

ON THE TRANSIENT BEHAVIOR OF EHRENFEST AND ENGSET PROCESSES

MATHIEU FEUILLET * ** AND
PHILIPPE ROBERT, * *** INRIA

Abstract

Two classical stochastic processes are considered, the Ehrenfest process, introduced in 1907 in the kinetic theory of gases to describe the heat exchange between two bodies, and the Engset process, one of the early (1918) stochastic models of communication networks. In this paper we investigate the asymptotic behavior of the distributions of hitting times of these two processes when the number of particles/sources goes to infinity. Results concerning the hitting times of boundaries in particular are obtained. We rely on martingale methods; a key ingredient is an important family of simple nonnegative martingales, an analogue, for the Ehrenfest process, of the exponential martingales used in the study of random walks or of Brownian motion.

Keywords: Ehrenfest process; Engset formula; exponential martingale; space–time harmonic function; hitting time

2010 Mathematics Subject Classification: Primary 60K25

Secondary 90B22

À la mémoire de Philippe Flajolet

1. Introduction

1.1. The Ehrenfest process

In this paper we consider the following continuous-time version of the classical Ehrenfest urn model. This process has been introduced to study the heat exchange between bodies. We assume that each particle of a set of N particles is located in one of two boxes (bodies), 0 and 1 say. A particle in box 0 goes into box 1 at rate ν and a particle in box 1 goes into box 0 at rate μ . We denote by $E_N(t)$ the number of particles in box 1 at time $t \geq 0$. This birth-and-death process can also be represented as $E_N(t) = Y_1(t) + \dots + Y_N(t)$, where $(Y_i(t), 1 \leq i \leq N)$ are N independent and identically distributed (i.i.d.) Markov jump processes with values in $\{0, 1\}$.

Initially, the model describes a discrete-time process $(Z_N(k))$ and at each unit of time a particle is taken at random (i.e. all particles are equally likely) to be moved from one box to the other; $Z_N(k)$ is the number of particles in box 1 at time k . This corresponds to the symmetrical case, $\mu = \nu$, and, clearly, $E_N(t)$ can be represented as $Z_N(\mathcal{N}_{N\mu}((0, t]))$ if $\mathcal{N}_{N\mu}$ is a Poisson process with rate $N\mu$. The process $(E_N(t))$ follows the same path as $(Z_N(k))$, but on a time scale with a factor $N\mu$. The Markov chain $(Z_N(k))$ is also a random walk on the graph of the *hypercube* $\{0, 1\}^N$, where edges connect elements that differ in only one coordinate. The equilibrium properties of $(E_N(t))$ and $(Z_N(k))$ are fairly well known, in particular a quite precise estimate of the duration of time to reach equilibrium is available. See [2].

Received 8 September 2011; revision received 20 October 2011.

* Postal address: INRIA Paris-Rocquencourt, Domaine de Voluceau, 78153 Le Chesnay, France.

** Email address: mathieu.feuillet@inria.fr

*** Email address: philippe.robert@inria.fr

Results on transient quantities of this process, like the distribution of the first time when the box 0 is empty, are rarer. There are generic results on birth-and-death processes which describe some of these distributions in terms of spectral characteristics of the associated infinitesimal generator: the spectral measure and a family of orthogonal polynomials. See [11] and [12]. In practice, the corresponding orthogonal polynomials and, sometimes, their spectral measure do not have a simple representation; this significantly complicates detailed investigations of these hitting times. In the symmetrical case, $\mu = \nu$, Palacios [16] gave closed-form expressions for the averages of hitting times, and Bingham [1] and Flajolet and Huillet [7] obtained a representation of their distributions. In the general case Crescenzo [3] and Flegg *et al.* [9] provided expressions of the densities. It turns out that the expressions obtained in these papers involve, in general, sums of combinatorial terms for which asymptotic results (when N goes to infinity) may be difficult to obtain. Note that this is nevertheless done in the symmetrical case in [7].

1.2. The Engset process

This is one of the oldest stochastic models of communication networks; see [4]. For this model, there are N sources of communication which are active (state 1) or inactive (state 0). An active source becomes inactive at rate μ . The total number of simultaneous active sources cannot exceed the quantity C_N , the number of circuits of the network. An inactive source can therefore become active only if there are already strictly less than C_N active sources, in which case it occurs at rate ν . If $X_N(t)$ is the number of active communications at time t , when $C_N = N$, the process is just the Ehrenfest process. Otherwise, $(X_N(t))$ can be described as a reflected version of $(E_N(t))$. At equilibrium the expression for the probability that X_N is equal to C_N is known as the *Engset formula*. For transient characteristics, the important quantity is the time it takes to have the full capacity C_N of the network used. For this reason, the distribution of the hitting time of C_N by $(X_N(t))$ is of special interest. To the best of the authors' knowledge, results concerning this hitting time are quite rare, in particular for possible asymptotics when N goes to infinity.

1.3. A storage system

Another, more recent, motivation for considering $(X_N(t))$ is the stochastic analysis of a storage system where files are duplicated on C_N servers. Each server breaks down independently at rate μ , in which case it is repaired but all its files are lost. As a simplified model, $X_N(t)$ is defined as the number of copies of a specified file. If $X_N(t) = x$ then a copy of the file is lost if one of the servers breaks down, i.e. at rate $x\mu$. If $X_N(t) = 0$, there is no copy of the file in the system, it is lost, so 0 is an absorbing point. If $0 < x < C_N$ then a new copy of the file may be added but at rate $N - x$; N is the maximal capacity of duplication of the system. It is easily seen that as long as $(X_N(t))$ does not hit 0, $(X_N(t))$ is precisely the Engset process. In this context it is of special interest to study the distribution of the first time when the file is lost, i.e. the hitting time of 0. See [6].

1.4. A collection of exponential martingales

In this paper we rely heavily on the use of martingales to derive explicit, simple, expressions of the Laplace transforms of the hitting times of a state of the system. We obtain expressions of these transforms as ratios of simple integrals for which various asymptotic results, when N goes to infinity, can be derived quite easily with standard technical tools. In particular, we do not need to cope with the asymptotic behavior of sums of combinatorial expressions. Quite surprisingly, up to now, martingales have not played a major role in the previous studies of

the Ehrenfest process. We mention the work by Simatos and Tibi [22], who used a martingale approach to estimate certain exit times for multidimensional Ehrenfest processes. It is one of the aims of this paper to show that a simple and important family of martingales allows a quite detailed investigation of this process, and also of its variants, such as the Engset process.

The key ingredient of this paper is a set of nonnegative martingales which will be called exponential martingales. If $(M(t))$ is a martingale on some probability space, the associated exponential martingale is the solution $(Z(t))$ of the stochastic differential equation (SDE)

$$dZ(t) = Z(t-) dM(t), \quad t \geq 0,$$

where $Y(t-)$ is the left limit of Y at t and $dY(t)$ is the limit on the right of t of $s \mapsto Y(s) - Y(t-)$. It is called the Doléans exponential of $(M(t))$. See Chapter IV of [19] for example. Although there is an exponential martingale for each martingale, a small subset of these martingales plays an important role. For the standard Brownian motion $(B(t))$ this is the martingale

$$\left(\exp\left(\beta B(t) - \frac{1}{2}\beta^2 t\right)\right)$$

for a fixed $\beta \in \mathbb{R}$. It is very helpful to derive explicit expressions for Laplace transforms of hitting times associated to Brownian motion. See [17]. For jump processes, this is less clear. It does not seem that a ‘classification’ of exponential martingales exists in general, even for birth-and-death processes. See Chapter V of [5] and [20] for related questions. Some examples of important processes are reviewed.

For $\xi \in \mathbb{R}_+$, \mathcal{N}_ξ denotes a Poisson process with rate ξ and $(\mathcal{N}_{\xi,i})$ denotes an i.i.d. sequence of such Poisson processes. All Poisson processes are assumed to be independent.

Random walks. The classical exponential martingale associated to the random walk $(S(t)) = (\mathcal{N}_\lambda([0, t]) - \mathcal{N}_\mu([0, t]))$ is given by

$$\left(\exp\left(-\beta S(t) - t(\lambda(1 - e^{-\beta}) + \mu(1 - e^\beta))\right)\right) \quad \text{for } \beta \in \mathbb{R}.$$

It is the exponential martingale associated to the martingale

$$(\beta(S(t) - (\lambda - \mu)t)).$$

The corresponding reflected process is the M/M/1 queue with input rate λ and service rate μ .

The M/M/∞ process. This is a classical Markov process on \mathbb{N} whose Q -matrix $Q = (q(x, y))$ is, for $x \in \mathbb{N}$, $q(x, x + 1) = \lambda$ and $q^N(x, x - 1) = \mu x$. It can also be seen as a kind of discrete Ornstein–Uhlenbeck process, defined as the solution $(L(t))$ of the SDE

$$dL(t) = \mathcal{N}_\lambda(dt) - \sum_{i=1}^{L(t-)} \mathcal{N}_{\mu,i}(dt).$$

The following martingale was introduced in [10]: for $\beta \in \mathbb{R}$,

$$\left((1 + \beta e^{\mu t})^{L(t)} \exp\left(-\frac{\beta e^{\mu t} \lambda}{\mu}\right) \right). \tag{1}$$

It is the exponential martingale associated to the martingale

$$\left(\int_0^t (1 + \beta e^{\mu s})[\mathcal{N}_\lambda(ds) - \lambda ds] - \sum_{i=1}^{+\infty} \int_0^t \frac{1}{1 + \beta e^{\mu s}} \mathbf{1}_{\{i < L(s-)\}}[\mathcal{N}_{\mu,i}(ds) - \mu ds] \right).$$

The Ehrenfest process. Such a process, $(E_N(t))$, with N particles can be seen as the solution of the SDE

$$dL(t) = \sum_{i=1}^{N-L(t-)} \mathcal{N}_{\nu,i}(dt) - \sum_{i=1}^{L(t-)} \mathcal{N}_{\mu,i}(dt).$$

It will be seen that the corresponding exponential martingale is given by

$$((1 - \beta\mu e^{(\mu+\nu)t})^{E_N(t)} (1 + \beta\nu e^{(\mu+\nu)t})^{N-E_N(t)}) \quad \text{for } \beta \in \mathbb{R}.$$

It is the exponential martingale associated to the martingale defined by, up to the multiplicative factor $\beta(\mu + \nu)$,

$$\left(\sum_{i=1}^N \int_0^t e^{(\mu+\nu)s} (\mathbf{1}_{\{Y_i(s-)=1\}} [\mathcal{N}_{\mu,i}(ds) - \mu ds] - \mathbf{1}_{\{Y_i(s-)=0\}} [\mathcal{N}_{\nu,i}(ds) - \nu ds]) \right),$$

where the $(Y_i(t))$ are such that $E_N(t) = Y_1(t) + \dots + Y_N(t)$. Recall that the Engset process is a reflected version of this process.

From these exponential martingales, explicit expressions of Laplace transforms of the distribution of hitting times associated with these processes can be derived. It may not be as straightforward as in the case of Brownian motion since the space variable t is not separated from the space variable, but a convenient integration with respect to the free parameter β solves the problem. See Chapters 5 and 6 of [18] for the M/M/1 and M/M/ ∞ processes, and Section 3 for the Ehrenfest process. See also [21].

From the point of view of potential theory, these martingales are associated to the set of extreme harmonic functions. This statement can be made precise in terms of the space–time Martin boundary. See [15] and the discussion in Section 3.

1.5. Organization of the paper

In Section 2, the two stochastic processes are defined precisely. In Section 3, the exponential martingale for the Ehrenfest process is introduced, and, based on it, several interesting martingales for the Ehrenfest process and the Engset process are constructed. As a corollary, closed-form expressions of the Laplace transform of the hitting time of a given state are obtained as the ratio of simple integrals. This holds in particular for the blocking time for the Engset process. The last three sections are devoted to the analysis of the asymptotic behavior of the distribution of the hitting time of C_N and 0 when N goes to infinity in such a way that $C_N \sim \eta N$ for some $0 < \eta \leq 1$. In each section we consider one of the following three possible regimes: *supercritical*, when the difference $C_N - X_N(t)$ converges to a finite process; *subcritical*, when $\nu < \eta$, the process ‘lives’ in the neighborhood of νN ; and, finally, *critical*, when $C_N - X_N(t)$ is of the order of \sqrt{N} . For each regime, by taking advantage of the simple expressions of the corresponding Laplace transforms obtained, various convergence-in-distribution results are derived.

2. The stochastic model

2.1. The Ehrenfest process

Let $(Y(t))$ be the simple Markov process on $\{0, 1\}$ whose Q -matrix Q_Y is given by

$$Q_Y = \begin{pmatrix} -\nu & \nu \\ \mu & -\mu \end{pmatrix}.$$

For $N \in \mathbb{N}$, if $(Y_i(t))$, $1 \leq i \leq N$, are N independent copies of $(Y(t))$, the Ehrenfest process $(E_N(t))$ is also a birth-and-death process, but on the state space $\{0, 1, \dots, N\}$, defined as

$$E_N(t) = Y_1(t) + Y_2(t) + \dots + Y_N(t). \tag{2}$$

The Q -matrix of $(E_N(t))$ will be denoted by $Q_{E_N} = (q_{E_N}(x, y))$, where, for $x \in \{0, \dots, N\}$,

$$q_{E_N}(x, x - 1) = \mu x \quad \text{and} \quad q_{E_N}(x, x + 1) = \nu(N - x).$$

2.2. The Engset process

For $1 \leq C_N \leq N$, the Engset process $(X_N(t))$ is a birth-and-death process on $\{1, \dots, C_N\}$ which can be seen as a reflected version of $(E_N(t))$ at the boundary C_N , i.e. its Q -matrix $Q_{X_N} = (q_{X_N}(x, y))$ is given by, for $0 \leq x \leq C_N$,

$$q_{X_N}(x, x - 1) = \mu x \quad \text{and} \quad q_{X_N}(x, x + 1) = \nu(N - x) \quad \text{if } x < C_N.$$

In particular, the process $(X_N(t))$ has the same distribution as the process $(E_N(t))$ constrained to the state space $\{0, \dots, C_N\}$. In particular, when $C_N = N$, the two processes $(X_N(t))$ and $(E_N(t))$ starting from the same initial state have the same distribution.

As ergodic birth-and-death processes, the Markov processes $(X_N(t))$ and $(E_N(t))$ are reversible and their stationary distribution at x is, up to a normalization constant, given by

$$\binom{N}{x} \left(\frac{\nu}{\mu}\right)^x,$$

if x is an element of their respective state space.

2.3. Normalization of the time scale

By considering the time scale $t \rightarrow t/(\nu + \mu)$ in the analysis of the processes $(E_N(t))$ and $(X_N(t))$, it can be assumed without any loss of generality that $\nu + \mu = 1$. This will be the case in this paper.

2.4. A limiting regime

In the following it will be assumed that the constant C_N is asymptotically of the order of N , i.e. that

$$\eta := \lim_{N \rightarrow +\infty} \frac{C_N}{N}, \tag{3}$$

holds for some $\eta \in (0, 1]$.

For large t , the probability that the variable $Y(t)$ defined above is at 1 is given by its equilibrium distribution at 1, that is, ν . It follows from the law of large numbers that $E_N(t)$ is of the order of $N\nu$. Roughly speaking, if $N\nu < C_N$ for large N , i.e. $\nu < \eta$, then the boundary at C_N should not play a significant role for first-order quantities related to $(X_N(t))$, and, therefore, the processes $(X_N(t))$ and $(E_N(t))$ should have the same behavior in the limit. However, if $\nu \geq \eta$, due to the reflecting boundary at C_N for $(X_N(t))$, the Ehrenfest and the Engset processes should behave differently. This phenomenon will be stated more precisely in the last three sections of the paper.

3. Positive martingales

Several families of positive martingales for the Ehrenfest and the Engset processes are introduced in this section. More specifically, when $(Z(t))$ is either $(X_N(t))$ or $(E_N(t))$, we identify a set of functions $f : \mathbb{N} \times \mathbb{R}_+ \mapsto \mathbb{R}_+$ such that the process $(f(Z(t), t))$ is a martingale, i.e. that, for $t \geq 0$, the relation

$$E(f(Z(t), t) \mid \mathcal{F}_t) = f(Z(s), s) \quad \text{for } s \leq t$$

holds almost surely, where (\mathcal{F}_t) is the natural filtration associated with $(Z(t))$.

If $Q_Z = (q_Z(\cdot, \cdot))$ is the Q -matrix of $(Z(t))$, this probabilistic property is equivalent to the fact that the function f is *space–time harmonic* with respect to Q_Z , i.e. that the relation

$$\frac{\partial}{\partial t} f(x, t) + Q_Z(f(\cdot, t))(x) = 0$$

holds for $x \in \mathbb{N}$ and $t \geq 0$, where, for $h : \mathbb{N} \mapsto \mathbb{R}_+$,

$$Q_Z(h)(x) = \sum_{y \in \mathbb{N}} q(x, y)h(y).$$

A space–time harmonic function of the Markov process $(Z(t))$ is just a harmonic function of the transient Markov process $((Z(t), t))$. See Appendix B of [18] for example. When $(Z(t))$ is $(E_N(t))$, we will prove that there is a family f_β , $\beta \in \mathbb{R}$, of such functions. As will be seen, these martingales can be interpreted as exponential martingales. In particular, they will give an explicit expression for the Laplace transform of the hitting times associated with both the $(E_N(t))$ and $(X_N(t))$ processes.

For a given birth-and-death process, there is already a complete description of all such positive martingales. This is the (space–time) Martin boundary of the birth-and-death process; see [15]. This description is expressed in terms of the orthogonal polynomials associated with the birth-and-death process which may be defined by an induction relation or by the measure with respect to which they are orthogonal; see [11] and Chapter 2 of [23]. As long as moments of some transient characteristics are investigated, these martingales can be used, but, in general, they do not seem to be particularly helpful to analyze the distributions of hitting times.

This situation is quite classical for, e.g. Brownian motion, for which there is a family of martingales indexed by $N \in \mathbb{N}$: if H_N is the Hermite polynomial of degree N then $(M_N(t)) = (t^{N/2}H_N(B(t)/\sqrt{t}))$ is a martingale. Another family of martingales is provided by the exponential martingale $(\exp(\beta B(t) - \beta^2 t/2))$ indexed by $\beta \in \mathbb{R}$. This exponential martingale can be expressed as a weighted sum of the martingales $(M_N(t))$, but to obtain explicit expressions of the distributions of hitting times, it is the most useful martingale. See Propositions 3.4 and 3.8 of [17, Chapter 3] for example. In the case of birth-and-death processes, a general result concerning the construction of such exponential martingales from the martingales associated with the orthogonal polynomials does not seem to exist.

3.1. Exponential martingales for the Ehrenfest process

Owing to the simple structure of the Ehrenfest process, these martingales are really elementary. Nevertheless, they play a fundamental role; most of the asymptotic results obtained in this paper are based on these martingales. A more general version in a multidimensional context has been introduced by Simatos and Tibi [22].

Proposition 1. (Exponential martingales.) For $\beta \in \mathbb{R}$, the process

$$(M_N^\beta(t)) = ((1 - \beta\mu e^t)^{E_N(t)}(1 + \beta\nu e^t)^{N-E_N(t)}) \tag{4}$$

is a martingale.

Proof. Define, for $N \geq 1, t \geq 0$, and $0 \leq x \leq N$,

$$h_N^\beta(x, t) = (1 - \beta\mu e^t)^x(1 + \beta\nu e^t)^{N-x}. \tag{5}$$

Clearly, the relations

$$\frac{\partial}{\partial t} h_1^\beta(0, t) = \beta\nu e^t = -Q_Y(h_1^\beta(\cdot, t))(0) \quad \text{and} \quad \frac{\partial}{\partial t} h_1^\beta(1, t) = -\beta\mu e^t = -Q_Y(h_1^\beta(\cdot, t))(1),$$

hold, where Q_Y is the Q -matrix of $(Y(t))$ introduced in Section 2. Consequently, the function h_1^β is space–time harmonic for the matrix Q_Y , or, equivalently, $(f(Y(t), t))$ is a martingale.

If $((Y_i(t)), 1 \leq i \leq N)$ are N independent copies of $(Y(t))$ then, by using the independence of the processes $(Y_i(t)), i = 1, \dots, N$, and (2), we find that the process

$$\left(\prod_{i=1}^N f(Y_i(t), t) \right) \stackrel{D}{=} (M_N^\beta(t)) = (h_N^\beta(E_N(t), t))$$

is also martingale with respect to the filtration $(\mathcal{F}_t) = (\sigma\langle Y_i(s), s \leq t, 1 \leq i \leq N \rangle)$. In particular, the function h_N^β is space–time harmonic with respect to Q_{E_N} .

3.2. Martingales associated with orthogonal polynomials

As remarked by Karlin and McGregor [11], [13], the polynomials associated with the Ehrenfest process are the $N + 1$ Krawtchouk polynomials $(\mathcal{K}_n^N, 0 \leq n \leq N)$ defined by

$$\mathcal{K}_n^N(x) = \binom{N}{n}^{-1} \sum_{\ell=0}^n (-1)^\ell \binom{x}{\ell} \binom{N-x}{n-\ell} \left(\frac{\mu}{\nu}\right)^\ell, \quad 0 \leq n, x \leq N.$$

These polynomials are orthogonal with respect to the binomial distribution

$$\left(\binom{N}{k} \nu^k \mu^{N-k}, 0 \leq k \leq N \right).$$

From the classical identity, see Karlin and McGregor [11] for example,

$$\sum_{\ell=0}^N \binom{N}{\ell} \mathcal{K}_n^N(x) u^\ell = (1 + u)^{N-x} \left(1 - \frac{\mu}{\nu} u \right)^x, \quad u \in \mathbb{R},$$

and the above proposition, it follows that

$$\left(\sum_{n=0}^N \binom{N}{\ell} \mathcal{K}_n^N(E_N(t)) \beta^n e^{nt} \right) = (M_N^{\beta/\nu}(t))$$

is a martingale. Consequently, for any $0 \leq n \leq N$, the process $(\mathcal{K}_n^N(E_N(t)) \exp(nt))$ is a martingale. The martingale $(M_N^{\beta/\nu}(t))$ can thus be seen as an encoding of these $N + 1$ martingales in the same way as the exponential martingale of the Brownian motion with the Hermite polynomials, or the martingale (1) with the Poisson–Charlier polynomials. See [18], [20], and [21].

Note that the space variable $E_N(t)$ and the time variable t are not separated in the exponential martingale (4). Provided that it can be used, Doob’s optional stopping theorem applied to some hitting time of some specified state x does not give useful information on the distribution of this variable. However, given that there is a free parameter $\beta \in \mathbb{R}$ in (4) and that the martingale property is clearly preserved by integration with respect to β , we can try to find a measure on \mathbb{R}_+ that will ‘separate’ the space and time variables. The following proposition uses such a method.

Proposition 2. For any $\alpha > 0$ and $t \geq 0$, if

$$\begin{aligned} I_\alpha^N(t) &= e^{-\alpha t} \int_0^1 (1-u)^{E_N(t)} \left(1 + \frac{\nu}{\mu}u\right)^{N-E_N(t)} u^{\alpha-1} du, \\ J_\alpha^N(t) &= e^{-\alpha t} \int_0^1 (1-u)^{N-E_N(t)} \left(1 + \frac{\mu}{\nu}u\right)^{E_N(t)} u^{\alpha-1} du, \end{aligned} \tag{6}$$

and

$$T_x^{E_N} = \inf\{t > 0: E_N(t) = x\}, \quad 0 \leq x \leq N,$$

then $(I_\alpha(t \wedge T_0^{E_N}))$ and $(J_\alpha(t \wedge T_N^{E_N}))$ are martingales.

Proof. Since $(N - E_N(t))$ is also an Ehrenfest process, but with the two parameters μ and ν exchanged, we need only prove that the process $(I_\alpha(t \wedge T_0^{E_N}))$ is a martingale.

Define

$$f_N(x, t) := \int_0^{e^{-t/\mu}} h_N^\beta(x, t) \beta^{\alpha-1} d\beta,$$

where h_N^β is defined in (5); then

$$\frac{\partial f_N}{\partial t}(x, t) = \int_0^{e^{-t/\mu}} \frac{\partial h_N^\beta}{\partial t}(x, t) \beta^{\alpha-1} d\beta - \frac{e^{-\alpha t}}{\mu^\alpha} h_N^{e^{-t/\mu}}(x, t).$$

Note that the last term of the above expression is 0 if $x \neq 0$. Consequently, for $x \neq 0$ and $t \geq 0$,

$$\frac{\partial f_N}{\partial t}(x, t) + Q_{E_N}(f_N)(x, t) = \int_0^{e^{-t/\mu}} \left[\frac{\partial h_N^\beta}{\partial t}(x, t) + Q_{E_N}(h_N^\beta)(x, t) \right] \beta^{\alpha-1} d\beta = 0,$$

because h_N^β is space–time harmonic with respect to Q_{E_N} , as in the proof of Proposition 1. In other words, the function f_N is space–time harmonic for the Q -matrix of the stopped process $(E_N(t \wedge T_0^{E_N}))$; hence,

$$(f_N(E_N(t \wedge T_0^{E_N}), t \wedge T_0^{E_N})) = (I_\alpha^N(t \wedge T_0^{E_N}))$$

is a martingale.

It is now easy to obtain a representation of the Laplace transform of the hitting times for the Ehrenfest process.

Proposition 3. (Laplace transform of the hitting times.) For $0 \leq x \leq y \leq N$ and if $T_x^{E_N} = \inf\{t > 0: E_N(t) = x\}$, the relations

$$E_y(\exp(-\alpha T_x^{E_N})) = \frac{B_x(\alpha)}{B_y(\alpha)} \quad \text{and} \quad E_x(\exp(-\alpha T_y^{E_N})) = \frac{D_x(\alpha)}{D_y(\alpha)} \tag{7}$$

hold, with

$$\begin{aligned}
 B_x(\alpha) &= \int_0^1 (1-u)^x \left(1 + \frac{\nu}{\mu}u\right)^{N-x} u^{\alpha-1} du, \\
 D_x(\alpha) &= \int_0^1 (1-u)^{N-x} \left(1 + \frac{\mu}{\nu}u\right)^x u^{\alpha-1} du.
 \end{aligned}
 \tag{8}$$

There is in fact only one result here since $(N - E_N(t))$ is, as already remarked, an Ehrenfest process with the parameters μ and ν exchanged. The second relation of (7) is therefore a consequence of the first relation.

Proof of Proposition 3. The martingale $I_\alpha(t \wedge T_0)$ is bounded and so uniformly integrable. Therefore, Doob’s optional stopping theorem gives the relation $E_y(I_\alpha(0)) = E_y(I_\alpha(T_x))$, and the first relation of (7) follows.

By expanding one of the terms of the integrand of (8), we obtain

$$\begin{aligned}
 B_x(\alpha) &= \sum_{k=0}^{N-x} \binom{N-x}{k} \left(\frac{\nu}{\mu}\right)^k \int_0^1 (1-u)^x u^{\alpha+k-1} du \\
 &= \sum_{k=0}^{N-x} \binom{N-x}{k} \left(\frac{\nu}{\mu}\right)^k \frac{\Gamma(x+1)\Gamma(\alpha+k)}{\Gamma(\alpha+x+k+1)},
 \end{aligned}$$

by Euler’s integral for beta and gamma functions. See [24, p. 254] for example. Laplace transforms of hitting times can therefore also be expressed as the ratio of such sums, as is quite common for hitting times of birth-and-death processes. See Equation (4.4) of [13] for example. Flajolet and Huillet [7] used this kind of representation in the symmetrical case. As will be seen, from the compact representation (8) with integrals, we will obtain asymptotic results for the distribution of these variables with standard techniques.

3.3. Martingales for the Engset process

It has been seen that the Engset process $(X_N(t))$ is a reflected version of the process $(E_N(t))$ at the boundary C_N . The two families of martingales in Proposition 2 cannot be used directly if the sample path of $(E_N(t))$ exceeds C_N , when the hitting times of 0 are analyzed for example. The idea is to construct a linear combination of the martingales $(I_\alpha(t))$ and $(J_\alpha(t))$ such that the space–time harmonicity of the corresponding function, which is valid when the space variables are in $\{1, \dots, C_N - 1\}$, also holds at the boundary C_N . This method has been used in [14] in the case of reflected random walks.

Proposition 4. For $\alpha > 0$, define

$$\begin{aligned}
 b_N(\alpha) &= \nu \int_0^1 (1-u)^{C_N} \left(1 + \frac{\nu}{\mu}u\right)^{N-C_N-1} u^\alpha du, \\
 d_N(\alpha) &= \mu \int_0^1 (1-u)^{N-C_N-1} \left(1 + \frac{\mu}{\nu}u\right)^{C_N} u^\alpha du,
 \end{aligned}$$

and

$$(K_\alpha^N(t)) = (d_N(\alpha)I_\alpha^N(t) + b_N(\alpha)J_\alpha^N(t)),$$

where $(I_\alpha^N(t))$ and $(J_\alpha^N(t))$ are defined by (6) with $E_N(t)$ replaced by $X_N(t)$. Then, if $T_0^{X_N}$ is the hitting time of 0 by $(X_N(t))$, the process $(K_\alpha^N(t \wedge T_0^{X_N}))$ is a martingale.

Proof. Define the function g_α such that, for all $t \geq 0$ and $0 \leq x \leq C_N$,

$$g_\alpha(x, t) = d_N(\alpha)e^{-\alpha t} \int_0^1 (1-u)^x \left(1 + \frac{\nu}{\mu}u\right)^{N-x} u^{\alpha-1} du + b_N(\alpha)e^{-\alpha t} \int_0^1 (1-u)^{N-x} \left(1 + \frac{\mu}{\nu}u\right)^x u^{\alpha-1} du.$$

The function g_α is space–time harmonic for the matrix Q_{X_N} on $\{1, \dots, C_N - 1\}$, that is,

$$\left[\frac{\partial g_\alpha}{\partial t} + Q_{X_N}(g_\alpha) \right](x, t) = 0, \quad 0 < x < C_N,$$

since the two matrices Q_{X_N} and Q_{E_N} are identical as long as the starting point is in $\{1, \dots, C_N - 1\}$, and $(I_\alpha^N(t \wedge T_0^{E_N}))$ and $(J_\alpha^N(t \wedge T_N^{E_N}))$ are martingales by Proposition 2.

The space–time harmonicity of g_α for the matrix Q_{E_N} at $C_N < N$ gives

$$\left[\frac{\partial g_\alpha}{\partial t} + Q_{X_N}(g_\alpha) \right](C_N, t) = -\nu(N - C_N)[g_\alpha(C_N + 1, t) - g_\alpha(C_N, t)].$$

For $0 \leq y \leq 1$, we have

$$(1-y)^{C_N+1} \left(1 + \frac{\nu}{\mu}y\right)^{N-C_N-1} - (1-y)^{C_N} \left(1 + \frac{\nu}{\mu}y\right)^{N-C_N} = -\frac{y}{\mu}(1-y)^{C_N} \left(1 + \frac{\nu}{\mu}y\right)^{N-C_N-1},$$

and, similarly,

$$\left[\frac{\partial g_\alpha}{\partial t} + Q_{X_N}(g_\alpha) \right](C_N, t) = \frac{N - C_N}{\mu}(d_N(\alpha)b_N(\alpha) - b_N(\alpha)d_N(\alpha)) = 0.$$

The function g_α is therefore space–time harmonic for the Q -matrix of the stopped process $(X_N(t \wedge T_0^{X_N}))$; hence, the process $(K_\alpha^N(t \wedge T_0^{X_N}))$ is a martingale.

Proposition 5. (Laplace transform of the hitting times for the Engset process.) *For $0 \leq x \leq y \leq C_N$, if $T_x^{X_N} = \inf\{s \geq 0: X_N(s) = x\}$ then, for $\alpha \geq 0$,*

$$E_x(\exp(-\alpha T_y^{X_N})) = \frac{D_x(\alpha)}{D_y(\alpha)} \quad \text{and} \quad E_y(\exp(-\alpha T_x^{X_N})) = \frac{d_N(\alpha)B_y(\alpha) + b_N(\alpha)D_y(\alpha)}{d_N(\alpha)B_x(\alpha) + b_N(\alpha)D_x(\alpha)},$$

using the notation given in Propositions 3 and 4.

Proof. The first identity comes from the fact that the two processes $(X_N(t))$ and $(E_N(t))$, starting from the same initial state, are identical in distribution as long as they do not reach C_N . In particular, if $X_N(0) = E_N(0) = x$, the variables $T_y^{E_N}$ and $T_y^{X_N}$ have the same distribution. The second identity is a direct consequence of the martingale property of $(K_\alpha^N(t \wedge T_0^{X_N}))$ proved in the above proposition.

4. A fluid picture

In this section we give a short description of the first-order properties of the Ehrenfest and Engset processes as N goes to infinity. Its purpose is mainly to introduce the three natural possible asymptotic regimes that will be investigated in detail in the last sections. The proofs of the asymptotic results are quite standard and will therefore be omitted.

From now on, it is assumed that relation (3) holds, that is,

$$\lim_{N \rightarrow +\infty} \frac{C_N}{N} = \eta > 0.$$

The Engset process $(X_N(t))$ can also be seen as the unique solution of the SDE

$$dX_N(t) = \mathbf{1}_{\{X_N(t-) < C_N\}} \sum_{i=1}^{N-X_N(t-)} \mathcal{N}_{\nu,i}(dt) - \sum_{i=1}^{X_N(t-)} \mathcal{N}_{\mu,i}(dt),$$

starting from $X_N(0)$, where, for $\xi > 0$, the $(\mathcal{N}_{\xi,k})$ are independent Poisson processes with rate ξ .

The initial state is assumed to satisfy

$$\lim_{N \rightarrow +\infty} \frac{X_N(0)}{N} = x_0 \in [0, \eta];$$

then, by complementing Poisson processes in order to get martingales, the above equation can be rewritten as

$$dX_N(t) = dM_N(t) + [\nu(N - X_N(t)) \mathbf{1}_{\{X_N(t) < C_N\}} - \mu X_N(t)] dt,$$

where $(M_N(t))$ is a martingale of the order of \sqrt{N} . In the same way as for the Erlang process, see Chapter 6 of [18] for example, we can prove the following convergence in distribution of processes:

$$\lim_{N \rightarrow +\infty} \left(\frac{X_N(t)}{N} \right) = (\min(\eta, \nu + (x_0 - \nu)e^{-t})).$$

This first-order description of the Engset process shows that there are three different asymptotic regimes.

Supercritical regime: $\nu > \eta$. Under this condition, the renormalized process is at the boundary C_N at time

$$t^* := \log\left(\frac{\nu - x_0}{\nu - \eta}\right). \tag{9}$$

A more detailed picture can be obtained by looking at the process

$$(Z_N(t)) = \left(C_N - X_N\left(\frac{t}{N}\right) \right)$$

of empty spaces with a ‘slow’ time scale. As N goes to infinity, it is easily seen that the Q -matrix of this birth-and-death process converges to the Q -matrix of an ergodic M/M/1 process with input rate η and service rate ν . In particular, this gives the asymptotic expression of the Engset formula: for $t \in \mathbb{R}_+$,

$$\lim_{N \rightarrow +\infty} \mathbb{P}(X_N(t) = C_N) = 1 - \frac{\eta}{\nu}.$$

Subcritical regime: $\nu < \eta$. In this case we have

$$\lim_{N \rightarrow +\infty} \left(\frac{X_N(t)}{N} \right) = (\nu + (x_0 - \nu)e^{-t}) = \lim_{N \rightarrow +\infty} \left(\frac{E_N(t)}{N} \right).$$

As expected, the boundary at C_N does not play a role: at first order, the Engset process and the Ehrenfest process are identical.

Critical regime: $\nu = \eta$. From the fluid limit picture we know that the system saturates ‘at infinity’, which, as we will see, is a too rough description of its evolution.

The next sections are devoted to the asymptotic analysis of the distributions of hitting times. For simplicity, it is assumed that the initial state is on the boundary, either 0 or C_N . Similar results could be obtained without any additional difficulty when the initial state is in the neighborhood of some $\lfloor zN \rfloor$ for $0 \leq z \leq \eta$.

5. Supercritical regime

As we have seen, under the condition $\nu > \eta$ and at time t^* defined in (9), the system is saturated in the fluid limit. This implies in particular that the hitting time of the boundary C_N ,

$$T_{C_N}^{X_N} = \inf\{s \geq 0; X_N(s) = C_N\},$$

converges in distribution to t^* . The following proposition gives a more precise asymptotic result. See [7] for a related result in the symmetrical case.

Proposition 6. *If $C_N = \eta N + O(1)$, $\eta < \nu$, and $X_N(0) = 0$, then the sequence of random variables*

$$\left(\sqrt{N} \left[T_{C_N}^{X_N} - \log \left(\frac{\nu}{\nu - \eta} \right) \right] \right)$$

converges in distribution to a centered normal random variable with variance

$$\frac{\sqrt{\eta(1 - \eta)}}{\nu - \eta}.$$

Proof. Proposition 3 gives

$$\begin{aligned} & E_0(\exp(-\alpha \sqrt{N} T_{C_N}^{X_N})) \\ &= \int_0^{\sqrt{N}} \left(1 - \frac{u}{\sqrt{N}} \right)^N u^{\sqrt{N}\alpha - 1} du \\ & \times \left(\int_0^{\sqrt{N}} \left(1 - \frac{u}{\sqrt{N}} \right)^{N - C_N} \left(1 + \frac{\mu}{\nu} \frac{u}{\sqrt{N}} \right)^{C_N} u^{\sqrt{N}\alpha - 1} du \right)^{-1} \end{aligned} \tag{10}$$

for $\alpha \geq 0$. The integrand in the numerator of (10) can be expressed as $\exp(f_N(u))$, with

$$f_N(u) = N \log \left(1 - \frac{u}{\sqrt{N}} \right) + (\sqrt{N}\alpha - 1) \log u.$$

The function has a unique maximum at

$$y_N = \frac{\alpha \sqrt{N} - 1}{\sqrt{N} + \alpha - 1/\sqrt{N}} = \alpha - \frac{1 + \alpha^2}{\sqrt{N}} + o\left(\frac{1}{\sqrt{N}}\right),$$

and

$$f_N(y_N) = (\alpha \log(\alpha) - \alpha)\sqrt{N} - \frac{\alpha^2}{2} - \log(\alpha) + o(1),$$

$$f''_N(y_N) = -\frac{1 + \alpha^2}{\alpha^2} - \frac{\sqrt{N}}{\alpha} + o(1).$$

Laplace’s method, see Section B.6 of [8] for example, therefore gives the relation

$$\int_0^{\sqrt{N}} e^{f_N(u)} du \sim \frac{\sqrt{2\pi}}{\sqrt{-f''_N(y_N)}} e^{f_N(y_N)} \sim \frac{\sqrt{2\pi\alpha}}{N^{1/4}} \exp\left((\alpha \log(\alpha) - \alpha)\sqrt{N} - \frac{\alpha^2}{2} - \log(\alpha)\right). \tag{11}$$

Similarly, the integrand in the denominator of (10) can be expressed as $\exp(g_N(u))$, with

$$g_N(u) = (N - C_N) \log\left(1 - \frac{u}{\sqrt{N}}\right) + C_N \log\left(1 + \frac{\mu}{v} \frac{u}{\sqrt{N}}\right) + (\sqrt{N}\alpha - 1) \log u.$$

This concave function on the interval $(0, \sqrt{N})$ has a unique maximum at $z_N = z_0 - \delta/\sqrt{N} + o(1/\sqrt{N})$, with

$$z_0 = \alpha \frac{v}{v - \eta} \quad \text{and} \quad \delta = v \frac{\alpha^2(v^2 + \eta - 2\eta v) + (v - \eta)^2}{(v - \eta)^3},$$

and, with some calculations, we find that

$$g_N(z_N) = \left(-\alpha + \alpha \log(\alpha) + \alpha \log\left(\frac{v}{v - \eta}\right)\right)\sqrt{N} + \frac{(2v\eta - v^2 - \eta) \alpha^2}{(v - \eta)^2} \frac{\alpha^2}{2} - \log(\alpha) - \log\left(\frac{v}{v - \eta}\right) + o(1)$$

and

$$g''_N(z_N) = -\left(1 - \eta + \eta \frac{(1 - v)^2}{v^2}\right) - \frac{\alpha\sqrt{N} - 1}{z_0^2} + o(1).$$

Again, by Laplace’s method we obtain

$$\exp\left(-\alpha\sqrt{N} \log \frac{v}{v - \eta}\right) \int_0^{\sqrt{N}} e^{g_N(u)} du \sim \frac{\sqrt{2\pi\alpha}}{N^{1/4}} \exp\left((\alpha \log(\alpha) - \alpha)\sqrt{N} + \frac{2v\eta - v^2 - \eta}{(v - \eta)^2} \frac{\alpha^2}{2} - \log(\alpha)\right). \tag{12}$$

Equation (10) together with (11) and (12) finally give

$$\lim_{N \rightarrow +\infty} \mathbb{E}\left(\exp\left(-\alpha\sqrt{N} \left[T_{C_N}^{X_N} - \log\left(\frac{v}{v - \eta}\right)\right]\right)\right) = \exp\left(\frac{\eta(1 - \eta) \alpha^2}{(v - \eta)^2} \frac{\alpha^2}{2}\right).$$

5.1. An informal proof

The limit theorem obtained in Proposition 6 is a consequence of some detailed, but simple, calculations that use Laplace’s method. We can quite quickly gain an idea of the possible limit with the help of the exponential martingale $(M_N^\beta(t))$ of Proposition 1 through a nonrigorous derivation. As will be seen, it gives the correct answer, but its justification seems to be difficult. The main problem comes from the fact that, in this martingale, the term e^t stopped at some random time may not be integrable. For example, it is easily seen that the first jump of the martingale $(M_1^\beta(t))$ is not a regular stopping time for this martingale, i.e. the optional stopping theorem is not valid here.

Define $Z_N = \sqrt{N}(\exp(T_{C_N}^{X_N}) - \exp(t^*))$, where t^* is, as before, $\log(v/(v - \eta))$. Using martingale (4) and assuming that the stopping time $T_{C_N}^{X_N}$ is regular for it, we obtain

$$E\left(\left(1 - \frac{\mu \exp(T_{C_N}^{X_N})\beta}{\sqrt{N}}\right)^{C_N} \left(1 + \frac{v \exp(T_{C_N}^{X_N})\beta}{\sqrt{N}}\right)^{N-C_N}\right) = \left(1 + \frac{v\beta}{\sqrt{N}}\right)^N.$$

This can be written as $E(\exp(U_N)) = 1$, with

$$U_N = C_N \log\left(1 - \frac{\mu \exp(T_{C_N}^{X_N})\beta}{\sqrt{N}}\right) + (N - C_N) \log\left(1 + \frac{v \exp(T_{C_N}^{X_N})\beta}{\sqrt{N}}\right) - N \log\left(1 + \frac{v\beta}{\sqrt{N}}\right);$$

hence,

$$U_N = -\beta(\eta - v)Z_N - ((\eta\mu^2 + (1 - \eta)v^2)e^{2t^*} - v^2)\frac{\beta^2}{2} + o\left(\frac{1}{N}\right) \\ = -\beta(\eta - v)Z_N - \frac{v^2\eta(1 - \eta)}{(v - \eta)^2}\frac{\beta^2}{2} + o\left(\frac{1}{N}\right),$$

provided that the limit can be taken under the integral. We finally obtain

$$\lim_{N \rightarrow +\infty} E(e^{-\beta Z_N}) = \exp\left(\frac{\beta^2 v^2 \eta (1 - \eta)}{2 (v - \eta)^4}\right).$$

Expressed as a limit theorem for $T_{C_N}^{X_N}$, this is precisely the above proposition.

6. Subcritical regime

It is assumed in this section that $v < \eta$, so that the Ehrenfest process ‘lives’ in the interior of the state space; the hitting time of the boundary C_N should therefore be quite large. The following propositions give asymptotic results concerning this phenomenon.

The first result concerns the time it takes for the Ehrenfest process to have all particles in one box when, initially, they are all in the other box. This is of course a very natural quantity for this process. In the discrete-time case, representations of the average of this quantity have been obtained in a symmetrical setting. See [1], [16], and the references therein.

Proposition 7. *If $v < 1$, $X_N(0) = 0$, and $C_N = N$, then the sequence of random variables $(Nv^N T_{C_N}^{X_N})$ converges in distribution to an exponentially distributed random variable with parameter $1 - v$.*

Proof. We use (7) to obtain, for $N \geq 1$,

$$E(\exp(-\alpha_N T_N^{X_N})) = \int_0^1 (1-u)^N u^{\alpha_N-1} du / \int_0^1 \left(1 + \frac{\mu}{v}u\right)^N u^{\alpha_N-1} du, \tag{13}$$

with $\alpha_N = \alpha N v^N$ for some $\alpha > 0$.

The numerator in this expression can be written, after an integration by parts, as

$$\begin{aligned} \int_0^1 (1-u)^N u^{\alpha_N-1} du &= \int_0^1 N(1-u)^{N-1} \frac{u^{\alpha_N}}{\alpha_N} du \\ &= N^{-\alpha_N} \int_0^N \left(1 - \frac{u}{N}\right)^{N-1} \frac{u^{\alpha_N}}{\alpha_N} du \\ &\sim \frac{1}{\alpha_N}, \end{aligned}$$

by dominated convergence.

By subtracting $1/\alpha_N$ from the denominator on the right-hand side of (13), we obtain

$$\begin{aligned} \Delta_N &:= \int_0^1 \left(\left(1 + \frac{\mu}{v}u\right)^N - 1 \right) u^{\alpha_N-1} du \\ &= \frac{N\mu}{v} \int_0^1 \left(1 + \frac{\mu}{v}u\right)^{N-1} \frac{1-u^{\alpha_N}}{\alpha_N} du \\ &= \frac{1}{v^N} \int_0^N \mu \left(1 - \mu \frac{u}{N}\right)^{N-1} \frac{1 - (1-u/N)^{\alpha_N}}{\alpha_N} du; \end{aligned}$$

hence,

$$\alpha_N \Delta_N = \frac{\alpha_N}{Nv^N} \int_0^N \mu \left(1 - \mu \frac{u}{N}\right)^{N-1} \frac{1 - (1-u/N)^{\alpha_N}}{\alpha_N/N} du \sim \alpha \int_0^{+\infty} \mu e^{-\mu u} u du = \frac{\alpha}{\mu}.$$

These two asymptotic results substituted into (13) give the desired convergence in distribution.

Theorem 2 of [1] provides a similar result in the symmetrical case, $\mu = v$, and in discrete time. In the present case, there is an additional factor N in the scaling of $T_{C_N}^{X_N}$ which is due to the fact that the continuous-time dynamics are N times faster than the discrete-time dynamics.

Proposition 8. *If $C_N = \eta N + O(1)$, $v < \eta < 1$, and $X_N(0) = 0$, then, if*

$$H = (1 - \eta) \log\left(\frac{1 - \eta}{1 - v}\right) + \eta \log\left(\frac{\eta}{v}\right),$$

the sequence of random variables

$$\left(\frac{\sqrt{\eta(1-\eta)}}{(\eta-v)\sqrt{2\pi}} \sqrt{N} e^{-NH} T_{C_N}^{X_N} \right) \tag{14}$$

converges in distribution to an exponentially distributed random variable with parameter 1.

We note that the exponential decay factor H in the above proposition is in fact a relative entropy of Bernoulli random variables with respective parameters η and v . Despite similar ‘entropy’ expressions appearing at several occasions in the study of these processes, we have not been able to find a simple explanation for the occurrences of these constants.

Proof of Proposition 8. For $\alpha > 0$, denote by α_N the product of α and the coefficient of $T_{C_N}^{X_N}$ in (14). Equation (7) is again used:

$$E(\exp(-\alpha_N T_N^{X_N})) = \int_0^1 (1-u)^N u^{\alpha_N-1} du / \int_0^1 (1-u)^{N-C_N} \left(1 + \frac{\mu}{v} u\right)^{C_N} u^{\alpha_N-1} du. \tag{15}$$

The asymptotic behavior of the numerator of this Laplace transform has already been obtained in the proof of the above proposition.

To study the denominator, we proceed as before. For $u \in [0, 1]$, write

$$f_N(u) = (N - C_N) \log(1 - u) + C_N \log\left(1 + \frac{\mu u}{v}\right).$$

This function has a unique maximum at

$$y_0 := \frac{C_N/N - v}{1 - v} = \frac{\eta - v}{1 - v} + O\left(\frac{1}{N}\right),$$

given by

$$f_N(y_0) = \left[(1 - \eta) \log\left(\frac{1 - \eta}{1 - v}\right) + \eta \log\left(\frac{\eta}{v}\right) \right] N + o(1)$$

and

$$f_N''(y_0) = -\frac{(1 - v)^2}{\eta(1 - \eta)} N + o(1).$$

The denominator in (15) is

$$\begin{aligned} & \int_0^1 \left[(1-u)^{N-C_N} \left(1 + \frac{\mu}{v} u\right)^{C_N} - 1 \right] u^{\alpha_N-1} du \\ &= \int_0^1 [e^{f_N(u)} - 1] u^{\alpha_N-1} du \\ &= \int_0^1 f_N'(u) e^{f_N(u)} \frac{y_0^{\alpha_N} - u^{\alpha_N}}{\alpha_N} du + \frac{1 - y_0^{\alpha_N}}{\alpha_N}, \end{aligned}$$

by integration by parts. The integral I_N on the right-hand side can be written as

$$I_N = \int_{-y_0\sqrt{N}}^{(1-y_0)\sqrt{N}} \frac{1}{\sqrt{N}} f_N' \left(y_0 + \frac{u}{\sqrt{N}} \right) e^{f_N(y_0 + u/\sqrt{N})} \frac{y_0^{\alpha_N} - (y_0 + u/\sqrt{N})^{\alpha_N}}{\alpha_N} du;$$

hence,

$$\begin{aligned} I_N &= y_0^{\alpha_N-1} \frac{-f_N''(y_0)}{N^{3/2}} e^{f_N(y_0)} \int_{-\infty}^{+\infty} u^2 \exp\left(\frac{f_N''(y_0)}{N} \frac{u^2}{2}\right) du + o\left(\frac{1}{N}\right) \\ &= \frac{1}{y_0} \sqrt{\frac{2\pi}{-f_N''(y_0)}} e^{f_N(y_0)} + o\left(\frac{1}{N}\right). \end{aligned}$$

Combining,

$$\lim_{N \rightarrow +\infty} E(\exp(-\alpha_N T_N^{X_N})) = \frac{1}{1 + \alpha}.$$

Proposition 9. (Hitting time of the empty state.) *Under the condition $\nu < \eta$ and if $C_N = \eta N + O(1)$ for $\eta > 0$ and $X_N(0) = C_N$, then the sequence of variables*

$$(N(1 - \nu)^N T_0^{X_N})$$

converges in distribution to an exponential random variable with parameter ν .

Note that this result can, informally, be justified by the result of Proposition 7. Without the boundary C_N , we could obtain the result by exchanging μ and ν and using Proposition 7. This result shows in particular that the boundary does not change the limiting behavior of $T_0^{X_N}$ in the subcritical regime.

Proof of Proposition 9. Define $\alpha_N = N(1 - \nu)^N$. Proposition 5 gives

$$E_{C_N}(\exp(-\alpha_N T_0^{X_N})) = \frac{d_N(\alpha_N)B_{C_N}(\alpha_N) + b_N(\alpha_N)D_{C_N}(\alpha_N)}{d_N(\alpha)B_0(\alpha_N) + b_N(\alpha_N)D_0(\alpha_N)}.$$

We start with the asymptotic behavior of $(d_N(\alpha_N))$:

$$\begin{aligned} d_N(\alpha_N) &= \mu \int_0^1 (1 - u)^{N-C_N-1} \left(1 + \frac{\mu}{\nu} u\right)^{C_N} u^{\alpha_N} du \\ &= \frac{\mu}{\sqrt{N}} \int_0^{\sqrt{N}} \left(1 - \frac{u}{\sqrt{N}}\right)^{N-C_N-1} \left(1 + \frac{\mu}{\nu} \frac{u}{\sqrt{N}}\right)^{C_N} du + o\left(\frac{1}{\sqrt{N}}\right) \\ &= \frac{1}{\sqrt{N}} \exp\left(\frac{\eta - \nu}{\nu} \sqrt{N}\right) \int_0^{+\infty} \exp\left(-\left(\frac{\eta}{\nu^2} + \frac{1 - \eta}{(1 - \nu)^2}\right) \frac{u^2}{2}\right) du + o\left(\frac{1}{\sqrt{N}}\right). \end{aligned}$$

The other coefficient $b_N(\alpha_N)$ is such that

$$\begin{aligned} b_N(\alpha_N) &= \nu \int_0^1 (1 - u)^{C_N} \left(1 + \frac{\nu}{\mu} u\right)^{N-C_N} u^{\alpha_N} du \\ &= \frac{\nu}{N} \int_0^N \left(1 - \frac{u}{N}\right)^{C_N} \left(1 + \frac{\nu}{\mu N} u\right)^{N-C_N-1} du + o\left(\frac{1}{N}\right) \\ &= \frac{\nu(1 - \nu)}{\nu - \eta} \frac{1}{N} + o\left(\frac{1}{N}\right). \end{aligned}$$

The proof of Proposition 7 provides the relations

$$D_0(\alpha_N) = \int_0^1 (1 - u)^N u^{\alpha_N-1} du \sim \frac{1}{\alpha_N}$$

and

$$B_0(\alpha_N) = \int_0^1 \left(1 + \frac{\nu}{\mu} u\right)^N u^{\alpha_N-1} du \sim \frac{1}{\alpha_N} + \frac{1}{\nu N(1 - \nu)^N}.$$

The two remaining terms to estimate are

$$\begin{aligned} B_{C_N}(\alpha_N) &= \int_0^1 (1 - u)^{C_N} \left(1 + \frac{\nu}{\mu} u\right)^{N-C_N} u^{\alpha_N-1} du, \\ D_{C_N}(\alpha_N) &= \int_0^1 (1 - u)^{N-C_N} \left(1 + \frac{\mu}{\nu} u\right)^{C_N} u^{\alpha_N-1} du. \end{aligned}$$

As in the proof of Proposition 8, we can show that $B_{C_N}(\alpha_N)$ and $D_{C_N}(\alpha_N)$ can be written as $1/\alpha_N + o(B_0(\alpha_N) - 1/\alpha_N)$. More informally, the term $(1 - u)^N$ under the integral for these two expressions reduces their asymptotic behavior by an exponential factor.

These various estimations finally give

$$\lim_{N \rightarrow +\infty} E(\exp(-\alpha_N T_0^{X_N})) = \lim_{N \rightarrow +\infty} \frac{B_{C_N}(\alpha_N)}{B_0(\alpha_N)} = \frac{1}{1 + \alpha/v}.$$

7. Critical regime

In this section, it is assumed that $C_N \sim \nu N$; if $X_N(0) = 0$, the fluid limit of the process is given by $\nu(1 - \exp(-t))$; the fluid boundary ν is reached at time $t = +\infty$. In fact, with a second-order description, the process $X_N(t)$ can be written as $X_N(t) \sim \nu(1 - \exp(-t))N + Y(t)\sqrt{N}$ for some ergodic diffusion process $(Y(t))$, so the hitting time $T_{C_N}^{X_N}$ of the boundary is such that

$$\exp(-T_{C_N}^{X_N}) \sim \frac{Y(T_{C_N}^{X_N})}{\nu\sqrt{N}},$$

which gives a rough estimation $T_{C_N}^{X_N} \sim \log(\sqrt{N})$. The following proposition shows that this approximation is in fact quite precise. See also Theorem 4 of [7].

Proposition 10. *If $C_N = \nu N + \delta\sqrt{N} + o(\sqrt{N})$ with $\nu < 1$, $\delta \in \mathbb{R}$, and $X_N(0) = 0$, then the sequence of random variables*

$$(T_{C_N}^{X_N} - \frac{1}{2} \log(N))$$

converges in distribution to a random variable Z on \mathbb{R} whose Laplace transform at $\alpha > 0$ is given by

$$E(e^{-\alpha Z}) = \Gamma(\alpha) \int_0^{+\infty} \exp\left(u \frac{\delta}{\nu} - \frac{u^2}{2} \frac{1 - \nu}{\nu}\right) u^{\alpha-1} du.$$

If $\delta = 0$ then the variable $Z - \log(\nu/(1 - \nu))/2$ has the following density on \mathbb{R} :

$$x \mapsto \sqrt{\frac{2}{\pi}} \exp\left(-x - \frac{e^{-2x}}{2}\right).$$

Note that the Laplace transform of the limit in distribution is the ratio of the Mellin transforms of the functions

$$u \mapsto \exp(-u) \quad \text{and} \quad u \mapsto \exp\left(-u \frac{\delta}{\nu} - \frac{u^2}{2} \frac{1 - \nu}{\nu}\right).$$

See Section B.7 of [8] on Mellin transforms.

Proof of Proposition 10. Proposition 3 gives

$$E_0(\exp(-\alpha T_{C_N}^{X_N})) = \int_0^1 (1 - u)^N u^{\alpha-1} du \int_0^1 (1 - u)^{N-C_N} \left(1 + \frac{\mu}{\nu} u\right)^{C_N} u^{\alpha-1} du$$

for $\alpha > 0$. The asymptotic behavior of the numerator is easy to obtain since

$$\int_0^1 (1 - u)^N u^{\alpha-1} du = \frac{1}{N^\alpha} \int_0^N \left(1 - \frac{u}{N}\right)^N u^{\alpha-1} du \sim \frac{\Gamma(\alpha)}{N^\alpha}.$$

The denominator can be expressed as

$$\frac{1}{N^{\alpha/2}} \int_0^{\sqrt{N}} e^{f_N(u)} u^{\alpha-1} du,$$

with

$$\begin{aligned} f_N(u) &= (N - C_N) \log\left(1 - \frac{u}{\sqrt{N}}\right) + C_N \log\left(1 + \frac{\mu}{v} \frac{u}{\sqrt{N}}\right) \\ &= -\frac{1-v}{v} \frac{u^2}{2} + \frac{\delta}{v} u + o\left(\frac{1}{\sqrt{N}}\right). \end{aligned}$$

By dominated convergence, we therefore obtain, for $\alpha > 0$,

$$\begin{aligned} \lim_{N \rightarrow +\infty} E_0\left(\exp\left(-\alpha \left[T_{C_N}^{X_N} - \log \frac{N}{2}\right]\right)\right) \\ = \Gamma(\alpha) \int_0^{+\infty} \exp\left(-\frac{1-v}{v} \frac{u^2}{2} + \frac{\delta}{v} u\right) u^{\alpha-1} du, \end{aligned}$$

completing the proof of the first part of the proposition.

For $\delta = 0$, a change of variable gives

$$\int_0^{+\infty} \exp\left(-\frac{u^2}{2} \frac{1-v}{v}\right) u^{\alpha-1} du = \frac{1}{2} \left(\frac{2v}{1-v}\right)^{\alpha/2} \Gamma\left(\frac{\alpha}{2}\right).$$

The Laplace transform of Z can therefore be expressed as

$$E(e^{-\alpha Z}) = 2 \left(\frac{1-v}{2v}\right)^{\alpha/2} \frac{\Gamma(\alpha)}{\Gamma(\alpha/2)} = \left(\frac{1-v}{v}\right)^{\alpha/2} \frac{2^{\alpha/2}}{\sqrt{\pi}} \Gamma\left(\frac{\alpha+1}{2}\right),$$

by using Legendre’s duplication formula for gamma functions (see, e.g. [24, p. 240]). Since

$$\begin{aligned} \frac{2^{\alpha/2}}{\sqrt{\pi}} \Gamma\left(\frac{\alpha+1}{2}\right) &= \frac{1}{\sqrt{\pi}} \int_0^{+\infty} \exp\left(\frac{\alpha \log(2u)}{2} - \frac{\log(u)}{2} - u\right) du \\ &= \sqrt{\frac{2}{\pi}} \int_{-\infty}^{+\infty} e^{-\alpha u} \exp\left(-u - \frac{e^{-2u}}{2}\right) du, \end{aligned}$$

with a change of variable, we obtain the desired result on the distribution of Z .

We conclude with the hitting time of the empty state. We remark that, at the correct time scale, the time is half the corresponding variable in the subcritical case. See Proposition 9. A simple, naive explanation is as follows. For the subcritical regime, the process lives in a region centered at νN and whose width is of the order of \sqrt{N} , and, therefore, makes many excursions in this region before reaching C_N . In the critical case the process lives near, but only on one side of, νN . In particular, it cannot go above C_N and, consequently, does not waste time on such excursions.

Proposition 11. (Hitting time of the empty state.) *If $C_N = \nu N + o(\sqrt{N})$ with $\nu > 0$ and $X_N(0) = C_N$, then the sequence of random variables*

$$(N(1-\nu)^N T_0^{X_N})$$

converges in distribution to an exponential random variable with parameter 2ν .

Proof. Define $\alpha_N = N(1 - \nu)^N$. Recall that

$$E_{C_N}(\exp(-\alpha_N T_0^{X_N})) = \frac{d_N(\alpha)B_{C_N}(\alpha) + b_N(\alpha)D_{C_N}(\alpha)}{d_N(\alpha)B_0(\alpha) + b_N(\alpha)D_0(\alpha)}.$$

We start with the asymptotic behavior of $(d_N(\alpha_N))$. By definition,

$$\begin{aligned} d_N(\alpha_N) &= \mu \int_0^1 (1 - u)^{N-C_N-1} \left(1 + \frac{\mu}{\nu} u\right)^{C_N} u^{\alpha_N} du \\ &= \mu \frac{1}{\sqrt{N}} \int_0^{\sqrt{N}} \left(1 - \frac{u}{\sqrt{N}}\right)^{N-C_N-1} \left(1 + \frac{\mu}{\nu} \frac{u}{\sqrt{N}}\right)^{C_N} du + o\left(\frac{1}{\sqrt{N}}\right). \end{aligned}$$

Since $\eta = \nu$ and $\mu = 1 - \nu$, $\nu(1 - \eta) = \mu\eta$, so

$$\begin{aligned} d_N(\alpha_N) &= (1 - \nu) \frac{1}{\sqrt{N}} \int_0^{+\infty} \exp\left(-\frac{1 - \nu}{\nu} \frac{u^2}{2}\right) du + o\left(\frac{1}{\sqrt{N}}\right) \\ &= \frac{1}{\sqrt{N}} \sqrt{\frac{\pi}{2}} \sqrt{\nu(1 - \nu)} + o\left(\frac{1}{\sqrt{N}}\right). \end{aligned}$$

Note that, up to a term -1 in an exponent which does not play a role in the limiting behavior, the quantity $b_N(\alpha)$ is almost $d_N(\alpha)$ with ν replaced by $1 - \nu$. Consequently, $b_N(\alpha)$ has the same asymptotic expansion as $d_N(\alpha)$.

The asymptotic behaviors of $B_0(\alpha_N)$, $D_0(\alpha_N)$, $B_{C_N}(\alpha_N)$, and $D_{C_N}(\alpha_N)$ are the same as those obtained in the proof of Proposition 9. By gathering these various estimations we find that

$$\lim_{N \rightarrow +\infty} E(\exp(-\alpha_N T_0^{X_N})) = \frac{2}{2 + \alpha/\nu},$$

as required.

References

- [1] BINGHAM, N. H. (1991). Fluctuation theory for the Ehrenfest urn. *Adv. Appl. Prob.* **23**, 598–611.
- [2] DIACONIS, P., GRAHAM, R. L. AND MORRISON, J. A. (1990). Asymptotic analysis of a random walk on a hypercube with many dimensions. *Random Structures Algorithms* **1**, 51–72.
- [3] DI CRESCENZO, A. (1998). First-passage-time densities and avoiding probabilities for birth-and-death processes with symmetric sample paths. *J. Appl. Prob.* **35**, 383–394.
- [4] ENGSET, T. O. (1998). On the calculation of switches in an automatic telephone system. *Teletronikk* **94**, 99–142.
- [5] FEINSILVER, P. J. (1978). *Special Functions, Probability Semigroups, and Hamiltonian Flows* (Lecture Notes Math. **696**). Springer, Berlin.
- [6] FEUILLET, M. AND ROBERT, P. (2012). The time scales of a stochastic network with failures. Preprint. In preparation.
- [7] FLAJOLET, P. AND HUILLET, T. (2008). Analytic combinatorics of the Mabinogion urn. In *Proc. 5th Colloquium on Mathematics and Computer Science* (Discrete Math. Theoret. Comput. Sci. Proc. **A1**), ed. U. Röslér, Association of Discrete Mathematics and Theoretical Computer Science, Nancy, pp. 549–571.
- [8] FLAJOLET, P. AND SEDGEWICK, R. (2009). *Analytic Combinatorics*. Cambridge University Press.
- [9] FLEGG, M. B., POLLETT, P. K. AND GRAMOTNEV, D. K. (2008). Ehrenfest model for condensation and evaporation processes in degrading aggregates with multiple bonds. *Phys. Rev. E* **78**, 031117, 9 pp.
- [10] FRICKER, C., ROBERT, P. AND TIBI, D. (1999). On the rates of convergence of Erlang’s model. *J. Appl. Prob.* **36**, 1167–1184.
- [11] KARLIN, S. AND MCGREGOR, J. (1957). The classification of birth and death processes. *Trans. Amer. Math. Soc.* **86**, 366–400.
- [12] KARLIN, S. AND MCGREGOR, J. (1959). Coincidence properties of birth and death processes. *Pacific J. Math.* **9**, 1109–1140.

- [13] KARLIN, S. AND MCGREGOR, J. (1965). Ehrenfest urn models. *J. Appl. Prob.* **2**, 352–376.
- [14] KENNEDY, D. P. (1976). Some martingales related to cumulative sum tests and single-server queues. *Stoch. Process. Appl.* **4**, 261–269.
- [15] LAMPERTI, J. AND SNELL, J. L. (1963). Martin boundaries for certain Markov chains. *J. Math. Soc. Japan* **15**, 113–128.
- [16] PALACIOS, J. L. (1993). Fluctuation theory for the Ehrenfest urn via electric networks. *Adv. Appl. Prob.* **25**, 472–476.
- [17] REVUZ, D. AND YOR, M. (1999). *Continuous Martingales and Brownian Motion* (Fundamental Principles Math. Sci. **293**), 3rd edn. Springer, Berlin.
- [18] ROBERT, P. (2003). *Stochastic Networks and Queues* (Appl. Math. **52**). Springer, Berlin.
- [19] ROGERS, L. C. G. AND WILLIAMS, D. (2000). *Diffusions, Markov Processes, and Martingales*, Vol. 2. Cambridge University Press.
- [20] SCHOUTENS, W. (2000). *Stochastic Processes and Orthogonal Polynomials* (Lecture Notes Statist. **146**). Springer, New York.
- [21] SCHOUTENS, W. AND TEUGELS, J. L. (1998). Lévy processes, polynomials and martingales. *Commun. Statist. Stoch. Models* **14**, 335–349.
- [22] SIMATOS, F. AND TIBI, D. (2010). Spatial homogenization in a stochastic network with mobility. *Ann. Appl. Prob.* **20**, 312–355.
- [23] SZEGŐ, G. (1975). *Orthogonal Polynomials*, 4th edn. American Mathematical Society, Providence, RI.
- [24] WHITTAKER, E. T. AND WATSON, G. N. (1996). *A Course of Modern Analysis*. Cambridge University Press.