



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Hautphenne, S;Latouche, G

Title:

Lyapunov Exponents for Branching Processes in a Random Environment: The Effect of Information

Date:

2016-04-01

Citation:

Hautphenne, S. & Latouche, G. (2016). Lyapunov Exponents for Branching Processes in a Random Environment: The Effect of Information. JOURNAL OF STATISTICAL PHYSICS, 163 (2), pp.393-410. <https://doi.org/10.1007/s10955-016-1474-3>.

Persistent Link:

<https://hdl.handle.net/11343/248508>

Sophie Hautphenne · Guy Latouche

Lyapunov exponents for branching processes in a random environment

The effect of information

Received: date / Accepted: date

Abstract We consider multitype branching processes evolving in a Markovian random environment. To determine whether or not the branching process becomes extinct almost surely is akin to computing the maximal Lyapunov exponent of a sequence of random matrices, which is a notoriously difficult problem. We define Markov chains associated to the branching process, and we construct bounds for the Lyapunov exponent. The bounds are obtained by adding or by removing information: to add information results in a lower bound, to remove information results in an upper bound, and we show that adding less information improves the lower bound. We give a few illustrative examples and we observe that the upper bound is generally more accurate than the lower bounds.

Keywords product of random matrices · Lyapunov exponent · multitype branching process · Markovian random environment · extinction criterion

Mathematics Subject Classification (2000) 60J80 · 60J22 · 60J27

S. Hautphenne
The University of Melbourne,
Department of Mathematics and Statistics,
3010, Melbourne, VIC, Australia
Tel.: +61 3 8344 9073
Fax: +61 3 8344 4599
E-mail: sophiemh@unimelb.edu.au

G. Latouche
Université libre de Bruxelles
Département d'informatique,
1050 Brussels, Belgium
E-mail: latouche@ulb.ac.be

1 Introduction

In his paper on subadditive ergodic theory, Kingman [12] proved that if $\{\mathbf{A}_n\}_{n=0}^\infty$ is an ergodic stationary process with values in the space $\mathbb{R}_+^{r \times r}$ of $r \times r$ matrices with strictly positive elements, and if $\mathbb{E}[\log^+ |(\mathbf{A}_0)_{ij}|] < \infty$ for $1 \leq i, j \leq r$, then

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \{ \mathbf{A}_0 \mathbf{A}_1 \dots \mathbf{A}_{n-1} \}_{ij} = \omega \quad (1)$$

exists w.p.1, where ω is a finite non random real number, independent of i and j . In addition, the limit ω takes different equivalent forms:

$$\omega = \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}[\log \{ \mathbf{A}_0 \mathbf{A}_1 \dots \mathbf{A}_{n-1} \}_{ij}] \quad \text{for all } i, j, \quad (2)$$

$$= \lim_{n \rightarrow \infty} \frac{1}{n} \log \| \mathbf{A}_0 \mathbf{A}_1 \dots \mathbf{A}_{n-1} \| \quad \text{a.s.}$$

$$= \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}[\log \| \mathbf{A}_0 \mathbf{A}_1 \dots \mathbf{A}_{n-1} \|], \quad (3)$$

independently of the matrix norm. In some contexts, the limit ω is called the *maximal Lyapunov exponent*, and the convergence property has been extended to real- and even complex-valued matrices under suitable regularity conditions (see for instance Oseledec [15] and Key [11]).

Lyapunov exponents come into play when studying the asymptotic behavior of stochastic dynamical systems. They appear in particular in an extinction criterion for multitype branching processes evolving in a random environment [18]. This specific application is the motivation behind our analysis and is our main focus here. Precisely, we consider an irreducible multitype continuous-time branching process with $r \in \mathbb{N}_0$ types of individuals, and we assume that its parameters vary over time according to a Markovian random environment $\{X(t) : t \in \mathbb{R}^+\}$ which is a continuous-time irreducible Markov chain on the finite state space $E = \{1, 2, \dots, m\}$.

In the absence of a random environment, the parameters of the branching process stay constant over time, and the (i, j) th entry of the mean population size matrix $\mathbf{A}(t)$ records the conditional expected number of individuals of type j alive at time t , given that the population starts at time 0 with one single individual of type i . In the particular context of a *Markovian branching process* (MBP), an individual of type i ($1 \leq i \leq r$) lives for an exponentially distributed amount of time with parameter λ_i , after which it generates a random number of children of each type j , $1 \leq j \leq r$. It is well known (Athreya and Ney [3]) that the mean population size matrix of an MBP takes the specific form $\mathbf{A}(t) = e^{\mathbf{\Omega}t}$ where $\mathbf{\Omega}$ is a matrix with entries

$$\Omega_{ij} = \lambda_i(m_{ij} - \delta_{ij}),$$

where m_{ij} is the expected number of children of type j from a parent of type i , and δ_{ij} is the Kronecker delta.

In presence of the random environment $\{X(t)\}$, the mean population size matrix at time t becomes random as it depends on the path of the environmental process during the time interval $[0, t]$. Let $\{\hat{X}_k : k \in \mathbb{N}\}$ denote the jump chain associated with $\{X(t)\}$, with transition probability matrix $\hat{\mathbf{P}}$ and stationary probability vector $\hat{\boldsymbol{\pi}}$. For $k \geq 0$, we denote by ξ_k the sojourn time in the k th environmental state; ξ_k is exponentially distributed with parameter c_ℓ if $\hat{X}_k = \ell$, $1 \leq \ell \leq m$. The $r \times r$ random matrix $\mathbf{A}_k := \mathbf{A}(\hat{X}_k, \xi_k)$ is the mean population size matrix corresponding to the interval of time between the k th and the $(k+1)$ st jump of $\{X(t)\}$. The matrices \mathbf{A}_k , $k \geq 0$, take their value in the non-countable state space $\mathcal{A} = \{\mathbf{A}(\ell, x) \in \mathbb{R}^{r \times r} : 1 \leq \ell \leq m, x \in \mathbb{R}^+\}$ of nonnegative matrices which depend on one discrete parameter ℓ (the environmental state) and one continuous parameter x (the sojourn time in the environmental state). In what follows, we assume that the sequence $\{\hat{X}_k\}$ is stationary, which ensures that the sequence $\{\mathbf{A}_k\}$ is ergodic. The matrices \mathbf{A}_k are then identically distributed, but they are not independent.

The next theorem follows from [18, Theorem 9.10] and gives a necessary and sufficient condition for the almost sure extinction of the branching process in a Markovian random environment. It holds under suitable regularity assumptions on the branching process, which are omitted here for the sake of concision; we refer to [18] for details.

Theorem 1 *There exists a constant ω such that*

$$\omega = \lim_{n \rightarrow \infty} \frac{1}{n} \log \{\mathbf{A}_0 \mathbf{A}_1 \cdots \mathbf{A}_{n-1}\}_{ij} \quad (4)$$

almost surely, independently of i and j . Extinction is almost sure if and only if $\omega \leq 0$.

The (i, j) th entry of the random matrix product in (4) is the conditional expected number of individuals of type j alive just before the n th environmental state change, given that the population starts at time 0 with one single individual of type i , and given the history of the environmental process. The limit ω may be interpreted as the asymptotic growth rate of the population.

Theorem 1 shows that being able to evaluate the limit ω is fundamental in order to characterise the criticality of a multitype branching process in a random environment. It is well known, however, that Lyapunov exponents are hard to compute (Kingman [12], Tsitsiklis and Blondel [19]), except under very special circumstances, such as in Key [10] where the random matrices in the family are assumed to be simultaneously diagonalizable, or in Pollicott [16] where the random matrices are independent and identically distributed, and take their values in a finite set. For a thorough survey on the basics of Lyapunov exponents, we refer to Watkins [20].

In the absence of an easily computed exact expression for ω , we look for upper and lower bounds, in an adaptation of the approach in [8] developed for

branching processes subject to binomial catastrophes at random times. Our bounds are obtained in several steps. First, we replace the branching process by a marked Markov chain on the state space $\{1, \dots, m\} \times \{1, \dots, r\}$. This has the advantage that we deal with the trajectories of a simple discrete-time Markov chain, instead of the conditional expected number of individuals of each type in the original branching process. Next, we take the expectation with respect to the \hat{X}_k s and the ξ_k s, and obtain an upper bound for ω . Finally, we add some information about the history of the associated Markov chain and so obtain a lower bound. In some cases, we have some leeway in the amount of information that we may add and we show that the lower bound is tighter when we add less information. Thus, roughly speaking, we may say that the conditional expectation in (4) is taken with respect to a finely balanced information; less information leads to an upper bound, while more information yields a lower bound.

We present our results in as general a setting as possible with respect to the structure of the random matrices \mathbf{A}_k . In places, we use the MBP model to obtain more precise results. In that case, the matrix $\hat{\Omega}$ becomes a function of the environmental state, taking the value $\Omega(\ell)$ when $\hat{X}_k = \ell$, and the matrix \mathbf{A}_k then has the explicit form $\mathbf{A}_k = e^{\Omega(\hat{X}_k)\xi_k}$.

Expressions such as in the right-hand side of (4) typically arise in the analysis of systems in random environments. For instance, Bolthausen and Goldsheid analyse in [4] a random walk on a strip subject to a random environment, and their parameter λ in [4, Equation 6] is a Lyapunov exponent. It will be readily seen that our approach may be adapted to the asymptotic analysis of such systems.

The paper is organised as follows. A marked process is defined in Section 2. Sections 3 and 4 are devoted to the construction of an upper and a lower bound for ω , respectively. We discuss in Section 5 an alternative definition of the associated Markov chain yielding a different lower bound. The constructions in Sections 3 and 4 stem from algebraic considerations, but it turns out that there is a physical interpretation for the two associated Markov chains in the case of an MBP; this is presented in Section 6. We give two illustrative examples in Section 7, showing that the precision of the lower bounds depends on the case at hand. We show in Section 8 how the lower bound may be made more precise in the special circumstance where the environmental process $\{X(t)\}$ is cyclic. Finally, we prove in Section 9 that our bounds are tight.

2 The single individual Markov chain

We assume that the matrices \mathbf{A}_k form an ergodic process and that they are nonnegative and irreducible for all $k \geq 0$, so that \mathbf{A}_k has a Perron-Frobenius (PF) eigenvalue ρ_k with corresponding strictly positive left- and

right-eigenvectors \mathbf{u}_k and \mathbf{v}_k , normalised such that $\mathbf{u}_k \mathbf{1} = 1$ and $\mathbf{u}_k \mathbf{v}_k = 1$. Due to the nature of the random matrices \mathbf{A}_k , the quantities ρ_k , \mathbf{u}_k and \mathbf{v}_k are all functions of \hat{X}_k and ξ_k .

We next proceed in two steps. First, we define the normalised random matrices

$$\mathbf{A}_k^* = \frac{1}{\rho_k} \mathbf{A}_k, \quad \text{for } k \geq 0,$$

which all have the PF eigenvalue 1. As $\{\mathbf{A}_k\}$ is ergodic, (4) can be rewritten as

$$\omega = \mathbb{E}[\log \rho_0] + \Psi, \quad (5)$$

where Ψ is a constant and

$$\Psi = \lim_{n \rightarrow \infty} \frac{1}{n} \log \{\mathbf{A}_0^* \mathbf{A}_1^* \cdots \mathbf{A}_{n-1}^*\}_{ij} \quad \text{a.s.} \quad (6)$$

In the second step, we define the random matrices $\mathbf{\Delta}_k = \text{diag}(\mathbf{v}_k)$ and

$$\mathbf{\Theta}_k = \mathbf{\Delta}_k^{-1} \mathbf{A}_k^* \mathbf{\Delta}_k, \quad \text{for } k \geq 0, \quad (7)$$

and we rewrite (6) as

$$\Psi = \lim_{n \rightarrow \infty} \frac{1}{n} \log \{\mathbf{\Delta}_0 \mathbf{\Theta}_0 \mathbf{\Delta}_0^{-1} \mathbf{\Delta}_1 \mathbf{\Theta}_1 \mathbf{\Delta}_1^{-1} \cdots \mathbf{\Delta}_{n-1} \mathbf{\Theta}_{n-1} \mathbf{\Delta}_{n-1}^{-1}\}_{ij} \quad \text{a.s.} \quad (8)$$

While the physical interpretation of the matrix product in (6) is not obvious, we may easily give one to the equivalent matrix product in (8). Indeed, it is easy to verify that the matrices $\mathbf{\Theta}_k$ are stochastic and irreducible. We have thus replaced the original process by a non-homogeneous Markov chain $\{\varphi_k : k \in \mathbb{N}\}$ on the state space $\{1 \dots r\}$ evolving in the random environment $\{\hat{X}_k, \xi_k\}$, with transition probability matrix $\mathbf{\Theta}_k = \mathbf{\Theta}(\hat{X}_k, \xi_k)$ at time k . We call $\{\varphi_k\}$ the *single individual Markov chain*. Note that similar transformations have been used in different contexts, as in [2] and [17] for instance, and are often associated with time-reversal of the Markov process, as we will consider in Section 5; here, however, the process $\{\varphi_k\}$ does not involve time-reversal.

Finally, we associate to the single individual Markov chain a random sequence $\{Z_k : k \in \mathbb{N}\}$ of marks, where

$$Z_k = (\mathbf{\Delta}_k)_{\varphi_k} (\mathbf{\Delta}_k^{-1})_{\varphi_{k+1}} = \frac{(\mathbf{v}_k)_{\varphi_k}}{(\mathbf{v}_k)_{\varphi_{k+1}}}, \quad \text{for } k \geq 0, \quad (9)$$

and we define the product $\mathcal{Z}_{n,j}$ by

$$\mathcal{Z}_{n,j} = Z_0 \cdots Z_{n-1} \mathbb{1}_{\{\varphi_n=j\}}, \quad (10)$$

for $n \geq 1$ and $1 \leq j \leq r$.

Lemma 1 *The Lyapunov exponent ω may be written as*

$$\omega = \mathbb{E}[\log \rho_0] + \Psi \quad (11)$$

where

$$\Psi = \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}[\mathcal{Z}_{n,j} | \varphi_0 = i, \boldsymbol{\xi}^{(n)}, \hat{\mathbf{X}}^{(n)}] \quad a.s., \quad (12)$$

independently of i and j , with $\boldsymbol{\xi}^{(n)} = (\xi_0, \dots, \xi_{n-1})$, and $\hat{\mathbf{X}}^{(n)} = (\hat{X}_0, \dots, \hat{X}_{n-1})$.

Proof We observe that

$$\mathbb{E}[Z_k \mathbb{1}_{\{\varphi_{k+1}=j\}} | \varphi_k = i, \boldsymbol{\xi}^{(k)}, \hat{\mathbf{X}}^{(k)}, \xi_k, \hat{X}_k] = (\boldsymbol{\Delta}_k \boldsymbol{\Theta}_k \boldsymbol{\Delta}_k^{-1})_{ij}, \quad k \geq 0,$$

and a simple calculation leads to

$$\mathbb{E}[\mathcal{Z}_{n,j} | \varphi_0 = i, \boldsymbol{\xi}^{(n)}, \hat{\mathbf{X}}^{(n)}] = (\boldsymbol{\Delta}_0 \boldsymbol{\Theta}_0 \boldsymbol{\Delta}_0^{-1} \cdots \boldsymbol{\Delta}_{n-1} \boldsymbol{\Theta}_{n-1} \boldsymbol{\Delta}_{n-1}^{-1})_{ij}, \quad (13)$$

which, by (7), is equivalent to

$$\mathbb{E}[\mathcal{Z}_{n,j} | \varphi_0 = i, \boldsymbol{\xi}^{(n)}, \hat{\mathbf{X}}^{(n)}] = (\mathbf{A}_0^* \cdots \mathbf{A}_{n-1}^*)_{ij}. \quad (14)$$

□

The advantage of (12) over (8) is that we now deal with a product of scalar random variables, $\mathcal{Z}_{n,j}$, instead of a product of random matrices.

Remark 1 In the case of a branching process in a random environment, we recognise in the first term of (11) the expected long-term growth rate of the population, while the second term reflects the fact that changes in the environment influence all individuals simultaneously; we return to this point in Section 6.

In the next two sections, we condition on less information to find an upper bound to Ψ , and we condition on more information to find a lower bound to Ψ .

3 Upper bound

Recall that the matrices $\{\mathbf{A}_k^*\}$ are random through both the environmental process and the sojourn time in each state; we may write explicitly $\mathbf{A}_k^* = \mathbf{A}^*(\hat{X}_k, \xi_k)$. Their conditional expectation, given that the environmental state is ℓ , are denoted as \mathbf{M}_ℓ with

$$\mathbf{M}_\ell = \int_0^\infty \mathbf{A}^*(\ell, x) c_\ell e^{-c_\ell x} dx \quad (15)$$

for each $1 \leq \ell \leq m$. We also define the $rm \times rm$ matrix

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_1 & & & \\ & \mathbf{M}_2 & & \\ & & \ddots & \\ & & & \mathbf{M}_m \end{bmatrix}, \quad (16)$$

and we use the notation $\text{sp}(\mathbf{A})$ for the spectral radius of any square matrix \mathbf{A} .

Theorem 2 *An upper bound for the Lyapunov exponent ω is given by ω_U with*

$$\omega_U = \mathbb{E}[\log \rho_0] + \log \text{sp}[\mathbf{M}(\hat{\mathbf{P}} \otimes \mathbf{I}_r)],$$

where \mathbf{I}_r is the identity matrix of size r , and \otimes denotes the Kronecker product.

Proof From (2), Ψ can be written as

$$\Psi = \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}[\log \{\mathbf{A}_0^* \mathbf{A}_1^* \cdots \mathbf{A}_{n-1}^*\}_{ij}], \quad (17)$$

which, by Jensen's Inequality, gives

$$\Psi \leq \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}[\{\mathbf{A}_0^* \mathbf{A}_1^* \cdots \mathbf{A}_{n-1}^*\}_{ij}].$$

By conditioning on \hat{X}_0 and ξ_0 , we have

$$\begin{aligned} \mathbb{E}[\{\mathbf{A}_0^*\}_{ij}] &= \sum_{1 \leq \ell \leq m} \hat{\pi}_\ell(\mathbf{M}_\ell)_{ij} \\ &= [(\hat{\boldsymbol{\pi}} \otimes \mathbf{I}_r) \mathbf{M}(\mathbf{1} \otimes \mathbf{I}_r)]_{ij}, \end{aligned}$$

where the vector $\mathbf{1}$ is a column vector of size m containing only 1's. By induction, one shows that for any $n \geq 1$,

$$\mathbb{E}[\{\mathbf{A}_0^* \mathbf{A}_1^* \cdots \mathbf{A}_{n-1}^*\}_{ij}] = [(\hat{\boldsymbol{\pi}} \otimes \mathbf{I}_r) [\mathbf{M}(\hat{\mathbf{P}} \otimes \mathbf{I}_r)]^n (\mathbf{1} \otimes \mathbf{I}_r)]_{ij},$$

so that

$$\begin{aligned} \Psi &\leq \lim_{n \rightarrow \infty} \frac{1}{n} \log \left[(\hat{\boldsymbol{\pi}} \otimes \mathbf{I}_r) [\mathbf{M}(\hat{\mathbf{P}} \otimes \mathbf{I}_r)]^n (\mathbf{1} \otimes \mathbf{I}_r) \right]_{ij} \\ &= \log \lim_{n \rightarrow \infty} \left[(\hat{\boldsymbol{\pi}} \otimes \mathbf{I}_r) [\mathbf{M}(\hat{\mathbf{P}} \otimes \mathbf{I}_r)]^n (\mathbf{1} \otimes \mathbf{I}_r) \right]_{ij}^{1/n} \\ &= \log \text{sp}[\mathbf{M}(\hat{\mathbf{P}} \otimes \mathbf{I}_r)], \end{aligned}$$

independently of $\hat{\boldsymbol{\pi}}$, i , and j . \square

Remark 2 In the MBP case, the matrices \mathbf{M}_ℓ have the following explicit form:

$$\begin{aligned}\mathbf{M}_\ell &= \int_0^\infty \mathbf{A}^*(\ell, x) c_\ell e^{-c_\ell x} dx \\ &= \int_0^\infty e^{[\boldsymbol{\Omega}(\ell) - \lambda(\ell)\mathbf{I}]x} c_\ell e^{-c_\ell x} dx \\ &= c_\ell [(c_\ell + \lambda(\ell))\mathbf{I} - \boldsymbol{\Omega}(\ell)]^{-1}.\end{aligned}\tag{18}$$

Remark 3 By (14),

$$\begin{aligned}\mathbb{E}[\{\mathbf{A}_0^* \mathbf{A}_1^* \cdots \mathbf{A}_{n-1}^*\}_{ij}] &= \mathbb{E}\left[\mathbb{E}[\mathcal{Z}_{n,j} \mid \varphi_0 = i, \boldsymbol{\xi}^{(n)}, \hat{\mathbf{X}}^{(n)}]\right] \\ &= \mathbb{E}[\mathcal{Z}_{n,j} \mid \varphi_0 = i].\end{aligned}$$

This relates the upper bound ω_U to the expectation of the product $\mathcal{Z}_{n,j}$ of the marks, *unconditional* on the path of the random environment. The upper bound is thus technically obtained by removing some information about the process.

4 Lower bound

To obtain a lower bound, we use the expression (12) of Ψ in terms of a product of marks, and we take the opposite direction to the one identified in Remark 3: we start from the conditional expectation of $\mathcal{Z}_{n,j}$ and give *more* information. We introduce two discrete-time Markov chains, $\{Y_k^{(1)} = (\hat{X}_k, \varphi_k) : k \in \mathbb{N}\}$ and $\{Y_k^{(2)} = (\hat{X}_k, \varphi_{k+1}) : k \in \mathbb{N}\}$, with respective transition probability matrices $\mathbf{P}^{(1)}$ and $\mathbf{P}^{(2)}$ of size $mr \times mr$. A transition of $\{Y_k^{(1)}\}$ from (ℓ, i) to (ℓ', j) occurs with probability

$$\mathbf{P}_{(\ell,i),(\ell',j)}^{(1)} = (N_\ell)_{ij} \hat{\mathbf{P}}_{\ell\ell'},$$

where

$$\mathbf{N}_\ell = \int_0^\infty \boldsymbol{\Theta}(\ell, x) c_\ell e^{-c_\ell x} dx.\tag{19}$$

We note for future use that the matrix \mathbf{N}_ℓ is irreducible. Similarly, $\mathbf{P}_{(\ell,i),(\ell',j)}^{(2)} = \hat{\mathbf{P}}_{\ell\ell'} (\mathbf{N}_{\ell'})_{ij}$, and we write in matrix form

$$\mathbf{P}^{(1)} = \mathbf{N}(\hat{\mathbf{P}} \otimes \mathbf{I}_r), \quad \mathbf{P}^{(2)} = (\hat{\mathbf{P}} \otimes \mathbf{I}_r)\mathbf{N},$$

where

$$\mathbf{N} = \begin{bmatrix} \mathbf{N}_1 & & & \\ & \mathbf{N}_2 & & \\ & & \ddots & \\ & & & \mathbf{N}_m \end{bmatrix}.$$

By irreducibility of $\hat{\mathbf{P}}$ and of the matrices $\mathbf{N}_1, \dots, \mathbf{N}_m$, both $\mathbf{P}^{(1)}$ and $\mathbf{P}^{(2)}$ are irreducible.

The lower bound obtained below is based on the assumption that the right eigenvector \mathbf{v}_k depends on the discrete random variable \hat{X}_k only, and not on the continuous random variable ξ_k . This is the case in particular for an MBP, as we show in Section 6. For each $1 \leq \ell \leq m$, we then define the vector $\mathbf{v}(\ell)$ as the right PF eigenvector of $\mathbf{\Omega}(\ell)$, so that $\mathbf{v}_k = \mathbf{v}(\ell)$ if $\hat{X}_k = \ell$, and we let

$$\bar{\mathbf{v}} = \begin{bmatrix} \mathbf{v}(1) \\ \mathbf{v}(2) \\ \vdots \\ \mathbf{v}(m) \end{bmatrix}.$$

Theorem 3 *If $\mathbf{v}_k = \mathbf{v}(\hat{X}_k)$ independently of ξ_k , then a lower bound for the Lyapunov exponent ω is given by ω_L with*

$$\omega_L = \mathbb{E}[\log \rho_0] + \boldsymbol{\pi}^{(1)}(\mathbf{I} - \mathbf{N}) \log \bar{\mathbf{v}}, \quad (20)$$

where $\boldsymbol{\pi}^{(1)}$ is the stationary probability vector of $\mathbf{P}^{(1)}$, and $\log \bar{\mathbf{v}}$ is taken entry-wise.

Proof We start from the conditional expectation of $\mathcal{Z}_{n,j}$ given the successive states $\boldsymbol{\varphi}^{(n)} = (\varphi_1, \dots, \varphi_{n-1})$ in addition to $\boldsymbol{\xi}^{(n)}$ and $\hat{\mathbf{X}}^{(n)}$. We readily find from (9) that

$$\begin{aligned} & \mathbb{E}[\mathcal{Z}_{n,j} \mid \varphi_0 = i, \boldsymbol{\varphi}^{(n)}, \boldsymbol{\xi}^{(n)}, \hat{\mathbf{X}}^{(n)}] \\ &= \frac{(\mathbf{v}_0)_i}{(\mathbf{v}_0)_{\varphi_1}} \frac{(\mathbf{v}_1)_{\varphi_1}}{(\mathbf{v}_1)_{\varphi_2}} \dots \frac{(\mathbf{v}_{n-2})_{\varphi_{n-2}}}{(\mathbf{v}_{n-2})_{\varphi_{n-1}}} \frac{(\mathbf{v}_{n-1})_{\varphi_{n-1}}}{(\mathbf{v}_{n-1})_j} (\boldsymbol{\Theta}_{n-1})_{\varphi_{n-1},j} \\ &= \prod_{\ell=1}^m \prod_{a=1}^r (\mathbf{v}(\ell))_a^{n_{(\ell,a)}^{(1)} - n_{(\ell,a)}^{(2)}} \frac{(\mathbf{v}_0)_i}{(\mathbf{v}_{n-1})_j} (\boldsymbol{\Theta}_{n-1})_{\varphi_{n-1},j}, \end{aligned}$$

where

$$n_{(\ell,a)}^{(1)} = \sum_{k=1}^{n-1} \mathbb{1}_{\{Y_k^{(1)} = (\ell,a)\}} \quad \text{and} \quad n_{(\ell,a)}^{(2)} = \sum_{k=0}^{n-2} \mathbb{1}_{\{Y_k^{(2)} = (\ell,a)\}}.$$

Note that $n_{(\ell,a)}^{(1)}$ and $n_{(\ell,a)}^{(2)}$ are functions of n but we omit this dependence in the notation. Observe that we use here the fact that $\mathbf{v}_k = \mathbf{v}(\hat{X}_k)$ independently of ξ_k . Then, using the ergodic theorem for finite Markov chains [14],

$$\begin{aligned} & \lim_{n \rightarrow \infty} \mathbb{E}[\mathcal{Z}_{n,j} \mid \varphi_0 = i, \boldsymbol{\varphi}^{(n)}, \boldsymbol{\xi}^{(n)}, \hat{\mathbf{X}}^{(n)}]^{1/n} \\ &= \lim_{n \rightarrow \infty} \prod_{\ell=1}^m \prod_{a=1}^r (\mathbf{v}(\ell))_a^{\frac{n_{(\ell,a)}^{(1)} - n_{(\ell,a)}^{(2)}}{n}} \left(\frac{(\mathbf{v}_0)_i}{(\mathbf{v}_{n-1})_j} (\boldsymbol{\Theta}_{n-1})_{\varphi_{n-1},j} \right)^{1/n} \\ &= \prod_{\ell=1}^m \prod_{a=1}^r (\mathbf{v}(\ell))_a^{\pi_{(\ell,a)}^{(1)} - \pi_{(\ell,a)}^{(2)}} \quad \text{a.s.} \end{aligned} \quad (21)$$

where $\boldsymbol{\pi}^{(1)}$ and $\boldsymbol{\pi}^{(2)}$ are the stationary distribution vectors of the irreducible stochastic matrices $\mathbf{P}^{(1)}$ and $\mathbf{P}^{(2)}$. Now,

$$\begin{aligned}\Psi &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}[\mathcal{Z}_{n,j} | \varphi_0 = i, \boldsymbol{\xi}^{(n)}, \hat{\mathbf{X}}^{(n)}] \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}_\varphi \left[\mathbb{E}[\mathcal{Z}_{n,j} | \varphi_0 = i, \boldsymbol{\varphi}^{(n)}, \boldsymbol{\xi}^{(n)}, \hat{\mathbf{X}}^{(n)}] \right],\end{aligned}$$

where the outermost expectation is with respect to $\boldsymbol{\varphi}^{(n)}$,

$$\begin{aligned}&\geq \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}_\varphi \left[\log \mathbb{E}[\mathcal{Z}_{n,j} | \varphi_0 = i, \boldsymbol{\varphi}^{(n)}, \boldsymbol{\xi}^{(n)}, \hat{\mathbf{X}}^{(n)}] \right] \\ &= \mathbb{E}_\varphi \left[\log \lim_{n \rightarrow \infty} \mathbb{E}[\mathcal{Z}_{n,j} | \varphi_0 = i, \boldsymbol{\varphi}^{(n)}, \boldsymbol{\xi}^{(n)}, \hat{\mathbf{X}}^{(n)}]^{1/n} \right]\end{aligned}\quad (22)$$

where we use Jensen's Inequality, followed by the Dominated Convergence Theorem after observing that $c^n < \mathcal{Z}_{n,j} < C^n$ for some constants $0 < c < C < \infty$.

From (11), (21) and (22), we find that $\omega \geq \omega_L$, where

$$\omega_L = \mathbb{E}[\log \rho_0] + (\boldsymbol{\pi}^{(1)} - \boldsymbol{\pi}^{(2)}) \log \bar{\mathbf{v}}.$$

Finally, it is easy to verify that $\boldsymbol{\pi}^{(2)} = \boldsymbol{\pi}^{(1)} \mathbf{N}$, which concludes the proof. \square

Remark 4 In the MBP case, $\mathbf{N}_\ell = \text{diag}(\mathbf{u}(\ell))^{-1} \mathbf{M}_\ell \text{diag}(\mathbf{u}(\ell))$, where \mathbf{M}_ℓ is given by (18).

5 A dual Markov chain

Instead of using the right-eigenvectors \mathbf{v}_k of the matrices \mathbf{A}_k as we do in (7), one may associate another Markov chain, starting from the left-eigenvectors \mathbf{u}_k , which we assume straightaway to depend on \hat{X}_k only. Define $\bar{\boldsymbol{\Delta}}_k = \text{diag}(\mathbf{u}_k)$ and

$$\bar{\boldsymbol{\Theta}}_k = \bar{\boldsymbol{\Delta}}_k^{-1} (\mathbf{A}_k^*)^\top \bar{\boldsymbol{\Delta}}_k \quad \text{for } k \geq 0. \quad (23)$$

Like the $\boldsymbol{\Theta}_k$ s, these are irreducible stochastic matrices and one shows that

$$\begin{aligned}\Psi &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \left\{ \bar{\boldsymbol{\Delta}}_{n-1} \bar{\boldsymbol{\Theta}}_{n-1} \bar{\boldsymbol{\Delta}}_{n-1}^{-1} \bar{\boldsymbol{\Delta}}_{n-2} \bar{\boldsymbol{\Theta}}_{n-2} \bar{\boldsymbol{\Delta}}_{n-2}^{-1} \cdots \right. \\ &\quad \left. \bar{\boldsymbol{\Delta}}_0 \bar{\boldsymbol{\Theta}}_0 \bar{\boldsymbol{\Delta}}_0^{-1} \right\}_{ji}, \quad \text{a.s.}\end{aligned}\quad (24)$$

We define $\{\tilde{X}_k\}$ as the *time-reversed* version of the environmental jump chain, with transition matrix

$$\tilde{\mathbf{P}} = \text{diag}(\hat{\boldsymbol{\pi}})^{-1} \hat{\mathbf{P}}^\top \text{diag}(\hat{\boldsymbol{\pi}}),$$

and we rewrite (24) as

$$\bar{\Psi} = \lim_{n \rightarrow \infty} \frac{1}{n} \log \left\{ \tilde{\Delta}_0 \tilde{\Theta}_0 \tilde{\Delta}_0^{-1} \tilde{\Delta}_1 \tilde{\Theta}_1 \tilde{\Delta}_1^{-1} \cdots \tilde{\Delta}_{n-1} \tilde{\Theta}_{n-1} \tilde{\Delta}_{n-1}^{-1} \right\}_{ji} \quad a.s.,$$

where $\tilde{\Delta}_k = \Delta(\tilde{X}_k)$, $\tilde{\Theta}_k = \Theta(\tilde{X}_k, \eta_k)$, and η_k is an exponential random variable with parameter $c_{\tilde{X}_k}$.

By following the same steps as in Section 4, we obtain a new lower bound; the proof is omitted.

Theorem 4 *If $\mathbf{u}_k = \mathbf{u}(\tilde{X}_k)$, then an alternative lower bound for the Lyapunov exponent ω is given by*

$$\tilde{\omega}_L = \mathbb{E}[\log \rho_0] + \tilde{\boldsymbol{\pi}}^{(1)} (\mathbf{I} - \tilde{\mathbf{N}}) \log \bar{\mathbf{u}}^\top,$$

where

$$\tilde{\mathbf{N}} = \begin{bmatrix} \tilde{\mathbf{N}}_1 & & & \\ & \tilde{\mathbf{N}}_2 & & \\ & & \ddots & \\ & & & \tilde{\mathbf{N}}_m \end{bmatrix},$$

$$\tilde{\mathbf{N}}_\ell = \int_0^\infty \tilde{\Theta}(\ell, x) c_\ell e^{-c_\ell x} dx,$$

the vector $\tilde{\boldsymbol{\pi}}^{(1)}$ is the stationary probability vector of $\tilde{\mathbf{P}}^{(1)} = \tilde{\mathbf{N}}(\tilde{\mathbf{P}} \otimes \mathbf{I}_r)$ and

$$\bar{\mathbf{u}} = [\mathbf{u}(1), \mathbf{u}(2), \dots, \mathbf{u}(m)].$$

□

6 The case of Markovian branching processes

In the MBP case, we may give a physical interpretation to the single individual Markov chain and to the dual Markov chain defined in (7) and (23), respectively, as well as to Lemma 1.

As mentioned in the introduction, $\mathbf{A}_k = e^{\boldsymbol{\Omega}(\tilde{X}_k)\xi_k}$, and so $\rho_k = e^{\lambda_k \xi_k}$ where λ_k is the maximal eigenvalue of $\boldsymbol{\Omega}(\tilde{X}_k)$. Thus,

$$\mathbb{E}[\log \rho_0] = \mathbb{E}[\lambda_0 \xi_0] = \sum_{1 \leq \ell \leq m} \hat{\pi}_\ell \lambda(\ell) / c_\ell,$$

where, for $1 \leq \ell \leq m$, $\lambda(\ell)$ is the PF eigenvalue of $\boldsymbol{\Omega}(\ell)$, and c_ℓ denotes the parameter of the exponential distribution of sojourn times in the environmental state ℓ . Furthermore, the vectors \mathbf{u}_k and \mathbf{v}_k correspond to the PF eigenvectors of $\boldsymbol{\Omega}(\tilde{X}_k)$ and are therefore independent of ξ_k . For any given environmental phase ℓ , the i th component of the vector $\mathbf{v}(\ell)$ is the long-term

fertility of one individual of type i . The left-eigenvector $\mathbf{u}(\ell)$ is the asymptotic distribution of types in the population, as time goes to infinity.

The matrix Θ_k may be written as

$$\Theta_k = \Delta_k^{-1} e^{(\mathbf{\Omega}(\hat{X}_k) - \lambda_k \mathbf{I}) \xi_k} \Delta_k = e^{\mathbf{G}(\hat{X}_k) \xi_k}$$

where $\mathbf{G}(\ell) = \text{diag}(\mathbf{v}(\ell))^{-1} (\mathbf{\Omega}(\ell) - \lambda(\ell) \mathbf{I}) \text{diag}(\mathbf{v}(\ell))$, for $\ell = 1, \dots, m$.

The matrix $\mathbf{G}(\ell)$ is the generator of the retrospective mutation chain (see Georgii and Baake [6]) in the environmental state ℓ : this process describes the change of type (or mutation) along an ancestral line in the family tree. If we denote by $\boldsymbol{\sigma}(\ell)$ the stationary probability vector of $\mathbf{G}(\ell)$, we have $\boldsymbol{\sigma}(\ell) = \mathbf{u}(\ell) \text{diag}(\mathbf{v}(\ell))$: $(\boldsymbol{\sigma}(\ell))_i$ is the asymptotic proportion of ancestors of type i for an individual chosen in the surviving population with the distribution $\mathbf{u}(\ell)$.

Thus, while $\mathbf{\Omega}(\ell)$ characterizes the future of the branching process, $\mathbf{G}(\ell)$ gives a picture of the family tree biased to represent the history of the survivors, and Θ_k is the transition matrix of the mutation along an ancestral line after a sojourn of length ξ_k in the environment \hat{X}_k .

In a similar manner, $\bar{\Theta}_k = e^{\bar{\mathbf{G}}(\hat{X}_k) \xi_k}$, where

$$\begin{aligned} \bar{\mathbf{G}}(\ell) &= \text{diag}(\mathbf{u}(\ell))^{-1} (\mathbf{\Omega}(\ell) - \lambda(\ell) \mathbf{I})^\top \text{diag}(\mathbf{u}(\ell)) \\ &= \text{diag}(\boldsymbol{\sigma}(\ell)) \mathbf{G}(\ell)^\top \text{diag}(\boldsymbol{\sigma}(\ell)) \end{aligned}$$

is the dual of $\mathbf{G}(\ell)$ with respect to the function $\text{diag}(\boldsymbol{\sigma}(\ell))$ (see Jansen and Kurt [9]). It is the generator of the Markov chain which results from choosing an individual in the asymptotic surviving population and following its line of descent backward in time.

We may rewrite the product $\mathcal{Z}_{n,j}$ of (10) as

$$\mathcal{Z}_{n,j} = (\mathbf{v}_0)_{\varphi_0} \frac{(\mathbf{v}_1)_{\varphi_1}}{(\mathbf{v}_0)_{\varphi_1}} \frac{(\mathbf{v}_2)_{\varphi_2}}{(\mathbf{v}_1)_{\varphi_2}} \dots \frac{(\mathbf{v}_{n-1})_{\varphi_{n-1}}}{(\mathbf{v}_{n-2})_{\varphi_{n-1}}} \frac{1}{(\mathbf{v}_{n-1})_{\varphi_n}} \mathbb{1}_{\{\varphi_n=j\}}$$

where each factor $(\mathbf{v}_k)_{\varphi_k} / (\mathbf{v}_{k-1})_{\varphi_k}$ represents the change of fertility, along an ancestral line, when the environment changes from \hat{X}_{k-1} to \hat{X}_k . Therefore, $\mathbb{E}[\mathcal{Z}_{n,j} | \varphi_0 = i, \boldsymbol{\xi}^{(n)}, \hat{\mathbf{X}}^{(n)}]$ is the expected growth of fertility due to environmental changes, on ancestral lines, and (11) gives us a breakdown of ω as the sum of two terms:

- $\mathbb{E}[\log \rho_0]$ is the expected future growth rate of the tree generated by an individual,
- Ψ is the biased growth rate along actual lines of descents.

Finally, it is worth observing that the single individual Markov chain is related as follows to the asymptotic mean population size in the branching process. For each environmental state ℓ , the mean population size matrix has an exponential asymptote,

$$e^{\mathbf{\Omega}(\ell)t} \sim e^{\lambda(\ell)t} \mathbf{v}(\ell) \mathbf{u}(\ell)$$

for large t , which follows from the PF theory. We now drop the dependency in ℓ in the notation for the clarity of the presentation. In order to characterize the rate at which $e^{\mathbf{\Omega}t}$ converges to its asymptote, we use the definition of the single individual Markov chain, and write

$$\frac{(e^{\mathbf{\Omega}t})_{ij}}{(e^{\lambda t} \mathbf{v} \mathbf{u})_{ij}} = \frac{[\text{diag}(\mathbf{v}) e^{\mathbf{G}t} \text{diag}(\mathbf{v})^{-1}]_{ij}}{(\mathbf{v} \mathbf{u})_{ij}} = \frac{v_i (e^{\mathbf{G}t})_{ij} v_j^{-1}}{v_i u_j} = \frac{(e^{\mathbf{G}t})_{ij}}{\sigma_j},$$

or similarly,

$$\log(e^{\mathbf{\Omega}t})_{ij} = \lambda t + \log(v_i u_j) + [\log(e^{\mathbf{G}t})_{ij} - \log \sigma_j].$$

So the rate at which the expected population size of the MBP converges to its asymptote $e^{\lambda t} \mathbf{v} \mathbf{u}$ is the same as the rate at which the single individual Markov chain converges to its asymptotic distribution $\boldsymbol{\sigma}$. This provides another relation between an MBP and its associated single individual Markov chain.

7 Numerical examples

The lower bounds ω_L and $\tilde{\omega}_L$ are obtained from different transformations, and numerical experimentation has shown that they may indeed be very different, without one being generally closer to ω . This is illustrated in two examples.

We have simulated 2000 paths of the random environment, $\{(\hat{X}_k, \xi_k), 0 \leq k \leq N-1\}$ with $N = 2048$, and we have computed

$$\omega_n = \frac{1}{n} \log \|e^{\mathbf{\Omega}(\hat{X}_0)\xi_0} e^{\mathbf{\Omega}(\hat{X}_1)\xi_1} \dots e^{\mathbf{\Omega}(\hat{X}_{n-1})\xi_{n-1}}\|,$$

for $n = 1, \dots, N$. On Figures 1 and 2 we plot the Cesaro averages of the sequence of means of ω_n over the 2000 simulations; in the various tables, ω_{sim} is the sample mean of ω_N .

Example 1 (A two-state random environment) In our first example, $m = 2$ and the generator of $\{X(t)\}$ is

$$\mathbf{Q} = \begin{bmatrix} -5 & 5 \\ 2 & -2 \end{bmatrix}.$$

Its stationary distribution is $\boldsymbol{\pi} = [0.2857, 0.7143]$. We take $r = 2$ and

$$\mathbf{\Omega}(1) = \begin{bmatrix} -15 & 12 \\ 9 & -29 \end{bmatrix}, \quad \mathbf{\Omega}(2) = \begin{bmatrix} -13 & 16 \\ 23 & -12 \end{bmatrix}. \quad (25)$$

The dominant eigenvalues of $\mathbf{\Omega}(1)$ and $\mathbf{\Omega}(2)$ are $\lambda(1) = -9.47$ and $\lambda(2) = 6.69$ respectively, so the branching process is subcritical in state 1 and supercritical in state 2. As $\pi_2 > \pi_1$, we expect the whole process to be supercritical and this is confirmed by our results, summarised in the table below and in Figure 1.

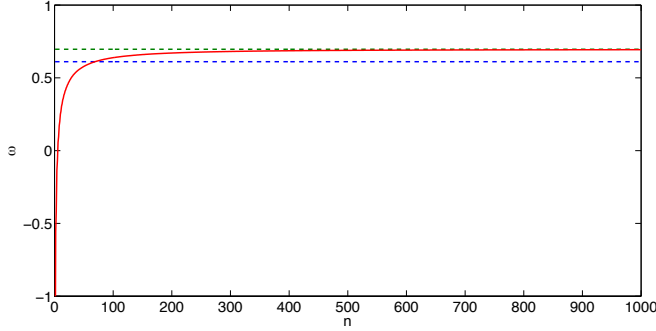


Fig. 1 Plain line: sample means of ω_n (Cesaro average) for successive values of n ; lower dashed line: ω_L ; upper dashed line: ω_U .

ω_L	$\tilde{\omega}_L$	ω_{sim}	ω_U
0.6107	0.6867	0.6933	0.6964

The simulation results are presented in Figure 1: the plain line is the average, over 2000 simulated paths, of the right-hand side of (4), one does see its convergence as n increases. The upper and lower dashed lines are for ω_U and ω_L . Clearly, ω is positive. For this example, ω_L is not very good and $\tilde{\omega}_L$ gives a better lower bound.

Example 2 (A three-state random environment)

We add a third state to the random environment and we assume the following generator

$$\mathbf{Q} = \begin{bmatrix} -4 & 2 & 2 \\ 1 & -1 & 0 \\ 2 & 4 & -6 \end{bmatrix},$$

with stationary distribution $\pi = [0.2143, 0.7143, 0.0714]$. The matrices $\mathbf{\Omega}(1)$ and $\mathbf{\Omega}(2)$ are the same as in Example 1, and the third is

$$\mathbf{\Omega}(3) = \begin{bmatrix} -39 & 12 \\ 3 & -17 \end{bmatrix} \quad (26)$$

with $\lambda(3) = -15.47$, so the branching process is very subcritical in the new environmental state. In this case, we obtain the following values.

$\tilde{\omega}_L$	ω_L	ω_{sim}	ω_U
0.5725	0.6442	0.7181	0.7475

We represent the bounds and the simulation in Figure 2. For this example, the better lower bound is ω_L .

The two examples have been chosen so that the environment spends asymptotically the same amount of time in the supercritical state 2 (the

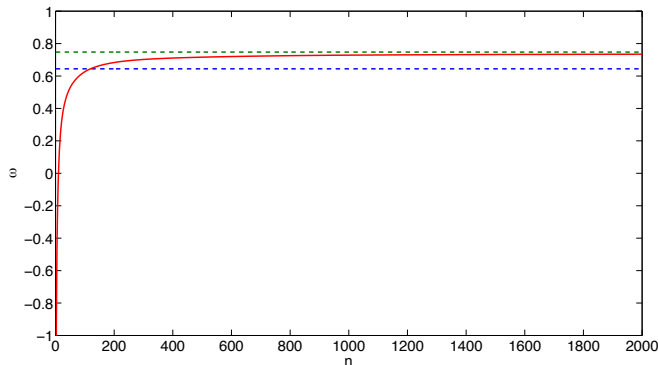


Fig. 2 Plain line: sample means of ω_n (Cesaro average) for successive values of n ; lower dashed line: ω_L ; upper dashed line: ω_U .

stationary probability is $\pi_2 = 0.7143$ in both cases). In addition, state 3 is more subcritical than state 1. Nevertheless, ω is greater, that is, the branching process is overall more supercritical, in the second example.

It is worth mentioning that we have generally observed in our experimentation that the upper bound is more accurate than the lower bounds.

8 Adding less information

Like in Section 4, we assume that the eigenvector \mathbf{v}_k depends on \hat{X}_k only. The lower bound ω_L is obtained by adding information in the form of *all* successive states of the single individual Markov chain $\{\varphi_k\}$. In general, it seems that ω_L is not a very close bound for ω , not as close as ω_U at any rate, and neither is $\tilde{\omega}_L$. We attempt, therefore, to obtain a better lower bound by adding less information. This is doable in the special case of a cyclic environmental process.

Assume that the jump chain $\{\hat{X}_k\}$ follows a deterministic cyclic path, and assume, without loss of generality, that $\hat{X}_k = (k \bmod m) + 1$. Observe that in this case, the matrix product in the limit (4) can be decomposed into i.i.d. blocks of length m ; this does not simplify the evaluation of the limit, however, since the blocks can take infinitely many values, and Lyapunov exponents are known to be hard to compute even in the general i.i.d. case [19]. Now, (12) is equivalent to

$$\begin{aligned} \Psi &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}[\mathcal{Z}_{n,j} | \varphi_0 = i, \boldsymbol{\xi}^{(n)}] \quad \text{a.s.} \\ &= \lim_{n \rightarrow \infty} \frac{1}{nm} \log \mathbb{E}_\rho[\mathbb{E}[\mathcal{Z}_{nm,j} | \varphi_0 = i, \boldsymbol{\xi}^{(nm)}, \boldsymbol{\rho}^{(n)}]] \quad \text{a.s.,} \end{aligned} \quad (27)$$

where the outermost expectation is with respect to $\boldsymbol{\rho}^{(n)}$ defined as $\boldsymbol{\rho}^{(n)} = (\varphi_m, \varphi_{2m}, \dots, \varphi_{(n-1)m})$. That is, in contrast with Theorem 3, we do not condition on the whole sequence of states of the single individual Markov chain, but only at the beginning of cycles for $\{\hat{X}_k\}$.

We redefine as follows the products $\mathcal{Z}_{n,j}$ in (9, 10):

$$\mathcal{Z}_{nm,j} = (\mathbf{v}(m))_{\varphi_0} R_0 R_1 \cdots R_{n-1} (\mathbf{v}(m))_{\varphi_{nm}}^{-1} \mathbb{1}_{\{\varphi_{nm}=j\}},$$

where

$$R_k = \frac{(\mathbf{v}(1))_{\varphi_{km}} (\mathbf{v}(2))_{\varphi_{km+1}} \cdots (\mathbf{v}(m))_{\varphi_{km+m-1}}}{(\mathbf{v}(m))_{\varphi_{km}} (\mathbf{v}(1))_{\varphi_{km+1}} \cdots (\mathbf{v}(m-1))_{\varphi_{km+m-1}}}, \quad (28)$$

and we use the fact that $(R_0, R_1, \dots, R_{k-1})$ is conditionally independent of (R_k, R_{k+1}, \dots) , given φ_{km} . Thus,

$$\begin{aligned} \mathbb{E}[\mathcal{Z}_{nm,j} | \varphi_0 = i, \boldsymbol{\xi}^{(nm)}, \boldsymbol{\rho}^{(n)}] \\ = (\mathbf{v}(m))_i \prod_{k=0}^{n-2} \mathbb{E}[R_k | \varphi_0 = i, \boldsymbol{\xi}^{(nm)}, \boldsymbol{\rho}^{(n)}] \cdot \\ \mathbb{E}[R_{n-1} (\mathbf{v}(m))_{\varphi_{nm}}^{-1} \mathbb{1}_{\{\varphi_{nm}=j\}} | \varphi_0 = i, \boldsymbol{\xi}^{(nm)}, \boldsymbol{\rho}^{(n)}] \end{aligned}$$

which we rewrite as

$$\begin{aligned} = (\mathbf{v}(m))_i \mathbb{E}[R_0 | \varphi_0 = i, \boldsymbol{\zeta}_0, \varphi_m] \prod_{k=1}^{n-2} \mathbb{E}[R_k | \varphi_{km}, \varphi_{(k+1)m}, \boldsymbol{\zeta}_k] \cdot \\ \mathbb{E}[R_{n-1} (\mathbf{v}(m))_{\varphi_{nm}}^{-1} \mathbb{1}_{\{\varphi_{nm}=j\}} | \varphi_{(n-1)m}, \boldsymbol{\zeta}_{n-1}] \end{aligned}$$

with $\boldsymbol{\zeta}_k = (\xi_{km}, \xi_{km+1}, \dots, \xi_{km+m-1})$

$$= f \prod_{k=1}^{n-2} \mathbb{E}[R_k | \varphi_{km}, \varphi_{(k+1)m}, \boldsymbol{\zeta}_k], \quad (29)$$

where we collect in

$$f = (\mathbf{v}(m))_i \mathbb{E}[R_0 | \varphi_0 = i, \boldsymbol{\zeta}_0, \varphi_m] \mathbb{E}[R_{n-1} (\mathbf{v}(m))_{\varphi_{nm}}^{-1} \mathbb{1}_{\{\varphi_{nm}=j\}} | \varphi_{(n-1)m}, \boldsymbol{\zeta}_{n-1}]$$

all the factors which play no role in the limit as $n \rightarrow \infty$. Observe that $\boldsymbol{\zeta}_0, \boldsymbol{\zeta}_1, \dots$ are independent and identically distributed random m -tuples, with the same distribution as $\boldsymbol{\zeta}^* = (\xi_1^*, \xi_2^*, \dots, \xi_m^*)$, where $\xi_1^*, \xi_2^*, \dots, \xi_m^*$ are independent, exponentially distributed random variables, respectively with

parameters c_1, c_2, \dots, c_m . Thus, from (27, 29),

$$\begin{aligned} e^\Psi &= \lim_{n \rightarrow \infty} \left\{ \mathbb{E}_\rho \left[f \prod_{k=1}^{n-2} \mathbb{E}[R_k | \varphi_{km}, \varphi_{(k+1)m}, \zeta_k] \right] \right\}^{1/nm} \\ &\geq \mathbb{E}_\rho \left[\lim_{n \rightarrow \infty} \left\{ f \prod_{k=1}^{n-2} \mathbb{E}[R_k | \varphi_{km}, \varphi_{(k+1)m}, \zeta_k] \right\}^{1/nm} \right] \end{aligned} \quad (30)$$

$$= \mathbb{E}_\rho \left[\lim_{n \rightarrow \infty} \left\{ \prod_{k=1}^{n-2} \mathbb{E}[R_k | \varphi_{km}, \varphi_{(k+1)m}, \zeta_k] \right\}^{1/nm} \right], \quad (31)$$

where (30) is justified by the Dominated Convergence Theorem followed by Jensen's inequality. We may now prove the following property.

Theorem 5 *A lower bound for the Lyapunov exponent ω is given by*

$$\omega_L^* = \mathbb{E}[\log \rho_0] + \frac{1}{m} \sum_{1 \leq i, j \leq r} \beta_{ij} \mathbb{E}[\log \mathbb{E}[R_0 | \varphi_0 = i, \varphi_m = j, \zeta^*]], \quad (32)$$

where

$$\beta_{ij} = \alpha_i (\mathbf{N}_1 \mathbf{N}_2 \cdots \mathbf{N}_m)_{ij}, \quad (33)$$

the matrices \mathbf{N}_ℓ are defined in (19) and α is the stationary vector of their product:

$$\alpha \mathbf{N}_1 \mathbf{N}_2 \cdots \mathbf{N}_m = \alpha, \quad \alpha \mathbf{1} = 1. \quad (34)$$

Proof We reorganise the product in (31) and group together the factors with equal values for φ_{km} and $\varphi_{(k+1)m}$, obtaining that

$$e^\Psi \geq \mathbb{E}_\rho \left[\lim_{n \rightarrow \infty} \left(\prod_{1 \leq i, j \leq r} \prod_{k=1}^{n_{ij}} R_{k;i,j} \right)^{1/nm} \right] \quad (35)$$

where $n_{ij} = \sum_{k=1}^{n-2} \mathbb{1}_{\{\varphi_{mk}=i, \varphi_{m(k+1)}=j\}}$ and, for fixed i and j , $\{R_{1;i,j}, R_{2;i,j}, \dots\}$ are i.i.d. random variables with the same distribution as $\mathbb{E}[R_0 | \varphi_0 = i, \varphi_m = j, \zeta^*]$.

By the Strong Law of Large Numbers, the limits $\beta_{ij} = \lim_{n \rightarrow \infty} n_{ij}/n$ exist and

$$\begin{aligned} \beta_{ij} &= \lim_{k \rightarrow \infty} \mathbb{P}[\varphi_{km} = i, \varphi_{(k+1)m} = j] \\ &= \lim_{k \rightarrow \infty} \mathbb{P}[\varphi_{km} = i] \mathbb{P}[\varphi_m = j | \varphi_0 = i] \\ &= \alpha_i (\mathbf{N}_1 \mathbf{N}_2 \cdots \mathbf{N}_m)_{ij}. \end{aligned}$$

Further, by the Strong Law of Large Numbers again, for fixed i and j ,

$$\lim_{n \rightarrow \infty} \left(\prod_{k=1}^{n_{ij}} R_{k;i,j} \right)^{1/n} = \exp(\beta_{ij} \mathbb{E}[\log \mathbb{E}[R_0 | \varphi_0 = i, \varphi_m = j, \zeta^*]])$$

so that the limit in (35) is independent of $\boldsymbol{\rho}^{(n)}$ and the inequality becomes $\Psi \geq \Psi^*$, where

$$\Psi^* = \frac{1}{m} \sum_{1 \leq i, j \leq r} \beta_{ij} \mathbb{E}[\log \mathbb{E}[R_0 | \varphi_0 = i, \varphi_m = j, \zeta^*]],$$

which proves that (32) is a lower bound for ω .

Now,

$$\Psi^* \geq \frac{1}{m} \sum_{1 \leq i, j \leq r} \beta_{ij} \mathbb{E}[\mathbb{E}[\log R_0 | \varphi_0 = i, \varphi_m = j, \zeta^*]]$$

by Jensen's inequality,

$$\begin{aligned} &= \frac{1}{m} \sum_{1 \leq i, j \leq r} \beta_{ij} \mathbb{E}[\log R_0 | \varphi_0 = i, \varphi_m = j] \\ &= \frac{1}{m} \sum_{1 \leq i \leq r} \alpha_i \mathbb{E}[\log R_0 | \varphi_0 = i]. \end{aligned} \quad (36)$$

From (28) we find that

$$\begin{aligned} &\mathbb{E}[\log R_0 | \varphi_0 = i] \\ &= \mathbb{E} \left[\log(\mathbf{v}(1))_i + \sum_{\ell=2}^m \log(\mathbf{v}(\ell))_{\varphi_{\ell-1}} - \log(\mathbf{v}(m))_i - \sum_{\ell=1}^{m-1} \log(\mathbf{v}(\ell))_{\varphi_\ell} \mid \varphi_0 = i \right] \\ &= \left(\sum_{\ell=1}^m \left(\prod_{t=1}^{\ell-1} \mathbf{N}_t \right) \log \mathbf{v}(\ell) \right)_i - \log(\mathbf{v}(m))_i - \left(\sum_{\ell=1}^{m-1} \left(\prod_{t=1}^{\ell} \mathbf{N}_t \right) \log \mathbf{v}(\ell) \right)_i \\ &= \left(\sum_{\ell=1}^m \left(\prod_{t=1}^{\ell-1} \mathbf{N}_t \right) (\mathbf{I} - \mathbf{N}_\ell) \log \mathbf{v}(\ell) \right)_i \end{aligned}$$

and so, by (36),

$$\Psi^* \geq \frac{1}{m} \boldsymbol{\alpha} \left(\sum_{\ell=1}^m \left(\prod_{t=1}^{\ell-1} \mathbf{N}_t \right) (\mathbf{I} - \mathbf{N}_\ell) \log \mathbf{v}(\ell) \right). \quad (37)$$

Finally, using the cyclic structure of $\hat{\mathbf{P}}$, one shows that the expression $\boldsymbol{\pi}^{(1)}(\mathbf{I} - \mathbf{N}) \log \bar{\mathbf{v}}$ defined in Theorem 3 is identical to the right-hand side of (37). Indeed, since $\hat{X}_k = (k \bmod m) + 1$, we have

$$\mathbf{P}^{(1)} = \begin{bmatrix} \mathbf{N}_1 & & & \\ & \mathbf{N}_2 & & \\ & & \ddots & \\ & & & \mathbf{N}_m \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{I} \\ & \ddots & \mathbf{I} \\ \mathbf{I} & & & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{N}_1 & & \\ & \mathbf{0} & \mathbf{N}_2 & \\ & & \ddots & \mathbf{N}_{m-1} \\ \mathbf{N}_m & & & \mathbf{0} \end{bmatrix}.$$

We decompose $\boldsymbol{\pi}^{(1)}$ into subvectors and write $\boldsymbol{\pi}^{(1)} = [\boldsymbol{\pi}^{(1)}(1), \dots, \boldsymbol{\pi}^{(1)}(m)]$ so that the system $\boldsymbol{\pi}^{(1)} = \boldsymbol{\pi}^{(1)}\mathbf{P}^{(1)}$ becomes

$$\begin{aligned}\boldsymbol{\pi}^{(1)}(1) &= \boldsymbol{\pi}^{(1)}(m)\mathbf{N}_m \\ \boldsymbol{\pi}^{(1)}(\ell) &= \boldsymbol{\pi}^{(1)}(\ell-1)\mathbf{N}_{\ell-1} = \boldsymbol{\pi}^{(1)}(m)\mathbf{N}_m\mathbf{N}_1 \cdots \mathbf{N}_{\ell-1}\end{aligned}$$

for $\ell = 2, \dots, m$. From the last equation, we obtain

$$\boldsymbol{\pi}^{(1)}(m)\mathbf{N}_m = \boldsymbol{\pi}^{(1)}(m)\mathbf{N}_m \prod_{t=1}^m \mathbf{N}_t,$$

and we conclude from (34) that $\boldsymbol{\pi}^{(1)}(m)\mathbf{N}_m = c\boldsymbol{\alpha}$ for some scalar c . Hence,

$$\boldsymbol{\pi}^{(1)}(\ell) = c\boldsymbol{\alpha} \prod_{t=1}^{\ell-1} \mathbf{N}_t, \quad \text{for } \ell = 1, \dots, m.$$

Since $\boldsymbol{\pi}^{(1)}\mathbf{1} = 1$ and the matrices \mathbf{N}_t are stochastic for all t , we find that $c = 1/m$, and so

$$\begin{aligned}\boldsymbol{\pi}^{(1)}(\mathbf{I} - \mathbf{N}) \log \bar{\mathbf{v}} &= \sum_{\ell=1}^m \boldsymbol{\pi}^{(1)}(\ell)(\mathbf{I} - \mathbf{N}_\ell) \log \mathbf{v}(\ell) \\ &= \frac{1}{m} \boldsymbol{\alpha} \left(\sum_{\ell=1}^m \left(\prod_{t=1}^{\ell-1} \mathbf{N}_t \right) (\mathbf{I} - \mathbf{N}_\ell) \log \mathbf{v}(\ell) \right) \\ &\leq \Psi^* \quad \text{by (37)}.\end{aligned}$$

This proves that $\omega_L^* \geq \omega_L$. \square

Denote by A_{ij} , $1 \leq i, j \leq r$, the expectations in (32):

$$A_{ij} = \mathbb{E}[\log \mathbb{E}[R_0 | \varphi_0 = i, \varphi_m = j, \boldsymbol{\zeta}^*]].$$

These need to be determined for ω_L^* to be of practical use. One verifies from first principles that

$$\mathbb{E}[R_0 | \varphi_0, \varphi_m, \boldsymbol{\zeta}^*] = \left(\prod_{1 \leq \ell \leq m} \boldsymbol{\Delta}_\ell^* \boldsymbol{\Theta}(\ell, \boldsymbol{\xi}_\ell^*) \right)_{\varphi_0, \varphi_m} / \left(\prod_{1 \leq \ell \leq m} \boldsymbol{\Theta}(\ell, \boldsymbol{\xi}_\ell^*) \right)_{\varphi_0, \varphi_m}$$

where $\boldsymbol{\Delta}_1^* = \boldsymbol{\Delta}(1)\boldsymbol{\Delta}(m)^{-1}$ and $\boldsymbol{\Delta}_\ell^* = \boldsymbol{\Delta}(\ell)\boldsymbol{\Delta}(\ell-1)^{-1}$, $\ell = 2, \dots, m$, so that

$$A_{ij} = \mathbb{E} \left[\log \left(\prod_{1 \leq \ell \leq m} \boldsymbol{\Delta}_\ell^* \boldsymbol{\Theta}(\ell, \boldsymbol{\xi}_\ell^*) \right)_{ij} \right] - \mathbb{E} \left[\log \left(\prod_{1 \leq \ell \leq m} \boldsymbol{\Theta}(\ell, \boldsymbol{\xi}_\ell^*) \right)_{ij} \right].$$

Although we do not have an explicit form for the expectations above, estimations by simulation are easily obtained, and this is what we did to compare ω_L^* and ω_L in the two examples below.

Example 3 (Continuation of Example 1) The random environment of Example 1 has only two states and is automatically cyclic. The new lower bound is $\omega_L^* = 0.6577$ and is indeed larger than $\omega_L = 0.6107$.

Example 4 (A three-state cyclic random environment) This is similar to Example 2: it is a three-state random environment with matrices $\mathbf{\Omega}(\ell)$ defined in (25, 26); the cyclic generator is

$$\mathbf{Q} = \begin{bmatrix} -4 & 4 & 0 \\ 0 & -1 & 1 \\ 6 & 0 & -6 \end{bmatrix},$$

and the stationary distribution is $\boldsymbol{\pi} = [0.1765, 0.7059, 0.1176]$. The bounds and the approximation are given in the table below

$\tilde{\omega}_L$	ω_L	ω_L^*	ω_{sim}	ω_U
0.1494	0.3480	0.3906	0.4218	0.4250

We see that in both examples, the difference $\omega_{sim} - \omega_L^*$ is less than half the difference $\omega_{sim} - \omega_L$. Needless to say, we might apply the same procedure to the dual of Section 5, thereby obtaining another lower bound, closer to ω than $\tilde{\omega}_L$. In the same manner as there is no systematic difference between $\tilde{\omega}_L$ and ω_L , we do not expect that there would be a systematic preference between this additional bound and ω_L^* .

9 Tightness of the bounds

The bounds are tight in that there exist branching processes for which the bounds are all equal and equal to ω . We show below that such is the case for branching processes with $r = 1$, and for processes such that the matrices \mathbf{A}_k commute for all $k \geq 0$.

Lemma 2 *If $r = 1$, then $\omega_L = \tilde{\omega}_L = \omega_L^* = \omega_U = \omega = \mathbb{E}[\log \rho_0]$.*

Proof If $r = 1$, all matrices reduce to scalars, so that $\mathbf{A}_k = \rho_k$ and $\mathbf{A}_k^* = 1$ a.s. for all $k \geq 0$. Therefore, $\Psi = 0$ by (6) and $\omega = \mathbb{E}[\log \rho_0]$ by (11).

Furthermore, $\mathbf{M}_\ell = 1$ for all ℓ , so that $\mathbf{M}(\hat{\mathbf{P}} \otimes I) = \hat{\mathbf{P}}$ is a stochastic matrix and $\log \text{sp}[\mathbf{M}(\hat{\mathbf{P}} \otimes I)] = 0$. We have thus proved that $\omega_U = \omega$.

Finally, $\Theta_k = 1$ a.s. for all $k \geq 0$, so that $\mathbf{N}_\ell = 1$, for all ℓ . Therefore, $\mathbf{I} - \mathbf{N} = 0$ and we conclude that $\omega_L = \omega$. The same argument gives us $\tilde{\omega}_L = \omega$. Since $\omega_L \leq \omega_L^* \leq \omega$, this shows that $\omega_L^* = \omega$ as well. \square

Lemma 3 *If the matrices \mathbf{A}_k are mutually commutative for all $k \geq 0$, then $\omega_L = \tilde{\omega}_L = \omega_L^* = \omega_U = \omega = \mathbb{E}[\log \rho_0]$.*

Proof Once more, it is enough to prove that $\text{sp}[\mathbf{M}(\hat{\mathbf{P}} \otimes \mathbf{I})] = 1$ and $\boldsymbol{\pi}^{(1)}(\mathbf{I} - \mathbf{N}) = \mathbf{0}$. Note that since the matrices \mathbf{A}_k are mutually commutative for all k , they all share the same strictly positive PF right-eigenvector, that we denote by \mathbf{v} . Then the matrices \mathbf{M}_ℓ defined in (15) are such that $\mathbf{M}_\ell \mathbf{v} = \mathbf{v}$ for all ℓ . Therefore,

$$\mathbf{M}(\hat{\mathbf{P}} \otimes \mathbf{I})(\mathbf{1} \otimes \mathbf{v}) = \mathbf{M}(\mathbf{1} \otimes \mathbf{v}) = \mathbf{1} \otimes \mathbf{v},$$

and since $\mathbf{1} \otimes \mathbf{v} > \mathbf{0}$, this proves that $\text{sp}[\mathbf{M}(\hat{\mathbf{P}} \otimes \mathbf{I})] = 1$ by the Perron-Frobenius Theorem.

Second, the mutual commutativity of the matrices \mathbf{A}_k implies the mutual commutativity of the matrices $\boldsymbol{\Theta}_k$, and we denote by \mathbf{w} the common strictly positive left-eigenvector of the matrices $\boldsymbol{\Theta}_k$ associated to their dominant eigenvalue 1, normalized such that $\mathbf{w}\mathbf{1} = 1$. The matrices \mathbf{N}_ℓ defined in (19) are then such that $\mathbf{w}\mathbf{N}_\ell = \mathbf{w}$ for all ℓ , and so

$$(\hat{\boldsymbol{\pi}} \otimes \mathbf{w})\mathbf{N}(\hat{\mathbf{P}} \otimes \mathbf{I}) = (\hat{\boldsymbol{\pi}} \otimes \mathbf{w})(\hat{\mathbf{P}} \otimes \mathbf{I}) = \hat{\boldsymbol{\pi}} \otimes \mathbf{w}.$$

We conclude, on the one hand, that $\boldsymbol{\pi}^{(1)} = \hat{\boldsymbol{\pi}} \otimes \mathbf{w}$ and, on the other hand, that $\boldsymbol{\pi}^{(1)}(\mathbf{I} - \mathbf{N}) = \mathbf{0}$. \square

Acknowledgements

Our paper has benefited from interesting questions and comments from two anonymous referees. Sophie Hautphenne is supported by the Australian Research Council (ARC), grants FL130100039 and DE150101044. Guy Latouche acknowledges the support of the Ministère de la Communauté française de Belgique through the ARC grant AUWB-08/13-ULB 5.

References

1. K. Athreya and S. Karlin. On branching processes with random environments: I: Extinction probabilities. *Ann. Math. Statist.*, 5:1499–1520, 1971.
2. L. Arnold, V. M. Gundlach and L. Demetrius. Evolutionary formalism for products of positive random matrices. *The Annals of Applied Probability*, 859–091, 1994.
3. K. B. Athreya and P. E. Ney. *Branching Processes*. Springer-Verlag, New York, 1972.
4. E. Bolthausen and I. Goldscheid. Recurrence and transience of random walks in random environments on a strip. *Communications in Mathematical Physics*, 214(2):429–447, 2000.
5. H. Furstenberg and H. Kesten. Products of random matrices. *Ann. Math. Statist.*, 31:457–469, 1960.
6. H.-O. Georgii and E. Baake. Supercritical multitype branching processes: The ancestral types of typical individuals. *Adv. Appl. Prob.*, 35:1090–1110, 2003.
7. R. Gharavi and V. Anantharam. An upper bound for the largest Lyapunov exponent of a Markovian random matrix product of nonnegative matrices. *Theoretical Computer Science*, 332:543–557, 2005.

-
8. S. Hautphenne, G. Latouche, and G. T. Nguyen. Markovian trees subject to catastrophes: would they survive forever? In *Matrix-Analytic Methods in Stochastic Models*, pages 87–106. Springer, 2013.
 9. S. Jansen and N. Kurt. On the notion(s) of duality for Markov processes. *Probability Surveys*, 11:59–120, 2014. doi: 10.1214/12-PS206.
 10. E. Key. Computable examples of the maximal Lyapunov exponent. *Probab. Th. Rel. Fields*, 75:97–107, 1987.
 11. E. Key. Lower bounds for the maximal Lyapunov exponent. *J. Theoret. Probab.*, 3(3):477–488, 1987.
 12. J. F. C. Kingman. Subadditive ergodic theory. *The Annals of Probability*, 1:883–909, 1973.
 13. R. Lima and M. Rahibe. Exact Lyapunov exponent for infinite products of random matrices. *J. Phys. A: Math. Gen.*, 27:3427–3437, 1994.
 14. J. R. Norris. *Markov chains*. Cambridge university press, Cambridge, 2008.
 15. V. I. Oseledec. A multiplicative ergodic theorem: Lyapunov characteristic numbers for dynamical systems. *Trans. Moscow Math. Soc.*, 19:197–231, 1968.
 16. M. Pollicott. Maximal Lyapunov exponents for random matrix products. *Inventiones mathematicae*, 181.1:209–226, 2010.
 17. E. Seneta. *Non-negative matrices and Markov chains*. Springer Science & Business Media, 2006.
 18. D. Tanny. On multitype branching processes in a random environment. *Adv. Appl. Prob.*, 13:464–497, 1981.
 19. J. Tsitsiklis and V. Blondel. The Lyapunov exponent and joint spectral radius of pairs of matrices are hard — when not impossible — to compute and to approximate. *Math. Control Signals Systems*, 10:31–40, 1997.
 20. J. C. Watkins. Limit theorems for products of random matrices. In J. E. Cohen, H. Kesten, and C. M. Newman, editors, *Random Matrices and Their Applications*, volume 50 of *Contemporary Mathematics*, pages 5–22. American Mathematical Society, Providence, 1986.