ASYMPTOTIC EXPANSION OF THE INVARIANT MEASURE FOR MARKOV-MODULATED ODES AT HIGH FREQUENCY

PIERRE MONMARCHÉ ⓘ,* *Sorbonne Université*
EDOUARD STRICKLER ⓘ,** *Université de Lorraine, CNRS, Inria, IECL*

Abstract

We consider time-inhomogeneous ordinary differential equations (ODEs) whose parameters are governed by an underlying ergodic Markov process. When this underlying process is accelerated by a factor ε^{-1}, an averaging phenomenon occurs and the solution of the ODE converges to a deterministic ODE as ε vanishes. We are interested in cases where this averaged flow is globally attracted to a point. In that case, the equilibrium distribution of the solution of the ODE converges to a Dirac mass at this point. We prove an asymptotic expansion in terms of ε for this convergence, with a somewhat explicit formula for the first-order term. The results are applied in three contexts: linear Markov-modulated ODEs, randomized splitting schemes, and Lotka–Volterra models in a random environment. In particular, as a corollary, we prove the existence of two matrices whose convex combinations are all stable but are such that, for a suitable jump rate, the top Lyapunov exponent of a Markov-modulated linear ODE switching between these two matrices is positive.

Keywords: Averaging; Lyapunov exponent; piecewise deterministic Markov process; cooperative linear system

2020 Mathematics Subject Classification: Primary 60J27; 60H10
Secondary 37H15

1. Introduction

Let $\mathcal{M}$ be a d-dimensional compact $\mathcal{C}^\infty$ manifold. Markov-modulated ODEs on $\mathcal{M}$ are dynamical systems $(x(t))_{t \geqslant 0}$ that are solutions to

$$x'(t) = F_{\sigma(t)}(x(t)), \tag{1.1}$$

where $(\sigma(t))_{t \geqslant 0}$ is a Markov process on some space $\mathcal{S}$ and, for all $s \in \mathcal{S}$, F_s is a vector field on $\mathcal{M}$.

We are interested in the high-frequency regime, namely when $\sigma(t)$ is replaced in (1.1) by $\sigma(t/\varepsilon)$ for some small ε. Provided σ is ergodic with respect to some probability measure π, the modulated process $x(t)$ is known to converge, as ε vanishes, to the solution of the averaged ODE

$$x'(t) = \bar{F}(x(t)), \qquad \text{where } \bar{F}(x) = \int_{\mathcal{S}} F_s(x) \, \pi(\mathrm{d}s). \tag{1.2}$$

In the case where this averaged ODE has a unique global attractor $\bar{x}$, we expect the solution of (1.1) to be close to $\bar{x}$ for t large and ε small. In other words, the invariant measure of the Markov process $(x(t), \sigma(t/\varepsilon))_{t \geqslant 0}$ converges, as ε vanishes, to $\delta_{\bar{x}} \otimes \pi$. The goal of the present work is to provide an infinitesimal expansion in terms of ε for this invariant measure. This is obtained by combining a similar expansion for the law of the process for a fixed $t > 0$ (following [24]) with a long-time convergence result for the limit process. The structure of the proof follows [25], which established a similar expansion for the invariant measure of Euler–Maruyam schemes for diffusion processes (the averaging estimates for a fixed $t > 0$ being in that case replaced by finite-time discretization error expansions).

One of our main motivations is to get the first-order term in the expansion of the top Lyapunov exponent of systems of switched linear ODEs [5, 9, 17], as detailed in Section 3.1. Our result also has applications, for instance, in some population models [4, 6, 22] or random splitting numerical schemes [1], cf. Section 3. More generally, Markov-modulated ODEs form a flexible class of Markov processes which appear in a variety of models, often to describe systems evolving in a randomly fluctuating environment, as in finance [13], biology [18] or reliability [20]. It is related to random ODEs; see, e.g., [2, 10] and references within.

In the specific context of piecewise-deterministic Markov processes (namely, when σ is a Markov chain on a finite set), the question of fast averaging has been addressed in [14], where a large-deviation principle is proved; in [24], where the expansion of the law of the process at fixed time t is given; or in [8], where the convergence of the invariant measure of the Markov process $(x(t), \sigma(t/\varepsilon))_{t \geqslant 0}$ is proved. In the recent paper [16], the authors deal with the case where σ is not ergodic.

The convergence of the invariant measure of the process in the fast-switching regime has been studied in some situations where the limit flow has a unique attractive point or cycle and possibly several unstable equilibria; see, for instance, in [5, 9, 12]. The analysis is then much more intricate than in our case (simply to get a convergence, without any asymptotic expansion) as the convergence toward the stable element cannot be uniform over the space. Some controllability conditions in the vicinity of unstable points are then necessary to ensure switching Markov process does not get stuck there, leading to some estimates on the exit times from small balls around these points or to the construction of suitable Lyapunov functions. Refining these arguments to get an asymptotic expansion in such more general situations is outside the scope of the present work.

The paper is organized as follows. In the remainder of this introduction we introduce our general notation and assumptions. Our main result, Theorem 2.1, is stated and proved in Section 2. Section 3 is devoted to examples of applications. Finally, an appendix gathers the proofs of some intermediate results.

1.1. Notation and assumptions

To conclude this section, let us clarify our setup. Denote by $(\mathcal{Q}_t)_{t \geqslant 0}$ the semigroup associated with $(\sigma_t)_{t \geqslant 0}$, namely $\mathcal{Q}_t f(\sigma) = \mathbb{E}_\sigma (f(\sigma_t))$, and by Q its infinitesimal generator.

Denote by $\mathcal{A}$ the set of bounded measurable functions from $\mathcal{M} \times \mathcal{S}$ to $\mathbb{R}$ which are $\mathcal{C}^\infty$ in their first variable x and such that all their derivatives in x are bounded over $\mathcal{M} \times \mathcal{S}$, and which, moreover, are such that, for all $x \in \mathcal{M}$ and all multi-index α, $s \mapsto \partial_x^\alpha f(x, s)$ is in the domain of Q. We consider on $\mathcal{A}$ the norms $\|\|f\|\|_j = \sum_{|\alpha| \leqslant j} \|\partial_x^\alpha f\|_\infty$.

Assumption 1.1.

(i) *Almost surely, the ODE* (1.1) *is well-defined for all times, with values in* $\mathcal{M}$. *Moreover, the coordinates of* $(x, s) \mapsto F_s(x)$ *are in* $\mathcal{A}$.

(ii) *The semigroup* $(\mathcal{Q}_t)_{t \geqslant 0}$ *admits a unique invariant probability measure* π. *Moreover, there exist* $C, \gamma > 0$ *such that*

$$\|\mathcal{Q}_t f - \pi f\|_\infty \leqslant C \mathrm{e}^{-\gamma t} \|f\|_\infty \tag{1.3}$$

for all bounded measurable f on $\mathcal{S}$ *and all* $t \geqslant 0$.

The uniform exponential ergodicity (1.3) implies that the operator Q^{-1} given by

$$Q^{-1}f := \int_0^\infty \mathcal{Q}_t \left(\pi f - f\right) \, \mathrm{d}t \tag{1.4}$$

is well-defined for all $f \in \mathcal{A}$, with $Q^{-1}f \in \mathcal{A}$ and $\|\|Q^{-1}f\|\|_j \leqslant C/\gamma \|\|f\|\|_j$ for all $j \in \mathbb{N}$. Moreover, using that $\partial_t \mathcal{Q}_t f = Q \mathcal{Q}_t f = \mathcal{Q}_t Q f$ for all $t \geqslant 0$, we get that $Q Q^{-1}f = Q^{-1}Qf = f - \pi f$ for $f \in \mathcal{A}$, i.e. Q^{-1} is the pseudo-inverse of Q.

Example 1.1. A case of interest is given by $Qf = \pi f - f$, which corresponds to $\sigma(t) = Y_{N_t}$ where $(N_t)_{t \geqslant 0}$ is a standard Poisson process with intensity 1, $Y_0 = \sigma(0)$, and $(Y_k)_{k \geqslant 1}$ is an indpendent and identically distributed sequence of random variables distributed according to π. It is readily checked that, in that case, $\mathcal{Q}_t f = \mathrm{e}^{-t}f + (1 - \mathrm{e}^{-t})\pi f$, and then $Q^{-1} = Q$.

Example 1.2. We will be particularly concerned with the case when $\mathcal{S}$ is a finite state space and $(\sigma_t)_{t \geqslant 0}$ is an irreducible continuous-time Markov chain on $\mathcal{S}$, with transition rate matrix Q. In that case, Q^{-1} is the *group inverse* of Q, defined as the unique matrix solution X to

$$QXQ = Q, \qquad XQX = X, \qquad XQ = QX. \tag{1.5}$$

[Such a matrix exists and is unique since Q has index 1, meaning that Q^2 and Q have the same rank, because 0 is a simple eigenvalue of Q; see [3, Chapter 4] for more details.] In the particular case when the cardinality of $\mathcal{S}$ is two, Q can be written as

$$Q = \begin{pmatrix} -p & p \\ q & -q \end{pmatrix}$$

for some $p, q > 0$. Moreover, we can easily check, using $Q^2 = -(p+q)Q$, that $X = (1/(p+q)^2)Q$ satisfies (1.5), and therefore $Q^{-1} = (1/(p+q)^2)Q$.

2. High-frequency expansion

The process $(X(t), \sigma(t/\varepsilon))_{t \geqslant 0}$ is a Markov process on $\mathcal{M} \times \mathcal{S}$ with generator $L = L_c + (1/\varepsilon)Q$, where Q is seen as an operator on functions on $\mathcal{M} \times \mathcal{S}$ acting only on the second variable, and $L_c f(x, s) = F_s(x) \cdot \nabla_x f(x, s)$ for all $f \in \mathcal{A}$. For $t \geqslant 0$, denote by P_t^ε the associated Markov semigroup, namely $P_t^\varepsilon f(x, s) = \mathbb{E}_{x,s} \left(f(X(t), \sigma(t/\varepsilon))\right)$.

We start by stating an expansion in ε of P_t^ε for a fixed t (the proof is postponed to the Appendix). This is essentially the result (and proof) of [24] but written in a dual form and with more explicit functional settings. Moreover, [24] only considers the case where $\mathcal{S}$ is a finite set, but it allows the jump rates of σ to depend on x.

Proposition 2.1. *There exist two families of operators* $P_t^{(k)}$, $S_t^{(k)}$ *indexed by* $t \geqslant 0$, $k \in \mathbb{N}$ *and acting on* $\mathcal{A}$, *with the following properties:*

- *For all $k, j \in \mathbb{N}$ and $T > 0$, there exists $C > 0$ such that, for all $f \in \mathcal{A}$,*

$$\sup_{t \in [0,T]} \left\| \left| P_t^{(k)} f \right| \right\|_j \leqslant C \| f \|_{j+2k}. \tag{2.1}$$

- *For all $k, j \in \mathbb{N}$, there exist $C, \gamma > 0$ such that, for all $t \geqslant 0$ and all $f \in \mathcal{A}$,*

$$\left\| \left| S_t^{(k)} f \right| \right\|_j \leqslant C \mathrm{e}^{-\gamma t} \| f \|_{j+2k}. \tag{2.2}$$

- *For all $n \in \mathbb{N}$, $T \geqslant 0$, and $j \in \mathbb{N}$, there exists $C > 0$ such that, for all $\varepsilon > 0$ and $f \in \mathcal{A}$, the remainder $R_{n,t}^{\varepsilon} f$ defined by*

$$R_{n,t}^{\varepsilon} f = P_t^{\varepsilon} f - \sum_{k=0}^{n} \varepsilon^k P_t^{(k)} f - \sum_{k=0}^{n} \varepsilon^k S_{t/\varepsilon}^{(k)} f \tag{2.3}$$

satisfies

$$\| R_{n,t}^{\varepsilon} f \|_j \leqslant C \varepsilon^{n+1} \| f \|_{j+3+2n}. \tag{2.4}$$

- *The first operators are given by*

$$P_t^{(0)} f(x, s) = \pi f(\bar{\varphi}_t(x)), \tag{2.5}$$
$$S_t^{(0)} f(x, s) = \mathcal{Q}_t f(x, s) - \pi f(x),$$
$$P_t^{(1)} f(x, s) = \mathbf{b}_1(t, x, s) + \int_0^{+\infty} \pi(L_c S_r^{(0)} f)(x) \, \mathrm{d}r + \int_0^t \pi(L_c \mathbf{b}_1)(r, \bar{\varphi}_{t-r}(x)) \, \mathrm{d}r,$$

where

$$\mathbf{b}_1(t, x, s) = Q^{-1}(\partial_t P_t^{(0)} f - L_c P_t^{(0)} f)(x, s). \tag{2.6}$$

Next, to address the question of the long-time behaviour of the process, we work under the following condition.

Assumption 2.1. *There is a globally attractive $\bar{x} \in \mathcal{M}$ for the averaged flow $\bar{F}$ given in* (1.2) *and, for all $j \in \mathbb{N}$, there exist $C, a > 0$ such that, for all $t \geqslant 0$ and $f \in \mathcal{C}^{\infty}(\mathcal{M})$,*

$$\| f \circ \varphi_t - f(\bar{x}) \|_j \leqslant C \mathrm{e}^{-at} \| f \|_{j+1},$$

where φ is the flow associated with $\bar{F}$.

Under Assumption 2.1, $\mu_0 = \delta_{\bar{x}} \otimes \pi$ is the unique invariant measure of $P_t^{(0)}$ given in (2.5). Our main result is the following expansion in terms of ε of any invariant measure of the process $(P_t^{\varepsilon})_{t \geqslant 0}$, in the spirit of Talay–Tubaro expansions in terms of the step size for discretization schemes [25].

Theorem 2.1. *Under Assumptions 1.1 and 2.1, for all $f \in \mathcal{A}$ there exist real sequences $(c_k)_{k \geqslant 1}$ and $(M_k)_{k \geqslant 1}$ such that, for all $n \in \mathbb{N}$ and any invariant measure μ_{ε} of $(P_t^{\varepsilon})_{t \geqslant 0}$,*

$$\left| \mu_{\varepsilon} f - \mu_0 f - \sum_{k=1}^{n} c_k \varepsilon^k \right| \leqslant M_n \varepsilon^{n+1}. \tag{2.7}$$

Moreover,

$$c_1 = \pi L_c Q^{-1} \left(L_c h - f \right) (\bar{x}) \tag{2.8}$$

where, for $(x, s) \in \mathcal{M} \times \mathcal{S}$,

$$h(x, s) = \int_0^\infty \pi \left(f(\bar{x}) - f(\varphi_r(x)) \right) \, \mathrm{d}r. \tag{2.9}$$

Remark 2.1. In dimension 1, it is possible to get an alternative expression for the function h appearing in (2.8). Indeed, for fixed $x \in \mathbb{R}$, with the change of variables $y = \varphi_r(x)$ we obtain, since $\mathrm{d}\varphi_r(x)/\mathrm{d}r = \bar{F}(\varphi_r(x))$,

$$h(x, s) = \int_0^{+\infty} \left(\pi f(\bar{x}) - \pi f(\varphi_r(x)) \right) \, \mathrm{d}r = \int_{\bar{x}}^x \frac{\pi f(y) - \pi f(\bar{x})}{\bar{F}(y)} \, \mathrm{d}y.$$

We will use this expression in Section 3.3.

Proof of Theorem 2.1. The proof is divided into two parts. In the first one we prove (2.7), and we give an expression for c_1 which depends on an arbitrary time $t > 0$. In the second part, we obtain the stated expression for c_1 by letting t vanish in this first expression.

Step 1. Fix $t > 0, f \in \mathcal{A}$, and let μ_ε be an invariant measure of $(P_s^\varepsilon)_{s \geq 0}$. By definition of the flow φ_t,

$$(P_t^{(0)})^n f(x, i) = \pi f \left(\varphi_t^{\circ n}(x) \right) = \pi f \left(\varphi_{tn}(x) \right).$$

Hence, $\left\| \left(P_t^{(0)} \right)^n (\mu_0 f - f) \right\|_j = \| \pi f(\bar{x}) - \pi (f \circ \varphi_{tm}) \|_j \leq C e^{-atn} \| f \|_{j+1}$. As a consequence, the function Ψ_t given by $\Psi_t = \sum_{n=0}^\infty \left(P_t^{(0)} \right)^n (\mu_0 f - f)$ is well-defined in $\mathcal{A}$ and such that, for all $j \geq 0$, $\| \Psi_t \|_j \leq C_j \| f \|_{j+1}$ for some $C_j > 0$ independent of f. Moreover,

$$\left(P_t^{(0)} - \mathrm{Id} \right) \Psi_t = \sum_{n=1}^\infty \left(P_t^{(0)} \right)^n (\mu_0 f - f) - \sum_{n=0}^\infty \left(P_t^{(0)} \right)^n (\mu_0 f - f) = f - \mu_0 f.$$

Then, using this Poisson equation and that μ_ε is invariant for P_t^ε,

$$\mu_\varepsilon f - \mu_0 f = \mu_\varepsilon \left(P_t^{(0)} - \mathrm{Id} \right) \Psi_t = \mu_\varepsilon \left(P_t^{(0)} - P_t^\varepsilon \right) \Psi_t.$$

To get the convergence of $\mu_\varepsilon f$ to $\mu_0 f$ at zeroth order, we simply bound, thanks to Proposition 2.1,

$$|\mu_\varepsilon f - \mu_0 f| \leq \left\| \left(P_t^{(0)} - P_t^\varepsilon \right) \Psi_t \right\|_\infty \leq C\varepsilon \| \Psi_t \|_3.$$

To get the higher-order expansion, write

$$\mu_\varepsilon f - \mu_0 f = \mu_\varepsilon \left(P_t^{(0)} - P_t^\varepsilon \right) \Psi_t = -\mu_\varepsilon \left(\sum_{k=1}^n \varepsilon^k P_t^{(k)} + \sum_{k=0}^n \varepsilon^k S_{t/\varepsilon}^{(k)} + R_{n,t}^\varepsilon \right) \Psi_t. \tag{2.10}$$

From Proposition 2.1 and the bounds on $\| \Psi_t \|_j$ for all $j \in \mathbb{N}$,

$$\left| \mu_\varepsilon \left(\sum_{k=0}^n \varepsilon^k S_{t/\varepsilon}^{(k)} + R_{n,t}^\varepsilon \right) \Psi_t \right| \leq C_n \varepsilon^{n+1}$$

for some C_n. Using that, from the convergence at order 0, $\left|\mu_\varepsilon P_t^{(k)} \Psi_t - \mu_0 P_t^{(k)} \Psi_t\right| \leqslant C_{0,k}\varepsilon$ for all $k \in \mathbb{N}$ for some $C_{0,k} > 0$, we get first with (2.10) at order $n = 1$ that

$$\left|\mu_\varepsilon f - \mu_0 f + \varepsilon \mu_0 P_t^{(1)} \Psi_t\right| \leqslant C_{1,0}\varepsilon^2 \tag{2.11}$$

for some $C_{1,0} > 0$. We can thus apply this first-order expansion with f replaced by $P_t^{(k)}\Psi_t$ to get $\left|\mu_\varepsilon P_t^{(k)}\Psi_t - c_{0,k} - c_{1,k}\varepsilon\right| \leqslant C_{1,k}\varepsilon^2$ for some $c_{0,k}, c_{1,k}, C_{1,k}$. Hence, considering the expansion (2.10) at order $n = 2$, we get

$$|\mu_\varepsilon f - \mu_0 f - \varepsilon c_{0,1} - \varepsilon^2 c_{1,1} - \varepsilon^2 c_{0,2}| \leqslant C_{2,0}\varepsilon^3$$

for some $C_{2,0} > 0$. Again, this can be used with f replaced by $P_t^{(k)}\Psi_t$, and a straightforward induction on n concludes the proof of (2.7). In particular, from (2.11), we see that, for all $t > 0$, c_1 can be written as $c_1 = -\mu_0 P_t^{(1)}\Psi_t$.

Step 2. The goal is now to let t vanish in this last expression. For a fixed $t > 0$, denote by $\mathbf{b}_1^\Psi$ the function given by (2.6) but with f replaced by Ψ_t. Then

$$P_t^{(1)}\Psi_t(x, s) = \mathbf{b}_1^\Psi(t, x, s) + \alpha(x) + \beta(t, x),$$

with

$$\alpha(x) = \int_0^{+\infty} \pi\left(L_c S_r^{(0)} \Psi_t\right)(x) \, \mathrm{d}r, \qquad \beta(t, x) = \int_0^t \pi\left(L_c \mathbf{b}_1^\Psi\right)(r, \bar{\varphi}_{t-r}(x)) \, \mathrm{d}r.$$

First, by definition of Q^{-1}, the average of $\mathbf{b}_1^\Psi$ with respect to π, hence with respect to μ_0, is zero. At this stage, we have obtained that $c_1 = -\mu_0\alpha - \mu_0\beta$. Recalling that $S_r^{(0)}f = \mathcal{Q}_r(f - \pi f)$, from

$$\Psi_t(x, s) = \pi f(\bar{x}) - f(x, s) + \sum_{n=1}^\infty \pi\left(f(\bar{x}) - f(\varphi_{tn}(x))\right),$$

integrating with respect to π, we get

$$\nabla_x S_r^{(0)}\Psi_t(\bar{x}, s) = \nabla_x \mathcal{Q}_r\left(\Psi_t - \pi\Psi_t\right)(\bar{x}, s) = \nabla_x \mathcal{Q}_r(\pi f - f)(\bar{x}, s).$$

As a consequence,

$$\mu_0(\alpha) = \alpha(\bar{x}) = \int_0^\infty \pi L_c S_r^{(0)}\Psi_t(\bar{x}) \, \mathrm{d}r = \int_\mathcal{S} F_s(\bar{x}) \cdot \left[\int_0^\infty \nabla_x \mathcal{Q}_r(\pi f - f) \, \mathrm{d}r\right](\bar{x}, s)\, \pi(\mathrm{d}s)$$

$$= \int_\mathcal{S} F_s(\bar{x}) \cdot \nabla_x Q^{-1} f(\bar{x}, s)\, \pi(\mathrm{d}s)$$

$$= \pi L_c Q^{-1} f(\bar{x}).$$

Also,

$$\mu_0(\beta) = \int_0^t \pi L_c \mathbf{b}_1^\Psi(r, \bar{x}) \, \mathrm{d}r = \pi\left[L_c \int_0^t \mathbf{b}_1^\Psi(r, \cdot) \, \mathrm{d}r\right](\bar{x})$$

$$= \pi\left[L_c \int_0^t Q^{-1}\left(\partial_r P_r^{(0)}\Psi_t - L_c P_r^{(0)}\Psi_t\right) \, \mathrm{d}r\right](\bar{x})$$

$$= -\pi\left[L_c Q^{-1} \int_0^t L_c P_r^{(0)}\Psi_t \, \mathrm{d}r\right](\bar{x}),$$

where we used that $\partial_r P_r^{(0)}\Psi_t$ does not depend on s and is thus in the kernel of Q^{-1}. Hence,

$$\int_0^t \pi L_c \mathbf{b}_1^\Psi(s, \bar{x}) \, \mathrm{d}s = -\pi \left[L_c Q^{-1} \frac{1}{t} \int_0^t L_c P_r^{(0)}(t\Psi_t) \, \mathrm{d}r \right](\bar{x}) \xrightarrow[t \to 0]{} -\pi \left[L_c Q^{-1} L_c(\pi h) \right](\bar{x}),$$

where we used that $P_0^{(0)} f = \pi f$ and that $t\Psi_t \to h$. Indeed, thanks to Assumption 2.1,

$$t\Psi_t(x) = t\pi f(\bar{x}) - tf(x, s) + t \sum_{n=1}^\infty \pi \left(f(\bar{x}) - f(\varphi_{tn}(x)) \right)$$

$$\xrightarrow[t \to 0]{} \int_0^\infty \pi \left(f(\bar{x}) - f\left(\varphi_r(x) \right) \right) \, \mathrm{d}r = h(x),$$

and the convergence $t\Psi \to h$ holds in all norms $\|\!\|\cdot\|\!\|_j$. The proof of (2.8) is concluded by letting t vanish in the equality $c_1 = -\mu_0 \alpha - \mu_0 \beta$ (since c_1 is independent of t). □

3. Applications

3.1. Top Lyapunov exponent for cooperative linear Markov-modulated ODEs

In this section, we consider linear Markov-modulated ODEs on $\mathbb{R}^d$ of the form

$$z'(t) = A\left(\sigma(t/\varepsilon)\right) z(t) \tag{3.1}$$

where, for all $\sigma \in \mathcal{S}$, $A(\sigma)$ is a $d \times d$ matrix. We work under Assumption 1.1 and write $\bar{A} = \int_{\mathcal{S}} A(s) \pi(\mathrm{d}s)$. Moreover, we focus on the setup of [7], given as the following assumptions.

Assumption 3.1

(i) *The Markov process $(\sigma(t))_{t \geq 0}$ is Feller.*

(ii) *For all $s \in \mathcal{S}$, $A(s)$ is a cooperative matrix, in the sense that its off-diagonal coefficients are non-negative.*

(iii) *The averaged matrix $\bar{A}$ is irreducible in the sense that, for all $i, j \in [\![1, d]\!]$ there exists a path $i = i_0, \ldots, i_q = j$ with $\bar{A}_{i_{k-1}, i_k} > 0$ for all $k \in [\![1, q]\!]$.*

The fact that the matrices are cooperative implies that $\mathbb{R}_+^d$ is fixed by (3.1) for $t \geq 0$ and that Assumption 2.1 holds. We decompose solutions of (3.1) on $\mathbb{R}_+^d$ as $z(t) = \rho(t)\theta(t)$, where $\rho(t) = \mathbb{1} \cdot z(t) > 0$ with $\mathbb{1} = (1, \ldots, 1) \in \mathbb{R}^d$ and $\theta(t) = z(t)/\rho(t) \in \Delta = \{x \in \mathbb{R}_+^d, x_1 + \cdots + x_d = 1\}$. The ODE (3.1) is then equivalent to

$$\rho'(t) = (\mathbb{1} \cdot A\left(\sigma(t/\varepsilon)\right) \theta(t))\rho(t), \quad \theta'(t) = F_{\sigma(t/\varepsilon)}(\theta(t)), \quad \text{with } F_s(\theta) = A(s)\theta - (\mathbb{1} \cdot A(s)\theta)\theta.$$

It is proved in [7] that, under Assumption 3.1, the Markov process $(\theta(t), \sigma(t/\varepsilon))_{t \geq 0}$ admits a unique invariant measure μ_ε on $\Delta \times \mathcal{S}$ and that, for all initial conditions $z \in \mathbb{R}_+^d \setminus \{0\}$, almost surely, $\lim_{t \to \infty} \rho(t)/t = \Lambda_\varepsilon$, where $\Lambda_\varepsilon := \int_{\Delta \times \mathcal{S}} \mathbb{1} \cdot A(s)\theta \, \mu_\varepsilon(\mathrm{d}\theta, \mathrm{d}s)$ is called the top Lyapunov exponent of the process. Denoting by $\lambda_{\max}(\bar{A})$ the principal eigenvalue of $\bar{A}$ (i.e. with maximal real part), [7, Proposition 4] entails that $\lim_{\varepsilon \to 0} \Lambda_\varepsilon = \lambda_{\max}(\bar{A})$. From Theorem 2.1 applied to the function $f : (\theta, s) \mapsto \mathbb{1} \cdot A(s)\theta$, we get an expansion of Λ_ε for small ε.

Proposition 3.1. *Under Assumptions 1.1 and 3.1, there exists a sequence $(c_k)_{k \geq 1}$ of real numbers such that, for all $n \geq 1$,*

$$\Lambda_\varepsilon = \lambda_{\max}(\bar{A}) + \sum_{k=1}^n c_k \varepsilon^k + \underset{\varepsilon \to 0}{o}(\varepsilon^n).$$

Moreover, denoting by $\bar{x}$ and $\bar{y}$, respectively, the right and left eigenvectors of $\bar{A}$ associated with $\lambda_{\max}(\bar{A})$ and such that $\bar{x} \in \Delta$ and $\bar{x} \cdot \bar{y} = 1$,

$$c_1 = \int_{\mathcal{S}} \bar{y}^\top Q^{-1}(A)(s) \left(\bar{x}\bar{y}^\top - I \right) A(s)\bar{x}\, \pi(\mathrm{d}s).$$

Proof. First, considering h given by (2.9), we claim that, for all $(x, s) \in \Delta \times \mathcal{S}$,

$$h(x, s) = -\ln(\bar{y} \cdot x). \tag{3.2}$$

Indeed, note that, for all $x \in \mathcal{M}$,

$$\pi f(\varphi_r(x)) = \frac{\mathbb{1} \cdot \bar{A}\mathrm{e}^{r\bar{A}}x}{\mathbb{1}_d \cdot \mathrm{e}^{r\bar{A}}x} = \frac{\mathrm{d}}{\mathrm{d}r} \ln\left(\mathbb{1} \cdot \mathrm{e}^{r\bar{A}}x\right).$$

Therefore, using that $\bar{x} = \varphi_r(\bar{x})$,

$$\int_0^\infty (\pi f(\bar{x}) - \pi f(\varphi_r(x)))\, \mathrm{d}r = \int_0^\infty \frac{\mathrm{d}}{\mathrm{d}r} \left(\ln\left(\frac{\mathbb{1} \cdot \mathrm{e}^{r\bar{A}}\bar{x}}{\mathbb{1} \cdot \mathrm{e}^{r\bar{A}}x} \right) \right)\, \mathrm{d}r$$

$$= \lim_{r \to \infty} \ln\left(\frac{\mathbb{1} \cdot \mathrm{e}^{r\bar{A}}\bar{x}}{\mathbb{1} \cdot \mathrm{e}^{r\bar{A}}x} \right) = -\ln(\bar{y} \cdot x),$$

where we have used the Perron–Frobenius theorem. From (3.2), we deduce that $L_c h(x, s) = -F_s(x) \cdot (\bar{y}/(\bar{y} \cdot x))$. Now, since $F_s(x) = A(s)x - f_s(x)x$, we get

$$L_c h(x, s) - f(x, s) = -\frac{A(s)x \cdot \bar{y}}{\bar{y} \cdot x},$$

and then

$$Q^{-1}(L_c h - f)(x, s) = -\frac{Q^{-1}(A)(s)x \cdot \bar{y}}{\bar{y} \cdot x},$$

from which we deduce that

$$\nabla_x Q^{-1}(L_c h - f)(\bar{x}, s) = \frac{Q^{-1}(A)(s)\bar{x} \cdot \bar{y}}{(\bar{y} \cdot \bar{x})^2}\bar{y} - \frac{[Q^{-1}(A)(s)]^\top \bar{y}}{\bar{y} \cdot \bar{x}}$$

$$= (Q^{-1}(A)(s)\bar{x} \cdot \bar{y})\bar{y} - [Q^{-1}(A)(s)]^\top \bar{y},$$

where we have used that $\bar{x} \cdot \bar{y} = 1$. As a consequence,

$$F_s(\bar{x}) \cdot \nabla_x Q^{-1}(L_c h - f)(\bar{x}, s)$$

$$= (A(s)\bar{x} - (\mathbb{1} \cdot A(s)\bar{x})\bar{x}) \cdot \left((Q^{-1}(A)(s)\bar{x} \cdot \bar{y})\bar{y} - [Q^{-1}(A)(s)]^\top \bar{y} \right)$$

$$= (Q^{-1}(A)(s)\bar{x} \cdot \bar{y})A(s)\bar{x} \cdot \bar{y} - A(s)\bar{x} \cdot [Q^{-1}(A)(s)]^\top \bar{y}$$

$$- (\mathbb{1} \cdot A(s)\bar{x})(Q^{-1}(A)(s)\bar{x} \cdot \bar{y})\bar{x} \cdot \bar{y} + (\mathbb{1} \cdot A(s)\bar{x})\bar{x} \cdot [Q^{-1}(A)(s)]^\top \bar{y}.$$

Since $\bar{x} \cdot \bar{y} = 1$, we see that the last line is equal to

$$-(\mathbb{1} \cdot A(s)\bar{x})Q^{-1}(A)(s)\bar{x} \cdot \bar{y} + (\mathbb{1} \cdot A(s)\bar{x})Q^{-1}(A)(s)\bar{x} \cdot \bar{y} = 0.$$

We end up with

$$F_s(\bar{x}) \cdot \nabla_x Q^{-1}(L_c h - f)(\bar{x}, s) = (Q^{-1}(A)(s)\bar{x} \cdot \bar{y})A(s)\bar{x} \cdot \bar{y} - Q^{-1}(A)(s)A(s)\bar{x} \cdot \bar{y}$$
$$= \bar{y}^\top Q^{-1}(A)(s)\left[\bar{x}\bar{y}^\top - I\right]A(s)\bar{x}.$$

Integrating with respect to π leads to

$$c_1 = \int_{\mathcal{S}} \bar{y}^\top Q^{-1}(A)(s)\left(\bar{x}\bar{y}^\top - I\right)A(s)\bar{x}\,\pi(\mathrm{d}s). \qquad \square$$

Example 3.1. If σ is a Markov chain on $\mathcal{S} = \{0, 1\}$ with matrix rates given by

$$Q = \begin{pmatrix} -p & p \\ 1-p & -(1-p) \end{pmatrix}, \qquad (3.3)$$

we get the following simple expression for c_1:

$$c_1 = p(1-p)\left[\bar{y}^\top(A_0 - A_1)^2\bar{x} - (\bar{y}^\top(A_0 - A_1)\bar{x})^2\right]. \qquad (3.4)$$

Indeed, recalling that $Q^{-1} = Q$ (see Example 1.2) and using that $\pi_s Q_{s,s'} = p(1-p)$ if $s \neq s'$ and $-p(1-p)$ if $s = s'$, we get $c_1 = \bar{y}^\top(A_0 P A_0 + A_1 P A_1 - A_0 P A_1 - A_1 P A_0)\bar{x}$, where $P = I - \bar{x}\bar{y}^\top$. Now,

$$\bar{y}^\top A_0 \bar{x}\bar{y}^\top A_1 \bar{x} + \bar{y}^\top A_1 \bar{x}\bar{y}^\top A_0 \bar{x} - \bar{y}^\top A_0 \bar{x}\bar{y}^\top A_0 \bar{x} - \bar{y}^\top A_1 \bar{x}\bar{y}^\top A_1 \bar{x} = -(\bar{y}^\top(A_0 - A_1)\bar{x})^2$$

and $\bar{y}^\top A_0 A_1 \bar{x} + \bar{y}^\top A_1 A_0 \bar{x} - \bar{y}^\top A_0 A_0 \bar{x} - \bar{y}^\top A_1 A_1 \bar{x} = -\bar{y}^\top(A_0 - A_1)^2\bar{x}$, which induces (3.4).

When considering the switching between matrices $(A_i)_{i\in\mathcal{S}}$, the following natural question arises: if all the matrices are stable (i.e. $\lambda_{\max}(A_i) < 0$ for all $i \in \mathcal{S}$), is the switched system (3.1) stable, in the sense that $\Lambda_\varepsilon < 0$? It is now known that this is not true. In the case of random switching between two 2×2 matrices, examples of stable matrices giving an unstable system can be founded in [5] and [21]. In the first reference, for $p = \frac{1}{2}$, $\lambda_{\max}(\bar{A}) < 0$ so that fast switching leads to an unstable system. In the second reference it is a bit more complicated, since $\lambda_{\max}(\bar{A}) > 0$ for all $p \in (0, 1)$, meaning that both slow and fast switching lead to a stable system. However, [21] showed that if switching neither too fast nor too slowly, the system is unstable. Yet, in [21], the matrices are not cooperative in the sense of Assumption 3.1. This is not surprising, since it is proved in [17] that switching between cooperative matrices of size 2×2 such that every matrix in the convex hull of the given matrices is stable, will *always* lead to a stable system. However, it is also shown in [17] that it is possible in some higher dimensions to construct an example where all the matrices in the convex hull are stable, and for which there exists a periodic switching such that the linear system explodes. Later, an explicit example in dimension 3 was given in [15]. Precisely, consider the matrices

$$A_0 = \begin{pmatrix} -1 & 0 & 0 \\ 10 & -1 & 0 \\ 0 & 0 & -10 \end{pmatrix}, \qquad A_1 = \begin{pmatrix} -10 & 0 & 10 \\ 0 & -10 & 0 \\ 0 & 10 & -1 \end{pmatrix}. \qquad (3.5)$$

It is shown in [15] that every convex combination of A_0 and A_1 is stable, and yet a switch of period 1 between A_0 and A_1 yields an explosion. In [8], the authors asked the question whether the same system but with Markovian switching can lead to explosion. Using numerical simulations, they suggest that this is true and the Lyapunov exponent is positive for neither too fast nor too slow switching. Now, using (3.4), we rigorously prove the following assertion.
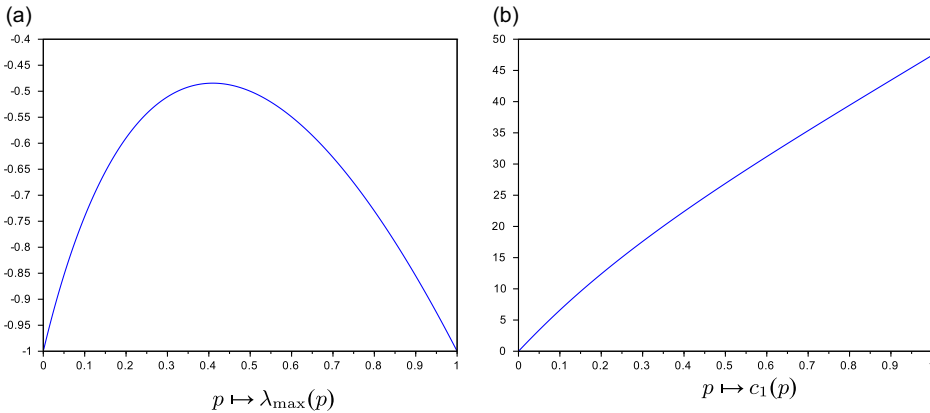
FIGURE 1. Plots of p from the proof of Proposition 3.2.

Proposition 3.2. *There exist two cooperative matrices B_0 and B_1 such that:*

- *For all $p \in [0, 1]$, $\lambda_{\max}((1 - p)B_0 + pB_1) < 0$.*

- *For some $\varepsilon > 0$ and $p \in (0, 1)$, the top Lyapunov exponent Λ_ε of the system (3.1) with σ generated by Q given in (3.3) is positive.*

Proof. To emphasize the dependency on p, we write $\bar{A}(p)$, $\Lambda_\varepsilon(p)$, and $c_1(p)$ for the mean matrix, the top Lyapunov exponent, and the first-order derivative, respectively, when the switching rates are given by the matrix Q in (3.3), and the matrices are (3.5). We also write $\lambda_{\max}(p)$ for $\lambda_{\max}(\bar{A}(p))$. Figure 1(a) shows $p \mapsto \lambda_{\max}(p)$, and we can see that the maximal value of this function is attained at some $p^* \in [0.3, 0.5]$ and is worth $\lambda^*_{\max} \in [-0.5, -0.45]$. We can see from Figure 1(b), which shows $p \mapsto c_1(p)$, that on the interval $[0.3, 0.5]$ we have $c_1(p) \geqslant 15$, so that, in particular, $c_1(p^*) \geqslant 15 > 0$.

Let $\delta > 0$ and $A_i^\delta = A_i + \delta I$. Replacing A_i by A_i^δ in (3.1), we can easily check that the Lyapunov exponent of the system is $\Lambda_\varepsilon^\delta(p) = \Lambda_\varepsilon(p) + \delta$. Moreover, $f_i^\delta(x) = \mathbb{1} \cdot (A_i + \delta I)x$, so that $F_i^\delta = F_i$, and the dynamic of the angular part of (3.1) is independent of δ. Hence, Proposition 3.1 entails that $\Lambda_\varepsilon^\delta(p) \geqslant \lambda_{\max}(p) + \delta + c_1(p)\varepsilon - M(p)\varepsilon^2$ for some constant $M(p)$ depending on p. Choosing $p = p^*$ and $\varepsilon = c_1(p^*)/2M(p^*) > 0$ yields

$$\Lambda_\varepsilon^\delta(p^*) \geqslant \lambda_{\max}(p^*) + \delta + \frac{c_1(p^*)^2}{4M(p^*)}.$$

Letting $\delta = -(c_1(p^*)^2/8M(p^*)) - \lambda_{\max}(p^*)$ concludes the proof, since

- $\Lambda_\varepsilon^\delta(p^*) \geqslant c_1^2(p^*)/8M(p^*) > 0$;

- for all $p \in [0, 1]$, $\lambda_{\max}(A_p^\delta) \leq \lambda_{\max}(A_{p^*}^\delta) = \lambda_{\max}(p^*) + \delta = -c_1^2(p^*)/8M(p^*) < 0$. □

Example 3.2. Assume as in Example 1.1 that $Qf = Q^{-1}f = \pi f - f$. Then we have $Q^{-1}(A)(s) = \bar{A} - A(s)$. Using $\bar{y}^\top \bar{A} = \lambda(\bar{A})\bar{y}^\top$ and $\bar{y}^\top \bar{x} = 1$, we get $\bar{y}^\top \bar{A}(\bar{x}\bar{y}^\top - I) = 0$, which yields $c_1 = \int_{\mathcal{S}} \bar{y}^\top A^2(s)\bar{x} - (\bar{y}^\top A(s)\bar{x})^2 \pi(\mathrm{d}s)$. Notice that, from Jensen's inequality,

$$\int_{\mathcal{S}} (\bar{y}^\top A(s)\bar{x})^2 \, \pi(\mathrm{d}s) \geqslant \left(\int_{\mathcal{S}} \bar{y}^\top A(s)\bar{x} \, \pi(\mathrm{d}s) \right)^2 = \lambda_{\max}(\bar{A})^2,$$

so that

$$c_1 \leqslant \int_{\mathcal{S}} \bar{y}^\top A^2(s)\bar{x} \, \pi(\mathrm{d}s) - \lambda_{\max}(\bar{A})^2 = \bar{y}^\top \overline{A^2}\bar{x} - \lambda_{\max}(\overline{A}^2) = \bar{y}^\top \left(\overline{A^2} - \overline{A}^2 \right)\bar{x}.$$

For further considerations and examples on the monotonicity of Lyapunov exponents in terms of the switching rate, we refer the interested reader to the follow-up work [23].

3.2. Richardson extrapolation for randomized splitting schemes

In [1], in order to approximate the solution of an ODE of the form

$$\dot{y}(t) = \bar{F}(y(t)), \tag{3.6}$$

the vector field $\bar{F}$ is split into $\bar{F}(x) = \sum_{k=1}^N F_k(x)$, where each flow $\dot{x} = F_k(x)$ can be solved exactly. More precisely, the two main examples of interest in [1] (to which we refer for details and motivation) are the so-called Lorenz-96 model and a finite-dimensional Galerkin projection of the vorticity formulation of two-dimensional Navier–Stokes. The trajectory y is then approximated by a discrete-time scheme of the form

$$x_{n+1}^\varepsilon = \Phi_{\tau_1/\varepsilon}^1 \circ \cdots \circ \Phi_{\tau_N/\varepsilon}^N(x_n^\varepsilon),$$

where Φ_t^i is the flow associated to F_i at time t and $\tau_1, \ldots, \tau_N$ are independent random variables distributed according to the standard exponential law. More precisely, for a fixed $t > 0$, $y(t)$ is approximated by x_n^ε with $n\varepsilon = t$, for a small ε. Notice that this does not enter directly our framework. However, consider the Markov-modulated ODE $\dot{x}_\varepsilon(t) = F_{\sigma(Nt/\varepsilon)}(x_\varepsilon(t))$, where σ is a cyclic chain on $[\![1, N]\!]$ with rate 1 (i.e. it is a standard Poisson process modulo N). We can take the times τ_i to define the jump times of σ, and thus x_n^ε is exactly $x_\varepsilon(\varepsilon S_n/N)$, where S_n is the (nN)th jump time of σ. In particular, if $n\varepsilon = t$, assuming that the vector fields F_k are bounded,

$$|x_\varepsilon(t) - x_n^\varepsilon| = |x_\varepsilon(t) - x_\varepsilon(\varepsilon S_n/N)| \leqslant \max_{k \in [\![1,N]\!]} \|F_k\|_\infty |\varepsilon S_n/N - t|,$$

$$\mathbb{E}|(\varepsilon S_n/N) - t|^2 = n\varepsilon^2/N = t\varepsilon/N.$$

In other words, $(x_\varepsilon(t))_{t \geqslant 0}$ (which can be sampled exactly if $(x_n^\varepsilon)_{n \in \mathbb{N}}$ can) can be seen as an alternative variation of the scheme of [1].

The convergence of x_n^ε to $y(t)$ as $\varepsilon \to 0$ for a fixed $t = n\varepsilon$, or more precisely of the corresponding Markov transition operators, is stated in [1, Theorem 4.1], which is thus similar to our Proposition 2.1 at order $n = 0$. Notice that, a priori, the ODE (3.6) is in $\mathbb{R}^d$ but [1, Theorem 4.1] is proved by restricting the study on the orbit of the process, which is assumed to be bounded (this is [1, Assumption 1]), so that our results apply in this context.

The higher-order expansion in Proposition 2.1 enables the use of a Richardson extrapolation to get better convergence rates.

Indeed, we get that, for any fixed $x_0 \in \mathbb{R}^d$, $t > 0$ and $f \in \mathcal{A}$, setting $c_t = P_t^{(1)} f(x_0)$, there exists $C > 0$ such that, for all $\varepsilon \in (0, 1]$, $|\mathbb{E}(f(\varphi_t(x_0)) - f(x_\varepsilon(t))) - c_t \varepsilon| \leqslant C\varepsilon^2$. As a consequence, for any $r > 1$, taking $\theta = r/(r-1)$ so that $\theta + (1-\theta)r = 0$, we get

$$\left| f(\varphi_t(x_0)) - \left[\theta \mathbb{E}(f(x_\varepsilon(t))) + (1-\theta)\mathbb{E}(f(x_{\varepsilon r}(t))) \right] \right| \leqslant (2\theta - 1) \, C\|f\|_3 \varepsilon^2,$$
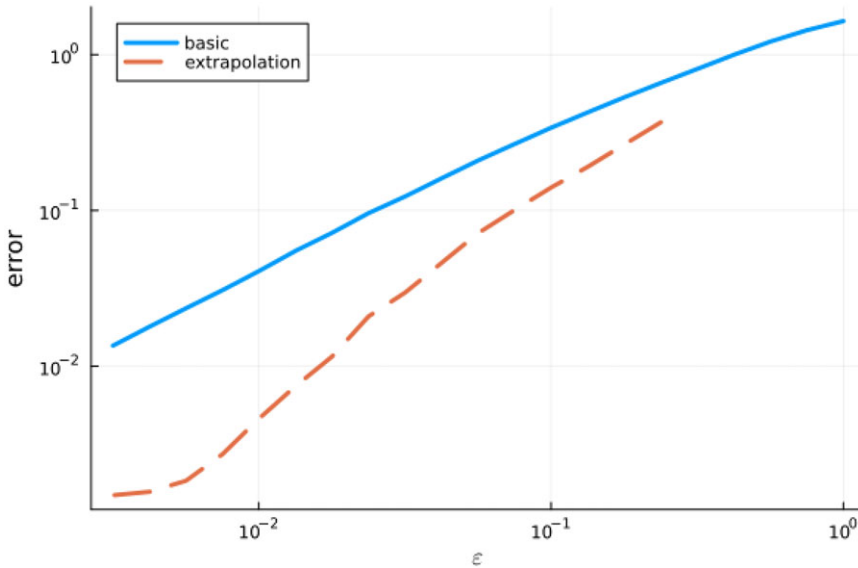
FIGURE 2. Weak error for the conservative Lorenz-96 model for the basic estimator $\hat{x}_\varepsilon(t)$ and the extrapolation $\hat{y}_\varepsilon(t)$.

providing an approximation of $f(\varphi_t(x_0))$ with a second-order weak error in terms of ε, at a numerical cost less than twice the simulation of $x_\varepsilon(t)$ (since $r > 1$, there are on average fewer jumps for $x_{\varepsilon r}(t)$).

As a simple illustration, we consider the conservative Lorenz-96 model from [1], which is (3.6) with, denoting by e_k the kth vector of the canonical basis of $\mathbb{R}^d$,

$$F_k(x) = (x_{k+1}e_k - x_k e_{k+1})x_{k-1},$$

where the indexes are periodized in the sense that $x_{n+1} = x_1$, $x_n = x_0$, and $x_{-1} = x_{n-1}$. Along the flow $\dot{x} = F_k(x)$, all coordinates but x_k and x_{k+1} are preserved, and

$$\begin{pmatrix} x_k(t) \\ x_{k+1}(t) \end{pmatrix} = \begin{pmatrix} \cos(x_{k-1}(0)t) & \sin(x_{k-1}(0)t) \\ -\sin(x_{k-1}(0)t) & \cos(x_{k-1}(0)t) \end{pmatrix} \begin{pmatrix} x_k(0) \\ x_{k+1}(0) \end{pmatrix}.$$

We run a trajectory up to time $t = 10$ starting from $x(0) = (1, \ldots, 1)$, which is a fixed point of $\bar{F}$ but not of any F_k. We compute $\hat{x}_\varepsilon(t)$, the average of $x_\varepsilon(t)$ over $m = 10^5$ independent experiments for each value of $\varepsilon \in \left\{ \left(\frac{3}{4}\right)^k, \ k \in [\![0, 20]\!] \right\}$. Taking $r = \left(\frac{4}{3}\right)^5 \simeq 4.21$ and $\theta = r/(r - 1)$ we define the extrapolation $\hat{y}_\varepsilon(t) = \theta \hat{x}_\varepsilon(t) + (1 - \theta)\hat{x}_{r\varepsilon}(t)$ for $\varepsilon \in \left\{ \left(\frac{3}{4}\right)^k, \ k \in [\![5, 20]\!] \right\}$. The errors $|\hat{x}(t) - x(0)|$ and $|\hat{y}(t) - x(0)|$ in terms of ε are plotted (in log–log scale) in Figure 2, where the speed-up of the extrapolation is clear.

3.3. Species invasion rate in a two-species Lotka–Volterra model in a random environment

We consider a model of two species in interaction, in a random environment, described by the following system of Lotka–Volterra equations:

$$\begin{cases} \dfrac{dX(t)}{dt} = X(t)\big(a_{10}^{\sigma(t)} - a_{11}^{\sigma(t)}X(t) + a_{12}^{\sigma(t)}Y(t)\big), \\ \dfrac{dY(t)}{dt} = Y(t)\big(a_{20}^{\sigma(t)} + a_{21}^{\sigma(t)}X(t) - a_{22}^{\sigma(t)}Y(t)\big). \end{cases} \tag{3.7}$$

We assume that $a_{11}^s, a_{22}^s > 0$ (intraspecific competition); the other parameters can be positive or negative. The coefficients a_{10}^s and a_{20}^s represent the growth rates of species 1 and 2 respectively, while the coefficients a_{12}^s and a_{21}^s model the interaction between the two species. If, for example, $a_{10}^s, a_{21}^s > 0$ while $a_{20}^s, a_{12}^s < 0$, we get a prey–predator system (species 1 being the prey, and species 2 the predator). Assumption 1.1 is also enforced in this whole section.

In the absence of species 2 and after accelerating the dynamics of σ by a factor $1/\varepsilon$, species 1 evolves according to a logistic equation in a random environment:

$$\frac{dX(t)}{dt} = X(t)\big(a_{10}^{\sigma(t/\varepsilon)} - a_{11}^{\sigma(t/\varepsilon)}X(t)\big). \tag{3.8}$$

We prove the following result in the Appendix.

Lemma 3.1. *Assume that* $p_0 := \inf_{s \in \mathcal{S}} a_{10}^s/a_{11}^s > 0$, $p_1 := \sup_{s \in \mathcal{S}} a_{10}^s/a_{11}^s < +\infty$, *and* $\inf_{s \in \mathcal{S}} a_s^{11} > 0$. *Then the process* $(X(t), \sigma(t/\varepsilon))_{t \geqslant 0}$ *with* X *given by* (3.8) *and* $X(0) \neq 0$ *will ultimately lie in* $E := [p_0, p_1] \times \mathcal{S}$ *and admits a unique stationary distribution* μ_ε *on* E. *Moreover,* $(X(t), \sigma(t/\varepsilon))$ *converges in law to* μ_ε *when* t *goes to infinity.*

The stationary distribution μ_ε represents the population of species 1 at equilibrium in the absence of species 2. The invasion growth rate of species 2 when species 1 is at equilibrium is then given by $\Lambda_y(\varepsilon) = \int_{\mathcal{M} \times \mathcal{S}} (a_{20}^s + a_{21}^s x)\, \mu_\varepsilon(dx, ds)$. Intuitively, we are interested in the sign of this quantity, since when the abundance of the second species is close to 0 we have, roughly,
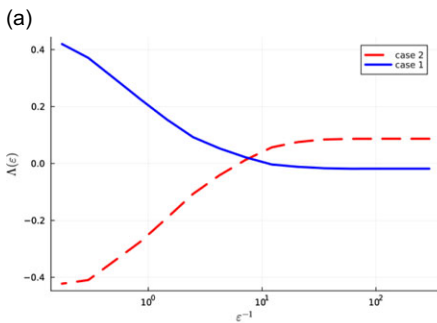
$$\frac{1}{t} \ln(Y(t)) \simeq \frac{1}{t} \int_0^t \big[a_{20}^{\sigma(u/\varepsilon)} + a_{21}^{\sigma(u/\varepsilon)}X(u)\big]\, du,$$

where X solves (3.8) and the right-hand side converges to $\Lambda_y(\varepsilon)$ when t goes to infinity, thanks to Lemma 3.1 and the ergodic theorem. It is made rigorous in [4] that, indeed, the sign of $\Lambda_y(\varepsilon)$ gives information on the local behaviour of (3.7) near the boundary $\{y = 0\}$. We illustrate this in Figure 3, considering two cases with two environments (i.e. $\sigma \in \{1, 2\}$), with the parameters detailed in Table 1. These parameters are chosen so that Λ_y is increasing in Case 1 and decreasing in Case 2, and changes sign at some point in both cases.

In Figure 3(a), $\Lambda_y(\varepsilon)$ is estimated by Monte Carlo with a long simulation of (3.8). We see that the Lyapunov exponent decreases (resp. increases) when the switching rate ε^{-1} increases in Case 1 (resp. Case 2). In Figure 3(b), (c), and (d), (3.7) is simulated with initial condition $(X_0, Y_0) = (1, 10^{-2})$. In Case 1 in Figure 3(b), the second species is persistent (i.e. invasion occurs) for $\varepsilon = 1$ (where $\Lambda_y(\varepsilon) > 0$), but as the switching rate increases, extinction occurs (for $\varepsilon = 10^{-3}$, the trajectory is close to the deterministic averaged flow, for which y_t vanishes exponentially fast). Conversely, in Figure 3(c) and (d) with the parameters of Case 2, slow switching ($\varepsilon \in \{1, 10\}$) gives $\Lambda_y(\varepsilon) < 0$ and extinction, while fast switching leads to persistence/invasion (as in the deterministic averaged flow, close to the $\varepsilon = 10^{-3}$ case).

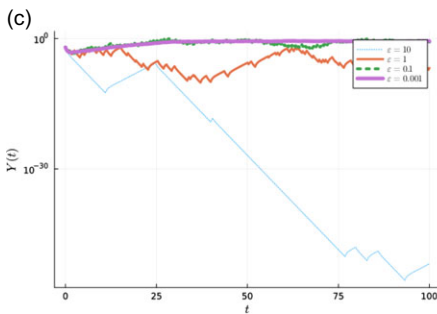TABLE 1. Values of the parameters of (3.7) in the two cases of Figure 3.

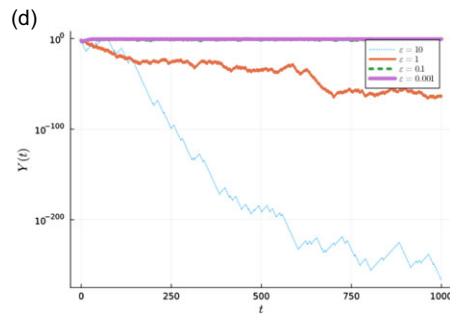|	σ	a_{10}^{σ}	a_{11}^{σ}	a_{12}^{σ}	a_{20}^{σ}	a_{21}^{σ}	a_{22}^{σ}
Case 1	1	2	1	0	-6	5	1
	2	1	1	0	-6	3	1
Case 2	1	3	1	0	-8	2	1
	2	2	1	0	-9	5	1

(a)

$\varepsilon^{-1} \mapsto \Lambda_y(\varepsilon)$ (x-axis in log-scale)

in cases 1 (blue) and 2 (red).
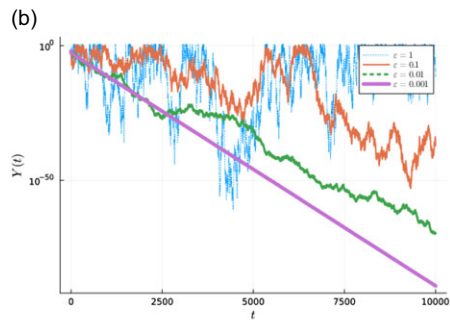
(b)

$t \mapsto Y_t$ (y-axis in log-scale)

in case 1 with $\varepsilon \in \{10^{-j}, j \in [\![0,3]\!]\}$.

(c)

$t \mapsto Y_t$ (y-axis in log-scale)

in case 2 with $\varepsilon \in \{10^{-3}, 0.1, 1, 10\}$.

(d)

Same as (c) but longer simulation time.

FIGURE 3. Simulation of Lotka–Volterra models with the parameters of Table 1.

Denoting by $\bar{\alpha} = \int_{\mathcal{S}} \alpha^s \pi(\mathrm{d}s)$ the mean with respect to π of a function α defined on $\mathcal{S}$, it is proved in [6], in the specific case when σ is a two-state Markov chain, that $\Lambda_y(\varepsilon)$ converges as ε goes to 0 to $\bar{a}_{20} + \bar{a}_{21}\bar{x}$. The following proposition extends this result to the general case of a process σ satisfying Assumption 1.1 and with a first-order expansion.

Proposition 3.3. *Assume that* $\inf_{s \in \mathcal{S}} a_{10}^s > 0$ *and* $\sup_{s \in \mathcal{S}} a_{11}^s < +\infty$. *Then, for any invariant probability measure* μ_ε *of* $(X_t, \sigma(t/\varepsilon))_{t \geq 0}$,

$$\Lambda_y(\varepsilon) = \bar{a}_{20} + \bar{a}_{21}\bar{x} + \varepsilon \frac{\bar{a}_{21}\bar{a}_{10}^2}{\bar{a}_{11}} \int_{\mathcal{S}} \left(\frac{a_{10}^s}{\bar{a}_{10}} - \frac{a_{11}^s}{\bar{a}_{11}}\right) Q^{-1} \left(\frac{a_{11}}{\bar{a}_{11}} - \frac{a_{21}}{\bar{a}_{21}}\right)(s)\, \pi(\mathrm{d}s) + \underset{\varepsilon \to 0}{o}(\varepsilon).$$

Proof. First, note that the unique equilibrium of $\bar{F}(x) = x(\bar{a}_{10} - \bar{a}_{11}x)$ is $\bar{x} := \bar{a}_{10}/\bar{a}_{11}$, and Assumption 2.1 holds (since, thanks to Lemma 3.1, it is sufficient to consider initial conditions in the compact set $[p_0, p_1]$). By Theorem 2.1, and since the marginal of μ_ε in s is independent of ε (so that $\mu_\varepsilon(a_{20}) = \bar{a}_{20}$), we have

$$\Lambda_y(\varepsilon) = \bar{a}_{20} + \bar{a}_{21}\bar{x} + \varepsilon c_1 + \underset{\varepsilon \to 0}{o}(\varepsilon),$$

with $c_1 = \pi L_c Q^{-1} (L_c h - f)(\bar{x})$ for $f(x, s) = a_{21}^s x$. Using the formula for h derived in Remark 2.1, we have

$$h(x, s) = \int_{\bar{x}}^x \frac{\bar{f}(y) - \bar{f}(\bar{x})}{\bar{F}(y)}\, \mathrm{d}y = \int_{\bar{x}}^x \bar{a}_{21} \frac{y - \bar{x}}{y(\bar{a}_{10} - \bar{a}_{11}y)}\, \mathrm{d}y = \frac{\bar{a}_{21}}{\bar{a}_{11}} \ln\left(\frac{\bar{x}}{x}\right),$$

which yields $L_c h(x, s) = -(a_{10}^s - a_{11}^s x)\bar{a}_{21}/\bar{a}_{11}$. Then, $\nabla_x(L_c h - f) = a_{11}\bar{a}_{21}/\bar{a}_{11} - a_{21}$, so that

$$L_c Q^{-1} (L_c h - f)(\bar{x}, s) = \bar{x}(a_{10}^s - a_{11}^s \bar{x})Q^{-1}\left(\frac{\bar{a}_{21}}{\bar{a}_{11}}a_{11} - a_{21}\right)(s),$$

and finally, integrating this with respect to π and using that $\bar{x} = \bar{a}_{10}/\bar{a}_{11}$,

$$c_1 = \frac{\bar{a}_{21}\bar{a}_{10}^2}{\bar{a}_{11}} \int_{\mathcal{S}} \left(\frac{a_{10}^s}{\bar{a}_{10}} - \frac{a_{11}^s}{\bar{a}_{11}}\right) Q^{-1} \left(\frac{a_{11}}{\bar{a}_{11}} - \frac{a_{21}}{\bar{a}_{21}}\right)(s)\, \pi(\mathrm{d}s), \tag{3.9}$$

which concludes the proof. $\qquad\square$

Let us discuss the sign of the expression (3.9) in some particular cases.

- If $a_{11}^s = \bar{a}_{11}$ and $a_{10}^s = \bar{a}_{10}$ for all s, then $\Lambda_y'(0) = 0$. This is expected, since in that case, the process X that is the solution to (3.8) is deterministic and does not depend on ε, and $\mu_\varepsilon = \delta_{\bar{a}_{10}/\bar{a}_{11}} \otimes \pi$ for all ε. More interestingly, if $a_{11}^s = \bar{a}_{11}$ and $a_{21}^s = \bar{a}_{21}$ but a_{10} is varying, we still have $\Lambda_y'(0) = 0$.

- If $a_{10}^s = \bar{a}_{10}$ and $a_{21}^s = \bar{a}_{21}$ for all s, using that $\pi Q^{-1} = Q^{-1}\mathbf{1} = 0$, we get

$$\begin{aligned}
c_1 &= -\frac{\bar{a}_{21}\bar{a}_{10}^2}{\bar{a}_{11}} \int_{\mathcal{S}} \frac{a_{11}^s}{\bar{a}_{11}} Q^{-1}\left(\frac{a_{11}}{\bar{a}_{11}}\right)(s)\, \pi(\mathrm{d}s) \\
&= -\frac{\bar{a}_{21}\bar{a}_{10}^2}{\bar{a}_{11}^3} \int_{\mathcal{S}} (a_{11}^s - \bar{a}_{11})Q^{-1}(a_{11} - \bar{a}_{11})(s)\, \pi(\mathrm{d}s).
\end{aligned}$$

This is always non-negative, as it can be interpreted as an asymptotic variance. Indeed, considering $g \in \mathcal{A}$ with $\pi g = 0$, for $t \geqslant 0$,

$$t\mathbb{E}_\pi \left| \frac{1}{t} \int_0^t g(\sigma(s))\,\mathrm{d}s \right|^2 = \frac{2}{t}\mathbb{E}_\pi \int_0^t \int_s^t g(\sigma(u))g(\sigma(s))\,\mathrm{d}s\,\mathrm{d}u$$

$$= \frac{2}{t} \int_0^t \int_s^t \pi(g\mathcal{Q}_{u-s}g)\,\mathrm{d}s\,\mathrm{d}u$$

$$= \frac{2}{t} \int_0^t \int_0^t \mathbb{1}_{v<t-s}\pi(g\mathcal{Q}_v g)\,\mathrm{d}s\,\mathrm{d}v$$

$$= \frac{2}{t} \int_0^t (t-v)\pi(g\mathcal{Q}_v g)\,\mathrm{d}v \underset{t\to\infty}{\longrightarrow} -2\pi\left(gQ^{-1}g\right),$$

where we used that $\int_0^t \mathcal{Q}_v g$ converges to $-Q^{-1}g$, and

$$\left| \int_0^t v\pi(g\mathcal{Q}_v g)\,\mathrm{d}v \right| \leqslant \int_0^\infty Cv\mathrm{e}^{-\gamma v}\|g\|_\infty^2\,\mathrm{d}v.$$

Example 3.3. We consider the case described in Example 1.1, where $Qf = Q^{-1}f = \pi f - f$. In that case, we get

$$Q^{-1}\left(\frac{a_{11}}{\bar{a}_{11}} - \frac{a_{21}}{\bar{a}_{21}}\right)(s) = \frac{a_{21}^s}{\bar{a}_{21}} - \frac{a_{11}^s}{\bar{a}_{11}},$$

which entails

$$c_1 = \frac{\overline{a_{10}a_{21}}}{\bar{a}_{10}\bar{a}_{21}} + \frac{\overline{a_{11}^2}}{(\bar{a}_{11})^2} - \frac{\overline{a_{10}a_{11}}}{\bar{a}_{10}\bar{a}_{11}} - \frac{\overline{a_{11}a_{21}}}{\bar{a}_{11}\bar{a}_{21}}.$$

Example 3.4. We now consider the two-state case, i.e. σ is a Markov chain on $\mathcal{S} = \{0, 1\}$, with matrix rates given by

$$Q = \begin{pmatrix} -p & p \\ 1-p & -(1-p) \end{pmatrix}$$

for some $p \in (0, 1)$. The behaviour of the solutions to (3.7) was studied in [6] through the signs of the invasion growth rates of species 1 and 2 in the competitive case (i.e. when a_{21}^s and a_{12}^s are negative). That study was complemented by [22], which proposed an alternative formula for the invasion growth rate that made it possible for them to understand the monotonicity of $\varepsilon \mapsto \Lambda_y(\varepsilon)$. In particular, it is a consequence of the proof of [22, Lemma 4.1] that $\varepsilon \mapsto \Lambda_y(\varepsilon)$ is increasing if $A_2 > 0$ while it is decreasing if $A_2 < 0$, where A_2 is the coefficient of the second-order term of the polynomial

$$P(x) = \left[\frac{a_{20}^1}{a_{10}^1}\left(1 + \frac{a_{21}^1}{a_{20}^1}x\right)\left(1 - \frac{a_{11}^0}{a_{10}^0}x\right) - \frac{a_{20}^0}{a_{10}^0}\left(1 + \frac{a_{21}^0}{a_{20}^0}x\right)\left(1 - \frac{a_{11}^1}{a_{10}^1}x\right) \right] \frac{a_{11}^1/a_{10}^1 - a_{11}^0/a_{10}^0}{|a_{11}^1/a_{10}^1 - a_{11}^0/a_{10}^0|},$$

that is, $A_2 = (a_{11}^1 a_{21}^0 - a_{11}^0 a_{21}^1)/(a_{10}^0 a_{10}^1)$, if we assume without loss of generality that $a_{11}^1 a_{10}^0 \geqslant a_{11}^0 a_{10}^0$. Let us prove that our results are in accordance with the results of [22] by studying the sign of the first-order term of the expansion of $\varepsilon \mapsto \Lambda_y(\varepsilon)$ when $\varepsilon \to 0$. Using Example 1.2,

we have $\pi = (1-p, p)$, $Q^{-1} = Q$, and $\pi_s Q_{s,s'} = p(1-p)$ if $s \neq s'$ and $-p(1-p)$ if $s = s'$, from which we compute

$$
\begin{aligned}
c_1 &= \frac{\bar{a}_{21}\bar{a}_{10}^2}{\bar{a}_{11}} \sum_{s,s' \in 0,1} \pi_s Q_{s,s'} \left(\frac{a_{10}^s}{\bar{a}_{10}} - \frac{a_{11}^s}{\bar{a}_{11}} \right) \left(\frac{a_{11}^{s'}}{\bar{a}_{11}} - \frac{a_{21}^{s'}}{\bar{a}_{21}} \right) \\
&= \frac{\bar{a}_{21}\bar{a}_{10}^2}{\bar{a}_{11}} p(1-p) \left[\left(\frac{a_{10}^0}{\bar{a}_{10}} - \frac{a_{11}^0}{\bar{a}_{11}} \right) \left(\frac{a_{11}^1}{\bar{a}_{11}} - \frac{a_{21}^1}{\bar{a}_{21}} \right) + \left(\frac{a_{10}^1}{\bar{a}_{10}} - \frac{a_{11}^1}{\bar{a}_{11}} \right) \left(\frac{a_{11}^0}{\bar{a}_{11}} - \frac{a_{21}^0}{\bar{a}_{21}} \right) \right. \\
&\qquad \left. - \left(\frac{a_{10}^0}{\bar{a}_{10}} - \frac{a_{11}^0}{\bar{a}_{11}} \right) \left(\frac{a_{11}^0}{\bar{a}_{11}} - \frac{a_{21}^0}{\bar{a}_{21}} \right) - \left(\frac{a_{10}^1}{\bar{a}_{10}} - \frac{a_{11}^1}{\bar{a}_{11}} \right) \left(\frac{a_{11}^1}{\bar{a}_{11}} - \frac{a_{21}^1}{\bar{a}_{21}} \right) \right] \\
&= \frac{\bar{a}_{21}\bar{a}_{10}^2}{\bar{a}_{11}} p(1-p) \left[\frac{a_{10}^0}{\bar{a}_{10}} - \frac{a_{11}^0}{\bar{a}_{11}} - \left(\frac{a_{10}^1}{\bar{a}_{10}} - \frac{a_{11}^1}{\bar{a}_{11}} \right) \right] \left[\frac{a_{11}^1}{\bar{a}_{11}} - \frac{a_{21}^1}{\bar{a}_{21}} - \left(\frac{a_{11}^0}{\bar{a}_{11}} - \frac{a_{21}^0}{\bar{a}_{21}} \right) \right] \\
&= \frac{\bar{a}_{21}\bar{a}_{10}^2}{\bar{a}_{11}} p(1-p) \left(\frac{a_{10}^0 - a_{10}^1}{\bar{a}_{10}} - \frac{a_{11}^0 - a_{11}^1}{\bar{a}_{11}} \right) \left(\frac{a_{11}^1 - a_{11}^0}{\bar{a}_{11}} - \frac{a_{21}^1 - a_{21}^0}{\bar{a}_{21}} \right) \\
&= \frac{p(1-p)a_{10}^0 a_{10}^1 \bar{a}_{10}}{(\bar{a}_{11})^3} \left(\frac{a_{11}^1}{a_{10}^1} - \frac{a_{11}^0}{a_{10}^0} \right) (a_{11}^1 a_{21}^0 - a_{11}^0 a_{21}^1).
\end{aligned}
$$

Since we have assumed that $a_{11}^1 / a_{10}^1 \geqslant a_{11}^0 / a_{10}^0$, the sign of $\Lambda_y'(0)$ is indeed the same as that of $a_{11}^1 a_{21}^0 - a_{11}^0 a_{21}^1$.

In Example 3.4 we see that the sign of c_1 only depends on the coefficients of the vector fields, through the sign of $a_{11}^1 a_{21}^0 - a_{11}^0 a_{21}^1$. The proof in [22] of the monotonicity of $\varepsilon \mapsto \Lambda_y(\varepsilon)$ relies on an explicit formula for μ_ε. Such an explicit formula is, however, in general only computable in the case of two environments. In contrast, our formula can provide some insights on the local behaviour of $\varepsilon \mapsto \Lambda_y(\varepsilon)$ near 0 for any number of environmental states. We now give an example where, contrary to the previous situation, the sign of c_1 depends also on the switching rates of σ.

Example 3.5. Let σ be an irreducible Markov chain on $\mathcal{S} = \{1, 2, 3\}$. We consider two different rate matrices for σ:

$$
Q = \begin{pmatrix} -(\frac{1}{2} + \delta) & \frac{1}{2} & \delta \\ \frac{1}{2} & -(\frac{1}{2} + \delta) & \delta \\ \frac{1}{2} & \frac{1}{2} & -1 \end{pmatrix}, \qquad Q' = \begin{pmatrix} -1 & \frac{1}{2} & \frac{1}{2} \\ \delta & -(\frac{1}{2} + \delta) & \frac{1}{2} \\ \delta & \frac{1}{2} & -(\frac{1}{2} + \delta) \end{pmatrix}
$$

for some small $\delta > 0$. Intuitively, if the rates matrix is Q, then σ will spend most of the time in states 1 and 2, while if the rates matrix is Q', then σ will spend most of the time in states 2 and 3. Hence, we can expect that in the first case, for small δ, c_1 will be close to $c_1(1, 2)$, given by

$$
c_1(1, 2) = \frac{a_{10}^1 a_{10}^2 \bar{a}_{10}}{4(\bar{a}_{11})^3} \left(\frac{a_{11}^2}{a_{10}^2} - \frac{a_{11}^1}{a_{10}^1} \right) (a_{11}^2 a_{21}^1 - a_{11}^1 a_{21}^2),
$$

which is the expression derived in Example 3.4 when σ is a Markov chain on $\{1, 2\}$ with $p = \frac{1}{2}$. In the second case, c_1 should be close to

$$c_1(2, 3) = \frac{a_{10}^2 a_{10}^3 \bar{a}_{10}}{4(\bar{a}_{11})^3} \left(\frac{a_{11}^3}{a_{10}^3} - \frac{a_{11}^2}{a_{10}^2} \right) (a_{11}^3 a_{21}^2 - a_{11}^2 a_{21}^3),$$

the expression when σ is a Markov chain on $\{2, 3\}$ with $p = \frac{1}{2}$. In particular, we can choose the coefficients $(a_{jk}^s)_{j,k,s}$ in such a way that $c_1(1, 2) < 0$ and $c_1(2, 3) > 0$, meaning that, with the same environmental states, different switching mechanisms can lead to opposite signs for the derivative at 0 of the growth rate. In other words, it can be as beneficial as it is detrimental for species y to increase the switching rates, depending on the switching structure.

Using (3.9) with $a_{11}^s = 1$ for all s and $a_{10}^1 = 3$, $a_{10}^2 = 2$, $a_{10}^3 = 1$, $a_{21}^1 = 2$, $a_{21}^2 = 5$, $a_{21}^3 = 3$, and $\delta = 0.0005$, we compute in the first case $c_1 \simeq -1.871\,628\,9 < 0$ and in the second case $c_1 \simeq 0.747\,377\,2 > 0$. As expected, these two values are respectively close to $c_1(1, 2) \simeq -1.874\,437\,8$ and $c_1(2, 3) \simeq 0.750\,374\,8$.

Notice that, for the two-environment processes, these parameters are those used in Table 1, and thus we get that the sign of c_1 in each case is consistent with the monotonicity displayed in Figure 3(a).

Appendix A. Proof of Proposition 2.1

Proof. This is essentially the proof from [24], in dual form since we work with $P_t f$ rather than the law of the process. It is organized in three steps. First, assuming formally that (2.3) gives an expansion in ε of P_t^ε, we deduce the expressions for $P_t^{(k)}$ and $S_t^{(k)}$. With these definitions, in the second step we establish the bounds in (2.1) and (2.2) on $P_t^{(k)}$ and $S_t^{(k)}$. Finally, from this, we obtain the bound in (2.4) on $R_{n,t}^\varepsilon$.

Step 1. Fix $f \in \mathcal{C}^\infty(\mathcal{M} \times \mathcal{S})$. For all $t \geqslant 0$,

$$\partial_t P_t^\varepsilon f = L^\varepsilon P_t^\varepsilon f. \tag{A.1}$$

Bearing in mind the ansatz that (2.3) gives an expansion in ε of $P_t^\varepsilon f$, developing each side of the above equality in ε, and equating terms of the same order (treating the terms $P_t^{(k)} f$ and $S_t^{(k)} f$ separately), we end up with the following equations for all $k \geqslant 0$:

$$Q P_t^{(0)} f = 0, \tag{A.2}$$

$$\partial_t P_t^{(k)} f = L_c \left(P_t^{(k)} f \right) + Q P_t^{(k+1)} f, \tag{A.3}$$

$$\partial_t S_t^{(0)} f = Q S_t^{(0)} f, \tag{A.4}$$

$$\partial_t S_t^{(k)} f = Q S_t^{(k)} f + L_c (S_t^{(k-1)} f). \tag{A.5}$$

To shorten the notation, for $k \in \mathbb{N}$ we write $u_k(t, x, s) = P_t^{(k)} f(x, s)$. Moreover, for a function $g(t, x, s)$ we use the notation $\pi g(t, x) = \int_{\mathcal{S}} g(t, x, s) \pi(\mathrm{d}s)$. Equation (A.2) is equivalent to say that there exists $\theta_0(t, x)$ such that, for all $s \in \mathcal{S}$, $u_0(t, x, s) = \theta_0(t, x)$. Now, if we want (A.3) for $k = 0$ to have a solution u_1, this will induce a constraint, fixing the value of θ_0. Indeed, this equation can be rewritten as $Q u_1 = \partial_t u_0 - L_c u_0$. The left-hand side averages to 0 with respect to π (for all fixed $x \in \mathcal{M} t \geqslant 0$), and thus we have to impose that $\pi(\partial_t u_0 - L_c u_0) = 0$. This is equivalent to

$$\partial_t \theta_0(t, x) - \bar{F}(x) \cdot \nabla \theta_0(t, x) = 0, \tag{A.6}$$

which is a transport equation with solution given by $\theta_0(t, x) = \theta_0(0, \varphi_t(x))$. In order to identify a suitable initial condition $\theta_0(0, x) = \pi u_0(0, x)$, we formally let ε and then t go to 0 in

$$\pi P_t^\varepsilon f(x) = \pi f(x) + \int_0^t \pi L_c P_r^\varepsilon f(x) \, dr,$$

which is obtained by integrating (A.1) and using that $\pi Q = 0$. This yields $\theta_0(0, x) = \pi f(x)$ and, as stated,

$$P_t^{(0)} f(x, s) = u_0(t, x, s) = \pi f(\varphi_t(x)). \tag{A.7}$$

Now, reasoning by induction, we assume that for some $k \geqslant 1$ we have constructed u_{k-1} satisfying $\pi(\partial_t u_{k-1} - L_c u_{k-1}) = 0$, from which we want to define u_k. Recalling the definition in (1.4) of the pseudo-inverse Q^{-1}, (A.3) is then equivalent to the existence of a function $\theta_k \colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ such that

$$u_k(t, x, s) = Q^{-1}(\partial_t u_{k-1} - L_c u_{k-1})(t, x, s) + \theta_k(t, x). \tag{A.8}$$

Once again, to find a suitable θ_k we use (A.3) at order $k + 1$, which we rewrite as $Q u_{k+1} = \partial_t u_k - L_c u_k$. As before, for fixed (t, x), since the left-hand side averages to 0 with respect to π, we impose

$$\pi(\partial_t u_k - L_c u_k)(t, x) = 0. \tag{A.9}$$

Notice that $\pi Q^{-1} = 0$, so that, averaging (A.8) with respect to π and differentiating with respect to time,

$$\pi \, \partial_t u_k = \partial_t \theta_k. \tag{A.10}$$

Injecting (A.10) and (A.8) into (A.9) yields

$$\partial_t \theta_k(t, x) - \bar{F}(x) \cdot \nabla \theta_k(t, x) = \pi L_c \mathbf{b}_k(t, x), \tag{A.11}$$

where we set $\mathbf{b}_k = Q^{-1}(\partial_t u_{k-1} - L_c u_{k-1})$. This is again a transport equation, similar to (A.6), but with a source term. The solution is given by

$$\theta_k(t, x) = \theta_k(0, \varphi_t(x)) + \int_0^t \pi L_c \mathbf{b}_k(r, \varphi_{t-r}(x)) \, dr. \tag{A.12}$$

For now, by construction, for any choice of the function $\theta_k(0, \cdot)$, the function u_k defined by (A.8) where θ_k is given by (A.12) satisfies (A.3) and (A.9). It remains to specify a suitable initial condition $\theta_k(0, x)$. To do so, we need to first look at the so-called boundary layer correction terms $S^{(j)}$.

Set $v_k(t, x, s) = S_t^{(k)} f(x, s)$. Then, (A.4) and (A.5) read

$$\partial_t v_0 = Q v_0, \tag{A.13}$$
$$\partial_t v_k = Q v_k + L_c v_{k-1}. \tag{A.14}$$

Equation (A.13) is simply solved as $v_0(t, x, s) = \mathcal{Q}_t v_0(0, x, s)$. Moreover, taking $t = 0$ and letting $\varepsilon \to 0$ in (2.3) yields $v_0(0, x, s) + u_0(0, x, s) = f(x, s)$, and hence we set

$$v_0(t, x, s) = \mathcal{Q}_t f(x, s) - \pi f(x), \tag{A.15}$$

which indeed solves (A.13). Next, assuming by induction that v_{k-1} has been defined for some $k \geqslant 1$, we can solve (A.14) to find

$$v_k(t, x, s) = \mathcal{Q}_t v_k(0, x, s) + \int_0^t \mathcal{Q}_{t-r} L_c v_{k-1}(r, x, s) \, dr.$$

It remains to choose a suitable initial condition $v_k(0, x)$, which will be determined by the requirement of the long-time decay (2.2). Indeed, since $\lim_{t \to \infty} \mathcal{Q}_t = \pi$, (2.2) can only hold if

$$0 = \pi \left(v_k(0, x) + \int_0^\infty \mathcal{Q}_r L_c v_{k-1}(r, x)\,\mathrm{d}r \right),$$

and therefore

$$\pi v_k(0, x) = -\int_0^\infty \pi \mathcal{Q}_r L_c v_{k-1}(r, x)\,\mathrm{d}r = -\int_0^\infty \pi L_c v_{k-1}(r, x)\,\mathrm{d}r. \qquad (\text{A.16})$$

Also, from (2.3) applied with $t = 0$, we get that $v_k(0, x, s) + u_k(0, x, s) = 0$ for all $k \geqslant 1$ and, integrating (A.8) at $t = 0$ with respect to π, $\theta_k(0, x) = \pi u_k(0, x)$. Hence, the only choice of $\theta_k(0, \cdot)$ that is compatible with (2.2) is

$$\theta_k(0, x) = \int_0^\infty pi L_c v_{k-1}(r, x)\,\mathrm{d}r.$$

Finally, $v_k(0, x, s) = -u_k(0, x, s)$ is determined, again using (A.8) at $t = 0$. At this point, we have completely determined a definition of $P_t^{(k)}$ and $S_t^{(k)}$, which is the one that we use in the rest of the proof. It is such that (A.2), (A.3), (A.4), and (A.5) hold.

Step 2. We prove that the bounds (2.1) and (2.2) on $P_t^{(k)}$ and $S_t^{(k)}$ hold. We work by induction on k, starting with $k = 0$. By compactness, recalling the definition of u_0 in (A.7), it is straightforward that for all $T > 0$ there exists $C_0(T, j, p)$ such that, for all $j \geqslant 0$ and all $f \in \mathcal{C}^\infty(\mathcal{M} \times \mathcal{S})$,

$$\sup_{t \in [0,T]} \||\partial_t^p u_0\||_j \leq C_0(T, j, p) \||f\||_{j+p}.$$

Let $j \geqslant 0$, $\alpha \in [\![1, N]\!]^j$, and $p \in \mathbb{N}$. Then, by (A.15), $\partial_x^\alpha v_0 = \mathcal{Q}_t(\partial_x^\alpha f - \pi(\partial_x^\alpha f))$. Using the ergodicity condition in (1.3), we get, for all $f \in \mathcal{C}^\infty$, $t \geqslant 0$, and $j \in \mathbb{N}$, $\||v_0(t, \cdot)\||_j \leq C\mathrm{e}^{-\gamma t}\||f\||_j$.

For $k \geqslant 1$, reasoning by induction, we assume that there exist constants $C_{k-1}(T, j, p)$ and $C_{k-1}(j)$ such that

$$\sup_{t \in [0,T]} \||\partial_t^p u_{k-1}(t, \cdot)\||_j \leq C_{k-1}(T, j, p) \||f\||_{j+p+2(k-1)}$$

and, for all $t \geqslant 0$,

$$\||v_{k-1}(t, \cdot)\||_j \leqslant C_{k-1}(j)\mathrm{e}^{-\gamma t}(1 + t^{k-1})\||f\||_{j+2(k-1)}.$$

Since $s, x \mapsto F_s(x)$ and its derivatives in x are uniformly bounded over $\mathcal{M} \times \mathcal{S}$, it is straightforward to check that that, for any $j, p \in \mathbb{N}$ and $T > 0$, there exist $C(j)$ and $C(T, j, p)$ such that, for any $g \in \mathcal{C}^\infty(\mathcal{M})$,

$$\||L_c g\||_j \leqslant C(j)\||g\||_{j+1}, \qquad \sup_{t \in [0,T]} \||\partial_t^p(g \circ \varphi_t)\||_j \leqslant C(T, j, p)\||g\||_{j+p},$$

which we use extensively in the rest of this step. In the following, T, p, j, k are fixed, and we denote by C a constant which may change from line to line and depends only on T, p, j, k.

First, for all $t \leq T$,

$$\||\partial_t^p(\theta_k(0, \cdot) \circ \varphi_t)\||_j \leqslant C\||\theta_k(0, \cdot)\||_{j+p}$$

$$\leqslant C \int_0^\infty \||L_c v_{k-1}(r, \cdot)\||_{j+p} \, \mathrm{d}r$$

$$\leqslant C \int_0^\infty \||v_{k-1}(r, \cdot)\||_{j+p+1} \, \mathrm{d}r \leq C\||f\||_{j+p+1+2(k-1)}$$

by induction. Next, using that $\partial_t^p \mathbf{b}_k = Q^{-1}(\partial_t^{p+1} u_{k-1} - L_c(\partial_t^p u_{k-1}))$, we get that, for all $t \leq T$,

$$\||\partial_t^p \mathbf{b}_k\||_j \leqslant C\big(\||\partial_t^{p+1} u_{k-1}\||_j + \||L_c(\partial_t^p u_{k-1})\||_j\big)$$

$$\leqslant C\big(\||\partial_t^{p+1} u_{k-1}\||_j + \||\partial_t^p u_{k-1}\||_{j+1}\big) \leq C\||f\||_{j+p+1+2(k-1)}$$

by induction. Now, from (A.12), for all $t \leq T$,

$$\||\theta_k(t, \cdot)\||_j \leqslant \||\theta_k(0, \cdot) \circ \varphi_t\||_j + \int_0^t \||L_c \mathbf{b}_k(s, \cdot) \circ \varphi_{t-s}\||_j \, \mathrm{d}s$$

$$\leqslant C\||\theta_k(0, \cdot)\||_j + C \int_0^t \||\mathbf{b}_k(s, \cdot)\||_{j+1} \, \mathrm{d}s \leq C\||f\||_{j+2+2(k-1)}.$$

To bound the derivatives in the time of θ_k, instead of using (A.12) it is simpler to work with (A.11), which gives, for $p \geqslant 1$,

$$\||\partial_t^p \theta_k(t, \cdot)\||_j \leqslant \||\pi L_c \partial_t^{p-1} \theta_k(t, \cdot)\||_j + \||\pi L_c \partial_t^{p-1} \mathbf{b}_k(t, \cdot)\||_j$$

$$\leqslant C\||\partial_t^{p-1} \theta_k(t, \cdot)\||_{j+1} + C\||\partial_t^{p-1} \mathbf{b}_k(t, \cdot)\||_{j+1}$$

$$\leqslant C\||\partial_t^{p-1} \theta_k(t, \cdot)\||_{j+1} + C\||f\||_{j+p+1+2(k-1)}.$$

Reasoning by induction on p yields $\||\partial_t^p \theta_k(t, \cdot)\||_j \leq C\||f\||_{j+p+2+2(k-1)}$. Gathering the previous bounds in (A.8) yields

$$\||\partial_t^p u_k(t, \cdot)\||_j \leqslant \||\partial_t^p \theta_k(t, \cdot)\||_{j,0} + \||\partial_t^p \mathbf{b}_k(t, \cdot)\||_j \leq C\||f\||_{j+p+2k}.$$

This concludes the study of u_k by induction. We now turn to the study of v_k. Using (A.16) and omitting the dependency on x, s to alleviate the notation, we decompose v_k as

$$v_k(t) = \mathcal{Q}_t v_k(0) + \int_0^t \mathcal{Q}_{t-r} L_c v_{k-1}(r) \, \mathrm{d}r$$

$$= \mathcal{Q}_t(v_k(0) - \pi v_k(0)) + \int_0^t \mathcal{Q}_{t-r}(L_c v_{k-1}(r) - \pi L_c v_{k-1}(r)) \, \mathrm{d}r - \int_t^\infty \pi L_c v_{k-1}(r) \, \mathrm{d}r.$$

We have $\||\mathcal{Q}_t(v_k(0) - \pi v_k(0))\||_j \leq Ce^{-\gamma t}\||v_k(0)\||_j \leq Ce^{-\gamma t}\||f\||_{j+2k}$, where we used that $v_k(0) = -u_k(0)$ and the previous bound on $u_k(0)$. Similarly,

$$\left\|\left\|\left\| \int_0^t \mathcal{Q}_{t-r}(L_c v_{k-1}(r) - \pi L_c v_{k-1}(r)) \, \mathrm{d}r \right\|\right\|\right\|_j \leqslant C \int_0^t e^{-\gamma(t-r)}\||L_c v_{k-1}(r)\||_j \, \mathrm{d}r$$

$$\leqslant C \int_0^t e^{-\gamma(t-r)}\||v_{k-1}(r)\||_{j+1} \, \mathrm{d}r$$

$$\leqslant Ce^{-\gamma t}(1 + t^k)\||f\||_{j+1+2(k-1)},$$

where we have used the induction hypothesis on v_{k-1}. Finally,

$$
\begin{aligned}
\int_t^\infty \||L_c v_{k-1}(r)\||_j \, \mathrm{d}r &\le C \int_t^\infty \||v_{k-1}(r)\||_{j+1} \, \mathrm{d}r \\
&\le C \int_t^\infty (1 + r^{k-1}) \mathrm{e}^{-\gamma r} \, \mathrm{d}r \||f\||_{j+1+2(k-1)} \\
&\le C(1 + t^{k-1}) \mathrm{e}^{-\gamma t} \||f\||_{j+1+2(k-1)}.
\end{aligned}
$$

This proves the desired bounds on v_k. In particular, at this stage we have proved (2.1) and (2.2).

Step 3. We now bound the remainder $R_{n,t}^\varepsilon$ given by (2.3) for a given $n \ge 1$. Introducing $\eta_t = (\partial_t - L_\varepsilon) R_{n,t}^\varepsilon f$ for a given $f \in \mathcal{C}^\infty$, we see that, thanks to (A.2), (A.3), (A.4), and (A.5), all the low-order terms in ε vanish (as designed) so that

$$
\eta_t = (\partial_t - L_\varepsilon)\left(P_t^\varepsilon f - \sum_{k=0}^n \varepsilon^k P_t^{(k)} f + \sum_{k=0}^n \varepsilon^k S_{t/\varepsilon}^{(k)} f \right) = \varepsilon^n \left(L_c P_t^{(n)} f - \partial_t P_t^{(n)} f + L_c S_{t/\varepsilon}^{(n)} f \right).
$$

The bounds obtained in the previous step yield $\sup_{t \in [0,T]} \||\eta_t\||_j \le C \varepsilon^n \||f\||_{j+1+2n}$ for some C depending only on T, j, n. Interpreting $R_{n,t}^\varepsilon f$ as the solution of the equation $\partial_t R_{n,t}^\varepsilon f = L_\varepsilon R_{n,t}^\varepsilon f + \eta_t$ with $R_{n,0}^\varepsilon f = 0$ yields the Feynman–Kac representation

$$
R_{n,t} f(x, s) = \mathbb{E}_{x,s}\left(\int_0^t \eta_r(X(r), \sigma(r)) \, \mathrm{d}r \right),
$$

see [11, Theorem 6.3]. The bounds on η_t immediately yield

$$
\sup_{t \in [0,T]} \|R_{n,t}^\varepsilon f\|_\infty \le C \varepsilon^n \|f\|_{1+2n}
$$

for some C depending only on T, n.

Now, to get similar bounds for the derivatives in space of $R_{n,t}^\varepsilon f$, we work by induction on j, assuming that we have already obtained that

$$
\sup_{t \in [0,T]} \||R_{n,t}^\varepsilon f\||_{j'} \le C \varepsilon^n \||f\||_{j'+1+2n} \tag{A.17}
$$

for all $j' < j$ for some $j \ge 1$ and some $C > 0$. Indeed, then, considering a multi-index $\alpha \in [\![1, d]\!]^j$, differentiating the equation $\partial_t R_{n,t}^\varepsilon = L_\varepsilon R_{n,t}^\varepsilon f + \eta_t$ yields

$$
\partial_t \partial_x^\alpha R_{n,t}^\varepsilon f = L_\varepsilon \partial_x^\alpha R_{n,t}^\varepsilon f + \partial_x^\alpha \eta_t + \sum_{|\alpha'| \le j} B_{\alpha'} \partial_x^{\alpha'} R_{n,t}^\varepsilon f, \tag{A.18}
$$

where, for all multi-index α', $B_{\alpha'}(x, s)$ is a matrix whose coefficients are some derivatives in x of F_s. In other words, considering a vector G_t whose coefficients are all the partial derivatives in x of $R_{n,t}^\varepsilon f$ of order j, then G_t solves an equation of the form

$$
\partial_t G_t = L_\varepsilon G_t + B' G_t + \eta_t', \tag{A.19}
$$

where L_ε acts coefficient-wise on G_t, the coefficients of the matrix $B'(x, s)$ are derivatives in x of F_s, and η_t' gathers $\partial_x^\alpha \eta_t$ and all the terms in (A.18) involving derivatives of $R_{n,t}^\varepsilon f$ of order

$|\alpha'| < j$. In particular, thanks to the bounds in (2.1) and (2.2) established in Step 2 and to the induction assumption, we have $\sup_{t\in[0,T]}\|\eta'_t\|_\infty \leq C\varepsilon^n \|f\|_{j+1+2n}$ for some $C > 0$ depending on T, n, j. We again use the Feynman–Kac formula for G_t based on (A.19), which reads

$$G_t(x, s) = \mathbb{E}_{x,s}\left(\int_0^t E_r \eta'_{t-r}(X(r), \sigma(r))\, \mathrm{d}r \right),$$

where $(E_t)_{t\geq 0}$ is the solution of the matrix-valued ODE $\partial_t E_t = E_t B'(X(t), \sigma(t))$ (for completeness, we provide a proof of this formula in Appendix B). Since B' is uniformly bounded, so is E_r over $r \in [0, T]$, and thus $\sup_{t\in[0,T]}\|G_t\|_\infty \leq C\varepsilon^n \|f\|_{j+1+2n}$ for some $C > 0$ depending on T, n, j. Since this holds for all $\alpha \in [\![1, d]\!]^j$, this concludes the proof by induction of (A.17) for all j'. To conclude, using that

$$R_{n,t}^\varepsilon f = R_{n+1,t}^\varepsilon f + \varepsilon^{n+1} P_t^{(n+1)} f + \varepsilon^{n+1} S_{t/\varepsilon}^{(n+1)} f,$$

we get that, for all $j \geq 0$,

$$\|R_{n,t}^\varepsilon f\|_j \leq \|R_{n+1,t}^\varepsilon f\|_j + \varepsilon^{n+1}\|P_t^{(n+1)} f\|_j + \varepsilon^{n+1}\|S_{t/\varepsilon}^{(n+1)} f\|_j \leq C\varepsilon^{n+1} \|f\|_{j+3+2n},$$

where we used (2.1), (2.2), and (A.17) (for $n + 1$). This concludes the proof of (2.4), and thus the proof of Proposition 2.1. $\qquad\square$

Appendix B. Feynman–Kac formula

Lemma B.1. *Let L be a Markov generator on a set $\mathcal{S}'$. For $n \geq 1$, consider $f_t(z) = (f_{1,t}(z), \ldots, f_{n,t}(z)) \in \mathbb{R}^n$ for $z \in \mathcal{S}'$, $t \geq 0$, a solution of*

$$\partial_t f_t(z) = L f_t(z) + A_t(z) f_t(z) + c_t(z), \quad z \in \mathcal{S}', \ t \geq 0,$$

where $(t, z) \mapsto A$ is a bounded $n \times n$ matrix field, $(t, z) \mapsto c$ is a bounded vector field, and $L f_t$ is to be understood component-wise. Then, for all $t \geq 0$ and $z \in \mathcal{S}'$,

$$f_t(z) = \mathbb{E}(E_t f_0(Z_t)) + \int_0^t \mathbb{E}(E_r c_{t-r}(Z_r))\ \mathrm{d}r,$$

where $(Z_t)_{t\geq 0}$ is a Markov process associated to L with $Z_0 = z$, and $(E_r)_{r\in[0,t]}$ is the solution of the $n \times n$ matrix-valued random ODE $\partial_r E_r = E_r A_{t-r}(Z_r)$ with $E_0 = I_n$, provided this is well-defined almost surely.

Proof. Fix $t > 0$, $z \in \mathcal{S}'$. Then $(Z_r, E_r)_{r\geq 0}$ given in the statement of the lemma is a time-inhomogeneous Markov process with generator

$$\mathcal{L}_r g(z, e) = e A_{t-r}(z) \cdot \nabla_e g(z, e) + L g(z, e),$$

and thus, considering the vector-valued function $h(r, z, e) = e f_{t-r}(z)$, we end up with

$$\begin{aligned}
\partial_r \mathbb{E}(E_r f_{t-r}(Z_r)) &= \partial_r \mathbb{E}(h(r, Z_r, E_r)) \\
&= \mathbb{E}((\partial_r + \mathcal{L}_r)\, h(r, Z_r, E_r)) \\
&= \mathbb{E}(E_s(-\partial_r + A_{t-r} + L)\, f_{t-r}(Z_r)) \\
&= -\mathbb{E}(E_r c_{t-r}(Z_r))\,.
\end{aligned}$$

Using that $E_0 = I$, the conclusion follows from

$$f_t(z) = \mathbb{E}\left(E_0 f_t(Z_0)\right) = \mathbb{E}\left(E_t f_0(Z_t)\right) - \int_0^t \partial_r \mathbb{E}\left(E_r f_{t-r}(Z_r)\right) \, \mathrm{d}r. \qquad \square$$

Appendix C. Proof of Lemma 3.1

Proof. We can write $\sigma^\varepsilon(t) = \sigma(t/\varepsilon)$, and then drop the superscript ε to alleviate the notation; in other words, without loss of generality it is sufficient to treat the case $\varepsilon = 1$ (up to replacing γ by γ/ε in Assumption 1.1).

For $s \in \mathcal{S}$, let $F_s(x) = x\left(a_{10}^s - a_{11}^s x\right)$. By the definition of p_0 and p_1 we have that, for all $s \in \mathcal{S}$, $F_s(x) > 0$ for $x \in (0, p_0)$ and $F_s(x) < 0$ for $x > p_1$, which implies that for all $X_0 \neq 0$ there exists a time T_0 such that, for all $t \geqslant T_0$, $X_t \in [p_0, p_1]$. Now, on the state space $E = [p_0, p_1] \times \mathcal{S}$, we introduce the metric

$$d((x, s), (y, s')) = 1_{s \neq s'} + 1_{s = s'} \frac{|x - y|}{p_1 - p_0},$$

and we consider on the space of probability measures on E the associated Wasserstein distance defined for all μ, ν by

$$\mathcal{W}_d(\mu, \nu) = \inf_\Gamma \int d((x, s), (y, s')) \, \mathrm{d}\Gamma((x, s), (x, s')),$$

where the infimum runs over all the coupling measures of μ and ν. We will show that there exist $C, \delta > 0$ such that, for all $(x, s), (y, s') \in E$ and all $t \geqslant 0$,

$$\mathcal{W}_d(\delta_{(x,s)}P_t, \delta_{(y,s')}P_t) \leq C e^{-\delta t} d((x, s), (y, s')), \tag{C.1}$$

where P_t stands for P_t^1, the semigroup of (X, σ) at time t. The above inequality implies Lemma 3.1 by classical arguments, since the space of probability measures on E endowed with the Wasserstein metric $\mathcal{W}_d$ is complete.

We first prove (C.1) for points (x, s) and (y, s') such that $s' = s$. In that case, a coupling of $\delta_{(x,s)}P_t$ and $\delta_{(y,s)}P_t$ is given by the random variables $((X_t^{x,s}, \sigma_t^s), (X_t^{y,s}, \sigma_t^s))$ where, for all $T \geqslant 0$,

$$X_T^{x,s} = x + \int_0^T F_{\sigma^s(u)}(X_u^{x,s}) \, \mathrm{d}u, \qquad X_T^{x,s} = y + \int_0^T F_{\sigma^s(u)}(X_u^{y,s}) \, \mathrm{d}u.$$

In other words, $X_t^{x,s}$ and $X_t^{y,s}$ denote the solution to (3.8) with respective initial conditions x and y and constructed with the same realisation of the process $(\sigma(u)_{u \geqslant 0}$ starting from s.

Lemma C.1. *Almost surely, for all $x, y \in [p_0, p_1]$, $s \in \mathcal{S}$, and $t \geqslant 0$,*

$$|X_t^{x,s} - X_t^{y,s}| \leq \frac{p_1}{p_0} e^{-\eta t} |x - y|, \tag{C.2}$$

where $\eta = p_0 \inf_{s \in \mathcal{S}} a_{11}^s > 0$. This immediately yields

$$\mathcal{W}_d(\delta_{(x,s)}P_t^\varepsilon, \delta_{(y,s)}P_t^\varepsilon) \leq \frac{p_1}{p_0} e^{-\eta t} d((x, s), (y, s)).$$

Proof. Let $s \in \mathcal{S}$ and write $\alpha_t = a_{10}^{\sigma^s(t)}$, $\beta_t = a_{11}^{\sigma^s(t)}$. Let $\varphi_t(z)$ be the solution of the differential equation

$$\dot{z}_t = \alpha_t - \beta_t e^{z_t}, \quad z_0 = z. \tag{C.3}$$

Then, for all $z, z' \geqslant \ln(p_0)$, and all $t \geqslant 0$, we have

$$|\varphi_t(z) - \varphi_t(z')| \leq \exp\left\{-p_0 \int_0^t \beta_u \, du\right\} |z - z'|. \tag{C.4}$$

Indeed, the associated variational equation is

$$\frac{d}{dt}(\partial_z \varphi_t(z)) = -\beta_u(\partial_z \varphi_t(z)) e^{\varphi_t(z)},$$

and thus

$$\partial_z \varphi_t(z) = \exp\left(-\int_0^t \beta_u e^{\varphi_u(z)} \, du\right) \leq \exp\left(-p_0 \int_0^t \beta_u \, du\right),$$

where the inequality comes from the fact that $z \geqslant \ln(p_0)$ implies $\varphi_t(z) \geqslant \ln(p_0)$ for all $t \geqslant 0$.

Now, consider the differential equation

$$\dot{x}_t = x_t(\alpha_t - \beta_t x_t), \quad x_0 = x, \tag{C.5}$$

and set $z_t = \ln(x_t)$. Then, z_t solves (C.3) with initial condition $\ln(x)$. Denoting by $\psi_t(x)$ the solution to (C.5), using (C.4), we have, for all $x, y \in [p_0, p_1]$ and all $t \geqslant 0$,

$$|\psi_t(x) - \psi_t(y)| = |e^{\varphi_t(\ln(x))} - e^{\varphi_t(\ln(y))}| \leq p_1 |\varphi_t(x) - \varphi_t(y)|$$

$$\leq p_1 \exp\left\{-p_0 \int_0^t \beta_u \, du\right\} |\ln(x) - \ln(y)|$$

$$\leq \frac{p_1}{p_0} \exp\left\{-p_0 \int_0^t \beta_u \, du\right\} |x - y|.$$

This entails the result since $X_t^{x,s}$ and $X_t^{y,s}$ solve (C.5) with the respective initial conditions x and y. □

We now tackle the case where the initial conditions of σ in the two chains are distinct: $s \neq s' \in \mathcal{S}$. Let $t > 0$. Thanks to Assumption 1.1, there exists a coupling $(\sigma_{t/2}^s, \sigma_{t/2}^{s'})$ of $\delta_s \mathcal{Q}_{t/2}$ and $\delta_{s'} \mathcal{Q}_{t/2}$ with $\mathbb{P}(\sigma_{t/2}^s \neq \sigma_{t/2}^{s'}) \leqslant C e^{-\gamma t}$. Fix $x, y \in [p_0, p_1]$, and let $X_{t/2}^{x,s}$ be such that $(\sigma_{t/2}^s, X_{t/2}^{x,s})$ is distributed according to $\delta_{x,s} P_{t/2}^\varepsilon$ (which is possible since the regular version of the conditional laws exists in Polish spaces; see, e.g., [19, Theorem 5.3]), and similarly for $X_{t/2}^{y,s'}$. Then, we define $(\sigma_r^s, X_r^{x,s}, \sigma_r^{s'}, X_r^{y,s'})_{r \geqslant t/2}$ as follows: under the event $\{\sigma_{t/2}^s = \sigma_{t/2}^{s'}\}$, we let $(\sigma_r^s)_{r \geqslant t/2}$ be a Markov process associated to $(\mathcal{Q}_r)_{r \geqslant 0}$ with initial condition $\sigma_{t/2}^s$, and we set $\sigma_r^{s'} = \sigma_r^s$ for all $r \geqslant t/2$; otherwise, we define these processes as two independent processes associated to $(\mathcal{Q}_r)_{r \geqslant 0}$ with respective initial conditions $\sigma_{t/2}^s$ and $\sigma_{t/2}^{s'}$. In both cases, we define $(X_r^{x,s}, X_r^{y,s'})_{r \geqslant t/2}$ as solutions to

$$X_r^{x,s} = X_{t/2}^{x,s} + \int_{t/2}^r F_{\sigma^s(u)}(X_u^{x,s}) \, du, \qquad X_r^{y,s'} = X_{t/2}^{y,s'} + \int_{t/2}^r F_{\sigma^{s'}(u)}(X_u^{y,s'}) \, du.$$

Then $\left(\sigma_t^s, X_t^{x,s}, \sigma_t^{s'}, X_t^{y,s'}\right)$ is a coupling of $\delta_{(x,s)} P_t^\varepsilon$ and $\delta_{(y,s)} P_t^\varepsilon$ with, using (C.2) and that $\left| X_{t/2}^{x,s} - X_{t/2}^{x,s} \right| \leqslant p_1 - p_0$ almost surely,

$$
\begin{aligned}
\mathbb{E}\left(d\left((X_t^{x,s}, \sigma_t^s), (X_t^{y,s'}, \sigma_t^{s'})\right)\right) &= \mathbb{P}\left(\sigma_{t/2}^s \neq \sigma_{t/2}^{s'}\right) + \mathbb{E}\left(\frac{|X_t^{x,s} - X_t^{y,s'}|}{p_1 - p_0} \mathbb{1}_{\sigma_{t/2}^s = \sigma_{t/2}^{s'}}\right) \\
&\leq C \mathrm{e}^{-\gamma t/2} + \frac{p_1}{p_0} \mathrm{e}^{-\eta t/2} \\
&\leq \left(C + \frac{p_1}{p_0}\right) \mathrm{e}^{-(\gamma \wedge \eta)t/2} d\left((x,s), (y, s')\right).
\end{aligned}
$$

As a conclusion, we have obtained that (C.1) holds for all $(x, s), (y, s') \in [p_0, p_1] \times \mathcal{S}$ for some constants $C, \delta > 0$, which completes the proof.   □

Acknowledgements

We wish to thank two anonymous referees for their useful comments.

Funding information

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] AGAZZI, A., MATTINGLY, J. C. AND MELIKECHI, O. (2023). Random splitting of fluid models: Unique ergodicity and convergence. *Commun. Math. Phys.* **401**, 497–549.

[2] ARNOLD, L., CRAUEL, H. AND ECKMANN, J.-P. (eds) (2006). *Lyapunov Exponents*. Springer, Berlin.

[3] BEN-ISRAEL, A. AND GREVILLE, T. N. E. (2003). *Generalized Inverses: Theory and Applications*, 2nd ed. Springer, New York.

[4] BENAÏM, M. (2018). Stochastic persistence. Preprint, arXiv:1806.08450.

[5] BENAÏM, M., LE BORGNE, S., MALRIEU, F. AND ZITT, P.-A. (2014). On the stability of planar randomly switched systems. *Ann. Appl. Prob.* **24**, 292–311.

[6] BENAÏM, M. AND LOBRY, C. (2016). Lotka–Volterra with randomly fluctuating environments or 'how switching between beneficial environments can make survival harder'. *Ann. Appl. Prob.* **26**, 3754–3785.

[7] BENAÏM, M., LOBRY, C., SARI, T. AND STRICKLER, E. (2023). A note on the top Lyapunov exponent of linear cooperative systems. Preprint, arXiv:2302.05874.

[8] BENAÏM, M. AND STRICKLER, E. (2019). Random switching between vector fields having a common zero. *Ann. Appl. Prob.* **29**, 326–375.

[9] CHITOUR, Y., MAZANTI, G., MONMARCHÉ, P. AND SIGALOTTI, M. (2021). On the gap between deterministic and probabilistic Lyapunov exponents for continuous-time linear systems. Preprint, arXiv:2112.07005.

[10] CHUESHOV, I. (2002). *Monotone Random Systems Theory and Applications* (Lect. Notes Math. **1779**). Springer, Berlin.

[11] DAVIS, M. H. A. (1984). Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models. *J. R. Statist. Soc, B* **46**, 353–388.

[12] DU, N. H., HENING, A., NGUYEN, D. H. AND YIN, G. (2021). Dynamical systems under random perturbations with fast switching and slow diffusion: Hyperbolic equilibria and stable limit cycles. *J. Differential Equat.* **293**, 313–358.

[13] ELLIOTT, R. J. AND SIU, T. K. (2009). On Markov-modulated exponential-affine bond price formulae. *Appl. Math. Finance* **16**, 1–15.

[14] FAGGIONATO, A., GABRIELLI, D. AND RIBEZZI CRIVELLARI, M. (2009). Non-equilibrium thermodynamics of piecewise deterministic Markov processes. *J. Statist. Phys.* **137**, 259–304.

[15] FAINSHIL, L., MARGALIOT, M. AND CHIGANSKY, P. (2009). On the stability of positive linear switched systems under arbitrary switching laws. *IEEE Trans. Automatic Control* **54**, 897–899.

[16] GODDARD, B., OTTOBRE, M., PAINTER, K. AND SOUTTAR, I. (2023). On the study of slow-fast dynamics, when the fast process has multiple invariant measures. Preprint, arXiv:2305.04632.

[17] GURVITS, L., SHORTEN, R. AND MASON, O. (2007). On the stability of switched positive linear systems. *IEEE Trans. Automatic Control* **52**, 1099–1103.

[18] HATZIKIROU, H., KAVALLARIS, N. I. AND LEOCATA, M. (2021). A novel averaging principle provides insights in the impact of intratumoral heterogeneity on tumor progression. *Mathematics* **9**, 2530.

[19] KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd ed. Springer, New York.

[20] KHAROUFEH, J. P. AND COX, S. M. (2005). Stochastic models for degradation-based reliability. *IIE Trans.* **37**, 533–542.

[21] LAWLEY, S. D., MATTINGLY, J. C. AND REED, M. C. (2014). Sensitivity to switching rates in stochastically switched ODEs. *Commun. Math. Sci.* **12**, 1343–1352.

[22] MALRIEU, F. AND ZITT, P.-A. (2017). On the persistence regime for Lotka–Volterra in randomly fluctuating environments. *ALEA Lat. Am. J. Prob. Math. Stat.* **14**, 733–749.

[23] MONMARCHÉ, P., SCHREIBER, S. J. AND STRICKLER, É. (2024). Impacts of tempo and mode of environmental fluctuations on population growth: Slow- and fast-limit approximations of Lyapunov exponents for periodic and random environments. Preprint, arXiv:2408.11179.

[24] PAKDAMAN, K., THIEULLEN, M. AND WAINRIB, G. (2012). Asymptotic expansion and central limit theorem for multiscale piecewise-deterministic Markov processes. *Stoch. Process. Appl.* **122**, 2292–2318.

[25] TALAY, D. AND TUBARO, L. (1990). Expansion of the global error for numerical schemes solving stochastic differential equations. *Stoch. Anal. Appl.* **8**, 483–509.