right), \quad n \geq 1$$

converges in the sense of finite dimensional distributions to the Gaussian bridge $G^{(\text{br})}$.

Remark 1.3. (i) Theorem 1.2(ii-iii) should be viewed as weaker analogs of Donsker's limit theorem for empirical distributions. It would be interesting to strengthen convergence in the sense of finite dimensional distributions here to a uniform central limit type theorem on appropriate classes of functions as in Dudley [9].

(ii) For instance, Simon's model with words given by uniform random variables discussed at the beginning of this introduction corresponds to the case where $S = [0, 1]$, and P_s is the law of a pair (C, ξ) where C is a Bernoulli variable with parameter p and $\xi = (1 - C)\delta_V$, with V a uniform variable on $[0, 1]$ independent of C . Hence P_s does not depend on s , we may take $\mu(ds) = (1 - p)ds$, $a \equiv 1$, $\lambda_1 = 1$ and $\lambda_2 = p$. One finds $\sigma^2(f) = \int_0^1 f^2(s)ds$, so $G(f)$ is simply the Wiener integral of $f \in \mathcal{L}^\infty$ with respect to a Brownian motion. Theorem 1.2 thus agrees with [8, Theorem 1.2].

(iii) It would also be interesting to further investigate fluctuations around the second order limit in the super-critical case $\rho > 1/2$. In the setting of Simon's model, Bertenghi [7] observed that these are asymptotically normal; see also [18].

The number C of copies of the sampled ball which are returned at a typical step does not play much of a role in (1.9), meaning that for such urn schemes with innovation, reinforcement does not impact the first order asymptotic of the empirical measures of colors in the urn. The effects of the reinforcement on the fluctuations can now be analyzed by inspection of the formula (1.11) for the covariance. For the sake of simplicity, let us focus here on the simpler case when the law P_s of the replacement pair (C, ξ) does not depend on the sampled color s , viz. $P_s = P$ for all $s \in S$. In short, consider first the case $C \equiv 0$ without reinforcement, where at each step one only returns in the urn the atoms of a point process ξ . By (1.9), the empirical distribution of colors still converges to $\bar{\mu}$, and by Theorem 1.2, the fluctuations are governed by a Gaussian bridge $G_0^{(\text{br})}$ with covariance

$$\mathbb{E}(G_0^{(\text{br})}(f)G_0^{(\text{br})}(g)) = E(\xi(f)\xi(g)) \quad \text{when } \mu(f) = \mu(g) = 0.$$

The comparison with (1.11) shows that the reinforced scheme with $C \neq 0$ induces an additional white noise component in the fluctuations, which is seen from the term $\bar{\mu}(fg)E(C^2)$. Keep in mind that the reinforcement variable C also impacts the eigenvalues λ_1 and λ_2 , and in particular that increasing C while keeping the innovation ξ unchanged also increases the ratio ρ .

It has been briefly mentioned at the beginning of this introduction that urns schemes with infinite colors, or measure-valued Pólya processes, have been considered in several recent works [3, 4, 5, 15, 16, 21, 22], and we shall now recast our contribution from this perspective. In [5] and [21], the authors consider urns processes as Markov chains $(U_n)_{n \geq 0}$ of random measures (not necessarily point processes) on a measurable space S , such that conditionally given U_0, \dots, U_n , a random color σ is sampled according to the probability measure proportional to U_n and then $U_{n+1} = U_n + R_\sigma$, where $(R_s)_{s \in S}$ is a replacement kernel of positive measures on S . A more general version where now the replacement kernels are random has been developed then in [15] and [22], and the framework there encompasses ours. Convergence of empirical distribution of colors has been established in [22, Theorem 1] or [15, Section 6], or also [3], under certain additional conditions which however need not to be fulfilled in our setting. Fluctuations for measure-valued Pólya processes have been considered very recently in [16], where results bearing the same flavor as our Theorem 5.1 are established. However [16] imposes a strong balance condition that would imply in our case that the total number of returned ball at each step is some constant, independently of the sampled color. Thus, although general urn schemes with random replacement kernels encompass the setting of the present work, and Theorems 1.1, 1.2 and 5.1 are close relatives to results there, the present contribution cannot be reduced to these works either. We have chosen here to consider random replacement kernels induced by point processes only rather than more general random measures as in the works that were just mentioned above, mostly to keep the connection with Pólya's original model as close as possible.

The general route that we will follow to establish Theorems 1.1 and 1.2 is well paved since the work by Athreya and Karlin [2] (see also [14]) for urn schemes with finitely many colors, or more generally for finite-dimensional stochastic approximation [6]. We shall exhibit natural martingales, investigate the asymptotic behavior of their quadratic variations and apply a version of the martingale central limit theorem. Here and there, we shall take a few shortcuts that were not fully available at the time when [2] was written, such as stable convergence (see [13, Section VIII.5c]). Another rather minor difference is that we do not use the embedding of urn schemes in a continuous time Markov branching process known as poissonization, but merely a Poisson subordination. More precisely, Theorem 1.2 will be deduced from Theorem 5.1, which is a more precise functional limit theorem for the continuous time version of the urn scheme with innovation, analogous to results of Janson [14] in the setting of urns with finitely many colors.

The plan for the rest of this text is as follows. Some further notation and preliminary observations on the replacement kernel and on Poisson subordination are presented in Section 2. The foundations of our analysis, namely the natural martingales, are laid down in Section 3; these yield Theorem 1.1. The large time behavior of quadratic variations is determined in Section 4. In Section 5, we establish a functional limit theorem (Theorem 5.1), and finally derive Theorem 1.2 from the latter.

2 Preliminaries

2.1 The replacement kernel

We introduce first some notation, referring to [17, Chapter 1] for background on the notions we shall use. We consider some Borel space endowed with its Borel sigma-

field (S, \mathcal{S}) , and write \mathcal{N}_S for the space of *finite* counting measures on S ; \mathcal{N}_S is also a Borel space. Recall that the replacement kernel $(P_s)_{s \in S}$ defines for every $s \in S$ a probability distribution on $\{-1, 0, 1, 2, \dots\} \times \mathcal{N}_S$ which determines the joint law of the number of copies C and the point process of innovation ξ when the sampled ball has color s . We assume throughout the rest of this text and without further mention that the requirements (1.2), (1.3) and (1.4) are fulfilled.

For every $s \in S$, $k \geq -1$ and $\Phi : \mathcal{N}_S \rightarrow \mathbb{R}_+$ measurable, there is the identity

$$E_s(\Phi(\xi)\mathbf{1}_{\{C=k\}}) = \int_{\mathcal{N}_S} \Phi(\nu)P_s(k, d\nu).$$

In particular, the intensity measure of ξ under P_s is given for every $f \in \mathcal{L}^\infty$ by

$$E_s(\xi(f)) = a(s)\mu(f) = \sum_{k \geq -1} \int_{\mathcal{N}_S} \nu(f)P_s(k, d\nu). \tag{2.1}$$

Observe also from the notation (1.5) that

$$\lambda_2 = E(C) = \sum_{k \geq -1} kP_s(k, \mathcal{N}_S) \quad \text{for any } s \in S. \tag{2.2}$$

We formalize the urn scheme with innovation depicted in the introduction and construct a sequence of point processes $(U_n)_{n \geq 0}$ on S as follows. We start from some arbitrary $U_0 \neq 0$ in \mathcal{N}_S . Next for every $n \geq 0$, we sample a color σ_n according to the current empirical distribution \bar{U}_n and independently of U_0, \dots, U_{n-1} , that is

$$\mathbb{P}(\sigma_n = s \mid U_0, \dots, U_n) = \bar{U}_n(\mathbf{1}_{\{s\}}) = U_n(\mathbf{1}_{\{s\}})/U_n(\mathbf{1}), \quad s \in S.$$

Then given $\sigma_n = s$, we set

$$U_{n+1} = U_n + C_{n+1}\delta_s + \xi_{n+1}, \tag{2.3}$$

where (C_{n+1}, ξ_{n+1}) is distributed according to P_s and independent of the preceding steps. The point process U_n gives the composition of the urn after n steps. The sequence $(U_n)_{n \geq 0}$ is a Markov chain on \mathcal{N}_S , with one-step transitions given by (2.3).

2.2 Poisson subordination

One of the most efficient and best know tools in the study of urn schemes is the so-called poissonization, which, roughly speaking, consists of embedding the urn scheme in a continuous time multitype branching process. See for instance, [2], [1, Section V.9], [20, Section 4.6] or [14]. Here, we shall rather use subordination² based on an independent Poisson process. The advantage is that this both circumvents the possibly delicate issue of depoissonization, and still makes some computations (notably for quadratic variations) much simpler than in discrete time. The drawback is that the independence between the counting processes of balls of different colors induced by poissonization is lost here; however this will not be an issue for us.

Specifically, let $N = (N(t))_{t \geq 0}$ denote a standard Poisson process, so the Stieltjes measure $dN(t)$ is a Poisson point process on $[0, \infty)$ with intensity the Lebesgue measure. We assume that N is independent of the urn scheme and use the jump times of N as the set of times at which a ball is sampled from the urn scheme. Specifically, we set

$$X_t = U_{N(t)}, \quad t \geq 0,$$

²Actually, subordination and poissonization are related one to the other by the so-called Lamperti transformation, as it will be briefly discussed in Remark 3.5 below. In this respect, subordination should be viewed as a mild variation of poissonization rather than a new technique in its own right. This enables us to circumvent some technical issues, such as the construction of the non-exploding branching process with general type space.

and think of $X = (X_t)_{t \geq 0}$ as the process describing the composition of the urn as time passes. We write respectively \mathbb{P} and \mathbb{E} for the probability measure and the mathematical expectation induced by this process, and also $(\mathcal{F}_t)_{t \geq 0}$ for its natural filtration. Occasionally, we may prefer to write \mathbb{P}_m and \mathbb{E}_m to stress that the initial composition of the urn is $U_0 = X_0 = m$ for some non-zero $m \in \mathcal{N}_S$.

We shall need some elementary bounds on the growth of the number of balls in the urn.

Lemma 2.1. *The following assertions hold:*

(i) *There exists some finite constant $\gamma < \infty$ such that*

$$\limsup_{t \rightarrow \infty} t^{-1} X_t(\mathbf{1}) \leq \gamma \quad \text{a.s.}$$

(ii) *There exists $\varepsilon > 0$ such that*

$$\lim_{t \rightarrow \infty} e^{\varepsilon t} \mathbb{P}(X_t \leq \varepsilon t) = 0.$$

Proof. (i) Observe from (1.4) and Markov's inequality that

$$\sup_{s \in S} P_s(C + \xi(\mathbf{1}) > b) = O(b^{-2}) \quad \text{as } b \rightarrow \infty.$$

We infer that there is some integrable variable, whose expectation is denoted by γ , that dominates stochastically the number $C + \xi(\mathbf{1})$ of balls which are added to the urn at any step. It follows from the law of large numbers that the total number of balls in the urn growth at most linearly,

$$\sup_{n \geq 1} n^{-1} U_n(\mathbf{1}) \leq \gamma \quad \text{a.s.}$$

and we conclude applying the law of large numbers for the Poisson process N (see, e.g., [10, Eq. (2.5) on page 456]).

(ii) Indeed, the process $(X_t(\mathbf{1}) - X_0(\mathbf{1}))_{t \geq 0}$ which counts the number of added balls can be bounded from below by some Poisson process with rate $\inf_{s \in S} P_s(C + \xi(\mathbf{1}) \geq 1)$, which is strictly positive by (1.4). The assertion is plain from elementary large deviations estimates. \square

Subordination turns Markov chains into Markov jump processes. In our setting, we get from (2.3) that the infinitesimal generator \mathcal{G} of X is given for any bounded measurable functional $\Phi : \mathcal{N}_S \rightarrow \mathbb{R}$ and any non-zero measure $m \in \mathcal{N}_S$ by

$$\begin{aligned} \mathcal{G}\Phi(m) &= \lim_{t \rightarrow 0^+} t^{-1} \mathbb{E}_m(\Phi(X_t) - \Phi(X_0)) \\ &= \int_S \bar{m}(ds) \sum_{k=-1}^{\infty} \int_{\mathcal{N}_S} (\Phi(m + k\delta_s + \nu) - \Phi(m)) P_s(k, d\nu), \end{aligned} \quad (2.4)$$

with the notation $\bar{m}(ds) = m(ds)/m(\mathbf{1})$.

The starting point of our analysis based on the following observation. We shall consider linear functionals on \mathcal{N}_S , that is of the type $\Phi(m) = m(f)$ for some $f \in \mathcal{L}^\infty$. Of course, such functionals are not bounded as soon as the support of f is infinite; nonetheless (2.4) still makes sense and we have for $m \neq 0$

$$\mathcal{G}\Phi(m) = \bar{m}(\mathcal{R}f),$$

where $\mathcal{R} : \mathcal{L}^\infty \rightarrow \mathcal{L}^\infty$ is the average replacement operator defined by (1.6).

In the next section, we will use these observations to exhibit simple martingales. Recall from (1.7) and (1.8) that λ_1 and λ_2 appear in this setting as the two eigenvalues of \mathcal{R} .

3 Key martingales

Recall that the function $a \in \mathcal{L}^\infty$ appearing in (1.3) modulates the intensity of innovations. We set

$$M_t(a) := X_t(a) \exp \left(-\lambda_1 \int_0^t \frac{dr}{\bar{X}_r(\mathbf{1})} \right), \quad t \geq 0.$$

We also introduce for every $f \in \mathcal{L}^\infty$ with $\mu(f) = 0$

$$M_t(f) := X_t(f) \exp \left(-\lambda_2 \int_0^t \frac{dr}{\bar{X}_r(\mathbf{1})} \right), \quad t \geq 0.$$

Beware that the factor in the exponential is λ_1 in the first definition and λ_2 in the second.

Lemma 3.1. *The processes $M(a)$ and $M(f)$ are square integrable martingales.*

Proof. Recall first from (1.4) and (2.3) that

$$\mathbb{E} (|U_{n+1}(\mathbf{1}) - U_n(\mathbf{1})|^2) \leq \sup_{s \in S} E_s \left((C + \xi(\mathbf{1}))^2 \right) < \infty.$$

It follows that $\mathbb{E}(X_t(\mathbf{1})^2) < \infty$ for every $t \geq 0$, and therefore processes $M(a)$ and $M(f)$ are square integrable.

We will henceforth focus on $M(a)$, as the proof of the martingale property for $M(f)$ relies on similar arguments. Our main task is to check that

$$\text{the process } \left(X_t(a) - \lambda_1 \int_0^t \bar{X}_r(a) dr \right)_{t \geq 0} \text{ is a martingale,} \quad (3.1)$$

since then it follows from the Itô formula that $M(a)$ is a martingale too. In this direction, we consider the bounded functional on \mathcal{N}_S given by $\Phi_b(m) = m(a) \wedge b$ for some $b > 0$. We know that

$$\left(\Phi_b(X_t) - \int_0^t \mathcal{G}\Phi_b(X_r) dr \right)_{t \geq 0} \text{ is a martingale,}$$

and we analyze the limit as $b \rightarrow \infty$.

Plainly, $\Phi_b(X_t)$ increases to $X_t(a) \in L^1(\mathbb{P})$ as b increases to ∞ , and therefore $\Phi_b(X_t)$ converges to $X_t(a)$ in $L^1(\mathbb{P})$. Furthermore, for every $m \in \mathcal{N}_S$,

$$|\Phi_b(m + k\delta_s + \nu) - \Phi_b(m)| \leq \|a\|_\infty (|k| + \nu(\mathbf{1})),$$

and since, thanks to (1.4),

$$\int_S \bar{m}(ds) \sum_{k=-1}^{\infty} \int_{\mathcal{N}_S} (|k| + \nu(\mathbf{1})) P_s(k, d\nu) \leq \sup_{s \in S} E_s (|C| + \xi(\mathbf{1})) < \infty,$$

we deduce from (2.4) by dominated convergence that for every $r \geq 0$, there is the almost-sure convergence

$$\begin{aligned} \lim_{b \rightarrow \infty} \mathcal{G}\Phi_b(X_r) &= \int_S \bar{X}_r(ds) \sum_{k=-1}^{\infty} \int_{\mathcal{N}_S} (ka(s) + \nu(a)) P_s(k, d\nu) \\ &= \bar{X}_r(a) (\lambda_2 + \mu(a)) \\ &= \lambda_1 \bar{X}_r(a), \end{aligned}$$

where for the second identity, we used (2.1) and (2.2).

On the other hand, the bounds above show that for any $m \in \mathcal{N}_S$,

$$|\mathcal{G}\Phi_b(m)| \leq \|a\|_\infty \sup_{s \in S} E_s(|C| + \xi(\mathbf{1})).$$

This enables us to apply dominated convergence, and for every $t \geq 0$

$$\lim_{b \rightarrow \infty} \int_0^t \mathcal{G}\Phi_b(X_r) dr = \lambda_1 \int_0^t \bar{X}_r(a) dr \quad \text{in } L^1(\mathbb{P}).$$

This completes the proof of (3.1). □

We next prepare the ground for some calculations involving quadratic variations. The urn scheme (2.3) yields that for any $f \in \mathcal{L}^\infty$,

$$\mathbb{E}((U_{n+1}(f) - U_n(f))^2 | U_0, \dots, U_n) = Q(U_n, f),$$

where for every non-zero $m \in \mathcal{N}_S$,

$$Q(m, f) := \int_S \bar{m}(ds) E_s((f(s)C + \xi(f))^2). \tag{3.2}$$

Since X is obtained from U by subordination via an independent Poisson process, its optional quadratic variation process

$$[X(f)]_t := \sum_{0 < r \leq t} |\Delta X_r(f)|^2, \quad t \geq 0,$$

has an absolutely continuous predictable compensator $\langle X(f) \rangle$ given by

$$d\langle X(f) \rangle_t = Q(X_t, f) dt. \tag{3.3}$$

See [13, §I.3b] for background.

The martingale $M(a)$ is positive, hence has a terminal value

$$M_\infty(a) := \lim_{t \rightarrow \infty} M_t(a).$$

Lemma 3.2. *The martingale $M(a)$ is a bounded in $L^2(\mathbb{P})$ and $M_\infty(a) > 0$, \mathbb{P} -a.s.*

Proof. We will check that the terminal value of the oblique bracket $\langle M(a) \rangle_\infty$ is integrable. To start with, $M(a)$ is a pure jump martingale with quadratic variation

$$d[M(a)]_t = \exp\left(-2\lambda_1 \int_0^t \frac{dr}{X_r(\mathbf{1})}\right) d[X(a)]_t.$$

The exponential term in the right-hand side is a continuous adapted process, so, using (3.3) for the second equality below, the predictable compensator $\langle M(f) \rangle$ of $[M(f)]$ is given by

$$\begin{aligned} d\langle M(a) \rangle_t &= \exp\left(-2\lambda_1 \int_0^t \frac{dr}{X_r(\mathbf{1})}\right) d\langle X(a) \rangle_t \\ &= Q(X_t, a) \exp\left(-2\lambda_1 \int_0^t \frac{dr}{X_r(\mathbf{1})}\right) dt, \end{aligned}$$

see [13, Theorem I.3.17]. Then observe from (3.2) that

$$Q(X_t, a) \leq \beta := \|a\|_\infty^2 \sup_{s \in S} E_s((|C| + \xi(\mathbf{1}))^2),$$

and recall from the assumption (1.4) that $\beta < \infty$. We have thus established the bound

$$\langle M(a) \rangle_t \leq \beta \int_0^t \exp\left(-2\lambda_1 \int_0^r \frac{du}{X_u(\mathbf{1})}\right) dr.$$

We next introduce the time substitution

$$T(t) := \inf \left\{ r \geq 0 : \int_0^r \frac{du}{X_u(\mathbf{1})} > t \right\}, \quad t \geq 0 \tag{3.4}$$

and stress that $T(t) < \infty$ a.s. for any $t > 0$, thanks to Lemma 2.1(i). Write

$$\beta^{-1} \langle M(a) \rangle_\infty \leq \int_0^\infty \exp\left(-2\lambda_1 \int_0^r \frac{du}{X_u(\mathbf{1})}\right) dr = \int_0^\infty X_{T(t)}(\mathbf{1}) e^{-2\lambda_1 t} dt.$$

We shall establish below that for any non-zero initial value $m \in \mathcal{N}_S$,

$$\mathbb{E}_m (X_{T(t)}(\mathbf{1})) \leq \frac{m(a)}{\bar{\mu}(a)} e^{\lambda_1 t}. \tag{3.5}$$

This entails the integrability of $\langle M(a) \rangle_\infty$, hence the boundedness of $M(a)$ in $L^2(\mathbb{P}_m)$, and more precisely that

$$\mathbb{E}_m (M_\infty(a)^2) \leq m(a)^2 + \frac{\beta m(a)}{\bar{\mu}(a)\lambda_1}. \tag{3.6}$$

Note first that since $T(t)$ is an \mathcal{F}_t -stopping time and $M(a)$ a nonnegative martingale, $\mathbb{E}_m(M_{T(t)}(a)) \leq m(a)$, and hence

$$\mathbb{E}_m (X_{T(t)}(a)) \leq m(a) e^{\lambda_1 t}.$$

Next, consider the function $f = a - \bar{\mu}(a)\mathbf{1}$, which has $\mu(f) = 0$. Again by optional sampling applied now to the martingale $M(f)$, we have for every bounded stopping time τ that

$$\bar{\mu}(a) \mathbb{E}_m \left(X_\tau(\mathbf{1}) \exp\left(-\lambda_2 \int_0^\tau \frac{dr}{X_r(\mathbf{1})}\right) \right) = \mathbb{E}_m \left(X_\tau(a) \exp\left(-\lambda_2 \int_0^\tau \frac{dr}{X_r(\mathbf{1})}\right) \right).$$

Apply this to $\tau = T(t) \wedge b$ and let $b \rightarrow \infty$. Since

$$X_{T(t) \wedge b}(a) \leq \|a\|_\infty X_{T(t) \wedge b}(\mathbf{1}) \leq \|a\|_\infty X_{T(t)}(\mathbf{1})$$

and $X_{T(t)}(\mathbf{1}) \in L^1(\mathbb{P}_m)$, we get by dominated convergence that

$$\begin{aligned} \lim_{b \rightarrow \infty} \mathbb{E}_m \left(X_{T(t) \wedge b}(a) \exp\left(-\lambda_2 \int_0^{T(t) \wedge b} \frac{dr}{X_r(\mathbf{1})}\right) \right) &= \mathbb{E}_m (X_{T(t)}(a) e^{-\lambda_2 t}) \\ &\leq m(a) e^{(\lambda_1 - \lambda_2)t}. \end{aligned}$$

By Fatou's Lemma, we get

$$\begin{aligned} e^{-\lambda_2 t} \mathbb{E}_m (X_{T(t)}(\mathbf{1})) &\leq \liminf_{b \rightarrow \infty} \mathbb{E}_m \left(X_{T(t) \wedge b}(\mathbf{1}) \exp\left(-\lambda_2 \int_0^{T(t) \wedge b} \frac{dr}{X_r(\mathbf{1})}\right) \right) \\ &\leq \frac{m(a)}{\bar{\mu}(a)} e^{(\lambda_1 - \lambda_2)t}. \end{aligned}$$

The proof of (3.5) is now complete, and we turn our attention to the strict positivity of the terminal value.

An application of Chebychev’s inequality in (3.6) yields

$$\mathbb{P}_m(M_\infty(a) = 0) \leq \frac{\beta}{\bar{\mu}(a)\lambda_1 m(a)}.$$

Observe that for any $t_0 > 0$, there is the equivalence

$$M_\infty(a) = 0 \iff \lim_{t \rightarrow \infty} X_{t+t_0}(a) \exp\left(-\lambda_1 \int_{t_0}^{t_0+t} \frac{dr}{X_r(\mathbf{1})}\right) = 0.$$

So by the Markov property at time t_0 ,

$$\mathbb{P}_m(M_\infty(a) = 0) \leq \frac{\beta}{\bar{\mu}(a)\lambda_1} \mathbb{E}_m(1/X_{t_0}(a)).$$

To conclude that $\mathbb{P}_m(M_\infty(a) = 0) = 0$, we only need to let $t_0 \rightarrow \infty$ and note that $\lim_{t_0 \rightarrow \infty} X_{t_0}(a) = \infty$ a.s., since a is bounded away from 0 and the total number of balls in the urn grows to ∞ as time goes to ∞ a.s. \square

We next turn our attention to the predictable quadratic variation of the martingale $M(f)$. Recall the notation (3.2).

Lemma 3.3. *Let $f \in \mathcal{L}^\infty$ with $\mu(f) = 0$. The oblique bracket of the square-integrable martingale $M(f)$ is absolutely continuous with,*

$$d\langle M(f) \rangle_t = Q(X_t, f) \exp\left(-2\lambda_2 \int_0^t \frac{dr}{X_r(\mathbf{1})}\right) dt.$$

Proof. The formula follows from the same calculation as at the beginning of the proof of Lemma 3.2. \square

Lemma 3.3 enables us to bound the square moments of $M(f)$, and yields the following.

Corollary 3.4. *For every $f \in \mathcal{L}^\infty$, we have*

$$\lim_{t \rightarrow \infty} X_t(f) \exp\left(-\lambda_1 \int_0^t \frac{dr}{X_r(\mathbf{1})}\right) = \frac{\mu(f)}{\mu(a)} M_\infty(a), \quad \text{a.s. and in } L^2(\mathbb{P}).$$

Proof. In the case $f = a$, the claim amounts to the convergence of the martingale $M(a)$ to its terminal value; see Lemma 3.2. By linearity, it thus suffices to check that for every $f \in \mathcal{L}^\infty$ with $\mu(f) = 0$, one has

$$\lim_{t \rightarrow \infty} M_t(f) \exp\left(-(\lambda_1 - \lambda_2) \int_0^t \frac{dr}{X_r(\mathbf{1})}\right) = 0 \quad \text{a.s. and in } L^2(\mathbb{P}).$$

In this direction, recall the time-substitution (3.4) and rewrite the above in the simpler form

$$\lim_{t \rightarrow \infty} e^{-(\lambda_1 - \lambda_2)t} M_{T(t)}(f) = 0 \quad \text{a.s. and in } L^2(\mathbb{P}). \tag{3.7}$$

The assumption (1.4) ensures that

$$q(f) := \sup_{m \in \mathcal{N}_S} Q(m, f) < \infty, \tag{3.8}$$

and we deduce from Lemma 3.3 that the time-changed process $(M_{T(t)}(f))_{t \geq 0}$ is still a square integrable martingale with oblique bracket

$$\begin{aligned} \langle M(f) \rangle_{T(t)} &\leq q(f) \int_0^{T(t)} \exp\left(-2\lambda_2 \int_0^v \frac{dr}{X_r(\mathbf{1})}\right) dv \\ &\leq q(f) \int_0^t e^{-2\lambda_2 r} X_{T(r)}(\mathbf{1}) dr. \end{aligned}$$

The process $e^{-(\lambda_1-\lambda_2)t}M_{T(t)}(f)$ is a supermartingale, and it therefore suffices to verify that the limit in (3.7) takes place in $L^2(\mathbb{P})$. We compute

$$\begin{aligned} \mathbb{E}(M_{T(t)}(f)^2) &= \mathbb{E}(X_0(f)^2) + \mathbb{E}(\langle M(f) \rangle_{T(t)}) \\ &\leq \mathbb{E}(X_0(f)^2) + q(f) \int_0^t e^{-2\lambda_2 r} \mathbb{E}(X_{T(r)}(\mathbf{1})) \, dr. \end{aligned}$$

We now see from (3.5) and the fact that $\lambda_1 > \lambda_2$ that

$$\lim_{t \rightarrow \infty} e^{-2(\lambda_1-\lambda_2)t} \mathbb{E}(M_{T(t)}(f)^2) = 0,$$

which completes the proof. □

Remark 3.5. The time-substitution $T(t)$ used in the proofs of Lemma 3.2 and Corollary 3.4 is a version of the Lamperti transformation; see [19, Chapter 12]. It turns the measure-valued process X into a continuous-time branching process with types in S . In this setting of branching processes, the process $M_{T(\cdot)}(a)$ should be viewed as a Malthusian martingale, and Lemma 3.2 and Corollary 3.4 are then well-known; see [1, Theorems 7.1 and 7.2].

Corollary 3.6. *For every $f \in \mathcal{L}^\infty$, we have*

$$\lim_{t \rightarrow \infty} t^{-1} X_t(f) = \lambda_1 \bar{\mu}(f), \quad \text{a.s.}$$

Proof. From Corollary 3.4, since the terminal value $M_\infty(a)$ is almost surely finite and strictly positive, it suffices to check that

$$\lim_{t \rightarrow \infty} t \exp\left(-\lambda_1 \int_0^t \frac{dr}{X_r(\mathbf{1})}\right) = \frac{M_\infty(a)}{\lambda_1 \bar{\mu}(a)}, \quad \text{a.s.} \tag{3.9}$$

For this, we apply Corollary 3.4 with $f = \mathbf{1}$ and get

$$\lim_{t \rightarrow \infty} \frac{1}{X_t(\mathbf{1})} \exp\left(\lambda_1 \int_0^t \frac{dr}{X_r(\mathbf{1})}\right) = \bar{\mu}(a)/M_\infty(a) \quad \text{a.s.}$$

As a consequence, we have

$$\lim_{t \rightarrow \infty} t^{-1} \int_0^t \frac{1}{X_u(\mathbf{1})} \exp\left(\lambda_1 \int_0^u \frac{dr}{X_r(\mathbf{1})}\right) \, du = \bar{\mu}(a)/M_\infty(a) \quad \text{as } t \rightarrow \infty.$$

It suffices to observe that

$$\int_0^t \exp\left(\lambda_1 \int_0^u \frac{dr}{X_r(\mathbf{1})}\right) \frac{du}{X_u(\mathbf{1})} = \frac{1}{\lambda_1} \left(\exp\left(\lambda_1 \int_0^t \frac{dr}{X_r(\mathbf{1})}\right) - 1 \right),$$

which yields (3.9). □

Obviously, Theorem 1.1 follows from Corollary 3.6 and the law of large numbers for the Poisson process N .

4 Asymptotic behaviors of quadratic variations

Limit theorems for martingales are of course the key ingredient for Theorem 1.2. Their applications require controlling quadratic variations, and for this we use mainly Lemmas 3.2 and 3.3 and (3.9). We start with a simple observation; recall the notation (1.12).

Corollary 4.1. *Suppose $\rho > 1/2$. Then for every $f \in \mathcal{L}^\infty$ with $\mu(f) = 0$, there is the convergence a.s.*

$$\lim_{t \rightarrow \infty} t^{-\rho} X_t(f) = L'_f;$$

and L'_f is non-degenerate except if for μ almost every $s \in S$, one has $Cf(s) + \xi(f) = 0$ P_s -almost surely.

Proof. Recall the notation (3.8) and apply Lemma 3.3 to get

$$\begin{aligned} \langle M(f) \rangle_\infty &\leq q(f) \int_0^\infty \exp\left(-2\lambda_2 \int_0^t \frac{dr}{X_r(\mathbf{1})}\right) dt \\ &\leq q(f) \int_0^\infty e^{-2\lambda_2 r} X_{T(r)}(\mathbf{1}) dr. \end{aligned}$$

Since $2\lambda_2 > \lambda_1$, we deduce from (3.5) that $\mathbb{E}(\langle M(f) \rangle_\infty) < \infty$. The martingale $M(f)$ thus converges a.s. and in $L^2(\mathbb{P})$. It only remains to use (3.9).

Since the convergence holds in $L^2(\mathbb{P})$, the terminal value is degenerate if and only if $\langle M(f) \rangle_\infty = 0$ \mathbb{P} -a.s. From Lemma 3.3, this is the case if and only if $Q(X_t, f) = 0$ for almost all $t \geq 0$, \mathbb{P} -a.s. Combining (3.2) and Theorem 1.1, we now see that this forces $P_s(Cf(s) + \xi(f) = 0) = 1$ for μ -a.e. $s \in S$. \square

We next turn our attention to the case $\rho \leq 1/2$ and first consider the growth of oblique brackets. Recall the notation (1.10).

Corollary 4.2. *The following assertions hold for every $f \in \mathcal{L}^\infty$ with $\mu(f) = 0$:*

(i) *If $\rho < 1/2$, then we have*

$$\lim_{t \rightarrow \infty} t^{2\rho-1} \langle M(f) \rangle_t = \frac{\sigma^2(f)}{1-2\rho} \left(\frac{M_\infty(a)}{\lambda_1 \bar{\mu}(a)} \right)^{2\rho} \quad \text{a.s.}$$

(ii) *If $\rho = 1/2$, then we have*

$$\lim_{t \rightarrow \infty} \frac{1}{\log t} \langle M(f) \rangle_t = \sigma^2(f) \frac{M_\infty(a)}{\lambda_1 \bar{\mu}(a)} \quad \text{a.s.}$$

Proof. Recall from the notation (3.2) that

$$Q(X_t, f) = \bar{X}_t(\varphi), \quad \text{with } \varphi(s) := E_s((f(s)C + \xi(f))^2) \text{ for } s \in S.$$

It then follows from Corollary 3.6 and (1.10) that

$$\lim_{t \rightarrow \infty} Q(X_t, f) = \bar{\mu}(\varphi) = \sigma^2(f) \quad \text{a.s.}$$

We deduce from Lemma 3.3 and (3.9) that almost surely

$$\frac{d\langle M(f) \rangle_t}{dt} \sim \sigma^2(f) \left(\frac{M_\infty(a)}{t \lambda_1 \bar{\mu}(a)} \right)^{2\rho} \quad \text{as } t \rightarrow \infty.$$

This yields our claims by integration. \square

We shall also need the following bounds on the p -variation of $M(f)$.

Lemma 4.3. Assume (1.13) and let $f \in \mathcal{L}^\infty$ with $\mu(f) = 0$. If $\rho < 1/2$, then we have for any $p \in (2, \alpha)$ with $p\rho < 1$ that

$$\mathbb{E} \left(\sum_{r \leq t} |\Delta M_r(f)|^p \right) = O(t^{1-p\rho}) \quad \text{as } t \rightarrow \infty.$$

If $\rho = 1/2$, then there is some $p > 2$ such that

$$\mathbb{E} \left(\sum_{t \geq 0} |\Delta M_t(f)|^p \right) < \infty.$$

Proof. Note first that for any $\varepsilon > 0$ and $t \geq 0$, there is the inequality

$$\exp \left(-2\lambda_1 \int_0^t \frac{du}{X_u(\mathbf{1})} \right) \leq \left(\frac{X_t(\mathbf{1})}{\varepsilon t} \right)^2 \exp \left(-2\lambda_1 \int_0^t \frac{du}{X_u(\mathbf{1})} \right) + \mathbf{1}_{\{X_t(\mathbf{1}) < \varepsilon t\}}.$$

Choosing ε sufficiently small and using Lemma 2.1 and Corollary 3.4 for $f = \mathbf{1}$, we get

$$\mathbb{E} \left(\exp \left(-2\lambda_1 \int_0^t \frac{du}{X_u(\mathbf{1})} \right) \right) = O(t^{-2}). \quad (4.1)$$

Next, computing predictable compensation in the same way as we did in the proof of Lemma 3.2, we get for any $p > 0$

$$\begin{aligned} & \mathbb{E} \left(\sum_{r \leq t} |\Delta M_r(f)|^p \right) \\ & \leq \|f\|_\infty^p \sup_{s \in S} E_s(|C| + \xi(\mathbf{1}))^p \int_0^t \mathbb{E} \left(\exp \left(-p\lambda_2 \int_0^r \frac{du}{X_u(\mathbf{1})} \right) \right) dr. \end{aligned}$$

Provided that $p\lambda_2 \leq 2\lambda_1$, that is $p\rho \leq 2$, we have by Jensen's inequality and (4.1) that

$$\mathbb{E} \left(\exp \left(-p\lambda_2 \int_0^r \frac{du}{X_u(\mathbf{1})} \right) \right) = O(t^{-p\rho}).$$

The statement is now clear when $\rho < 1/2$, and in the case $\rho = 1/2$, we may just take any $p \in (2, \alpha]$. \square

The asymptotic behavior of optional quadratic variations can be deduced in turn from Corollary 4.2 and the following observation.

Lemma 4.4. Assume (1.13) and that $\rho \leq 1/2$. Then for every $f \in \mathcal{L}^\infty$ with $\mu(f) = 0$, we have

$$\lim_{t \rightarrow \infty} \frac{[M(f)]_t}{\langle M(f) \rangle_t} = 1 \quad \text{in probability.}$$

Proof. Suppose first that $\rho < 1/2$. By Corollary 4.2, it suffices to verify that

$$\lim_{t \rightarrow \infty} t^{2\rho-1} \mathbb{E}(|D_t(f)|) = 0, \quad (4.2)$$

where $D(f) := [M(f)] - \langle M(f) \rangle$. Since $\langle M(f) \rangle$ is continuous, D is a pure jump martingale whose jumps coincide with those of $[M(f)]$, so $|\Delta D|^2 = |\Delta M(f)|^4$. We can bound the quadratic variation of D for any $2 < p \leq 4$ by

$$[D]_t \leq \left(\sum_{r \leq t} |\Delta M_r(f)|^p \right)^{4/p}.$$

Choosing p as in Lemma 4.3, and applying the Jensen and the Burkholder-Davis-Gundy inequalities, we get

$$\mathbb{E}(|D_t|) = O(t^{2(1-p\rho)/p}) = o(t^{1-2\rho});$$

so (4.2) is verified.

In the case $\rho = 1/2$, the same argument shows that the martingale D is bounded in $L^{p/2}(\mathbb{P})$ and hence converges a.s. \square

5 Applying stable limit theorems

The purpose of this section is to establish the following functional limit theorem from which Theorem 1.2(ii–iii) will readily follow. The space $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ of càdlàg real-valued paths is endowed with the Skorokhod topology; see [13, Section VI.1]. Recall the notation (1.10), and let also $W = (W(t))_{t \geq 0}$ denote a standard real Brownian motion.

Theorem 5.1. *Assume (1.13) and take any $f \in \mathcal{L}^\infty$ with $\mu(f) = 0$.*

(i) *If $\rho < 1/2$, then there is the convergence in distribution on $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ in the sense of Skorokhod*

$$\lim_{n \rightarrow \infty} \left(\frac{X_{tn}(f)}{\sqrt{n}} \right)_{t \geq 0} = \left(\frac{\sigma(f)}{\sqrt{1-2\rho}} t^\rho W(t^{1-2\rho}) \right)_{t \geq 0}.$$

(ii) *If $\rho = 1/2$, then there is the convergence in distribution on $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ in the sense of Skorokhod*

$$\lim_{n \rightarrow \infty} \left(\frac{X_{e^{nt}}(f)}{\sqrt{ne^{nt}}} \right)_{t \geq 0} = (\sigma(f)W(t))_{t \geq 0}.$$

Proof of Theorem 5.1. Again, we shall only prove (i) as the argument in the critical case $\rho = 1/2$ requires only simple modifications. We first focus on the case $\sigma^2(f) > 0$. The proof relies on general results on the so-called stable convergence for martingales (see [13, Section VIII.5.c] for background) and the technical results we established so far.

We shall apply [13, Theorem VIII.5.50] for the martingale Y resulting from $M(f)$ after the deterministic time change $t \mapsto t^{1/(1-2\rho)}$, namely

$$Y_t := M_{t^{1/(1-2\rho)}}(f), \quad t \geq 0.$$

More precisely, the statement there requests, for the sake of simplicity, boundedness of the jumps of Y (see [13, Eq. VIII.5.48]); however its proof only uses the weaker condition VIII.3.14 for the sequence $n^{-1/2}Y_{nt}$, namely that the sequence $n^{-1/2} \max_{r \leq nt} |\Delta Y_r|$ is uniformly integrable for any fixed t . We may of course choose $t = 1$ without loss of generality, and the weaker condition VIII.3.14 follows from Lemma 4.3. Indeed, we pick some $p > 2$ and write first

$$\max_{r \leq n} |\Delta Y_r|^p = \max_{r \leq n^{1/(1-2\rho)}} |\Delta M_r(f)|^p \leq \sum_{r \leq n^{1/(1-2\rho)}} |\Delta M_r(f)|^p.$$

Assuming further that $p < \alpha$ and $p\rho < 1$, which is always an option, we get from Lemma 4.3 that as $n \rightarrow \infty$,

$$\mathbb{E} \left(\max_{r \leq n} |\Delta Y_r|^p \right) = O(n^{(1-p\rho)/(1-2\rho)}) = O(n^{p/2}),$$

and the sought uniform integrability is checked.

Next, we have from Corollary 4.2(i) and Lemma 4.4,

$$\lim_{n \rightarrow \infty} n^{-1}[Y]_{nt} = \lim_{n \rightarrow \infty} n^{-1}[M]_{(nt)^{1/(1-2\rho)}} = \eta^2 t \quad \text{in probability,}$$

where

$$\eta = \frac{\sigma(f)}{\sqrt{1-2\rho}} \left(\frac{M_\infty(a)}{\lambda_1 \bar{\mu}(a)} \right)^\rho.$$

This is the condition VIII.5.49 in [13].

Theorem VIII.5.50 and Proposition VIII.5.33(ii) from [13] now entail that the sequence of pairs $(\eta, (n^{-1/2}Y_{nt})_{t \geq 0})$ converges in law to $(\eta, \eta W)$, where W is a standard Brownian motion independent of η . Recall also from Lemma 3.2 that $\eta > 0$ a.s. It follows that

$$\frac{n^{-(1-2\rho)/2}}{\eta} Y_{(nt)^{1-2\rho}} = \frac{X_{nt}(f)}{\sqrt{n}} \times \frac{n^\rho}{\eta} \exp \left(-\lambda_2 \int_0^{nt} \frac{dr}{X_r(\mathbf{1})} \right), \quad t \geq 0,$$

converges in law to the time-changed Brownian motion $(W(t^{1-2\rho}))_{t \geq 0}$. To complete the proof, it suffices to recall from (3.9) that

$$\lim_{n \rightarrow \infty} \frac{n^\rho}{\eta} \exp \left(-\lambda_2 \int_0^{nt} \frac{dr}{X_r(\mathbf{1})} \right) = \frac{1}{\eta} \left(\frac{M_\infty(a)}{t \lambda_1 \bar{\mu}(a)} \right)^\rho = \frac{\sqrt{1-2\rho}}{\sigma(f)} t^{-\rho},$$

and apply Slutsky's lemma.

Last, we turn our attention to the degenerate case when $\sigma^2(f) = 0$. Even though our claim should be intuitively clear from Corollary 4.2 and (3.9), making intuition rigorous might be not quite straightforward and we shall actually show a sharper estimate.

Consider the subset of colors

$$S_f := \{s \in S : P_s(Cf(s) + \xi(f) \neq 0) > 0\}.$$

The condition $\sigma^2(f) = 0$ implies that S_f is a μ -null set. We set $g = \mathbf{1}_{S_f}$, so $\mu(g) = 0$ and we know from Lemma 3.1 that $M_t(g)$ is a nonnegative martingale. Therefore, we have

$$\lim_{t \rightarrow \infty} X_t(g) \exp \left(-\lambda_2 \int_0^t \frac{dr}{X_r(\mathbf{1})} \right) < \infty \quad \text{a.s.}$$

(note that when $\lambda_2 < 0$, this implies that there are eventually no more balls with color in S_f left in the urn, a.s.). It follows from (3.9) and Corollary 3.6 that *a fortiori* as $t \rightarrow \infty$,

$$\frac{X_t(g)}{X_t(\mathbf{1})} = O(t^{\rho-1}) \quad \text{a.s.} \tag{5.1}$$

On the other hand, we see from the urn dynamics and the definition of S_f , that if one samples from the urn a ball with color in $S \setminus S_f$ at time t , then $X_t(f) = X_{t-}(f)$ a.s. That is, changes for $X_t(f)$ only occur at times t when a ball with color in S_f is sampled. By an argument similar to that in the proof of Lemma 3.1, we get that if we set

$$c_f := \sup_{s \in S} E_s(|Cf(s) + \xi(f)|),$$

which is a finite quantity thanks to (1.4), then the process

$$|X_t(f)| - c_f \int_0^t \frac{X_r(g)}{X_r(\mathbf{1})} dr, \quad t \geq 0,$$

is a super-martingale. Recall from (5.1) that as $t \rightarrow \infty$, the integral term in the difference is, $O(t^\rho)$ when $\rho > 0$, $O(\log t)$ when $\rho = 0$, or bounded when $\rho < 0$. As a consequence, if we set $X_t^*(f) := \sup_{0 \leq r \leq t} |X_r(f)|$, then the family of variables $\{t^{-\rho} X_t^*(f) : t \geq 1\}$ is tight when $\rho > 0$; the same holds for $\{X_t^*(f) : t \geq 0\}$ when $\rho < 0$, and also for $\{X_t^*(f)/\log t : t \geq 2\}$ when $\rho = 0$. In particular, when $\rho < 1/2$, $t^{-1/2} X_t^*(f)$ converges in probability to 0 as $t \rightarrow \infty$. \square

Theorem 1.2 now follows readily from Corollary 3.6 and Theorems 5.1 and 1.1.

Proof of Theorem 1.2. (i) By linearity, it suffices to consider the case when $\mu(f) = 0$. We write first

$$n^{1-\rho}\bar{U}_n(f) = \frac{n}{U_n(\mathbf{1})} \times n^{-\rho}U_n(f),$$

and recall from Theorem 1.1 that the first term in the product in the right-hand side converges to $1/\lambda_1$ as $n \rightarrow \infty$. We next rewrite Corollary 4.1 in the form

$$\lim_{t \rightarrow \infty} \left(\frac{N(t)}{t} \right)^\rho N(t)^{-\rho} U_{N(t)}(f) = L'_f, \quad \text{a.s.}$$

Since $N(t) \sim t$ a.s., this yields $n^{-\rho}U_n(f) \sim L'_f$ and the claim is proved with $L_f = \lambda^{-1}L'_f$.

(ii) Fix first $f \in \mathcal{L}^\infty$ with $\mu(f) = 0$. Write $0 < \gamma(1) < \gamma(2) < \dots$ for the increasing sequence of the jump times of the Poisson process N , so $N(\gamma(n)) = n$ and $U_n(f) = X_{\gamma(n)}(f)$. Since $\lim_{n \rightarrow \infty} n^{-1}\gamma(n) = 1$ a.s., we deduce from Theorem 5.1(i) that there is the convergence

$$\lim_{n \rightarrow \infty} \left(\frac{1-2\rho}{n} \right)^{1/2} U_n(f) = G(f) = G^{(\text{br})}(f) \quad \text{in distribution,}$$

where G is the centered Gaussian process on \mathcal{L}^∞ with covariance given by (1.11) and $G^{(\text{br})}$ its bridge version. It now follows from Theorem 1.1 and Slutsky's lemma that

$$\lim_{n \rightarrow \infty} \lambda_1 ((1-2\rho)n)^{1/2} \bar{U}_n(f) = G^{(\text{br})}(f) \quad \text{in distribution.}$$

Since \bar{U}_n and $G^{(\text{br})}$ are both translation invariant (in the sense that $\bar{U}_n(f+c) = \bar{U}_n(f)$ and $G^{(\text{br})}(f+c) = G^{(\text{br})}(f)$ for any $c \in \mathbb{R}$), the convergence in distribution above holds for any fixed $f \in \mathcal{L}^\infty$, without the restriction that $\mu(f) = 0$. Finally, that the latter holds more generally jointly for any finite family of functions in \mathcal{L}^∞ stems from the Cramér-Wold device.

(iii) The argument is similar to (ii). □

References

- [1] K. B. Athreya and P. E. Ney, *Branching processes*, Dover Publications, Inc., Mineola, NY, 2004, Reprint of the 1972 original [Springer, New York; MR0373040]. MR2047480
- [2] Krishna B. Athreya and Samuel Karlin, *Embedding of urn schemes into continuous time Markov branching processes and related limit theorems*, *Ann. Math. Statist.* **39** (1968), 1801–1817. MR232455
- [3] Antar Bandyopadhyay, Svante Janson, and Debleena Thacker, *Strong convergence of infinite color balanced urns under uniform ergodicity*, *J. Appl. Probab.* **57** (2020), no. 3, 853–865. MR4148061
- [4] Antar Bandyopadhyay and Debleena Thacker, *Pólya urn schemes with infinitely many colors*, *Bernoulli* **23** (2017), no. 4B, 3243–3267. MR3654806
- [5] Antar Bandyopadhyay and Debleena Thacker, *A new approach to Pólya urn schemes and its infinite color generalization*, *Ann. Appl. Probab.* **32** (2022), no. 1, 46–79. MR4386521
- [6] Albert Benveniste, Michel Métivier, and Pierre Priouret, *Adaptive algorithms and stochastic approximations*, *Applications of Mathematics* (New York), vol. 22, Springer-Verlag, Berlin, 1990, Translated from the French by Stephen S. Wilson. MR1082341
- [7] Marco Bertenghi, *Asymptotic normality of superdiffusive step-reinforced random walks*, arXiv:2101.00906.
- [8] Jean Bertoin, *How linear reinforcement affects Donsker's theorem for empirical processes*, *Probab. Theory Related Fields* **178** (2020), no. 3-4, 1173–1192. MR4168396

- [9] R. M. Dudley, *Uniform central limit theorems*, Cambridge Studies in Advanced Mathematics, vol. 63, Cambridge University Press, Cambridge, 1999. MR1720712
- [10] Stewart N. Ethier and Thomas G. Kurtz, *Markov processes. Characterization and convergence.*, Wiley Ser. Probab. Stat., Hoboken, NJ: John Wiley & Sons, 2005 (English). MR0838085
- [11] Fred M. Hoppe, *The sampling theory of neutral alleles and an urn model in population genetics*, J. Math. Biol. **25** (1987), 123–159 (English). MR0896430
- [12] Iacopo Iacopini, Sta ša Milojević, and Vito Latora, *Network dynamics of innovation processes*, Phys. Rev. Lett. **120** (2018), 048301.
- [13] Jean Jacod and Albert N. Shiryaev, *Limit theorems for stochastic processes*, second ed., Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 288, Springer-Verlag, Berlin, 2003. MR1943877
- [14] Svante Janson, *Functional limit theorems for multitype branching processes and generalized Pólya urns*, Stochastic Process. Appl. **110** (2004), no. 2, 177–245. MR2040966
- [15] Svante Janson, *Random replacements in Pólya urns with infinitely many colours*, Electron. Commun. Probab. **24** (2019), Paper No. 23, 11. MR3953824
- [16] Svante Janson, Cécile Mailler, and Denis Villemonais, *Fluctuations of balanced urns with infinitely many colours*, Electronic Journal of Probability **28** (2023), 1–72. MR4609444
- [17] Olav Kallenberg, *Random measures, theory and applications*, Probability Theory and Stochastic Modelling, vol. 77, Springer, Cham, 2017. MR3642325
- [18] Naoki Kubota and Masato Takei, *Gaussian fluctuation for superdiffusive elephant random walks*, J. Stat. Phys. **177** (2019), no. 6, 1157–1171 (English). MR4034803
- [19] Andreas E. Kyprianou, *Fluctuations of Lévy processes with applications*, second ed., Universitext, Springer, Heidelberg, 2014, Introductory lectures. MR3155252
- [20] Hosam M. Mahmoud, *Pólya urn models*, Texts in Statistical Science Series, CRC Press, Boca Raton, FL, 2009. MR2435823
- [21] Cécile Mailler and Jean-François Marckert, *Measure-valued Pólya urn processes*, Electron. J. Probab. **22** (2017), Paper No. 26, 33. MR3629870
- [22] Cécile Mailler and Denis Villemonais, *Stochastic approximation on noncompact measure spaces and application to measure-valued Pólya processes*, Ann. Appl. Probab. **30** (2020), no. 5, 2393–2438. MR4149532
- [23] M. V. Simkin and V. P. Roychowdhury, *Re-inventing Willis*, Phys. Rep. **502** (2011), no. 1, 1–35. MR2788549
- [24] Herbert A. Simon, *On a class of skew distribution functions*, Biometrika **42** (1955), no. 3/4, 425–440. MR0073085
- [25] Francesca Tria, Vittorio Loreto, Vito Domenico Pietro Servedio, and Steven H. Strogatz, *The dynamics of correlated novelties*, arXiv:1310.1953 (2013).

Acknowledgments. I would like to thank warmly two anonymous referees for their constructive comments, their very careful reading, and for correcting several errors in my original draft.

Electronic Journal of Probability

Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS¹)
- Easy interface (EJMS²)

Economical model of EJP-ECP

- Non profit, sponsored by IMS³, BS⁴, ProjectEuclid⁵
- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund⁶ (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

¹LOCKSS: Lots of Copies Keep Stuff Safe <http://www.lockss.org/>

²EJMS: Electronic Journal Management System: <https://vtex.lt/services/ejms-peer-review/>

³IMS: Institute of Mathematical Statistics <http://www.imstat.org/>

⁴BS: Bernoulli Society <http://www.bernoulli-society.org/>

⁵Project Euclid: <https://projecteuclid.org/>

⁶IMS Open Access Fund: <https://imstat.org/shop/donation/>