Title: Exponential ergodicity in the bounded-Lipschitz distance for some piecewise-deterministic Markov processes with random switching between flows

Author: Dawid Czapla, Katarzyna Horbacz, HannaWojewódka-Ściążko

Citation style: Czapla Dawid, Horbacz Katarzyna, Wojewódka-Ściążko Hanna. (2022). Exponential ergodicity in the bounded-Lipschitz distance for some piecewise-deterministic Markov processes with random switching between flows. "Nonlinear Analysis" Vol. 215 (2022), s. 1-32, art. no. 112678, doi: 10.1016/j.na.2021.112678

Nonlinear Analysis 215 (2022) 112678

Contents lists available at ScienceDirect

Nonlinear Analysis

www.elsevier.com/locate/na

Exponential ergodicity in the bounded-Lipschitz distance for some piecewise-deterministic Markov processes with random switching between flows

Dawid Czapla [a,*], Katarzyna Horbacz [a], Hanna Wojewódka-Ściążko [a,b]

[a] Institute of Mathematics, University of Silesia in Katowice, Bankowa 14, Katowice, 40-007, Poland
[b] Institute of Theoretical and Applied Informatics, Polish Academy of Sciences, Bałtycka 5, Gliwice, 44-100, Poland

ARTICLE INFO

Article history:
Received 2 February 2021
Accepted 25 October 2021
Communicated by Francesco Maggi

MSC:
60J25
60J05
37A30
37A25

Keywords:
Piecewise-deterministic Markov process
Switching semiflows
Exponential ergodicity
Fortet–Mourier distance
Coupling
Gene expression

ABSTRACT

In this paper, we study a subclass of piecewise-deterministic Markov processes with a Polish state space, involving deterministic motion punctuated by random jumps that occur at exponentially distributed time intervals. Over each of these intervals, the process follows a flow, selected randomly among a finite set of all possible ones. Our main goal is to provide a set of verifiable conditions guaranteeing the exponential ergodicity for such processes (in terms of the bounded Lipschitz distance), which would refer only to properties of the flows and the transition law of the Markov chain given by the post-jump locations. Moreover, we establish a simple criterion on the exponential ergodicity for a particular instance of these processes, applicable to certain biological models, where the jumps result from the action of an iterated function system with place-dependent probabilities.

1. Introduction

Piecewise-deterministic Markov processes (PDMPs), first introduced by Davis [18] in 1984 (cf. also [19,20]), constitute a general class of non-diffusive Markov processes, for which randomness stems only from the jump mechanism, including the jumping times, the post-jump locations and other changes occurring at the moments of jumps. This huge family of processes is extensively used for modelling purposes in many applied subjects, like biology [8,9,13,34,39], storage modelling [7] or internet traffic [24].

In this paper, we are concerned with the PDMPs that evolve on a Polish space through jumps arriving according to a Poisson process. This means that the span of time between consecutive jumps is exponentially

* Corresponding author.
E-mail addresses: dawid.czapla@us.edu.pl (D. Czapla), katarzyna.horbacz@us.edu.pl (K. Horbacz), hanna.wojewodka@us.edu.pl (H. Wojewódka-Ściążko).

https://doi.org/10.1016/j.na.2021.112678

distributed with a constant rate λ. Between any two adjacent jumps, the dynamics of these processes is driven by one of the semiflows, randomly selected from a finite set $\{S_i : i \in I\}$ of possible ones, according to a given stochastic matrix $[\pi_{ij}]_{i,j\in I}$. The state right after a jump (usually called the post-jump location) depends randomly on the one immediately preceding this jump, and its probability distribution is governed by a given Markov transition function (a stochastic kernel) $(y, B) \mapsto J(y, B)$.

More specifically, given an arbitrary Polish metric space Y, we shall investigate a stochastic process $\Psi := \{(Y(t), \xi(t))\}_{t\geq 0}$ with values in $X := Y \times I$, whose motion can be described as follows. Starting from some initial value (y_0, i_0), the process evolves deterministically in such a way that $\{Y(t)\}_{t\geq 0}$ follows $t \mapsto S_{i_0}(t, y_0)$ until the first jump time, say $t_1 > 0$. At this moment the trajectory of the first coordinate jumps to another point of Y, say y_1, so that the probability it will fall into a Borel set $B \subset Y$ is $J(S_{i_0}(t_1, y_0), B)$. At the same time the index of the "active" semiflow, determined by $\{\xi(t)\}_{t\geq 0}$, is randomly switched from i_0 to another (or the same) one i_1 with probability $\pi_{i_0 i_1}$. Then the motion restarts from the new state (y_1, i_1) and proceeds as before. Formally, the process Ψ can be therefore defined by setting

$$Y(t) := S_{\xi_n}(t - \tau_n, Y_n) \quad \text{and} \quad \xi(t) := \xi_n \quad \text{for} \quad t \in [\tau_n, \tau_{n+1}), \;\; n \in \mathbb{N}_0,$$

where $\bar{\Phi} := \{(Y_n, \xi_n, \tau_n)\}_{n\in\mathbb{N}\cup\{0\}}$ is a time-homogeneous Markov chain with state space $X \times [0, \infty)$ and transition law satisfying

$$\mathbb{P}(\bar{\Phi}_{n+1} \in B \times \Xi \times T |\, \bar{\Phi}_n = (y, i, s)) = \sum_{j\in\Xi} \pi_{ij} \int_{T\cap\,[s,\infty)} \lambda e^{-\lambda(t-s)} J(S_i(t, y), B)\, dt$$

for any $n \in \mathbb{N} \cup \{0\}$, $y \in Y, i \in I, s \geq 0$ and Borel sets $B \subset Y$, $\Xi \subset I$, $T \subset [0, \infty)$. Obviously, all the randomness of the PDMP Ψ is contained in the chain $\bar{\Phi}$. What is more, the sequence $\Phi := \{(Y_n, \xi_n)\}_n$ of the post-jump locations itself is an X-valued Markov chain (with respect to its natural filtration). Clearly, on the family of rectangles $B \times \Xi$ (where $B \subset Y$ is a Borel set, and $\Xi \subset I$), the transition law of this chain takes the form

$$P((y, i), B \times \Xi) := \mathbb{P}(\Phi_{n+1} \in B \times \Xi \mid \Phi_n = (y, i)) = \sum_{j\in\Xi} \pi_{ij} \int_0^\infty \lambda e^{-\lambda t} J(S_i(t, y), B)\, dt.$$

The subclass of the PDMPs considered here somewhat resembles those investigated in [1–4,11,12]. All these papers, however, focus on processes evolving on finite-dimensional (and thus locally compact) spaces. While proving the existence of invariant distributions and ergodicity (usually in the total variation norm) in such a setup, one can use various adaptations of conventional methods of Meyn and Tweedie [36,37], based mainly on the Harris recurrence (assured e.g. by Hörmander-type bracket conditions, just as in [3]) or some criteria referring to the so-called drift towards a petite set. These techniques, however, are mostly valid only for ψ-irreducible processes, which is, obviously, not the case in our framework. On the other hand, [10], for instance, deals with a large class of regime switching Markov processes (a much more general family than that of PDMPs), which take values in a Polish space. Nevertheless, the criteria on the exponential ergodicity (in the Wasserstein distance) provided in that work are based on fairly general assumptions, such as the "exponential contractivity" of the given Markov semigroups or a Lyapunov–Foster type condition in the continuous-time context, which might be difficult to verify in practice (at least in a direct way).

The main goal of this paper is to provide relatively easy to check conditions on the kernel J and the semiflows S_i which would guarantee that both the transition operator of the chain Φ and the transition semigroup of the process Ψ are exponentially ergodic in the bounded Lipschitz distance (equivalent to the one induced by the Dudley norm [22]). Such a metric, also known as the Fortet–Mourier distance (see e.g. [31,33]), is defined on the cone of non-negative finite Borel measures on X, and induces the topology of weak convergence of such measures [6]. Roughly speaking, the aforementioned form of ergodicity means that the process under consideration admits a unique stationary (invariant) distribution, to which its distribution converges at an exponential rate in the Fortet–Mourier distance, independently of the initial state. The rigorous meaning of this term is given in Definitions 2.1 and 2.2. The general strategy of our approach is as follows:

(I) We begin with showing that, whenever J enjoys some strengthened form of the Feller property, there exists a one-to-one correspondence between the set of invariant distributions of the process Ψ and those of the associated chain Φ (Theorem 5.1).

(II) Next, we note that the existence of an appropriate coupling $(\Phi^{(1)}, \Phi^{(2)})$ between two copies of Φ, such that the mean distance between them decreases geometrically with time, in conjunction with the so-called Foster–Lyapunov drift condition (see, e.g., [21, Definition 6.23]) and the Feller property imposed on P, ensures the exponential ergodicity of Φ (Lemma 6.1).

(III) The essential step in our analysis is proving that, for a given coupling $(\Phi^{(1)}, \Phi^{(2)})$ of the chain Φ enjoying the property indicated in (II), the corresponding coupling of the process Ψ has an analogous property, provided that the semiflows S_i fulfil a certain Lipschitz-type condition (Lemma 6.2). The key idea here is partially inspired by the techniques used in the proof of [10, Theorem 1.4].

(IV) From the results discussed in steps (I)–(III) we can conclude that, under suitable assumptions on the semiflows S_i and the kernel J, providing all the requirements mentioned above, the existence of an appropriate coupling of Φ implies the exponential ergodicity of the process Ψ (Theorem 6.1).

(V) Finally, we employ some additional hypotheses which, together with the previous ones, ensure that the coupling mentioned in (II) exists. This leads us to the main result of the paper, stated as Theorem 7.1. In particular, at this stage we require the existence of a substochastic kernel Q_J on Y^2 with certain specific properties (in the spirit of [15,29]), such that

$$Q_J((y_1, y_2), \cdot \times Y) \le J(y_1, \cdot) \quad \text{and} \quad Q_J((y_1, y_2), Y \times \cdot) \le J(y_2, \cdot),$$

which further enables us to construct a substochastic kernel Q_P on X^2, having the analogous properties with respect to P (Lemma 7.1). The transition function of the desired coupling can be then defined as the sum of Q_P and a suitable complementary kernel (Proposition 7.1). Such a construction is inspired by the ideas of Hairer [25], regarding the so-called asymptotic coupling technique (also used e.g. in [41,43]).

What is especially noteworthy here is the fact that this approach also elucidates the way in which the exponential ergodicity of the PDMP Ψ is inherited from the same property for the associated chain Φ. This is visible in steps (I) and (III).

The obtained general result (i.e. Theorem 7.1) is further applied to derive a simple criterion on the exponential ergodicity (in the Fortet–Mourier distance) in the case where the jump kernel J is a transition law of a random iterated function system (Proposition 7.2). This is done by taking advantage of the fact that the kernel Q_J, playing a key role in step (V), can be defined explicitly in such a model. More specifically, we discuss the case in which J is given by

$$J(y, B) = \int_\Theta \mathbb{1}_B(w_\theta(y)) p_\theta(y)\, \vartheta(d\theta) \quad \text{for each } \; y \in Y \; \text{ and any Borel set } \; B \subset Y,$$

where $\{w_\theta : \theta \in \Theta\}$ is an arbitrary family of continuous transformations from Y to itself, indexed by the elements of a measure space (Θ, ϑ), and $\Theta \ni \theta \mapsto p_\theta(y)$, $y \in Y$, is the associated set of state-dependent probability density functions with respect to ϑ. In this setting, the model under consideration may serve as a framework for analysing the dynamics of gene expression in prokaryotes (see e.g. [5,14,34]), discussed in more detail within Example 7.3. Moreover, if $\vartheta(\Theta) = 1$ and $p_\theta \equiv 1$ for every $\theta \in \Theta$, then $\{Y(t)\}_{t\ge 0}$ can be treated as the solution to a stochastic evolution equation with Poisson noise (see, e.g., [17,26,30,32,35]). An interpretation of Proposition 7.2 in this setup is presented in Example 7.2.

The discrete-time dynamical system Φ with the above-specified shape of the jump kernel J, even in a more general setting, wherein the probabilities π_{ij} depend on the state, has been more widely examined (in terms of ergodicity and classical limit theorems) in our previous articles [14–17]. For instance, in [14], the exponential ergodicity of Φ has been used to prove the strong law of large numbers for the chain $\{f(\Phi_n)\}_{n\in\mathbb{N}\cup\{0\}}$ (with a Lipschitz continuous function $f : X \to \mathbb{R}$), which, in turn, has enabled us to derive the analogous law for the process $\{f(\Psi(t))\}_{t\geq 0}$ (without using the ergodicity of Ψ). The result provided in the present paper should prove to be useful in establishing also the central limit theorem for this process, which would be rather difficult to achieve based only on the properties of Φ.

The organization of the paper is as follows. In Section 2, we introduce notation and some basic concepts regarding Markov semigroups acting on measures, including the employed definition of ergodicity. Section 3 provides a detailed description of the subclass of the PDMPs under study. In Section 4, we list and discuss all the assumptions underlying our main results. Section 5 is devoted to establishing a one-to-one correspondence between invariant distributions of the processes Ψ and Φ, that is, the realization of step (I). The essential part of our analysis, referring to the coupling argument, which has been described within steps (II)–(IV), is contained in Section 6. Step (V), including the construction of a suitable coupling for Φ, is included in Section 7. Finally, also in this part of the paper, we state the main result and discuss some special cases of the model, for which the jumps are determined by a random iterated function system.

2. Preliminaries

Consider a complete separable metric space (E, ρ), endowed with its Borel σ-field $\mathcal{B}(E)$. By $B_E(x, r)$ we will denote the open ball in E centred at x of radius $r > 0$. The symbol $\mathbb{1}_A$ will be used to denote the indicator function of a subset A of E (or any other space, which should be clear from the context). Additionally, we set $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ and $\mathbb{R}_+ := [0, \infty)$.

Let $B_b(E)$ stand for the Banach space of all real-valued, Borel measurable functions on E, equipped with the supremum norm $\|f\|_\infty := \sup_{x\in E} |f(x)|$. By $C_b(E)$ we shall denote the subspace of $B_b(E)$ consisting of all continuous functions. In addition to this, we also define the set $\mathrm{Lip}_{b,1}(E)$ as follows:

$$\mathrm{Lip}_{b,1}(E) := \left\{ f \in C_b(E) :\ 0 \leq f \leq 1,\ \sup_{x\neq y} \frac{|f(x) - f(y)|}{\rho(x, y)} \leq 1 \right\}.$$

Moreover, we will write $\mathcal{M}(E)$ and $\mathcal{M}_{prob}(E)$ to denote the cone of all finite non-negative, Borel measures on E, and its subset consisting of all probability measures, respectively. Further, given any Borel measurable function $V : E \to [0, \infty)$, we shall consider the subset $\mathcal{M}^V_{prob}(E)$ of $\mathcal{M}_{prob}(E)$ consisting of all measures with finite first moment with respect to V, that is,

$$\mathcal{M}^V_{prob}(E) := \left\{ \mu \in \mathcal{M}_{prob}(E) :\ \int_E V(x)\, \mu(dx) < \infty \right\}.$$

For brevity of notation, the Lebesgue integral $\int_E f\, d\mu$ of a Borel measurable function $f : E \to \mathbb{R}$ with respect to a signed Borel measure μ – if exists – will be sometimes denoted by $\langle f, \mu\rangle$. Furthermore, we will write δ_x for the Dirac measure at $x \in E$ on $\mathcal{B}(E)$.

To describe the distance between measures, we will use the Fortet–Mourier metric (equivalent to the metric induced by the Dudley norm [22]), which on $\mathcal{M}(E)$, is defined by

$$d_{FM,\rho}(\mu, \nu) := \sup_{f\in \mathrm{Lip}_{b,1}(E)} |\langle f, \mu - \nu\rangle| \quad \text{for any} \quad \mu, \nu \in \mathcal{M}(E).$$

It is well-known that, as long as E is separable (which is the case here), the metric $d_{FM,\rho}$ induces the topology of weak convergence of measures on $\mathcal{M}(E)$ (cf. [22, Theorems 6 and 8] or [6, Theorem 8.3.2]). Let us recall here that a sequence $\{\mu_n\}_{n\in\mathbb{N}} \subset \mathcal{M}(E)$ of measures is called weakly convergent to a measure $\mu \in \mathcal{M}(E)$ if $\langle f, \mu_n\rangle \to \langle f, \mu\rangle$, as $n \to \infty$, for any $f \in C_b(E)$. Moreover, if (E, ρ) is complete (which is also the case in our setting), then so is the space $(\mathcal{M}_{prob}(E), d_{FM,\rho})$ (see [22, Theorem 9]).

Before further discussion, it is also useful to recall several basic concepts in the theory of Markov operators.

A function $P : E \times \mathcal{B}(E) \to [0, 1]$ is called a *(sub)stochastic kernel* if for each $A \in \mathcal{B}(E)$, $x \mapsto P(x, A)$ is a Borel measurable map on E, and for each $x \in E$, $A \mapsto P(x, A)$ is a (sub)probability Borel measure on $\mathcal{B}(E)$. The composition of two such kernels, say P and Q, is defined by

$$PQ(x, A) := \int_X Q(y, A)P(x, dy) \quad \text{for any} \quad x \in E,\ A \in \mathcal{B}(E). \tag{2.1}$$

According to this rule, we can also define recursively the so-called n-step kernel P^n, by setting $P^1 := P$ and $P^{n+1} := P^n P$ for every $n \in \mathbb{N}$.

For any (sub)stochastic kernel P, we can consider two operators (which will be denoted by the same symbol), one acting on $\mathcal{M}(E)$, and the second one acting on $B_b(E)$, defined by

$$\mu P(A) := \int_E P(x, A)\, \mu(dx) \quad \text{for} \quad \mu \in \mathcal{M}(E),\ A \in \mathcal{B}(E), \tag{2.2}$$

$$Pf(x) := \int_E f(y)\, P(x, dy) \quad \text{for} \quad f \in B_b(E),\ x \in E. \tag{2.3}$$

Note that these operators are related to each other in the following way:

$$\langle f, \mu P\rangle = \langle Pf, \mu\rangle \quad \text{for any} \quad f \in B_b(E),\ \mu \in \mathcal{M}(E).$$

Obviously, the nth iterations $(\cdot)P^n$ and $P^n(\cdot)$ are induced by the n-step kernel P^n. If P is a stochastic kernel, then $P : \mathcal{M}(E) \to \mathcal{M}(E)$, given by (2.2), is called a *(regular) Markov operator*, and $P : B_b(E) \to B_b(E)$, defined by (2.3), is said to be its dual operator. Let us stress that formula (2.3) will be sometimes applied, with a slight abuse of notation, to unbounded above functions as well; for example, we shall write $P\rho(\cdot, x^*)$ (for a fixed $x^* \in E$).

A family of stochastic kernels $\{P_t\}_{t\in\mathbb{R}_+}$ (or the induced family of Markov operators) is called a *Markov semigroup* if $P_sP_t = P_{s+t}$ for any $s, t \geq 0$ in the sense of (2.1), and $P_0(x, \cdot) = \delta_x$ for every $x \in E$. In terms of Markov operators, this is obviously equivalent to saying that $P_s \circ P_t = P_{s+t}$ for all $s, t \geq 0$, and $(\cdot)P_0$ is the identity map.

Given a stochastic kernel P on $E \times \mathcal{B}(E)$ and $\mu \in \mathcal{M}_{prob}(E)$, by a *time-homogeneous Markov chain* with (one-step) transition law P and initial distribution μ we mean a sequence of E-valued random variables $\Phi := \{\Phi_n\}_{n\in\mathbb{N}_0}$, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P}_\mu)$, such that, for any $A \in \mathcal{B}(E)$ and $n \in \mathbb{N}$,

$$\mathbb{P}_\mu(\Phi_0 \in A) = \mu(A), \tag{2.4}$$

$$\mathbb{P}_\mu(\Phi_{n+1} \in A \mid \mathcal{F}_n) = \mathbb{P}(\Phi_{n+1} \in A \mid \Phi_n) = P(\Phi_n, A), \tag{2.5}$$

where $\mathcal{F}_n$ is the σ-field generated by $\Phi_0, \dots, \Phi_n$. The expectation operator with respect to $\mathbb{P}_\mu$ is then denoted by $\mathbb{E}_\mu$. In the case where $\mu = \delta_x$ with some $x \in E$, we simply write $\mathbb{P}_x$ and $\mathbb{E}_x$ rather than $\mathbb{P}_{\delta_x}$ and $\mathbb{E}_{\delta_x}$, respectively. Obviously $\mathbb{P}_x = \mathbb{P}_\mu(\cdot|\Phi_0 = x)$ for any $x \in E$.

One can easily check that, for every $k \in \mathbb{N}$, the k-step transition probabilities of Φ are determined by the kernels P^k, i.e. $\mathbb{P}(\Phi_{n+k} \in A \mid \Phi_n) = P^k(\Phi_n, A)$. Consequently, it follows that the Markov operator $(\cdot)P$ describes the evolution of the distribution of Φ, i.e. $\mu_{n+1} = \mu_n P$ for any $n \in \mathbb{N}_0$, where μ_n is the distribution of Φ_n. In this connection, it is reasonable to call $(\cdot)P$ the *transition operator* of Φ. Furthermore, it is also worth noting that the dual operator of $(\cdot)P^n$ can be expressed as

$$P^n f(x) = \mathbb{E}_x[f(\Phi_n)] \quad \text{for any} \quad x \in E,\ f \in B_b(E),\ n \in \mathbb{N}. \tag{2.6}$$

On the other hand, it is well-known that, for any stochastic kernel P on $E\times\mathcal{B}(E)$ and any $\mu\in\mathcal{M}_{prob}(E)$, on some probability space $(\Omega,\mathcal{F},\mathbb{P}_\mu)$, there exists a time-homogeneous Markov chain Φ with transition law P and initial measure μ (see e.g. [38]). In practice, it is convenient to assume that $\Omega:=E^{\mathbb{N}_0}$, $\mathcal{F}:=\mathcal{B}(E^{\mathbb{N}_0})$ (where $E^{\mathbb{N}_0}$ is endowed with the product topology), and that $\{\Phi_n\}_{n\in\mathbb{N}_0}$ is a sequence of canonical projections from Ω to E, that is, $\Phi_n(\omega)=x_n$ for any $\omega=(x_0,x_1,\ldots)\in\Omega$, with $x_0,x_1,\ldots\in E$. Then, for each $\mu\in\mathcal{M}_{prob}(E)$, one can construct a probability measure $\mathbb{P}_\mu$ on $\mathcal{F}$ such that

$$\mathbb{P}_\mu(F)=\int_E\int_E\ldots\int_E \mathbb{1}_{A_0\times\cdots\times A_n}(x_0,\ldots,x_n)P(x_{n-1},dx_n)\ldots P(x_0,dx_1)\mu(dx_0) \tag{2.7}$$

for any $n\in\mathbb{N}_0$ and $F=\{\Phi_0\in A_0,\ldots,\Phi_n\in A_n\}$, where $A_0,\ldots,A_n\in\mathcal{B}(E)$. It then follows easily that Φ obeys (2.4) and (2.5) for every $\mu\in\mathcal{M}_{prob}(E)$, and, what is more, we have

$$\mathbb{P}_\mu(F)=\int_X \mathbb{P}_x(F)\,\mu(dx)\quad\text{for any}\quad F\in\mathcal{F},\ \mu\in\mathcal{M}_{prob}(E). \tag{2.8}$$

The Markov chain constructed in this way is called a *canonical* one.

Given a Markov semigroup $\{P_t\}_{t\in\mathbb{R}_+}$ of stochastic kernels on $E\times\mathcal{B}(E)$ and a measure $\mu\in\mathcal{M}_{prob}(E)$, by a *time-homogeneous Markov process* with transition semigroup $\{P_t\}_{t\in\mathbb{R}_+}$ and initial distribution μ we mean a family of E-valued random variables $\Psi:=\{\Psi(t)\}_{t\in\mathbb{R}_+}$ on some probability space $(\Omega,\mathcal{F},\mathbb{P}_\mu)$ such that, for any $A\in\mathcal{B}(E)$ and $s,t\geq 0$,

$$\begin{gathered}\mathbb{P}_\mu(\Psi(0)\in A)=\mu(A),\\ \mathbb{P}_\mu(\Psi(s+t)\in A\mid\mathcal{F}(s))=\mathbb{P}_\mu(\Psi(s+t)\in A\mid\ \Psi(s))=P_t(\Psi(s),A),\end{gathered} \tag{2.9}$$

where $\mathcal{F}(s)$ is the σ-field generated by $\{\Psi(h):h\leq s\}$. From (2.9) it obviously follows that $\mu(s+t)=\mu(s)P_t$ for any $s,t\geq 0$, where $\mu(t)$ stands for the distribution of $\Psi(t)$ for every $t\geq 0$. Moreover, analogously as in the discrete case, the dual operator of P_t can be expressed as

$$P_tf(x)=\mathbb{E}_x[f(\Psi(t))]\quad\text{for any}\quad x\in E,\ f\in B_b(E),\ t\geq 0. \tag{2.10}$$

Let us now briefly recall some notions concerning the ergodicity of Markov operators, which will be used throughout the paper.

First of all, a Markov operator P is called *Feller* if its dual operator preserves continuity, that is, $P(C_b(E))\subset C_b(E)$. Furthermore, a Markov semigroup $\{P_t\}_{t\in\mathbb{R}_+}$ is called *Feller* if P_t is Feller for any $t\geq 0$.

A measure $\mu_*\in\mathcal{M}(E)$ is said to be *invariant* for a Markov operator P if $\mu_*P=\mu_*$. By analogy, we say that $\tilde{\mu}_*\in\mathcal{M}(E)$ is *invariant* for a Markov semigroup $\{P_t\}_{t\in\mathbb{R}_+}$ whenever $\tilde{\mu}_*P_t=\tilde{\mu}_*$ for every $t\geq 0$.

We finalize this section with the definitions of two properties that will be verified in the main results of this paper.

Definition 2.1. Let P be a transition operator of an E-valued Markov chain Φ. Given a Borel measurable function $V:E\to[0,\infty)$, we shall say that P (or the chain Φ) is V-*exponentially ergodic* in $d_{FM,\rho}$ if it admits a unique invariant probability measure μ_*^{Φ}, such that $\mu_*^{\Phi}\in\mathcal{M}_{prob}^V(E)$, and there exists a constant $q\in(0,1)$ such that, for every $\mu\in\mathcal{M}_{prob}^V(E)$ and some $C(\mu)<\infty$, we have

$$d_{FM,\rho}\left(\mu P^n,\mu_*^{\Phi}\right)\leq C(\mu)q^n\quad\text{for any}\quad n\in\mathbb{N}.$$

Definition 2.2. Let $\{P_t\}_{t\in\mathbb{R}_+}$ be a transition semigroup of an E-valued Markov process Ψ. Given a Borel measurable function $V:E\to[0,\infty)$, we shall say that $\{P_t\}_{t\in\mathbb{R}_+}$ (or the process Ψ) is V-*exponentially ergodic* in $d_{FM,\rho}$ if it admits a unique invariant probability measure μ_*^{Ψ}, such that $\mu_*^{\Psi}\in\mathcal{M}_{prob}^V(E)$, and there exists a constant $\gamma>0$ such that, for every $\mu\in\mathcal{M}_{prob}^V(E)$ and some $\bar{C}(\mu)<\infty$, we have

$$d_{FM,\rho}\left(\mu P_t,\mu_*^{\Psi}\right)\leq\bar{C}(\mu)e^{-\gamma t}\quad\text{for any}\quad t\geq 0.$$

3. The model under study

Let (Y, ρ_Y) be a complete separable metric space, and let I be a finite set endowed with the discrete metric $\mathbf{d}$, i.e. $\mathbf{d}(i,j) = 1$ if $i \neq j$ and $\mathbf{d}(i,j) = 0$ otherwise. In what follows, we shall also refer to the spaces

$$X := Y \times I \quad \text{and} \quad \bar{X} := X \times \mathbb{R}_+,$$

considered with the product topologies. Additionally, we assume that X is endowed with a metric $\rho_{X,c}$ of the form

$$\rho_{X,c}(x_1, x_2) = \rho_Y(y_1, y_2) + c\,\mathbf{d}(i_1, i_2) \quad \text{for} \quad x_1 = (y_1, i_1),\, x_2 = (y_2, i_2) \in X, \tag{3.1}$$

where c is a given positive constant, whose value will be relevant in Section 7. The Fortet–Mourier distance in $\mathcal{M}(X)$, induced by the metric $\rho_{X,c}$, will be simply denoted by $d_{FM,c}$ (rather than $d_{FM,\rho_{X,c}}$). Throughout the paper, we will also refer to the standard bounded metric induced by $\rho_{X,c}$, that is,

$$\bar{\rho}_{X,c}(x_1, x_2) := \rho_{X,c}(x_1, x_2) \wedge 1 \quad \text{for any} \quad x_1, x_2 \in X, \tag{3.2}$$

where $\wedge$ stands for the minimum.

Consider a collection $\{S_i : i \in I\}$ of jointly continuous semiflows acting from $\mathbb{R}_+ \times Y$ to Y. By calling S_i a semiflow we mean, as usual, that

$$S_i(s, S_i(t, y)) = S_i(s + t, y) \quad \text{and} \quad S_i(0, y) = y \quad \text{for any} \quad s, t \in \mathbb{R}_+,\; y \in Y.$$

Furthermore, suppose that we are given a right stochastic matrix $\{\pi_{ij} : i, j \in I\}$, i.e.

$$\pi_{ij} \in \mathbb{R}_+ \;\text{ for any }\; i, j \in I, \quad \text{and} \quad \sum_{j \in I} \pi_{ij} = 1 \;\text{ for every }\; i \in I,$$

a positive constant λ, as well as an arbitrary stochastic kernel $J : Y \times \mathcal{B}(Y) \to [0,1]$.

Let us now define a stochastic kernel $\bar{P} : \bar{X} \times \mathcal{B}(\bar{X}) \to [0,1]$ by setting

$$\bar{P}((y,i,s), \bar{A}) = \sum_{j \in I} \pi_{ij} \int_0^\infty \lambda e^{-\lambda h} \int_Y \mathbb{1}_{\bar{A}}(u, j, h+s)\, J(S_i(h,y), du)\, dh \tag{3.3}$$

for any $y \in Y$, $i \in I$, $s \in \mathbb{R}_+$ and $\bar{A} \in \mathcal{B}(\bar{X})$. Moreover, let $P : X \times \mathcal{B}(X) \to [0,1]$ denote the kernel given by

$$P((y,i), A) := \bar{P}((y,i,0), A \times \mathbb{R}_+) \quad \text{for} \quad y \in Y,\; i \in I,\; A \in \mathcal{B}(X), \tag{3.4}$$

where $\bar{P}$ is given by (3.3).

Remark 3.1. Taking into account the continuity of the maps $Y \ni y \mapsto S_i(t,y)$, $t \geq 0$, $i \in I$, it is easy to see that P is Feller whenever so is the kernel J.

By $\bar{\Phi} := \{(Y_n, \xi_n, \tau_n)\}_{n \in \mathbb{N}_0}$ we will denote a time-homogeneous Markov chain with state space $\bar{X}$ and transition law $\bar{P}$, wherein Y_n, ξ_n, τ_n take values in Y, I, $\mathbb{R}_+$, respectively. More precisely, $\bar{\Phi}$ will be regarded as the canonical Markov chain, constructed on the coordinate space $\Omega := \bar{X}^{\mathbb{N}_0}$, equipped with the σ-field $\mathcal{F} := \mathcal{B}\left(\bar{X}^{\mathbb{N}_0}\right)$ and a suitable family $\{\mathbb{P}_\nu : \nu \in \mathcal{M}_{prob}(\bar{X})\}$ of probability measures on $\mathcal{F}$, where the subscript ν indicates the initial distribution of $\bar{\Phi}$. For every $\nu \in \mathcal{M}_{prob}(\bar{X})$, we therefore have

$$\mathbb{P}_\nu(\bar{\Phi}_0 \in \bar{A}) = \nu(\bar{A}) \quad \text{for any} \quad \bar{A} \in \mathcal{B}(\bar{X}),$$
$$\mathbb{P}_\nu(\bar{\Phi}_{n+1} \in \bar{A} \mid \bar{\Phi}_n = (y,i,s)) = \bar{P}((y,i,s), \bar{A}) \quad \text{for any} \quad (y,i,s) \in \bar{X},\; \bar{A} \in \mathcal{B}(\bar{X}),\; n \in \mathbb{N}_0.$$

Obviously, the sequences $\Phi := \{(Y_n, \xi_n)\}_{n\in\mathbb{N}_0}$, $\{\xi_n\}_{n\in\mathbb{N}_0}$ and $\{\tau_n\}_{n\in\mathbb{N}_0}$ are Markov chains with respect to their own natural filtrations, and their transition laws satisfy

$$\mathbb{P}_\nu(\Phi_{n+1} \in A \mid \Phi_n = (y,i)) = P((y,i),A) \quad \text{for} \quad (y,i) \in X,\ A \in \mathcal{B}(X),$$

$$\mathbb{P}_\nu(\xi_{n+1} = j \mid \xi_n = i) = \pi_{ij} \quad \text{for} \quad i,j \in I \tag{3.5}$$

$$\mathbb{P}_\nu(\tau_{n+1} \le t \mid \tau_n = s) = \mathbb{1}_{[s,\infty)}(t)\left(1 - e^{-\lambda(t-s)}\right) \quad \text{for} \quad s,t \in \mathbb{R}_+. \tag{3.6}$$

Moreover, note that the increments $\Delta\tau_n := \tau_n - \tau_{n-1}$, $n \in \mathbb{N}$, form a sequence of independent, exponentially distributed random variables with the same rate parameter λ, and thus $\tau_n \uparrow \infty$, as $n \to \infty$, $\mathbb{P}_\nu$-a.s. (due to the strong law of large numbers).

The main focus of our study will be a PDMP $\Psi := \{(Y(t), \xi(t))\}_{t\in\mathbb{R}_+}$ with jump times τ_n, $n \in \mathbb{N}_0$, defined via interpolation of the chain Φ, so that:

$$Y(t) := S_{\xi_n}(t - \tau_n, Y_n), \quad \xi(t) = \xi_n \quad \text{whenever} \quad t \in [\tau_n, \tau_{n+1}) \ \text{ for } \ n \in \mathbb{N}_0. \tag{3.7}$$

The transition semigroup of this process will be denoted by $\{P_t\}_{t\in\mathbb{R}_+}$. Obviously, the discrete-time model Φ with transition law P, determined by (3.4), can be viewed as the Markov chain given by the post-jump locations of Ψ, since

$$\Phi_n = (Y_n, \xi_n) = (Y(\tau_n), \xi(\tau_n)) = \Psi(\tau_n) \quad \text{for every} \quad n \in \mathbb{N}_0.$$

Looking at the shape of the kernel $\bar{P}$, one can say (somewhat informally) that the probability of visiting a given set B right after the $(n+1)$th jump, given $S_{\xi_n}(\Delta\tau_{n+1}, Y_n) = y$, is equal to $J(y,B)$.

Remark 3.2. As has been already mentioned in the introduction, the above-described model (Φ, Ψ) is a generalization of that considered in [14] (apart from the probabilities π_{ij}, which are constant here); cf. also [17]. More specifically, in [14], the kernel J is a transition law of some randomly perturbed iterated function system, i.e. it has the form:

$$J(y,B) = \int_{\operatorname{supp}\nu} \int_{\Theta} \mathbb{1}_B(w_\theta(y) + v)\, p_\theta(y)\, \vartheta(d\theta)\nu(dv) \quad \text{for} \quad B \in \mathcal{B}(Y).$$

In that case, Y is a closed subset of a Banach space H, $\nu \in \mathcal{M}_{prob}(H)$ is a probability measure with bounded support, Θ stands for an arbitrary topological space, endowed with a Borel measure ν, $\{w_\theta : \theta \in \Theta\}$ is a given family of continuous transformations from Y to itself, such that $w_\theta(Y) + v \subset Y$ for any $v \in \operatorname{supp}\nu$, and $\Theta \ni \theta \mapsto p_\theta(y) \in \mathbb{R}_+$, $y \in Y$, are the associated (state-dependent) probability density functions with respect to ϑ. We shall come back to this particular case in Section 7.2.

4. Assumptions on the model components

Let us begin this section by listing all the conditions that will be used throughout the remainder of the paper. The list can be naturally divided into two parts. The first one contains assumptions referring to the deterministic part of the model (i.e., the flows S_i and the probabilities π_{ij}), which read as follows:

(S1) For some $y^* \in Y$, we have

$$\int_0^\infty e^{-\lambda t}\rho_Y(S_i(t,y^*), y^*)\, dt < \infty \quad \text{for any} \quad i \in I. \tag{4.1}$$

(S2) There exist $L > 0$ and $\alpha \in \mathbb{R}$ such that

$$\rho_Y(S_i(t,y_1), S_i(t,y_2)) \le Le^{\alpha t}\rho_Y(y_1,y_2) \quad \text{for} \quad y_1,y_2 \in Y,\ i \in I,\ t \ge 0.$$

(S3) There exist a Lebesgue measurable function $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ and a function $\mathcal{L} : Y \to \mathbb{R}_+$, bounded on bounded sets, such that

$$K_\varphi := \int_0^\infty e^{-\lambda t}\varphi(t)\,dt < \infty$$

and

$$\rho_Y(S_i(t,y), S_j(t,y)) \le \varphi(t)\mathcal{L}(y) \quad \text{for any} \quad t \ge 0,\ y \in Y,\ i,j \in I.$$

(S4) There exists $j_0 \in I$ such that $\min_{i\in I} \pi_{ij_0} > 0$.

The second part of the list includes certain conditions on the kernel J, governing the post-jump locations. They are as follows:

(J1) There exist $y^* \in Y$ and constants $\tilde{a} > 0$, $\tilde{b} \ge 0$ for which J satisfies

$$J\rho_Y(\cdot, y^*)(y) \le \tilde{a}\rho_Y(y, y^*) + \tilde{b} \quad \text{for any} \quad y \in Y. \tag{4.2}$$

(J2) There exists a substochastic kernel $Q_J : Y^2 \times \mathcal{B}(Y^2) \to [0,1]$ such that

$$Q_J((y_1,y_2), B \times Y) \le J(y_1, B) \quad \text{and} \quad Q_J((y_1,y_2), Y \times B) \le J(y_2, B) \tag{4.3}$$

for any $y_1, y_2 \in Y$ and $B \in \mathcal{B}(Y)$, which enjoys the following properties:

$$\int_{Y^2} \rho_Y(u,v)\, Q_J((y_1,y_2), du \times dv) \le \tilde{a}\rho_Y(y_1,y_2) \quad \text{for any} \quad y_1, y_2 \in Y, \tag{4.4}$$

$$\inf_{(y_1,y_2)\in Y^2} Q_J((y_1,y_2), \widetilde{U}(\tilde{a}\rho_Y(y_1,y_2))) \ge \eta \quad \text{for some} \quad \eta > 0, \tag{4.5}$$

where

$$\widetilde{U}(r) := \{(u,v) \in Y^2 :\ \rho_Y(u,v) \le r\} \quad \text{for} \quad r > 0, \tag{4.6}$$

and $\tilde{a}$ is the constant for which (J1) holds, as well as there exists $\tilde{l} > 0$ such that

$$Q_J((y_1,y_2), Y^2) > 1 - \tilde{l}\rho_Y(y_1,y_2) \quad \text{for any} \ \ y_1, y_2 \in Y. \tag{4.7}$$

(J3) For every function $g \in C_b(Y \times \mathbb{R}_+)$, the map $Y \times \mathbb{R}_+ \ni (y,t) \mapsto Jg(\cdot,t)(y)$ is jointly continuous.

In the main results, these two kinds of assumptions will be linked with each other by requiring that the constants $L > 0$, $\alpha \in \mathbb{R}$ and $\tilde{a} > 0$, appearing in (S2) and (J1), respectively, satisfy the inequality

$$\tilde{a}L + \frac{\alpha}{\lambda} < 1, \tag{4.8}$$

which, in turn, guarantees that the dual of the Markov operator P, induced by (3.4), enjoys a property known as the Foster–Lyapunov drift condition (see Lemma 4.1, given below). Such a condition is commonly used while studying the ergodic properties of Markov processes (see e.g. [36,37]). It is worth nothing here that, in particular, the inequality above yields that $\alpha < \lambda$. An assumption similar to (4.8) appears, e.g., in [31, Proposition 5.1], where a Poisson driven stochastic differential equation is considered.

Remark 4.1. Obviously, condition (J3) is a strengthened form of the Feller property for J.

Remark 4.2. Note that, if (4.1) is fulfilled for some $y^* \in Y$, then it is also valid for every $y^* \in Y$, whenever (S2) holds with some $\alpha < \lambda$. Hence, while considering the conjunction of conditions (S1), (S2) and (J1), we may and we will always assume that (4.1) and (4.2) hold with the same y^*.

Lemma 4.1. *Suppose that (S1), (S2) and (J1) hold with $L, \alpha, \tilde{a}$ satisfying inequality (4.8). Then P fulfils the Foster–Lyapunov condition with the function $V : X \to [0, \infty)$ of the form*

$$V(x) := \rho_Y(y, y^*) \quad \textit{for} \quad x = (y, i) \in X, \tag{4.9}$$

where y^* is specified by (J1), that is, there exist $a \in (0, 1)$ and $b \in [0, \infty)$ such that

$$PV(x) \le aV(x) + b \quad \textit{for all} \quad x \in X. \tag{4.10}$$

Moreover, the constants a, b can be chosen as

$$a := \frac{\tilde{a}\lambda L}{\lambda - \alpha} \quad \textit{and} \quad b := \tilde{a}\lambda \max_{i \in I} \int_0^\infty e^{-\lambda t} \rho_Y(S_i(t, y^*), y^*)\, dt + \tilde{b} \tag{4.11}$$

with $L, \alpha, \tilde{a}, \tilde{b}$ and y^* determined by (S2) and (J1).

Proof. Let a, b be defined as above. Then $b \ge 0$ (since $\tilde{a}, \tilde{b} \ge 0$), and (4.8) yields that $a \in (0, 1)$. Finally, we are led to the conclusion by the following estimates:

$$\begin{aligned}
PV(y, i) &= \int_0^\infty \lambda e^{-\lambda h} J\rho_Y(\cdot, y^*)(S_i(h, y))\, dh \le \tilde{a} \int_0^\infty \lambda e^{-\lambda h} \rho_Y(S_i(h, y), y^*)\, dh + \tilde{b} \\
&\le \tilde{a} \int_0^\infty \lambda e^{-\lambda h} \left[\rho_Y(S_i(h, y), S_i(h, y^*)) + \rho_Y(S_i(h, y^*), y^*)\right] dh + \tilde{b} \\
&\le \tilde{a} \int_0^\infty \lambda e^{-\lambda h} \left[L e^{\alpha h} \rho_Y(y, y^*) + \rho(S_i(h, y^*), y^*)\right] dh + \tilde{b} \\
&= \tilde{a}\lambda L \left(\int_0^\infty e^{-(\lambda - \alpha)h}\, dh\right) V(y, i) + \tilde{a}\lambda \int_0^\infty e^{-\lambda h} \rho(S_i(h, y^*), y^*)\, dh + \tilde{b} \\
&\le aV(y, i) + b \quad \text{for any} \quad y \in Y,\ i \in I. \quad \square
\end{aligned}$$

In Section 7, we also assume that the constant c, appearing in (3.1), is sufficiently large. More precisely, having imposed conditions (S1)–(S3), (J1) and (4.8), we require that

$$c \ge \frac{\lambda - \alpha}{L} \left(M_{\mathcal{L}} K_\varphi + \frac{M_{\mathcal{L}} M_\varphi}{\lambda}\right) + 1, \tag{4.12}$$

where

$$M_{\mathcal{L}} := \sup\{\mathcal{L}(y) : \rho_Y(y, y^*) \le 4b/(1 - a)\}, \tag{4.13}$$

$$M_\varphi := \sup\left\{\varphi(t) : t \le \lim_{s \to \alpha} s^{-1} \ln\left(\lambda(\lambda - s)^{-1}\right)\right\}, \tag{4.14}$$

and the constants a and b are given by (4.11).

Conditions (S1)–(S3) are satisfied, for example, for the flows generated by some classes of dissipative differential equations. This rests on the following observation (cf. [28]):

Remark 4.3. Suppose that Y is a closed subset of a Hilbert space H, endowed with an inner product $\langle \cdot | \cdot \rangle$, inducing the norm $\|\cdot\|$. Let $A_i : Y \to H$, $i \in I$, be a finite collection of α-dissipative operators with some $\alpha \le 0$, i.e.

$$\langle A_i y_1 - A_i y_2 \mid y_1 - y_2 \rangle \le \alpha \|y_1 - y_2\|^2 \quad \text{for any} \quad y_1, y_2 \in Y,\ i \in I. \tag{4.15}$$

Furthermore, assume that the so-called *range condition* holds, that is, there exists a positive constant T such that

$$Y \subset \operatorname{Range}(\mathrm{id}_Y - tA_i) \quad \text{for all} \quad t \in (0, T),\ i \in I. \tag{4.16}$$

Then, according to [28, Theorem 5.11], for any $i \in I$ and $y \in Y$, the initial value problem

$$u'(t) = A_i\, u(t) \;\text{ for }\; t \geq 0, \quad u(0) = y,$$

has a unique (strong) solution $\mathbb{R}_+ \ni t \mapsto S_i(t,y) \in Y$, which obviously generates a semiflow. What is more, by virtue of [28, Theorem 5.3 and Corollary 5.4], the semiflows S_i satisfy

$$\|S_i(t,y_1) - S_i(t,y_2)\| \leq e^{\alpha t}\,\|y_1 - y_2\| \quad \text{for any} \quad y_1, y_2 \in Y,$$

$$\|S_i(t,y) - y\| \leq t\,\|A_i y\| \quad \text{for any} \quad y \in Y.$$

This, in turn, implies that conditions (S1)–(S3) hold for such S_i, $i \in I$, with an arbitrarily fixed $y^* \in Y$, the dissipativity constant α,

$$L = 1, \quad \mathcal{L}(y) = 2\max_{i\in I}\|A_i y\| \quad \text{and} \quad \varphi(t) = t,$$

provided that A_i, $i \in I$, are bounded on bounded sets. Obviously, if I contains only one element, then (S3) is satisfied trivially (with $\mathcal{L} \equiv \varphi \equiv 0$), and thus the assumption of the boundedness on bounded sets is superfluous in this case.

On the other hand, the following simple example (inspired by [3, Example 5.2]) demonstrates that (S1)–(S3) may also hold for the flows generated by α-dissipative equations with a positive α.

Example 4.1. Suppose that $(Y, \|\cdot\|)$ is a Banach space, and let $z \in Y$, $\alpha \in \mathbb{R}$ be fixed. Consider the semiflows $S_1, S_2 : \mathbb{R}_+ \times Y \to Y$ induced by the differential equations (in Y)

$$u'(t) = \alpha u(t) \quad \text{and} \quad u'(t) = \alpha(u(t) - z) \quad \text{for} \quad t \geq 0,$$

respectively, that is

$$S_1(t,y) = e^{\alpha t} y, \quad S_2(t,y) = e^{\alpha t}(y - z) + z \quad \text{for} \quad t \geq 0,\; y \in Y.$$

Then conditions (S1)–(S3) are satisfied for $\{S_1, S_2\}$ with the given α, $L = 1$, $\mathcal{L} \equiv 1$ and $\varphi(t) = |1 - e^{\alpha t}|\,\|z\|$.

A significant example of a stochastic kernel J enjoying hypotheses (J1)–(J3) is the transition law of an iterated function system (or, more generally, the kernel specified in Remark 3.2), provided that the involved transformations and the associated probability densities satisfy certain usual conditions. This case will be discussed in detail within Section 7.2.

5. Correspondence between invariant distributions of Ψ and Φ

In the first part of the study, we shall establish a one-to-one correspondence between invariant probability measures for the transition operator P of the chain Φ, induced by (3.4), and those for the transition semigroup $\{P_t\}_{t\in\mathbb{R}_+}$ of the process Ψ, specified by (3.7). For this aim, let us consider the Markov operators $G, W : \mathcal{M}(X) \to \mathcal{M}(X)$ generated by the stochastic kernels given by

$$G((y,i),A) = \int_0^\infty \lambda e^{-\lambda t}\mathbb{1}_A(S_i(t,y),i)\,dt, \tag{5.1}$$

$$W((y,i),A) := \sum_{j\in I}\pi_{ij}\int_Y \mathbb{1}_A(u,j)J(y,du) \tag{5.2}$$

for any $(y,i) \in X$ and $A \in \mathcal{B}(X)$, where J stands for the kernel appearing in (3.3). It is easy to check that $GW = P$.

Having defined such operators, we can state the following result:

Theorem 5.1. *Let P and $\{P_t\}_{t\in\mathbb{R}_+}$ denote the transition operator of the chain Φ and the transition semigroup of the PDMP Ψ, respectively. Further, suppose that the kernel J, appearing in (3.3), satisfies (J3). Then*

(a) if μ_*^{Φ} is an invariant probability measure of P, then the measure $\mu_*^{\Psi} := \mu_*^{\Phi} G$ is invariant for $\{P_t\}_{t\in\mathbb{R}_+}$, and we have $\mu_*^{\Psi} W = \mu_*^{\Phi}$;
(b) if μ_*^{Ψ} is an invariant probability measure of $\{P_t\}_{t\in\mathbb{R}_+}$, then the measure $\mu_*^{\Phi} := \mu_*^{\Psi} W$ is invariant for P, and we have $\mu_*^{\Phi} G = \mu_*^{\Psi}$.

Theorem 5.1 can be proved exactly in the same way as [14, Theorem 4.4], which refers to the case where the kernel J is defined explicitly (as mentioned in Remark 3.2). In order to adapt this proof to our setting (with an arbitrary stochastic kernel J), one only needs to establish the properties collected in the lemma given below.

Lemma 5.1. *The following statements hold for the transition semigroup $\{P_t\}_{t\in\mathbb{R}_+}$ of the process Ψ:*

(i) If J is Feller, then $\{P_t\}_{t\in\mathbb{R}_+}$ is Feller.
(ii) For any $f \in B_b(X)$, there exists a bounded Borel measurable map $u_f : X\times\mathbb{R}_+\to\mathbb{R}$ such that $\lim_{t\to 0} \|u_f(\cdot,t)\|_\infty /t = 0$, and

$$P_t f(y,i) = e^{-\lambda t} f(S_i(t,y),i) + \lambda e^{-\lambda t}\int_0^t \psi_f((y,i),s,t)\,ds + u_f((y,i),t),$$

for any $(y,i)\in X$ and $t>0$, where

$$\psi_f((y,i),s,t) := \sum_{j\in I}\pi_{ij}\int_Y f(S_j(t-s,u),j)J(S_i(s,y),du) \quad \textit{for} \quad s\in[0,t],\ t>0.$$

(iii) $\{P_t\}_{t\in\mathbb{R}_+}$ is stochastically continuous, i.e.

$$\lim_{t\to 0} P_t f(y,i) = f(y,i) \quad \textit{for any} \quad (y,i)\in X,\ f\in C_b(X).$$

(iv) If (J3) holds then, for any function $s:\mathbb{R}_+\to\mathbb{R}_+$ satisfying $0\le s(t)\le t$ for all $t\ge 0$, the map $t\mapsto \psi_f((y,i),s(t),t)$ is continuous at $t=0$ for every $f\in C_b(X)$ and any $(y,i)\in X$. Moreover, $\psi_f(\cdot,0,0) = Wf$, where W is the operator induced by (5.2).

Proof. Throughout the proof, we will write $\bar{x} := (x,0)$ for any given $x\in X$. Moreover, for every $i\in I$, we put $f(S_i(h,\cdot)) := 0$ if $h<0$.

Let $t\in\mathbb{R}_+$ and $f\in B_b(X)$. Then, according to (2.10), for every $x\in X$, we can write

$$\begin{aligned} P_t f(x) = \mathbb{E}_{\bar{x}} f(Y(t),\xi(t)) &= \sum_{n=0}^{\infty} \mathbb{E}_{\bar{x}}\left[\mathbb{1}_{[\tau_n,\tau_{n+1})}(t) f(S_{\xi_n}(t-\tau_n,Y_n),\xi_n)\right] \\ &= \sum_{n=0}^{\infty} \mathbb{E}_{\bar{x}}\left[\mathbb{1}_{[0,t]}(\tau_n) f(S_{\xi_n}(t-\tau_n,Y_n),\xi_n)\cdot \mathbb{1}_{(t,\infty)}(\tau_{n+1})\right]. \end{aligned} \tag{5.3}$$

Taking into account (2.7), it is clear that, for any $g,h\in B_b(\bar{X})$ and $n\in\mathbb{N}_0$,

$$\begin{aligned} \mathbb{E}_{\bar{x}}[g(\bar{\Phi}_n)h(\bar{\Phi}_{n+1})] &= \int_{\bar{X}}\int_{\bar{X}} g(w)h(z)\,\bar{P}(w,dz)\bar{P}^n(\bar{x},dw) \\ &= \int_{\bar{X}} g(w)\bar{P}h(w)\bar{P}^n(\bar{x},dw) = \bar{P}^n(g\bar{P}h)(\bar{x}). \end{aligned} \tag{5.4}$$

Hence, defining

$$g_t(u,j,s) := \mathbb{1}_{[0,t]}(s) f(S_j(t-s,u),j) \quad \text{and} \quad h_t(u,j,s) := \mathbb{1}_{(t,\infty)}(s) \tag{5.5}$$

for any $u \in Y$, $j \in I$ and $s \in \mathbb{R}_+$, we see that

$$\begin{aligned} P_t f(x) &= \sum_{n=0}^{\infty} \mathbb{E}_{\bar{x}}[g_t(Y_n,\xi_n,\tau_n) h_t(Y_{n+1},\xi_{n+1},\tau_{n+1})] \\ &= \sum_{n=0}^{\infty} \bar{P}^n(g_t \bar{P} h_t)(\bar{x}) \quad \text{for every} \quad x \in X. \end{aligned} \tag{5.6}$$

(i): Suppose that the kernel J is Feller, and that $f \in C_b(X)$. To prove that $P_t f$ is continuous, we first observe that, for any function $\varphi \in B_b(\bar{X})$ such that $X \ni x \mapsto \varphi(x,s)$ is continuous for every $s \geq 0$ (which is the case for g_t and h_t), the map $X \ni x \mapsto \bar{P}\varphi(x,s)$ is continuous for any $s \geq 0$ as well, since

$$\bar{P}\varphi(x,s) = \sum_{j \in I} \pi_{ij} \int_0^{\infty} \lambda e^{-\lambda h} \, J\varphi(\cdot, j, h+s)(S_i(h,y))\, dh \;\text{ for any }\; x=(y,i) \in X,\; s \geq 0.$$

This implies that all the maps $X \ni x \mapsto \bar{P}^n(g_t \bar{P} h_t)(\bar{x})$, $t \geq 0$, $n \in \mathbb{N}_0$, are continuous. Further, considering the Poisson process $(N(s))_{s \in \mathbb{R}_+}$ of the form

$$N(s) := \max\{n \in \mathbb{N}_0 : \tau_n \leq s\}, \quad s \geq 0, \tag{5.7}$$

and bearing in mind (5.4), we see that

$$\begin{aligned} |\bar{P}^n(g_t \bar{P} h_t)(\bar{x})| &= \left|\mathbb{E}_{\bar{x}}[\mathbb{1}_{\{N(t)=n\}} f(S_{\xi_n}(t-\tau_n, Y_n), \xi_n)]\right| \leq \|f\|_\infty \, \mathbb{P}_{\bar{x}}(N(t)=n) \\ &= \|f\|_\infty \, e^{-\lambda t} \frac{(\lambda t)^n}{n!} \quad \text{for any} \quad \bar{x} \in \bar{X},\; n \in \mathbb{N}_0. \end{aligned} \tag{5.8}$$

Using (5.6), (5.8) and the discrete analogue of the Lebesgue dominated convergence theorem, we can therefore conclude that $P_t f$ is indeed continuous.

(ii): Let us define

$$u_f(x,t) := \sum_{n=2}^{\infty} \bar{P}^n(g_t \bar{P} h_t)(\bar{x}) \quad \text{for} \quad x \in X.$$

From (5.6) it now follows that

$$P_t f(x) = g_t \bar{P} h_t(\bar{x}) + \bar{P}(g_t \bar{P} h_t)(\bar{x}) + u_f(x,t) \quad \text{for any} \quad x \in X. \tag{5.9}$$

Referring to (5.8), we get

$$\begin{aligned} \frac{|u_f(x,t)|}{t} &\leq \|f\|_\infty \frac{1}{t} e^{-\lambda t} \sum_{n=2}^{\infty} \frac{(\lambda t)^n}{n!} = \|f\|_\infty \frac{1}{t} e^{-\lambda t}(e^{\lambda t} - 1 - \lambda t) \\ &= \|f\|_\infty \left(\frac{1-e^{-\lambda t}}{t} - \lambda e^{-\lambda t} \right) \end{aligned} \tag{5.10}$$

for any $x \in X$, which obviously gives $\lim_{t \to 0} \|u_f(\cdot,t)\|_\infty / t = 0$. Further, having in mind (3.3) and (5.5), we obtain

$$\begin{aligned} g_t \bar{P} h_t(y,i,s) &= \mathbb{1}_{[0,t]}(s) f(S_i(t-s,y),i) \int_0^{\infty} \lambda e^{-\lambda h} \mathbb{1}_{(t,\infty)}(h+s)\, dh \\ &= \mathbb{1}_{[0,t]}(s) f(S_i(t-s,y),i) e^{-\lambda(t-s)} \quad \text{for any} \quad y \in Y,\; i \in I,\; s \geq 0, \end{aligned} \tag{5.11}$$

and, in particular (for $s=0$),

$$g_t \bar{P} h_t(\bar{x}) = g_t \bar{P} h_t(y,i,0) = e^{-\lambda t} f(S_i(t,y),i) \quad \text{for every} \quad x=(y,i) \in X.$$

Furthermore, appealing to (3.3) and (5.11), we can also conclude that

$$\begin{aligned}\bar{P}(g_t\bar{P}h_t)(\bar{x}) &= \sum_{j\in I}\pi_{ij}\int_0^\infty \lambda e^{-\lambda h}\int_Y g_t(u,j,h)\bar{P}h_t(u,j,h)\,J(S_i(h,y),du)\,dh\\ &= \sum_{j\in I}\pi_{ij}\int_0^\infty \lambda e^{-\lambda h}\int_Y \mathbb{1}_{[0,t]}(h)f(S_j(t-h,u),j)e^{-\lambda(t-h)}\,J(S_i(h,y),du)\,dh\\ &= \lambda e^{-\lambda t}\int_0^t\sum_{j\in I}\pi_{ij}\int_Y f(S_j(t-h,u),j)\,J(S_i(h,y),du)\,dh\\ &= \lambda e^{-\lambda t}\int_0^t \psi_f((y,i),h,t)\,dh \quad \text{for all} \quad x=(y,i)\in X.\end{aligned} \tag{5.12}$$

Finally, assertion (ii) follows from (5.9)–(5.12).

(iii): Condition (iii) follows immediately from (ii) and the boundedness of ψ_f.

(iv): For the proof of (iv), fix $f\in C_b(X)$, $(y,i)\in X$, and define

$$g(u,t) := \sum_{j\in I}\pi_{ij}f(S_j(t-s(t),u),j) \quad \text{for} \quad (u,t)\in Y\times\mathbb{R}_+.$$

Since S_j are jointly continuous, so is g, and thus $g\in C_b(Y\times\mathbb{R}_+)$. Moreover, we have

$$Jg(\cdot,t)(S_i(s(t),y)) = \psi_f((y,i),s(t),t) \quad \text{for any} \quad t\geq 0.$$

Hence, $s\mapsto \psi_f((y,i),s(t),t)$ is continuous at 0 whenever (J3) holds and $s(t)\to 0$, as $t\to 0$. The identity $\psi_f(\cdot,0,0) = Wf$ is just a consequence of the definition of W. □

6. A coupling argument (involving Φ) for establishing the exponential ergodicity of both Φ and Ψ

The main goal of this section is to prove that the existence of an appropriate coupling between two copies of the chain Φ, which brings them closer to each other on average at a geometric rate, combined with the Foster–Lyapunov condition on P (cf. Lemma 4.1) and suitable assumptions on the flows implies that both Φ and Ψ are exponentially ergodic in the Fortet–Mourier metric.

More specifically, we shall consider a coupling between two copies of the chain Φ, enhanced with a copy of the sequence $\{\tau_n\}_{n\in\mathbb{N}_0}$, that is, a time-homogeneous Markov chain $\widehat{\Phi} := \{(\Phi_n^{(1)},\Phi_n^{(2)},\widetilde{\tau}_n)\}_{n\in\mathbb{N}_0}$ evolving on the space $Z := X^2\times\mathbb{R}_+$, whose transition law $\widehat{P}: Z\times\mathcal{B}(Z)\to[0,1]$ satisfies

$$\begin{gathered}\widehat{P}((x_1,x_2,s),A\times X\times\mathbb{R}_+) = P(x_1,A),\quad \widehat{P}((x_1,x_2,s),X\times A\times\mathbb{R}_+) = P(x_2,A),\\ \widehat{P}((x_1,x_2,s),X^2\times T) = \int_0^\infty \lambda e^{-\lambda t}\mathbb{1}_T(t+s)\,dt =: E_\lambda(s,T)\end{gathered} \tag{6.1}$$

for any $x_1,x_2\in X$, $s\geq 0$, $A\in\mathcal{B}(X)$ and $T\in\mathcal{B}(\mathbb{R}_+)$.

Such an augmented coupling $\widehat{\Phi}$ will be regarded as a canonical Markov chain, defined on the coordinate space $(\widehat{\Omega},\widehat{\mathcal{F}})$, with $\widehat{\Omega} := Z^{\mathbb{N}_0}$ and $\widehat{\mathcal{F}} := \mathcal{B}\left(Z^{\mathbb{N}_0}\right)$, equipped with an appropriate collection $\left\{\widehat{\mathbb{P}}_{(\mu_1,\mu_2)} : \mu_1,\mu_2\in\mathcal{M}_{prob}(X)\right\}$ of probability measures on $\widehat{\mathcal{F}}$ such that

$$\begin{gathered}\widehat{\mathbb{P}}_{(\mu_1,\mu_2)}(\widehat{\Phi}_0\in D) = (\mu_1\otimes\mu_2\otimes\delta_0)(D) \quad \text{for any} \quad D\in\mathcal{B}(Z),\\ \widehat{\mathbb{P}}_{(\mu_1,\mu_2)}(\widehat{\Phi}_{n+1}\in D\mid \widehat{\Phi}_n = z) = \widehat{P}(z,D) \quad \text{for any} \quad z\in Z,\ D\in\mathcal{B}(Z),\ n\in\mathbb{N}_0,\end{gathered}$$

where δ_0 stands for the Dirac measure at 0 on $\mathcal{B}(\mathbb{R}_+)$. The expectation operator corresponding to $\widehat{\mathbb{P}}_{(\mu_1,\mu_2)}$ will be denoted by $\widehat{\mathbb{E}}_{(\mu_1,\mu_2)}$. In the case where $\mu_1=\delta_{x_1}$ and $\mu_2=\delta_{x_2}$ with some $x_1,x_2\in X$, we will write (x_1,x_2) instead of $(\delta_{x_1},\delta_{x_2})$ in the subscripts.

We begin the analysis by establishing a general connection between the exponential ergodicity of P and the existence of an appropriate coupling of Φ, based on the Lyapunov condition. It is worth noting here that, in fact, the result below does not depend on the shape of the transition law P.

Lemma 6.1. *Suppose that P is Feller. Furthermore, assume that the transition law $\widehat{P}$ of the chain $\widehat{\Phi}$, satisfying (6.1), can be constructed so that*

$$\widehat{\mathbb{E}}_{(x_1,x_2)}\left[\bar{\rho}_{X,c}\left(\Phi_n^{(1)},\Phi_n^{(2)}\right)\right] \le C_0(V(x_1)+V(x_2)+1)q^n \quad \text{for all } n\in\mathbb{N},\ x_1,x_2\in X, \tag{6.2}$$

where $\bar{\rho}_{X,c}$ is given by (3.2), $q\in(0,1)$, $C_0<\infty$, and $V:X\to[0,\infty)$ is an arbitrary continuous function such that (4.10) holds for some $a\in(0,1)$ and some $b\in[0,\infty)$. Then there exists a unique invariant distribution μ_*^{Φ} for P such that $\mu_*^{\Phi}\in\mathcal{M}_{prob}^V(X)$ and

$$d_{FM,c}(\mu P^n,\mu_*^{\Phi}) \le C_0(\langle V,\mu\rangle + \langle V,\mu_*^{\Phi}\rangle + 1)q^n \quad \text{for any } n\in\mathbb{N},\ \mu\in\mathcal{M}_{prob}(X). \tag{6.3}$$

Proof. First of all, note that, for any $\mu_1,\mu_2\in\mathcal{M}_{prob}(X)$ and $n\in\mathbb{N}$,

$$d_{FM,c}(\mu_1 P^n,\mu_2 P^n) \le C_0(\langle V,\mu_1\rangle + \langle V,\mu_2\rangle + 1)q^n. \tag{6.4}$$

To see this, it suffices to observe that, for every $f\in \mathrm{Lip}_{b,1}(X)$,

$$\begin{aligned}
|\langle f,\mu_1 P^n-\mu_2 P^n\rangle| &= \left|\widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[f(\Phi_n^{(1)})-f(\Phi_n^{(2)})\right]\right| \le \widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left|f(\Phi_n^{(1)})-f(\Phi_n^{(2)})\right|\\
&\le \widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\bar{\rho}_{X,c}\left(\Phi_n^{(1)},\Phi_n^{(2)}\right)\right]\\
&= \int_{X^2}\widehat{\mathbb{E}}_{(x_1,x_2)}\left[\bar{\rho}_{X,c}\left(\Phi_n^{(1)},\Phi_n^{(2)}\right)\right](\mu_1\otimes\mu_2)(dx_1,dx_2)\\
&\le C_0\left(\int_X\int_X (V(x_1)+V(x_2)+1)\,\mu_1(dx_1)\mu_2(dx_2)\right)q^n\\
&= C_0(\langle V,\mu_1\rangle+\langle V,\mu_2\rangle+1)q^n,
\end{aligned}$$

where the first equality follows from (2.6), and the second one is due to (2.8).

The next step is to prove that P admits an invariant probability measure. For this purpose, let us fix arbitrarily $x_0\in X$ and notice that $\{\delta_{x_0}P^n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in the metric space $(\mathcal{M}_{prob}(X),d_{FM,c})$. Indeed, applying (6.4) with $\mu_1=\delta_{x_0}$ and $\mu_2=\delta_{x_0}P^k$ (for each $k\in\mathbb{N}_0$), together with (4.10), we can conclude that

$$\begin{aligned}
d_{FM,c}(\delta_{x_0}P^n,\delta_{x_0}P^{k+n}) &\le C_0(V(x_0)+P^kV(x_0)+1)q^n\\
&\le C_0\left(V(x_0)+a^kV(x_0)+\frac{b}{1-a}+1\right)q^n\\
&\le C_0\left(2V(x_0)+\frac{b}{1-a}+1\right)q^n \quad \text{for any} \quad n\in\mathbb{N},\ k\in\mathbb{N}_0.
\end{aligned}$$

Consequently, since $(\mathcal{M}_{prob}(X),d_{FM,c})$ is complete, $\{\delta_{x_0}P^n\}_{n\in\mathbb{N}}$ is weakly convergent to some $\mu_*\in\mathcal{M}_{prob}(X)$. From the Feller property it follows that μ_* is invariant for P, since, for any $f\in C_b(X)$,

$$\langle f,\mu_*P\rangle = \langle Pf,\mu_*\rangle = \lim_{n\to\infty}\langle Pf,\delta_{x_0}P^n\rangle = \lim_{n\to\infty}\left\langle f,\delta_{x_0}P^{n+1}\right\rangle = \langle f,\mu_*\rangle.$$

Obviously, (6.4), together with the invariance of μ_*, ensure that (6.3) holds with $\mu_*^{\Phi}:=\mu_*$.

In order to show that $\mu_*\in\mathcal{M}_{prob}^V(X)$, consider $V_k(x):=\min(V(x),k)$ for any $x\in X$ and $k\in\mathbb{N}$. Then $\{V_k\}_{k\in\mathbb{N}}$ is a non-decreasing sequence of functions of $C_b(X)$. From (4.10) it follows that

$$P^nV_k \le P^nV \le a^nV + \frac{b}{1-a} \quad \text{for all} \quad k,n\in\mathbb{N},$$

whence

$$\langle V_k, \mu_* \rangle = \lim_{n\to\infty} \langle V_k, \delta_{x_0} P^n \rangle = \lim_{n\to\infty} P^n V_k(x_0) \le \frac{b}{1-a} \quad \text{for every} \quad k \in \mathbb{N}.$$

By using the Lebesgue monotone convergence theorem, we therefore obtain

$$\langle V, \mu_* \rangle = \lim_{k\to\infty} \langle V_k, \mu_* \rangle \le \frac{b}{1-a} < \infty. \tag{6.5}$$

What is left is to show that there are no other invariant measures for P. To this end, it is enough to know

$$\lim_{n\to\infty} d_{FM,c}(\mu P^n, \mu_*) = 0 \quad \text{for any} \quad \mu \in \mathcal{M}_{prob}(X), \tag{6.6}$$

but this can be easily derived from (6.3) and (6.5). More precisely, these conditions guarantee that (6.6) holds for any $\mu \in \mathcal{M}_{prob}^V(X)$, and so, in particular, we have

$$P^n f(x) = \langle f, \delta_x P^n \rangle \to \langle f, \mu_* \rangle, \quad \text{as} \quad n \to \infty, \quad \text{for any} \quad x \in X \quad \text{and} \quad f \in C_b(X).$$

Now, letting μ be an arbitrary probability measure and applying the Lebesgue dominated convergence theorem, we obtain

$$\langle f, \mu P^n \rangle = \langle P^n f, \mu \rangle \to \langle f, \mu_* \rangle, \quad \text{as} \quad n \to \infty, \quad \text{for any} \quad f \in C_b(X),$$

which obviously gives (6.6) and completes the proof. □

Given an augmented coupling $\widehat{\Phi} := \{(\Phi_n^{(1)}, \Phi_n^{(2)}, \widetilde{\tau}_n)\}_{n\in\mathbb{N}_0}$ between any two copies of Φ, and writing

$$\Phi_n^{(i)} = (Y_n^{(i)}, \xi_n^{(i)}) \quad \text{for} \quad n \in \mathbb{N}_0, \; i \in \{1, 2\},$$

to indicate the coordinates $Y_n^{(i)}$ and $\xi_n^{(i)}$ with values in Y and I, respectively, we can consider the two corresponding copies $\Psi^{(1)}$ and $\Psi^{(2)}$ of the process Ψ, defined as follows:

$$\Psi^{(i)}(t) := (Y^{(i)}(t), \xi^{(i)}(t)) \quad \text{for} \quad t \ge 0, \; i \in \{1, 2\},$$

where

$$Y^{(i)}(t) := S_{\xi_n^{(i)}}\left(t - \widetilde{\tau}_n, Y_n^{(i)}\right), \quad \xi^{(i)}(t) = \xi_n^{(i)} \;\text{ whenever }\; t \in [\widetilde{\tau}_n, \widetilde{\tau}_{n+1}), \; n \in \mathbb{N}_0, \; i \in \{1, 2\}.$$

Our aim now is to show that any two copies of the process Ψ, defined as above on the path space of $\widehat{\Phi}$, get closer to each other on average at an exponential rate, whenever $\widehat{\Phi}$ satisfies (6.2) and the flows enjoy (S2) (with $\alpha < \lambda$). This will be a crucial step in deriving the exponential ergodicity of the semigroup $\{P_t\}_{t\in\mathbb{R}_+}$. The proof of this result, given below, is inspired by certain ideas developed in [10].

Lemma 6.2. *Let $\mathcal{R} \subset \mathcal{M}_{prob}(X)$, and suppose that the transition law $\widehat{P}$ of the chain $\widehat{\Phi}$, satisfying (6.1), can be constructed so that there exist $q \in (0, 1)$ and $\mathcal{C} : \mathcal{R}^2 \to \mathbb{R}_+$ such that*

$$\widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\bar{\rho}_{X,c}\left(\Phi_n^{(1)}, \Phi_n^{(2)}\right)\right] \le \mathcal{C}(\mu_1, \mu_2) q^n \quad \textit{for all} \quad \mu_1, \mu_2 \in \mathcal{R}, \; n \in \mathbb{N}. \tag{6.7}$$

Furthermore, assume that condition (S2) holds with some $L > 0$ and some $\alpha < \lambda$. Then there exist $\gamma > 0$ and $\bar{\mathcal{C}} : \mathcal{R}^2 \to \mathbb{R}_+$ such that

$$\widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\bar{\rho}_{X,c}\left(\Psi^{(1)}(t), \Psi^{(2)}(t)\right)\right] \le \bar{\mathcal{C}}(\mu_1, \mu_2)\, e^{-\gamma t} \quad \textit{for all} \quad \mu_1, \mu_2 \in \mathcal{R}, \; t \ge 0. \tag{6.8}$$

Proof. Fix $\mu_1, \mu_2 \in \mathcal{R}$, and let κ be the coupling time for $\left\{\left(\xi_n^{(1)}, \xi_n^{(2)}\right)\right\}_{n\in\mathbb{N}_0}$, that is,

$$\kappa := \inf\{n \in \mathbb{N}_0 : \xi_n^{(1)} = \xi_n^{(2)}\}.$$

From (6.7) we infer that, for every $n \in \mathbb{N}$,

$$\begin{aligned}\widehat{\mathbb{P}}_{(\mu_1,\mu_2)}(\kappa > n) &\le \widehat{\mathbb{P}}_{(\mu_1,\mu_2)}\left(\xi_n^{(1)} \ne \xi_n^{(2)}\right) = \widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\mathbb{1}_{\left\{\xi_n^{(1)}\ne\xi_n^{(2)}\right\}}\right] \\ &= \widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\mathbf{d}(\xi_n^{(1)}, \xi_n^{(2)})\right] \le \widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\bar{\rho}_{X,c}(\Phi_n^{(1)}, \Phi_n^{(2)})\right] \\ &\le \mathcal{C}(\mu_1,\mu_2)q^n,\end{aligned} \tag{6.9}$$

which, in particular, shows that $\mathbb{P}_{(\mu_1,\mu_2)}(\kappa < \infty) = 1$.

In what follows, the processes $\{\xi_n^{(2)}\}_{n\in\mathbb{N}_0}$ and $\{\Psi^{(2)}(t)\}_{t\in\mathbb{R}_+}$ will be identified with their copies $\{\widetilde{\xi}_n^{(2)}\}_{n\in\mathbb{N}_0}$ and $\{\widetilde{\Psi}^{(2)}(t)\}_{t\in\mathbb{R}_+}$, respectively, defined as follows:

$$\widetilde{\xi}_n^{(2)} := \begin{cases} \xi_n^{(2)} & \text{if} \quad n < \kappa, \\ \xi_n^{(1)} & \text{if} \quad n \ge \kappa, \end{cases}$$

$$\widetilde{\Psi}^{(2)}(t) := \left(S_{\widetilde{\xi}_n^{(2)}}\left(t - \widetilde{\tau}_n, Y_n^{(2)}\right), \widetilde{\xi}_n^{(2)}\right), \quad \text{whenever} \quad t \in [\widetilde{\tau}_n, \widetilde{\tau}_{n+1}) \quad \text{for} \quad n \in \mathbb{N}_0.$$

By $(\widetilde{N}(s))_{s\in\mathbb{R}_+}$ we will denote the Poisson process given by

$$\widetilde{N}(s) := \max\{n \in \mathbb{N}_0 : \widetilde{\tau}_n \le s\} \quad \text{for} \quad s \ge 0.$$

Moreover, we will write $\{\mathcal{F}_n\}_{n\in\mathbb{N}_0}$ for the natural filtration of $\widehat{\Phi} = \left\{\left(\Phi_n^{(1)}, \Phi_n^{(2)}, \widetilde{\tau}_n\right)\right\}_{n\in\mathbb{N}_0}$.

Let $n \in \mathbb{N}_0$ and $t \ge 0$ be arbitrary. Keeping in mind that $\bar{\rho}_{X,c} \le 1$, we can write

$$\begin{aligned}&\widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\bar{\rho}_{X,c}\left(\Psi^{(1)}(t), \Psi^{(2)}(t)\right)\mathbb{1}_{\{\widetilde{N}(t)=n\}} \mid \mathcal{F}_n\right] \\ &\quad\le \widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\bar{\rho}_{X,c}\left(\Psi^{(1)}(t), \Psi^{(2)}(t)\right)^{1/2}\mathbb{1}_{\{\widetilde{N}(t)=n\}} \mid \mathcal{F}_n\right] \\ &\quad\le \widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\bar{\rho}_{X,c}\left(\Psi^{(1)}(t), \Psi^{(2)}(t)\right)^{1/2}\mathbb{1}_{\{\widetilde{N}(t)=n\}}\mathbb{1}_{\{\kappa\le n\}} \mid \mathcal{F}_n\right] \\ &\qquad+ \widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\mathbb{1}_{\{\widetilde{N}(t)=n\}}\mathbb{1}_{\{\kappa>n\}} \mid \mathcal{F}_n\right].\end{aligned} \tag{6.10}$$

Defining

$$\bar{L} = \max\{L, 1\} \quad \text{and} \quad \bar{\alpha} = \max\{\alpha, 0\}$$

we see that (S2) gives

$$\rho_Y(S_i(t,u), S_i(t,v)) \wedge 1 \le \bar{L}e^{\bar{\alpha}t}(\rho_Y(u,v)\wedge 1) \quad \text{for any} \quad u,v \in Y,\ i \in I,\ t \ge 0. \tag{6.11}$$

Taking into account that $\{\widetilde{N}(t) = n\} = \{\widetilde{\tau}_n \le t < \widetilde{\tau}_{n+1}\}$, and that $\xi_n^{(1)} = \xi_n^{(2)}$ whenever $n \ge \kappa$ (due the adopted identification $\xi^{(2)} = \widetilde{\xi}^{(2)}$), we may apply (6.11) to estimate the first term on the right-hand side of (6.10) as follows:

$$\begin{aligned}&\widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\bar{\rho}_{X,c}\left(\Psi^{(1)}(t), \Psi^{(2)}(t)\right)^{1/2}\mathbb{1}_{\{\widetilde{N}(t)=n\}}\mathbb{1}_{\{\kappa\le n\}} \mid \mathcal{F}_n\right] \\ &= \mathbb{1}_{\{\widetilde{\tau}_n\le t\}}\mathbb{1}_{\{\kappa\le n\}}\left[\rho_Y\left(S_{\xi_n^{(1)}}\left(t-\widetilde{\tau}_n, Y_n^{(1)}\right), S_{\xi_n^{(2)}}\left(t-\widetilde{\tau}_n, Y_n^{(2)}\right)\right)\wedge 1\right]^{1/2} \\ &\quad\times \widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\mathbb{1}_{\{\widetilde{\tau}_{n+1}>t\}} \mid \mathcal{F}_n\right] \\ &\le \mathbb{1}_{\{\widetilde{\tau}_n\le t\}}\mathbb{1}_{\{\kappa\le n\}}\bar{L}^{1/2}e^{\bar{\alpha}(t-\widetilde{\tau}_n)/2}\left[\rho_Y\left(Y_n^{(1)}, Y_n^{(2)}\right)\wedge 1\right]^{1/2}\widehat{\mathbb{P}}_{(\mu_1,\mu_2)}(\widetilde{\tau}_{n+1} > t \mid \mathcal{F}_n) \\ &\le \mathbb{1}_{\{\widetilde{\tau}_n\le t\}}\mathbb{1}_{\{\kappa\le n\}}\bar{L}^{1/2}e^{\bar{\alpha}(t-\widetilde{\tau}_n)/2}\,\bar{\rho}_{X,c}\left(\Phi_n^{(1)}, \Phi_n^{(2)}\right)^{1/2}\widehat{\mathbb{P}}_{(\mu_1,\mu_2)}(\widetilde{\tau}_{n+1} > t \mid \mathcal{F}_n).\end{aligned} \tag{6.12}$$

Since, according to (3.6),

$$\widehat{\mathbb{P}}_{(\mu_1,\mu_2)}(\widetilde{\tau}_{n+1} > t \mid \mathcal{F}_n) = \widehat{\mathbb{P}}_{(\mu_1,\mu_2)}(\widetilde{\tau}_{n+1} > t \mid \tau_n) = e^{-\lambda(t-\widetilde{\tau}_n)} \quad \text{on} \quad \{\widetilde{\tau}_n \leq t\},$$

it follows that

$$\begin{aligned}\widehat{\mathbb{E}}_{(\mu_1,\mu_2)}&\left[\bar{\rho}_{X,c}\left(\Psi^{(1)}(t),\,\Psi^{(2)}(t)\right)^{1/2}\mathbb{1}_{\{\widetilde{N}(t)=n\}}\mathbb{1}_{\{\kappa\leq n\}} \mid \mathcal{F}_n\right]\\ &\leq \mathbb{1}_{\{\widetilde{\tau}_n\leq t\}}\mathbb{1}_{\{\kappa\leq n\}}\bar{L}^{1/2}\,e^{-(\lambda-\bar{\alpha}/2)t}e^{(\lambda-\bar{\alpha}/2)\widetilde{\tau}_n}\bar{\rho}_{X,c}\left(\Phi_n^{(1)},\,\Phi_n^{(2)}\right)^{1/2}.\end{aligned} \tag{6.13}$$

Consequently, due to (6.10) and (6.13), we obtain

$$\begin{aligned}\widehat{\mathbb{E}}_{(\mu_1,\mu_2)}&\left[\bar{\rho}_{X,c}\left(\Psi^{(1)}(t),\,\Psi^{(2)}(t)\right)\mathbb{1}_{\{\widetilde{N}(t)=n\}} \mid \mathcal{F}_n\right]\\ &\leq \mathbb{1}_{\{\widetilde{\tau}_n\leq t\}}\mathbb{1}_{\{\kappa\leq n\}}\bar{L}^{1/2}e^{-(\lambda-\bar{\alpha}/2)t}e^{(\lambda-\bar{\alpha}/2)\widetilde{\tau}_n}\bar{\rho}_{X,c}\left(\Phi_n^{(1)},\,\Phi_n^{(2)}\right)^{1/2}\\ &\quad+\widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\mathbb{1}_{\{\widetilde{N}(t)=n\}}\mathbb{1}_{\{\kappa>n\}} \mid \mathcal{F}_n\right].\end{aligned} \tag{6.14}$$

Taking the expectation of both sides of the last inequality and using the Cauchy–Schwarz inequality gives

$$\begin{aligned}\widehat{\mathbb{E}}_{(\mu_1,\mu_2)}&\left[\bar{\rho}_{X,c}\left(\Psi^{(1)}(t),\,\Psi^{(2)}(t)\right)\mathbb{1}_{\{\widetilde{N}(t)=n\}}\right]\\ &\leq \bar{L}^{1/2}\,e^{-(\lambda-\bar{\alpha}/2)t}\,\widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\mathbb{1}_{\{\widetilde{\tau}_n\leq t\}}e^{(\lambda-\bar{\alpha}/2)\widetilde{\tau}_n}\bar{\rho}_{X,c}\left(\Phi_n^{(1)},\,\Phi_n^{(2)}\right)^{1/2}\right]\\ &\quad+\widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\mathbb{1}_{\{\widetilde{N}(t)=n\}}\mathbb{1}_{\{\kappa>n\}}\right]\\ &\leq \bar{L}^{1/2}\,e^{-(\lambda-\bar{\alpha}/2)t}\left(\widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\mathbb{1}_{\{\widetilde{\tau}_n\leq t\}}e^{(2\lambda-\bar{\alpha})\widetilde{\tau}_n}\right]\right)^{1/2}\\ &\quad\times\left(\widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\bar{\rho}_{X,c}\left(\Phi_n^{(1)},\,\Phi_n^{(2)}\right)\right]\right)^{1/2}+\widehat{\mathbb{P}}_{(\mu_1,\mu_2)}(\widetilde{N}(t)=n)^{1/2}\,\widehat{\mathbb{P}}_{(\mu_1,\mu_2)}(\kappa>n)^{1/2}.\end{aligned} \tag{6.15}$$

What is left is to estimate the right-hand side of (6.15). To do this, we first observe that, for any $\lambda_0 > 0$,

$$\begin{aligned}\widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\mathbb{1}_{\{\widetilde{\tau}_n\leq t\}}e^{(2\lambda-\bar{\alpha})\widetilde{\tau}_n}\right] &= \int_0^t e^{(2\lambda-\bar{\alpha})s}e^{-\lambda s}\frac{\lambda^n s^{n-1}}{(n-1)!}\,ds\\ &= \lambda_0\left(\frac{\lambda}{\lambda_0}\right)^n\int_0^t e^{(\lambda-\bar{\alpha})s}\frac{(\lambda_0 s)^{n-1}}{(n-1)!}\,ds\\ &\leq \lambda_0\left(\frac{\lambda}{\lambda_0}\right)^n\int_0^t e^{(\lambda-\bar{\alpha})s}e^{\lambda_0 s}\,ds\\ &= \lambda_0\left(\frac{\lambda}{\lambda_0}\right)^n\int_0^t e^{(\lambda+\lambda_0-\bar{\alpha})s}\,ds\\ &\leq \frac{\lambda_0}{\lambda+\lambda_0-\bar{\alpha}}\left(\frac{\lambda}{\lambda_0}\right)^n e^{(\lambda+\lambda_0-\bar{\alpha})t},\end{aligned} \tag{6.16}$$

where the first equality follows from the fact that $\widetilde{\tau}_n$ has the Erlang distribution with rate λ. Consequently, applying (6.16), together with hypothesis (6.7), we see that

$$\begin{aligned}&\bar{L}^{1/2}\,e^{-(\lambda-\bar{\alpha}/2)t}\left(\widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\mathbb{1}_{\{\widetilde{\tau}_n\leq t\}}e^{(2\lambda-\bar{\alpha})\widetilde{\tau}_n}\right]\right)^{1/2}\left(\widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\bar{\rho}_{X,c}\left(\Phi_n^{(1)},\,\Phi_n^{(2)}\right)\right]\right)^{1/2}\\ &\leq \mathcal{C}(\mu_1,\mu_2)^{1/2}\left(\frac{\bar{L}\lambda_0}{\lambda+\lambda_0-\bar{\alpha}}\right)^{1/2}\left(\frac{q\lambda}{\lambda_0}\right)^{n/2}e^{-(\lambda-\lambda_0)t/2}.\end{aligned} \tag{6.17}$$

Moreover, from (6.9) it follows that, for any $\bar{q} > 0$,

$$\begin{aligned} \widehat{\mathbb{P}}_{(\mu_1,\mu_2)}(\widetilde{N}(t)=n)^{1/2}\,\widehat{\mathbb{P}}_{(\mu_1,\mu_2)}(\kappa>n)^{1/2} &\le \left(e^{-\lambda t}\frac{(\lambda t)^n}{n!}\cdot \mathcal{C}(\mu_1,\mu_2)q^n\right)^{1/2} \\ &= \mathcal{C}(\mu_1,\mu_2)^{1/2}e^{-\lambda t/2}\left(\frac{(\lambda q\bar{q}^{-1}t)^n}{n!}\right)^{1/2}\bar{q}^{n/2} \\ &\le \mathcal{C}(\mu_1,\mu_2)^{1/2}e^{-\lambda t/2}e^{\lambda q\bar{q}^{-1}t/2}\,\bar{q}^{n/2} \\ &= \mathcal{C}(\mu_1,\mu_2)^{1/2}e^{-\lambda(1-q\bar{q}^{-1})t/2}\,\bar{q}^{n/2}. \end{aligned} \tag{6.18}$$

Let us now take $\bar{q} \in (q,1)$ and choose $\lambda_0 \in (0,\lambda)$ so that $q\lambda\lambda_0^{-1} < 1$. This choice guarantees that

$$\gamma := \min\left\{\frac{\lambda-\lambda_0}{2}, \frac{\lambda(1-q\bar{q}^{-1})}{2}\right\} > 0 \quad \text{and} \quad r := \max\left\{\left(\frac{q\lambda}{\lambda_0}\right)^{1/2}, \bar{q}^{1/2}\right\} \in (0,1).$$

From (6.15), (6.17) and (6.18) we can now conclude that

$$\widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\bar{\rho}_{X,c}\left(\Psi^{(1)}(t),\,\Psi^{(2)}(t)\right)\mathbb{1}_{\{\widetilde{N}(t)=n\}}\right] \le \mathcal{C}(\mu_1,\mu_2)^{1/2}\left[\left(\frac{\bar{L}\lambda_0}{\lambda+\lambda_0-\alpha}\right)^{1/2}+1\right]r^n e^{-\gamma t}.$$

Finally, defining $\widetilde{\mathcal{C}}(\mu_1,\mu_2) := \mathcal{C}(\mu_1,\mu_2)^{1/2}[(\bar{L}\lambda_0)^{1/2}(\lambda+\lambda_0-\alpha)^{-1/2}+1]$, we infer that

$$\widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\bar{\rho}_{X,c}\left(\Psi^{(1)}(t),\,\Psi^{(2)}(t)\right)\right] \le \sum_{n=0}^{\infty}\widetilde{\mathcal{C}}(\mu_1,\mu_2)r^n e^{-\gamma t} = \frac{\widetilde{\mathcal{C}}(\mu_1,\mu_2)}{1-r}e^{-\gamma t}.$$

Hence (6.8) holds with $\bar{\mathcal{C}}(\mu_1,\mu_2) := \widetilde{\mathcal{C}}(\mu_1,\mu_2)(1-r)^{-1}$ for $\mu_1,\mu_2 \in \mathcal{R}$. The proof is now complete. □

Lemmas 6.1, 6.2 and Theorem 5.1 enable us to prove the main result of this section, which reads as follows:

Theorem 6.1. *Let P be the transition operator of the chain Φ, induced by (3.4), and let $\{P_t\}_{t\in\mathbb{R}_+}$ denote the transition semigroup of the process Ψ, defined by (3.7). Further, suppose that conditions (S1), (S2) and (J1) are fulfilled with $L, \alpha, \tilde{a}$ satisfying (4.8), and that the following holds:*

- (C) *There exists an augmented coupling $\widehat{\Phi} = \{(\Phi_n^{(1)}, \Phi_n^{(2)}, \widetilde{\tau}_n)\}_{n\in\mathbb{N}_0}$ of the chain Φ, with transition law $\widehat{P}$ satisfying (6.1), such that (6.2) holds with certain constants $q \in (0,1)$, $C_0 < \infty$ and the function V defined by (4.9).*

Then, if J is Feller, the operator P is V-exponentially ergodic in $d_{FM,c}$ (in the sense of Definition 2.1). Moreover, assuming that J enjoys the strengthened Feller property specified by (J3), also the semigroup $\{P_t\}_{t\in\mathbb{R}_+}$ is V-exponentially ergodic in $d_{FM,c}$ (in the sense of Definition 2.2).

Proof. First of all, from Lemma 4.1 it follows that P satisfies condition (4.10) with V given by (4.9) and the constants $a \in (0,1)$, $b \ge 0$ specified in (4.11). Consequently, if J is Feller (and thus so is P, due to Remark 3.1), then, by virtue of Lemma 6.1, the operator P is exponentially ergodic in $d_{FM,c}$.

It therefore remains to prove that $\{P_t\}_{t\in\mathbb{R}_+}$ is also exponentially ergodic, provided that the Feller property of J is strengthened to condition (J3).

Let $\mu_*^{\Phi} \in \mathcal{M}_{prob}^V(X)$ be the unique invariant probability measure of P (which exists by Lemma 6.1). Then, upon assuming (J3), Theorem 5.1 guarantees the existence of exactly one invariant probability measure μ_*^{Ψ} for $\{P_t\}_{t\in\mathbb{R}_+}$, which can be expressed as $\mu_*^{\Psi} = \mu_*^{\Phi}G$, where G is the Markov operator induced by (5.1). Moreover, conditions (S1) and (S2) yield that $\mu_*^{\Psi} \in \mathcal{M}_{prob}^V(X)$.

Now, to complete the proof, it suffices to show that there exists a constant $\gamma > 0$ and $\bar{\mathcal{C}} : \mathcal{M}_{prob}^V(X)^2 \to \mathbb{R}_+$ such that

$$d_{FM,c}(\mu_1 P_t, \mu_2 P_t) \le \bar{\mathcal{C}}(\mu_1, \mu_2) e^{-\gamma t} \quad \text{for any} \quad \mu_1, \mu_2 \in \mathcal{M}_{prob}^V(X),\ t \ge 0. \tag{6.19}$$

As we have already seen in the proof of Lemma 6.1, from hypothesis (C) it follows that

$$\widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\bar{\rho}_{X,c}(\Phi_n^{(1)}, \Phi_n^{(2)})\right] \le \mathcal{C}(\mu_1, \mu_2) q^n \quad \text{for any} \quad \mu_1, \mu_2 \in \mathcal{M}_{prob}^V(X),\ n \in \mathbb{N},$$

where

$$\mathcal{C}(\mu_1, \mu_2) := C_0(\langle V, \mu_1 \rangle + \langle V, \mu_2 \rangle + 1) \quad \text{with some} \quad C_0 \in \mathbb{R}.$$

In view of Lemma 6.2, this, together with (S2), implies the existence of $\gamma > 0$ and $\bar{\mathcal{C}} : \mathcal{M}_{prob}^V(X)^2 \to \mathbb{R}_+$ such that

$$\widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\bar{\rho}_{X,c}\left(\Psi^{(1)}(t), \Psi^{(2)}(t)\right)\right] \le \bar{\mathcal{C}}(\mu_1, \mu_2)\, e^{-\gamma t} \quad \text{for all} \quad \mu_1, \mu_2 \in \mathcal{M}_{prob}^V(X),\ t \ge 0.$$

Finally, it suffices to observe that, for any $\mu_1, \mu_2 \in \mathcal{M}_1^V(X)$ and $f \in \mathrm{Lip}_{b,1}(X)$,

$$\begin{aligned}
|\langle f, \mu_1 P_t - \mu_2 P_t \rangle| &= \left|\widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[f\left(\Psi^{(1)}(t)\right) - f\left(\Psi^{(2)}(t)\right)\right]\right| \\
&\le \widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left|f\left(\Psi^{(1)}(t)\right) - f\left(\Psi^{(2)}(t)\right)\right| \\
&\le \widehat{\mathbb{E}}_{(\mu_1,\mu_2)}\left[\bar{\rho}_{X,c}\left(\Psi^{(1)}(t), \Psi^{(2)}(t)\right)\right] \le \bar{\mathcal{C}}(\mu_1, \mu_2) e^{-\gamma t},
\end{aligned}$$

which obviously assures (6.19), and thus ends the proof. □

Remark 6.1. It is worth noting that condition (6.19) is achieved by using only (C) and (S2).

7. Sufficient conditions for the exponential ergodicity

This section is intended to construct a suitable coupling of Φ, for which condition (C) of Theorem 6.1 is satisfied, which, in turn, will enable us to state a verifiable criterion for the exponential ergodicity of P and $\{P_t\}_{t\in\mathbb{R}_+}$ in the Fortet–Mourier distance. To do this, we shall need to employ all the hypotheses stated in Section 4 (except condition (J3)) and assume that the constant c, associated with the metric $\rho_{X,c}$, defined by (3.1), is large enough to assure (4.12).

7.1. The main result

Assuming that $Q_j : Y^2 \times \mathcal{B}(Y^2) \to [0,1]$ is a substochastic kernel satisfying (4.3), let us consider $\bar{Q}_P : Z \times \mathcal{B}(Z) \to [0,1]$ (where $Z = X^2 \times \mathbb{R}_+$), given by

$$\begin{aligned}
\bar{Q}_P((x_1, x_2, s), D) := \sum_{j\in I} (\pi_{i_1,j} \wedge \pi_{i_2,j}) \int_0^\infty \lambda e^{-\lambda h} \int_{Y^2} \mathbb{1}_D((u_1, j), (u_2, j), h + s) \\
\times Q_J((S_{i_1}(h, y_1), S_{i_2}(h, y_2)), du_1 \times du_2)\, dh
\end{aligned} \tag{7.1}$$

for any $x_1 := (y_1, i_1), x_2 := (y_2, i_2) \in X$, $s \in \mathbb{R}_+$ and $D \in \mathcal{B}(Z)$.

It is then easy to see that Q_P is also a substochastic kernel, and, for any $x_1, x_2 \in X$, $s \in \mathbb{R}_+$, $A \in \mathcal{B}(X)$ and $T \in \mathcal{B}(\mathbb{R}_+)$, we have

$$\begin{aligned}
\bar{Q}_P((x_1, x_2, s), A \times X \times \mathbb{R}_+) &\le P(x_1, A), \\
\bar{Q}_P((x_1, x_2, s), X \times A \times \mathbb{R}_+) &\le P(x_2, A), \\
\bar{Q}_P((x_1, x_2, s), X^2 \times T) &\le E_\lambda(s, T),
\end{aligned} \tag{7.2}$$

where $E_\lambda(s,\cdot)$ denotes the distribution with density $t \mapsto \mathbb{1}_{[s,\infty)}(t)\lambda e^{\lambda(t-s)}$. This enables us to define a substochastic kernel $\bar{R}_P : Z \times \mathcal{B}(Z) \to [0,1]$ so that, on cubes $D := A_1 \times A_2 \times T$, where $A_1, A_2 \in \mathcal{B}(X)$ and $T \in \mathcal{B}(\mathbb{R}_+)$, the measure $\bar{R}_P((x_1,x_2,s),\cdot)$ is given by

$$\begin{aligned}\bar{R}_P((x_1,x_2,s),D) := {} & \frac{1}{\left[1-\bar{Q}_P((x_1,x_2,s),Z)\right]^2} \\ & \times \left[P(x_1,A_1) - \bar{Q}_P((x_1,x_2,s),A_1 \times X \times \mathbb{R}_+)\right] \\ & \times \left[P(x_2,A_2) - \bar{Q}_P((x_1,x_2,s),X \times A_2 \times \mathbb{R}_+)\right] \\ & \times \left[E_\lambda(s,T) - \bar{Q}_P((x_1,x_2,s),X^2 \times T)\right]\end{aligned} \tag{7.3}$$

when $\bar{Q}_P((x_1,x_2,s),Z) < 1$, and $\bar{R}_P((x_1,x_2,s),D) := 0$ otherwise.

A simple computation shows that $\widehat{P} : \mathcal{B}(Z) \times Z \to [0,1]$ given by

$$\widehat{P}((x_1,x_2,s),D) := \bar{Q}_P((x_1,x_2,s),D) + \bar{R}_P((x_1,x_2,s),D) \tag{7.4}$$

for any $x_1, x_2 \in X$, $s \geq 0$ and $D \in \mathcal{B}(Z)$ defines a stochastic kernel satisfying conditions (6.1). In other words, the kernel defined in this way can play the role of transition law of the augmented coupling $\widehat{\Phi} := \{(\Phi_n^{(1)}, \Phi_n^{(2)}, \widetilde{\tau}_n)\}_{n\in\mathbb{N}_0}$ discussed in Section 6.

What is more, one can show that, under suitable assumptions, such a coupling fulfils hypothesis (C) of Theorem 6.1, which is stated precisely in the following result:

Proposition 7.1. *Suppose that conditions (S1)–(S4), (J1) and (J2) hold with constants $L, \alpha, \tilde{a}$ satisfying (4.8). Then the coupling $\widehat{\Phi}$ with transition law $\widehat{P}$ defined by (7.4) satisfies (6.2) with V given by (4.9), sufficiently large c, specified by (4.12), and certain constants $q \in (0,1)$ and $C_0 < \infty$.*

The proof of this statement proceeds almost in the same way as that of [15, Lemma 2.3], provided that hypotheses (B1)–(B5) stated in [15, Section 2] are fulfilled for the operator P, given by (3.4), and the kernel $Q_P : X^2 \times \mathcal{B}(X^2) \to [0,1]$ of the form

$$Q_P((x_1,x_2),C) := \bar{Q}_P((x_1,x_2,0),C \times \mathbb{R}_+), \quad x_1,x_2 \in X,\ C \in \mathcal{B}(X^2), \tag{7.5}$$

where $\bar{Q}_P$ is defined by (7.1). These hypotheses (also assumed in [29, Theorem 2.1]) can be derived quite easily from the assumptions of Proposition 7.1. The proof of this claim, as well as a suitable adaptation of the reasoning employed in [15], which eventually proves Proposition 7.1, are postponed to Section 7.3.

In view of Proposition 7.1, we can replace hypothesis (C) of Theorem 6.1 with conditions (S3), (S4) and (J2), which, together with (S1), (S2) and (J1), guarantee the existence of a suitable coupling of Φ. This leads us to the main result of the paper:

Theorem 7.1. *Suppose that conditions (S1)–(S4), (J1) and (J2) hold with $L, \alpha, \tilde{a}$ satisfying (4.8). Further, let V be given by (4.9), and let c be large enough to assure (4.12). Then, if J is Feller, the transition operator P of the chain Φ, induced by (3.4), is V-exponentially ergodic in $d_{FM,c}$. Moreover, if (J3) holds as well, then also the transition semigroup $\{P_t\}_{t\in\mathbb{R}_+}$ of the process Ψ, defined by (3.7), is V-exponentially ergodic in $d_{FM,c}$.*

7.2. Application to the model with jumps generated by random iterations

Let us look closer at the case that has already been mentioned in Remark 3.2. For simplicity of notation, we will skip the perturbations (the linear structure of Y is then not required). In such a case, the kernel J is the transition law of an iterated function system, consisting of an arbitrary set $\{w_\theta : \theta \in \Theta\}$ of continuous transformations from Y to itself and the associated collection of probability densities $\Theta \ni \theta \mapsto p_\theta(y) \in \mathbb{R}_+$, $y \in Y$, with respect to ϑ, for which $\int_\Theta p_\theta(y)\,\vartheta(d\theta) = 1$ for any $y \in Y$. Here, $(\Theta, \mathcal{B}(\Theta), \vartheta)$ is a topological space with a non-trivial Borel measure ϑ. Moreover, we assume that the maps $(y,\theta) \mapsto w_\theta(y)$ and $(y,\theta) \mapsto p_\theta(y)$ are product measurable.

In the above-described setting, J is given by

$$J(y,B) = \int_\Theta \mathbb{1}_B(w_\theta(y))\,p_\theta(y)\,\vartheta(d\theta) \quad \text{for} \quad y \in Y,\ B \in \mathcal{B}(Y), \tag{7.6}$$

and P takes the form

$$P((y,i),A) = \sum_{j\in I} \pi_{ij} \int_0^\infty \lambda e^{-\lambda h} \int_\Theta \mathbb{1}_A(w_\theta(S_i(h,y)),j))\,p_\theta(S_i(h,y))\,\vartheta(d\theta)\,dh \tag{7.7}$$

for any $y \in Y$, $i \in I$ and $A \in \mathcal{B}(X)$. Moreover, note that, in this framework, the first coordinate of the chain Φ can be expressed explicitly by the recursive formula:

$$Y_{n+1} = w_{\theta_{n+1}}(Y(\tau_{n+1}-)) = w_{\theta_{n+1}}(S_{\xi_n}(\Delta\tau_{n+1}, Y_n)) \ \text{ a.s} \quad \text{for} \quad n \in \mathbb{N}_0,$$

where $\{\theta_n\}_{n\in\mathbb{N}}$ is an appropriate sequence of random variables with values in Θ, such that

$$\mathbb{P}_\nu(\theta_{n+1} \in D \mid S_{\xi_n}(\Delta\tau_{n+1}, Y_n) = y) = \int_D p_\theta(y)\vartheta(d\theta) \ \text{ for } \ D \in \mathcal{B}(\Theta),\ y \in Y,\ n \in \mathbb{N}, \tag{7.8}$$

and the transition laws of $\{\xi_n\}_{n\in\mathbb{N}_0}$ and $\{\tau_n\}_{n\in\mathbb{N}_0}$ are determined by (3.5) and (3.6), respectively.

We shall impose the following assumptions on the transformations $y \mapsto w_\theta(y)$ and densities $\theta \mapsto p_\theta(y)$ (in the spirit of those made in [29, Proposition 3.1]; cf. also [41] and [42, Theorem 3.1]): there exist $y^* \in Y$, for which

$$\tilde{b} := \sup_{y\in Y} \int_\Theta \rho_Y(w_\theta(y^*), y^*) p_\theta(y)\,\vartheta(d\theta) < \infty, \tag{7.9}$$

and positive constants $\tilde{a}$, η and $\tilde{l}$ such that, for any $y_1, y_2 \in Y$,

$$\int_\Theta \rho_Y(w_\theta(y_1), w_\theta(y_2))\,p_\theta(y_1)\vartheta(d\theta) \le \tilde{a}\rho_Y(y_1,y_2), \tag{7.10}$$

$$\int_{\Theta(y_1,y_2)} p_\theta(y_1) \wedge p_\theta(y_2)\,\vartheta(d\theta) \ge \eta, \tag{7.11}$$

where

$$\Theta(y_1,y_2) := \{\theta \in \Theta :\ \rho_Y(w_\theta(y_1), w_\theta(y_2)) \le \tilde{a}\rho_Y(y_1,y_2)\},$$

and

$$\int_\Theta |p_\theta(y_1) - p_\theta(y_2)|\,\vartheta(d\theta) \le \tilde{l}\rho_Y(y_1,y_2). \tag{7.12}$$

Remark 7.1. Note that (7.9) is trivially satisfied in the case where Θ is compact, and $\theta \mapsto w_\theta(y^*)$ is continuous for some $y^* \in Y$.

Theorem 7.1 allows us to establish the following result:

Proposition 7.2. *Suppose that the kernel J is of the form (7.6), and the transformations w_θ are continuous. Further, assume that there exist $y^* \in Y$ and positive constants $\tilde{a}, \tilde{b}, \eta, \tilde{l}$ such that conditions (7.9)–(7.12) hold. Then (J1)–(J3) are satisfied with the same $y^*, \tilde{a}, \tilde{b}, \eta$ and $\tilde{l}$. If, additionally, conditions (S1)–(S4) are fulfilled with L and α such that (4.8) holds, then both the transition operator P of Φ (induced by (7.7) in this case) and the transition semigroup $\{P_t\}_{t\in\mathbb{R}_+}$ of Ψ, defined by (3.7), are V-exponentially ergodic in $d_{FM,c}$ with V given by (4.9) and sufficiently large c, specified by (4.12).*

Proof. In view of Theorem 7.1, it suffices to show that conditions (J1)–(J3) hold.

First of all, note that (J1) follows immediately from (7.9) and (7.10), since

$$J\rho_Y(\cdot,y^*)(y)=\int_\Theta \rho_Y(w_\theta(y),y^*)p_\theta(y)\,\vartheta(d\theta)\le \tilde{a}\rho_Y(y^*,y)+\tilde{b}\quad\text{for all}\quad y\in Y.$$

Now, we will show that (J2) is fulfilled with $Q_J : X^2\times\mathcal{B}(X^2)\to[0,1]$ given by

$$Q_J((y_1,y_2),C):=\int_\Theta \mathbb{1}_C(w_\theta(y_1),w_\theta(y_2))(p_\theta(y_1)\wedge p_\theta(y_2))\,\vartheta(d\theta) \tag{7.13}$$

for any $y_1,y_2\in Y$ and $C\in\mathcal{B}(Y^2)$. Obviously, Q_J is a substochastic kernel satisfying (4.3). Condition (7.10) yields that, for any $y_1,y_2\in Y$,

$$Q_J\rho_Y(y_1,y_2)=\int_\Theta \rho_Y(w_\theta(y_1),w_\theta(y_2))(p_\theta(y_1)\wedge p_\theta(y_2))\,\vartheta(d\theta)\le\tilde{a}\rho_Y(y_1,y_2),$$

which gives (4.4). Further, (7.11) implies (4.5), since, for any $y_1,y_2\in Y$, we have

$$Q_J((y_1,y_2),\widetilde{U}(\tilde{a}\rho_Y(y_1,y_2)))=\int_{\Theta(y_1,y_2)} p_\theta(y_1)\wedge p_\theta(y_2)\,\vartheta(d\theta)\ge\eta>0,$$

with $\widetilde{U}(\cdot)$ defined by (4.6). Finally, (4.7) can be easily concluded from hypothesis (7.12) and the inequality $s\wedge t\ge s-|s-t|$, which is valid for any $s,t\in\mathbb{R}$.

What is left is to show that (J3) holds. To this end, let $g\in C_b(Y\times\mathbb{R}_+)$ and fix $(y_0,t_0)\in Y\times\mathbb{R}_+$. Then, again using (7.12), we see that

$$\begin{aligned}|Jg(\cdot,t_0)(y_0)-Jg(\cdot,t)(y)|&\le\int_\Theta |g(w_\theta(y_0),t_0)p_\theta(y_0)-g(w_\theta(y),t)p_\theta(y)|\,\vartheta(d\theta)\\&\le\int_\Theta |g(w_\theta(y_0),t_0)-g(w_\theta(y),t)|p_\theta(y_0)\,\vartheta(d\theta)\\&\quad+\|g\|_\infty\,\tilde{l}\rho_Y(y_0,y)\quad\text{for any}\quad(y,t)\in Y\times\mathbb{R}_+.\end{aligned}$$

Consequently, having in mind the continuity of g and w_θ, we can conclude that $Y\times\mathbb{R}_+\ni(y,t)\mapsto Jg(\cdot,t)(y)$ is jointly continuous, by applying the Lebesgue dominated convergence theorem. The use of Theorem 7.1 now ends the proof. □

Remark 7.2. Obviously, Proposition 7.2 remains valid if J is defined exactly as in Remark 3.2 (provided that Y is a closed subset of a Banach space). The proof is then almost the same as that given above. In that case, however, one needs to consider Q_J of the form

$$Q_J((y_1,y_2),C):=\int_{\operatorname{supp}\nu}\int_\Theta \mathbb{1}_C(w_\theta(y_1)+v,w_\theta(y_2)+v)(p_\theta(y_1)\wedge p_\theta(y_2))\,\vartheta(d\theta)\,\nu(dv).$$

Example 7.1. A simple consequence of Proposition 7.2 is that conditions (J1)–(J3) are satisfied with $\tilde{a}=1$ for every stochastic kernel J defined as the shift of a Borel probability measure on a separable Banach space $(Y,\|\cdot\|)$ with finite first moment with respect to $\|\cdot\|$. In particular, the conclusions of Theorem 7.1 are then valid whenever (S1)–(S4) hold with $L+\alpha\lambda^{-1}<1$. More specifically, let ϑ be a Borel probability measure on Y such that

$$\int_Y \|y\|\,\vartheta(dy)<\infty, \tag{7.14}$$

and consider $J(y,B)=\vartheta(B-y)$ for $y\in Y$, $B\in\mathcal{B}(Y)$. Then, taking $\Theta=Y$, we can write J in the form (7.6) with $w_\theta(y)=y+\theta$ and $p_\theta(y)=1$ for any $y,\theta\in Y$. Having this in mind, one only needs to note that (7.9)–(7.12) hold with $y^*=0$, $\tilde{a}=1$, $\eta=1$ and any $\tilde{l}>0$. It is also worth mentioning here that, according to the celebrated Fernique theorem [23], condition (7.14) holds, e.g., for every centred Gaussian measure on Y.

In the case where J is defined by (7.6) with $p_\theta(y) = 1$ for all $\theta \in \Theta$, $y \in Y$ (assuming that ϑ is a probability measure on Θ), the process $\{Y(t)\}_{t\in\mathbb{R}_+}$ can be viewed as the solution to a Poisson-driven stochastic differential equation (see [17]), close in spirit to those examined, e.g., in [26,30,32]. The following example provides an interpretation of Proposition 7.2 in this particular case.

Example 7.2 (*A Poisson-driven stochastic differential equation*)**.** Suppose that Y is a closed subset of a separable Hilbert space H, endowed with an inner product $\langle \cdot \mid \cdot \rangle$, and that $\|\cdot\|$ is the norm induced by $\langle \cdot \mid \cdot \rangle$. Further, let ϑ be a Borel probability measure on Θ.

Consider a Poisson random measure $\mathbf{N}$ on $\mathcal{B}(\mathbb{R}_+ \times \Theta)$ with compensator $\ell_1 \otimes \lambda\vartheta$ (where ℓ_1 stands for the Lebesgue measure restricted to $\mathcal{B}(\mathbb{R}_+)$), i.e., such that the expectation of $\mathbf{N}(t,D) := \mathbf{N}([0,t]\times D)$ is equal to $t\lambda\vartheta(D)$ for any $t \geq 0$ and $D \in \mathcal{B}(\Theta)$. By [40, Theorem 54 and Corollary 55] (cf. also [27]) we know that such a random measure do exist on some probability space, and, what is more, the proof of this result shows that it can be written as

$$\mathbf{N}(t,D) = \sum_{n=1}^{\infty} \mathbb{1}_{[0,t]\times D}(\tau_n, \theta_n) \quad \text{for any} \quad t \geq 0,\ D \in \mathcal{B}(\Theta),$$

where $\{\tau_n\}_{n\in\mathbb{N}}$ and $\{\theta_n\}_{n\in\mathbb{N}}$ are independent sequences of random variables taking values in $\mathbb{R}_+$ and Θ, respectively, such that the increments $\Delta\tau_n := \tau_n - \tau_{n-1}$, $n \in \mathbb{N}$ (where $\tau_0 := 0$) are mutually independent and exponentially distributed with the same rate λ, while θ_n, $n \in \mathbb{N}$, are identically distributed with common density h. Then, for any given $D \in \mathcal{B}(\Theta)$, the variables τ_n can be viewed as the jump times of the Poisson process $\{\mathbf{N}(t,D)\}_{t\in\mathbb{R}_+}$.

Finally, given continuous functions $\sigma : Y \times \Theta \to Y$, $A_i : Y \to H$, $i \in I$, and a sequence $\{\xi_n\}_{n\in\mathbb{N}_0}$ of I-valued random variables, consider the stochastic differential equation of the form

$$dY(t) = A_{\xi(t)}(Y(t))dt + \int_\Theta \sigma(Y(t-),\theta)\,\mathbf{N}(dt,\,d\theta), \quad t \geq 0, \tag{7.15}$$

for an unknown process $\{Y(t)\}_{t\geq 0}$ taking values in Y, with initial condition $Y(0) = Y_0$, where $\xi(t) := \xi_n$ whenever $\mathbf{N}(t,\Theta) = n$ (or equivalently, $t \in [\tau_n, \tau_{n+1})$) for every $n \in \mathbb{N}$. By a solution of this initial value problem we mean a càdlàg process $\{Y(t)\}_{t\in\mathbb{R}_+}$ satisfying

$$Y(t) = Y_0 + \int_0^t A_{\xi(s)}(Y(s))ds + \int_0^t \int_\Theta \sigma(Y(s-),\theta)\,\mathbf{N}(ds,\,d\theta), \quad t \geq 0.$$

As we have mentioned in Remark 4.3, upon assuming that A_i are dissipative and enjoy the range condition, that is, (4.15) and (4.16) hold with some $\alpha \leq 0$, for every $i \in I$, there exists a semiflow $S_i : \mathbb{R}_+ \times Y \to Y$ such that $\mathbb{R}_+ \ni t \mapsto S_i(t,y)$ is the unique solution of the problem $u'(t) = A_i u(t)$, $u(0) = y$ for any $y \in Y$. Then, following the reasoning in [17, §5.2], one can show that the solution of (7.15) has the form

$$Y(t) = \begin{cases} S_{\xi_n}(t-\tau_n, Y(\tau_n)) & \text{if } t \in [\tau_n, \tau_{n+1}), \\ w_{\theta_{n+1}}(Y(\tau_{n+1}-)) & \text{if } t = \tau_{n+1}, \end{cases} \tag{7.16}$$

where $w_\theta(y) := y + \sigma(y,\theta)$ for $y \in Y$ and $\theta \in \Theta$, and $\{\theta_n\}_{n\in\mathbb{N}}$ is a sequence of Θ-valued random variables with conditional distributions specified by (7.8) with $p_\theta(y) = 1$ for all $y \in Y$ and $\theta \in \Theta$. This means that the Markov process $\Psi := \{(Y(t),\xi(t))\}_{t\in\mathbb{R}_+}$ introduced in Section 3, for which the kernel J is given by (7.6) with w_θ and p_θ defined as above, solves (7.15). We can therefore conclude from Proposition 7.2 that Ψ is V-exponentially ergodic in $d_{FM,c}$ with $V(y,i) := \|y\|$ (for $(y,i) \in Y \times I$) and sufficiently large c, provided that the following conditions hold:

(i) For some $y^* \in Y$ we have $\int_\Theta \|\sigma(y^*,\theta)\|\, \vartheta(d\theta) < \infty$;
(ii) There exists $a_\sigma > 0$ such that

$$\int_\Theta \|\sigma(y_1,\theta) - \sigma(y_2,\theta)\|\, \vartheta(d\theta) \le a_\sigma \|y_1 - y_2\| \quad \text{for any} \quad y_1, y_2 \in Y;$$

(iii) There exists $\eta > 0$ such that $\vartheta(\{\theta \in \Theta : \|\sigma(y_1,\theta) - \sigma(y_2,\theta)\| \le a_\sigma \|y_1 - y_2\|\}) \ge \eta$ for all $y_1, y_2 \in Y$;
(iv) For every $i \in I$, the operator A_i is bounded on bounded sets, it satisfies (4.15) with $\alpha < -\lambda a_\sigma$, and (4.16) holds.
(v) $\min_{i\in I} \pi_{ij_0} > 0$ for some $j_0 \in I$.

To see this, it suffices to observe that hypotheses (i)–(iii) imply that (7.9)–(7.11) are fulfilled with $y^* = 0$, $\tilde{a} := 1 + a_\sigma$ and the given η, whilst (7.12) holds trivially. Moreover, using (iv) we can deduce (by referring to Remark 4.3) that (S1)–(S3) are satisfied with the given α, $L = 1$, $\mathcal{L}(y) = 2\max_{i\in I} \|A_i y\|$ and $\varphi(t) = t$, and that $\tilde{a} + \alpha/\lambda = 1 + a_\sigma + \alpha/\lambda < 1$, which yields (4.8). Finally, condition (S4) is just assumed in (v).

It is worth stressing here that, in fact, the assumption of the boundedness on bounded sets in (iii), imposed on A_i, is only needed in the presence of switching, i.e., if I contains more than one element (since otherwise (S3) is satisfied trivially).

Remark 7.3. Let us note that Eq. (7.15) (without switching between dynamics) also resembles the one considered in [35], except the latter involves the so-called compensated Poisson measure instead of $\mathbf{N}$ itself. In the setup of Example 7.2, the equation from [35] takes the form

$$dY(t) = A(Y(t))dt + \int_\Theta \sigma(Y(t-),\theta)\,(\mathbf{N} - \ell_1 \otimes \lambda\vartheta)(dt, d\theta). \tag{7.17}$$

In [35], the authors assume that Y is a reflexive Banach space such that $Y \hookrightarrow H \hookrightarrow Y^*$ (where Y^* is the dual of Y) with dense and continuous embeddings, where $Y \hookrightarrow H$ is also compact, $A : Y \to Y^*$ is a strongly-weakly closed operator, and $\sigma : H \times \Theta \to H$ is a measurable map such that

$$\int_\Theta \|\sigma(y,\theta)\|^2\, \vartheta(d\theta) < \infty \quad \text{for any} \quad y \in Y. \tag{7.18}$$

Within this framework, they prove (see [35, Theorem 2.4]), among others, that the initial value problem associated with (7.17) admits a unique solution process (with almost all values in Y), which is V-exponentially ergodic (with $V(y) = \|y\|$) in the Fortet–Mourier distance. This is done (by entirely different methods than those used here) under condition (7.18), certain *coercivity and growth* hypotheses (involving the maps A and σ), and the assumption that

$$2\langle Ay_1 - Ay_2 \mid y_1 - y_2\rangle + \lambda \int_\Theta \|\sigma(y_1,\theta) - \sigma(y_1,\theta)\|^2\, \vartheta(d\theta) \le \tilde{\alpha} \|y_1 - y_2\|^2 \tag{7.19}$$

for any $y_1, y_2 \in Y$ and some $\tilde{\alpha} < 0$. In connection with Example 7.2, it is worth observing that, if we strengthen hypotheses (i) and (ii) by requiring, instead, that

(i′) $\int_\Theta \|\sigma(y^*,\theta)\|^2\, \vartheta(d\theta) < \infty$ for some $y^* \in Y$,
(ii′) There exists $a_\sigma > 0$ such that

$$\int_\Theta \|\sigma(y_1,\theta) - \sigma(y_2,\theta)\|^2\, \vartheta(d\theta) \le a_\sigma \|y_1 - y_2\|^2 \quad \text{for any} \quad y_1, y_2 \in Y,$$

then we will get (7.18), and by applying (ii′) in conjunction with (iv), we can also derive (7.19) with $\tilde{\alpha} = 2\alpha + \lambda a_\sigma$, where $\tilde{\alpha} < 0$ due to the assumption that $\alpha < -\lambda a_\sigma$. This comparison shows that the exponential ergodicity for the solution of (7.17) can also be attained by replacing (7.18) and (7.19) by conditions similar to those in Example 7.2, i.e. (i′), (ii′) and (iv) (at least under the coercivity and growth assumptions from [35]).

At the end of this section, let us illustrate the usefulness of Proposition 7.2 by investigating a simple model of gene expression, which has already been mentioned in the introduction.

Example 7.3 (*Gene Expression*)**.** Let $Y(t) \in \mathbb{R}_+$ describe the concentration of some protein (encoded by a single structural gene) in a prokaryotic cell at time t. The protein molecules undergo degradation, which is interrupted by transcription occurring in the so-called bursts, followed by variable periods of inactivity. We assume that the bursts appear at random times $\tau_1 < \tau_2 < \cdots$, and that the inactivity periods $\varDelta\tau_n$ are exponentially distributed with a common intensity λ, which is reasonable from the biological perspective. Since a prokaryotic mRNA can be efficiently transcribed and translated at the same time, the variables τ_n determine the moments of the protein production as well.

Moreover, we assume that the degradation rate depends linearly on the current amount of the gene product, and the proportionality coefficient changes under the influence of the bursts (which may cause occasional fluctuations in the environment, somehow perturbing the degradation dynamics). More precisely, given a finite collection $\{\alpha_i : i \in I\}$ of negative real numbers, we assume that the dynamics of the degradation process in the time interval $[\tau_n, \tau_{n+1})$, provided that $Y(\tau_n) = y$, is governed by one of the initial-value problems

$$u'(t) = \alpha_i\, u(t), \quad u(0) = y, \quad \text{where} \quad i \in I,$$

which generate the semiflows of the form $S_i(t, y) = e^{\alpha_i t} y$ for $t, y \geq 0$ and $i \in I$. Furthermore, we assume that the index i of the dynamics in the interval $[\tau_n, \tau_{n+1})$ is specified by an I-valued random variable ξ_n, which depends only on ξ_{n-1}, so that (3.5) holds. Consequently, we get $Y(t) = S_{\xi_n}(t - \tau_n, y)$ for $t \in [\tau_n, \tau_{n+1})$, whenever $Y(\tau_n) = y$.

Finally, we let the amount of the protein produced at time τ_n be a random variable θ_n with values in some compact interval $\varTheta := [0, M]$, and we assume that it depends on $Y(\tau_n-)$ in accordance with (7.8). The process $\{Y(t)\}_{t\geq 0}$ then changes from $Y(\tau_n-)$ to $Y(\tau_n) = Y(\tau_n-) + \theta_n$ for every $n \in \mathbb{N}$.

In view of the above, defining $w_\theta(y) := y + \theta$ for $y \in \mathbb{R}_+$, $\theta \in \varTheta$, and putting $\tau_0 := 0$, we see that $\{Y(t)\}_{t\in\mathbb{R}_+}$ (evolving on $Y = \mathbb{R}_+$) is of the form (7.16). Hence the process $\varPsi := \{(Y(t), \xi(t))\}_{t\in\mathbb{R}_+}$, where $\xi(t) := \xi_n$ for $t \in [\tau_n, \tau_{n+1})$, can be viewed as an instance of the PDMP introduced in Section 3, for which the kernel J is of the form (7.6). Furthermore, it is easily seen that conditions (7.9), (7.10) and (S1)–(S3) hold for this model with $y^* = 0$, $\tilde{a} = 1$, $L = 1$, $\alpha = \max_{i\in I} \alpha_i$, $\varphi(t) = 2e^{\alpha t}$ and $\mathcal{L}(y) = y$. Since $\alpha < 0$, we also get $\tilde{a}L + \alpha\lambda^{-1} = 1 + \alpha\lambda^{-1} < 1$, which ensures (4.8). Consequently, according to Proposition 7.2, the process $\varPsi$ is V-exponentially ergodic in $d_{FM,c}$ with $V(y, i) = y$ and sufficiently large c, provided that $\min_{i\in I} \pi_{ij_0} > 0$ for some $j_0 \in I$ and the densities $\theta \mapsto p_\theta(y)$, $y \in Y$, are such that (7.11) and (7.12) hold. One can check that the latter is attained whenever $c \geq 8M(-\alpha)^{-1}(\lambda - \alpha)(2\lambda - \alpha) + 1$.

7.3. Proof of Proposition 7.1

Let P and Q_P be the kernels defined by (3.4) and (7.5), respectively. Moreover, consider the augmented coupling $\widehat{\varPhi}$ of the chain $\varPhi$ (constructed in Section 6) with transition law $\widehat{P}$ defined by (7.4). In particular, $\{(\varPhi_n^{(1)}, \varPhi_n^{(2)})\}_{n\in\mathbb{N}_0}$ itself is then governed by the kernel

$$\widetilde{P}((x_1, x_2), C) := \widehat{P}((x_1, x_2, 0), C \times \mathbb{R}_+) \quad \text{for} \quad x_1, x_2 \in X,\ C \in \mathcal{B}(X^2). \tag{7.20}$$

In order to prove Proposition 7.1, we first need to derive hypotheses (B1)–(B5), used in [15, Section 2] (and also assumed in [29, Theorem 2.1]), from conditions (S1)–(S4), (J1) and (J2), imposed in combination with (4.8) and (4.12). To begin, let a, b be the constants specified in (4.11), and let V stand for the function given by (4.9). Further, define $F := G \cup K$, where

$$\begin{gathered} G := \{((y_1, i_1), (y_2, i_2)) \in X^2 : i_1 = i_2\}, \\ K := \{((y_1, i_1), (y_2, i_2)) \in X^2 : V(y_1, i_1) + V(y_2, i_2) < R\} \quad \text{with} \quad R := \frac{4b}{1-a}. \end{gathered} \tag{7.21}$$

Lemma 7.1. *Suppose that conditions (S1)–(S4), (J1) and (J2) hold with $L, \alpha, \tilde{a}$ satisfying (4.8) and that c is large enough to assure (4.12). Then the following statements are fulfilled:*

(B1) *We have $a \in (0,1)$, $b \geq 0$, and $PV(x) \leq aV(x) + b$ for all $x \in X$.*

(B2) $\operatorname{supp} Q_P((x_1, x_2), \cdot) \subset F$ *and*

$$\int_{X^2} \rho_{X,c}(w_1, w_2)\, Q_P((x_1, x_2), dw_1 \times dw_2) \leq a\rho_{X,c}(x_1, x_2) \quad \textit{for any} \quad (x_1, x_2) \in F.$$

(B3) *Defining $U(r) := \{(w_1, w_2) \in X^2 : \rho_{X,c}(w_1, w_2) \leq r\}$ for any $r > 0$, we have*

$$\inf_{(x_2, x_2) \in F} Q_P((x_1, x_2), U(a\rho_{X,c}(x_1, x_2))) > 0,$$

(B4) *There exists $l > 0$ such that*

$$Q_P((x_1, x_2), X^2) \geq 1 - l\rho_{X,c}(x_1, x_2) \quad \textit{for any} \quad (x_1, x_2) \in F.$$

(B5) *There exist $\gamma \in (0,1)$ and $C_\gamma > 0$ such that*

$$\widehat{\mathbb{E}}_{(x_1, x_2)}(\gamma^{-\sigma_K}) \leq C_\gamma \quad \textit{whenever} \;\; V(x_1) + V(x_2) < R,$$

where $\sigma_K := \inf\{n \in \mathbb{N} : (\Phi_n^{(1)}, \Phi_n^{(2)}) \in K\}$.

Proof. First of all, note that condition (B1) is guaranteed by Lemma 4.1.

The proof of the first part of (B2) goes as follows. Let $(x_1, x_2) := ((y_1, i_1), (y_2, i_2)) \in X^2$. Since X^2 is endowed with the product topology, we may consider it as a metric space with the distance

$$\rho_{X^2,c}((w_1, w_2), (z_1, z_2)) := \rho_{X,c}(w_1, z_1) + \rho_{X,c}(w_2, z_2) \quad \text{for} \quad (w_1, w_2), (x_1, x_2) \in X^2.$$

The support of $Q_P(x_1, x_2, \cdot)$ can be then expressed as

$$\operatorname{supp} Q_P((x_1, x_2), \cdot) = \{(z_1, z_2) \in X^2 : Q_P((x_1, x_2), B_{X^2}((z_1, z_2), \varepsilon)) > 0 \;\text{ for any }\; \varepsilon > 0\},$$

where $B_{X^2}((z_1, z_2), \varepsilon)$ is the open ball in $(X^2, \rho_{X^2,c})$ centred at (z_1, z_2) with radius ε. Let

$$(z_1, z_2) := ((u_1, j_1), (u_2, j_2)) \in X^2 \backslash F.$$

Then, in particular, $(z_1, z_2) \notin G$, and thus $j_1 \neq j_2$. This implies that, for any $(w_1, w_2) := ((v_1, j), (v_2, j)) \in G$, we have

$$\rho_{X^2,c}((w_1, w_2), (z_1, z_2)) \geq c(\mathbf{d}(j, j_1) + \mathbf{d}(j, j_2)) \geq c,$$

whence $B_{X^2}((z_1, z_2), c) \cap G = \emptyset$. Taking into account the definition of Q_P, given in (7.5), we therefore obtain

$$Q_P((x_1, x_2), B_{X^2}((z_1, z_2), c)) = Q_P((x_1, x_2), B_{X^2}((z_1, z_2), c) \cap G) = 0,$$

which yields that $(z_1, z_2) \in X^2 \backslash \operatorname{supp} Q_P((x_1, x_2), \cdot)$.

Passing to the proof of the second part of (B2), let $(x_1, x_2) = ((y_1, i_1), (y_2, i_2)) \in F$. Then $i_1 = i_2$ or $y_1, y_2 \in B_Y(y^*, R)$ (due to the definition of V). Hence, from (S2) and (S3) it follows that

$$\begin{aligned} \rho_Y(S_{i_1}(t, y_1), S_{i_2}(t, y_2)) &\le \rho_Y(S_{i_1}(t, y_1), S_{i_1}(t, y_2)) + \rho_Y(S_{i_1}(t, y_2), S_{i_2}(t, y_2)) \\ &\le Le^{\alpha t}\rho_Y(y_1, y_2) + \varphi(t)\mathcal{L}(y_2)\mathbf{d}(i_1, i_2) \\ &\le Le^{\alpha t}\rho_Y(y_1, y_2) + \varphi(t)M_{\mathcal{L}}\mathbf{d}(i_1, i_2), \end{aligned} \tag{7.22}$$

where $M_{\mathcal{L}}$ is given by (4.13). Consequently, referring to the definition of Q_P, (4.4) and (4.12), we can conclude that

$$\begin{aligned} &\int_{X^2} \rho_{X,c}(w_1, w_2)\, Q_P((x_1, x_2), dw_1 \times w_2) \\ &= \sum_{j\in I}(\pi_{i_1,j} \wedge \pi_{i_2,j}) \int_0^\infty \lambda e^{-\lambda h} \int_{Y^2} \rho_Y(v_1, v_2)\, Q_J((S_{i_1}(h, y_1), S_{i_2}(h, y_2)), dv_1 \times dv_2)\, dh \\ &\le \tilde{a}\lambda \int_0^\infty e^{-\lambda h}\rho_Y(S_{i_1}(h, y_1), S_{i_2}(h, y_2))\, dh \\ &\le \tilde{a}\lambda L \left(\int_0^\infty e^{-(\lambda-\alpha)h}\, dh\right) \rho_Y(y_1, y_2) + \tilde{a}\lambda M_{\mathcal{L}} \left(\int_0^\infty e^{-\lambda h}\varphi(h)\, dh\right) \mathbf{d}(i_1, i_2) \\ &= \frac{\tilde{a}\lambda L}{\lambda-\alpha}\rho_Y(y_1, y_2) + \frac{\tilde{a}\lambda L}{\lambda-\alpha}\, \frac{(\lambda-\alpha)M_{\mathcal{L}}K_\varphi}{L}\, \mathbf{d}(i_1, i_2) \le a(\rho_Y(y_1, y_2) + c\mathbf{d}(i_1, i_2)) \\ &= a \cdot \rho_{X,c}(x_1, x_2), \end{aligned}$$

which is the desired claim.

We now proceed to show condition (B3). First, define $t_0 := \lim_{s\to\alpha} s^{-1} \ln\left(\lambda(\lambda - s)^{-1}\right)$, which is obviously positive, and observe that

$$\tilde{a}Le^{\alpha t} \le \frac{\tilde{a}\lambda L}{\lambda-\alpha} = a \quad \text{for any} \quad t \le t_0. \tag{7.23}$$

Now, let $(x_1, x_2) = ((y_1, i_1), (y_2, i_2)) \in F$, and note that, for any $u_1, u_2 \in Y$, $j \in I$, and $0 \le t \le t_0$, we have

$$(u_1, u_2) \in \widetilde{U}(\tilde{a}\rho_Y(S_{i_1}(t, y_1), S_{i_2}(t, y_2))) \quad\Rightarrow\quad ((u_1, j), (u_2, j)) \in U(a\rho_{X,c}(x_1, x_2)), \tag{7.24}$$

where $\widetilde{U}(\cdot)$ and $U(\cdot)$ are defined as in (4.6) and (B3), respectively. To see this, it suffices to apply conditions (7.22), (7.23) and (4.12), which ensure that, for any $t \le t_0$ and every $(u_1, u_2) \in U(\tilde{a}\rho_Y(S_{i_1}(t, y_1), S_{i_2}(t, y_2)))$,

$$\begin{aligned} \rho_{X,c}((u_1, j), (u_2, j)) = \rho_Y(u_1, u_2) &\le \tilde{a}\rho_Y(S_{i_1}(t, y_1), S_{i_2}(t, y_2)) \\ &\le \tilde{a}Le^{\alpha t}\rho_Y(y_1, y_2) + \tilde{a}M_{\mathcal{L}}\varphi(t)\mathbf{d}(i_1, i_2) \\ &\le a\rho_Y(y_1, y_2) + \tilde{a}M_{\mathcal{L}}M_\varphi\mathbf{d}(i_1, i_2) \\ &= a\left(\rho_Y(y_1, y_2) + \frac{M_{\mathcal{L}}M_\varphi(\lambda-\alpha)}{\lambda L}\mathbf{d}(i_1, i_2)\right) \le a\rho_{X,c}(x_1, x_2), \end{aligned}$$

where M_φ given by (4.14), whence $((u_1, j), (u_2, j)) \in U(a\rho_{X,c}(x_1, x_2))$. Now, using (7.24), together with (4.5), we obtain

$$\begin{aligned} &\int_{Y^2} \mathbb{1}_{U(a\rho_{X,c}(x_1,x_2))}((u_1, j), (u_2, j))Q_J((S_{i_1}(h, y_1), S_{i_2}(h, y_2)), du_1 \times du_2) \\ &\ge \int_{Y^2} \mathbb{1}_{\widetilde{U}\left(\tilde{a}\rho_Y(S_{i_1}(h,y_1),S_{i_2}(h,y_2))\right)}(u_1, u_2)Q_J((S_{i_1}(h, y_1), S_{i_2}(h, y_2)), du_1 \times du_2) \\ &= Q_J((S_{i_1}(h, y_1), S_{i_2}(h, y_2)), \widetilde{U}(\tilde{a}\rho_Y(S_{i_1}(h, y_1), S_{i_2}(h, y_2)))) \ge \eta \quad \text{for any} \quad h \le t_0. \end{aligned}$$

Finally, from (S4) it follows that

$$\begin{aligned}
&Q_P((x_1,x_2),U(a\rho_{X,c}(x_1,x_2)))\\
&\qquad\ge \sum_{j\in I}(\pi_{i_1,j}\wedge\pi_{i_2,j})\int_0^{t_0}\lambda e^{-\lambda h}\int_{Y^2}\mathbb{1}_{U(a\rho_{X,c}(x_1,x_2))}((u_1,j),(u_2,j))\\
&\qquad\quad\times Q_J((S_{i_1}(h,y_1),S_{i_2}(h,y_2)),du_1\times du_2)\,dh\\
&\qquad\ge (\min_{i\in I}\pi_{ij_0})\eta(1-e^{-\lambda t_0})>0,
\end{aligned}$$

with η defined by (4.5), which gives (B3).

Now, we shall establish condition (B4). To do this, fix $(x_1,x_2):=((y_1,i_1),(y_2,i_2))\in F$, and note that, due to (4.7),

$$\begin{aligned}
Q_P((x_1,x_2),X^2)&=\sum_{j\in I}(\pi_{i_1,j}\wedge\pi_{i_2,j})\int_0^\infty\lambda e^{-\lambda h}Q_J((S_{i_1}(h,y_1),S_{i_2}(h,y_2)),Y^2)\,dh\\
&\ge\sum_{j\in I}(\pi_{i_1,j}\wedge\pi_{i_2,j})-\tilde{l}\lambda\int_0^\infty e^{-\lambda h}\rho_Y(S_{i_1}(h,y_1),S_{i_2}(h,y_2))\,dh.
\end{aligned}\tag{7.25}$$

On other hand, referring again to (7.22), we get

$$\begin{aligned}
&\int_0^\infty e^{-\lambda h}\rho_Y(S_{i_1}(h,y_1),S_{i_2}(h,y_2))\,dh\\
&\quad\le L\left(\int_0^\infty e^{-(\lambda-\alpha)h}\,dh\right)\rho_Y(y_1,y_2)+M_{\mathcal{L}}\left(\int_0^\infty e^{-\lambda h}\varphi(h)\,dh\right)\mathbf{d}(i_1,i_2)\\
&\quad\le\frac{L}{\lambda-\alpha}\rho_Y(y_1,y_2)+M_{\mathcal{L}}K_\varphi\mathbf{d}(i_1,i_2).
\end{aligned}\tag{7.26}$$

Moreover, we can write

$$\sum_{j\in I}\pi_{i_1,j}\wedge\pi_{i_2,j}\ge 1-\mathbf{d}(i_1,i_2).\tag{7.27}$$

Consequently, taking into account (7.25), (7.26), (7.27) and (4.12), we infer that

$$\begin{aligned}
Q_P((x_1,x_2),X^2)&\ge 1-\frac{\tilde{l}\lambda L}{\lambda-\alpha}\rho_Y(y_1,y_2)-(1+\tilde{l}\lambda M_{\mathcal{L}}K_\varphi)\mathbf{d}(i_1,i_2)\\
&\ge 1-\left(\frac{\tilde{l}\lambda L}{\lambda-\alpha}+1+\tilde{l}\lambda M_{\mathcal{L}}K_\varphi\right)\rho_{X,c}(x_1,x_2),
\end{aligned}$$

which completes the proof of (B4).

What remains is to show that (B5) holds. To do this, we shall apply [29, Lemma 2.2] to the coupling $\{(\Phi_n^{(1)},\Phi_n^{(2)})\}_{n\in\mathbb{N}_0}$ with transition law $\widetilde{P}$, given by (7.20). For this purpose, it suffices to observe that, letting

$$\widetilde{V}(x_1,x_2):=V(x_1)+V(x_2)=\rho_Y(y_1,y^*)+\rho_Y(y_2,y^*)\quad\text{for}\quad x_1=(y_1,i_1),x_2=(y_2,i_2)\in X^2,$$

we have

$$\widetilde{P}\widetilde{V}(x_1,x_2)\le a\widetilde{V}(x_1,x_2)+2b\quad\text{for any}\quad(x_1,x_2)\in X^2,$$

which follows directly from (B1) and (7.22). The proof of Lemma 7.1 is now complete. □

Having established Lemma 7.1, we can prove Proposition 7.1 by arguing as in the proof of [15, Lemma 2.1]. First of all, we need to be able to distinguish the case where the next step of the chain $\widehat{\Phi}$ is drawn only according to $\bar{Q}_P$ from the case when it is determined only by $\bar{R}_P$. For this aim, we consider $Z^\star:=Z\times\{0,1\}$, which can be viewed as a copy of $Z = X^2 \times \mathbb{R}_+$, split into two disjoint subsets $Z \times \{0\}$ and $Z \times \{1\}$. Then we define a new stochastic kernel $\widehat{P}^\star : Z^\star \times \mathcal{B}(Z^\star) \to [0,1]$ by setting

$$\widehat{P}^\star((x_1,x_2,s,k),H) = (\bar{Q}_P((x_1,x_2,s),\cdot) \otimes \delta_1^\star)(H) + (\bar{R}_P((x_1,x_2,s),\cdot) \otimes \delta_0^\star)(H)$$

for any $x_1, x_2 \in X$, $s \in \mathbb{R}_+$, $k \in \{0,1\}$ and $H \in \mathcal{B}(Z^\star)$, where $\delta_0^\star$ (resp. $\delta_1^\star$) stands for the Dirac measure at 0 (resp. at 1) on $2^{\{0,1\}}$. Obviously, for any $(x_1,x_2,s,k) \in Z^\star$ and $D \in \mathcal{B}(Z)$, we have

$$\begin{aligned}
\widehat{P}^\star((x_1,x_2,s,k), D \times \{1\}) &= \bar{Q}_P((x_1,x_2,s),D),\\
\widehat{P}^\star((x_1,x_2,s,k), D \times \{0\}) &= \bar{R}_P((x_1,x_2,s),D),\\
\widehat{P}^\star((x_1,x_2,s,k), D \times \{0,1\}) &= \widehat{P}((x_1,x_2,s),D).
\end{aligned}$$

Further, we introduce the canonical Markov chain $\widehat{\Phi}^\star := \{(\widehat{\Phi}'_n, \kappa_n)\}_{n\in\mathbb{N}_0}$ with transition law $\widehat{P}^\star$, wherein $\widehat{\Phi}' = \{(\Phi_n'^{(1)}, \Phi_n'^{(2)}, \tau_n')\}_{n\in\mathbb{N}_0}$ is an appropriate copy of $\widehat{\Phi}$, and κ_n takes values in $\{0,1\}$. We therefore assume that $\widehat{\Phi}^\star$ is defined on the space $(\Omega^\star, \mathcal{F}^\star) := \left((Z^\star)^{\mathbb{N}_0}, \mathcal{B}\big((Z^\star)^{\mathbb{N}_0}\big)\right)$, equipped with an appropriate family $\left\{\widehat{\mathbb{P}}^\star_{(x_1,x_2)} : x_1, x_2 \in X\right\}$ of probability measures on $\mathcal{F}^\star$, such that $\widehat{\Phi}^\star$ starts at $((x_1,x_2,0),0)$ almost surely with respect to $\widehat{\mathbb{P}}^\star_{(x_1,x_2)}$. The symbol $\widehat{\mathbb{E}}^\star_{(x_1,x_2)}$ will denote the expectation operator corresponding to $\widehat{\mathbb{P}}^\star_{(x_1,x_2)}$.

Let us now fix arbitrarily $(x_1,x_2) \in X^2$ and $n, M, N \in \mathbb{N}$ such that $n > M > N$. Further, consider the random times

$$\sigma_K^{(N)} := \inf\left\{m \geq N : \left(\Phi_m'^{(1)}, \Phi_m'^{(2)}\right) \in K\right\}, \quad \zeta := \inf\{m \in \mathbb{N} : \kappa_i = 1 \text{ for any } i \geq m\},$$

where K is given by (7.21), and define

$$H_{N,n} := \{\kappa_N = \kappa_{N+1} = \cdots = \kappa_n = 1\}, \quad H_{N,n}^c := \Omega^\star \backslash H_{N,n}.$$

Taking into account that $\widehat{\mathbb{P}}^\star(H_{N,n}^c) \leq \widehat{\mathbb{P}}^\star(\zeta > N)$, and that $\bar{\rho}_{X,c}(w_1,w_2) \leq 1$ for any $w_1, w_2 \in X$, we can write the following estimate:

$$\begin{aligned}
\widehat{\mathbb{E}}_{(x_1,x_2)}\left[\bar{\rho}_{X,c}\left(\Phi_n^{(1)}, \Phi_n^{(2)}\right)\right] &= \widehat{\mathbb{E}}^\star_{(x_1,x_2)}\left[\bar{\rho}_{X,c}\left(\Phi_n'^{(1)}, \Phi_n'^{(2)}\right)\right]\\
&= \int_{X^2} \bar{\rho}_{X,c}(w_1,w_2)\, \widehat{\mathbb{P}}^\star_{(x_1,x_2)}\left((\Phi_n'^{(1)}, \Phi_n'^{(2)}) \in dw_1 \times dw_2\right)\\
&\leq \int_{X^2} \bar{\rho}_{X,c}(w_1,w_2)\, \widehat{\mathbb{P}}^\star_{(x_1,x_2)}|_{\left\{\sigma_K^{(N)} \leq M\right\} \cap H_{N,n}}\left((\Phi_n'^{(1)}, \Phi_n'^{(2)}) \in dw_1 \times dw_2\right)\\
&\quad + \widehat{\mathbb{P}}^\star_{(x_1,x_2)}(\sigma_K^{(N)} > M) + \widehat{\mathbb{P}}^\star_{(x_1,x_2)}(\zeta > N),
\end{aligned}$$

with the convention that $\widehat{\mathbb{P}}^\star_{(x_1,x_2)}|_H(\cdot) := \widehat{\mathbb{P}}^\star_{(x_1,x_2)}(H \cap \cdot)$.

Since, according to Lemma 7.1, hypotheses (B1)–(B5) are fulfilled, we can now apply [15, Lemma 2.2] to conclude that there exist constants $C_1, C_2, C_3 \geq 0$, $q_1, q_2, q_3 \in (0,1)$ and $p \geq 1$ such that

$$\widehat{\mathbb{E}}_{(x_1,x_2)}\left[\bar{\rho}_{X,c}\left(\Phi_n^{(1)}, \Phi_n^{(2)}\right)\right] \leq (C_1 q_1^{n-M} + C_2 q_2^{M-pN} + C_3 q_3^N)(1 + V(x_1) + V(x_2)).$$

Finally, letting $n \geq \lceil 4p \rceil$ and taking $N = \lfloor n/(4p) \rfloor$, $M = \lfloor n/2 \rfloor$, we obtain

$$\widehat{\mathbb{E}}_{(x_1,x_2)}\left[\bar{\rho}_{X,c}\left(\Phi_n^{(1)}, \Phi_n^{(2)}\right)\right] \leq \widetilde{C}_0(V(x_1) + V(x_2) + 1)q^n$$

with $q := \{q_1^{1/2}, q_2^{1/4}, q_3^{1/(4p)}\}$ and $\widetilde{C}_0 = \max\{q_1^{-1}, q_3^{-1}\}(C_1 + 2C_2 + 2C_3)$. Obviously, since $\bar{\rho}_{X,c} \leq 1$, this inequality holds, in fact, for all $n \in \mathbb{N}$ with $C_0 := q^{-\lceil 4p \rceil} \max\{\widetilde{C}_0, 1\}$ in the place of $\widetilde{C}_0$. The proof of Proposition 7.1 is now complete.

Funding

The work of Hanna Wojewódka-Ściążko has been partly supported by the National Science Centre of Poland, grant number 2018/02/X/ST1/01518.

References

[1] M. Benaïm, T. Hurth, E. Strickler, A user-friendly condition for exponential ergodicity in randomly switched environments, Electron. Commun. Probab. 23 (2018) 1–12.
[2] M. Benaïm, S. Le Borgne, F. Malrieu, P.-A. Zitt, Quantitative ergodicity for some switched dynamical systems, Electron. Commun. Probab. 17 (0) (2012) 1–14.
[3] M. Benaïm, S. Le Borgne, F. Malrieu, P.-A. Zitt, Qualitative properties of certain piecewise deterministic Markov processes, Ann. Inst. H. Poincaré Probab. Statist. 51 (3) (2015) 1040–1075.
[4] M. Benaïm, E. Strickler, Random switching between vector fields having a common zero, Ann. Appl. Probab. 29 (1) (2019) 326–375.
[5] W. Biedrzycka, M. Tyran-Kamińska, Existence of invariant densities for semiflows with jumps, J. Math. Anal. Appl. 435 (1) (2016) 61–84.
[6] V. Bogachev, Measure Theory, Vol. 2, Springer-Verlag, Berlin, 2007.
[7] O. Boxma, H. Kaspi, O. Kella, D. Perry, On/off storage systems with state-dependent inpout, Probab. Engrg. Inform. Sci. 19 (1) (2005) 1–14.
[8] E. Buckwar, M. Riedler, An exact stochastic hybrid model of excitable membranes including spatio-temporal evolution, J. Math. Biol. 63 (6) (2011) 1051–1093.
[9] B. Cloez, R. Dessalles, A. Genadot, F. Malrieu, A. Marguet, R. Yvinec, Probabilistic and piecewise deterministic models in biology, ESAIM: Proc. Surv. 60 (2017) 225–245.
[10] B. Cloez, M. Hairer, Exponential ergodicity for Markov processes with random switching, Bernoulli 21 (1) (2015) 505–536.
[11] O. Costa, F. Dufour, Stability of piecewise deterministic Markov processes, SIAM J. Control Optim. 37 (5) (1999) 1483–1502.
[12] O. Costa, F. Dufour, Stability and ergodicity of piecewise deterministic Markov processes, in: 2008 47th IEEE Conference on Decision and Control, IEEE, 2008.
[13] A. Crudu, A. Debussche, A. Muller, O. Radulescu, Convergence of stochastic gene networks to hybrid piecewise deterministic processes, Ann. Appl. Probab. 22 (5) (2012) 1822–1859.
[14] D. Czapla, K. Horbacz, H. Wojewódka-Ściązko, Ergodic properties of some piecewise-deterministic Markov process with application to gene expression modelling, Stoch. Proc. Appl. 130 (5) (2020a) 2851–2885.
[15] D. Czapla, K. Horbacz, H. Wojewódka-Ściązko, A useful version of the central limit theorem for a general class of Markov chains, J. Math. Anal. Appl. 484 (1) (2020b) 123725.
[16] D. Czapla, K. Horbacz, H. Wojewódka-Ściązko, The strassen invariance principle for certain non-stationary Markov-Feller chains, Asymptot. Anal. 121 (1) (2021) 1–34.
[17] D. Czapla, J. Kubieniec, Exponential ergodicity of some Markov dynamical systems with application to a Poisson driven stochastic differential equation, Dyn. Syst. 34 (1) (2019) 130–156.
[18] M. Davis, Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models, J. R. Stat. Soc. Ser. B Stat. Methodol. 46 (3) (1984) 353–388.
[19] M. Davis, Markov Models and Optimization, Chapman and Hall, 1993.
[20] M. Davis, Point Process Theory and Applications: Marked Point and Piecewise Deterministic Processes, Birkhäuser-Verlag, 2006.
[21] R. Douc, E. Moulines, D. Stoffer, Nonlinear Time Series: Theory, Methods and Applications with R Examples, CRC Press, Boca Raton, 2014.
[22] R. Dudley, Convergence of baire measures, Studia Math. 27 (1966) 251–268.
[23] X. Fernique, Intégrabilité des vecteurs gaussiens, C. R. Acad. Sci. Sér. A-B 270 (1) (1970) A1698–A1699.
[24] C. Graham, P. Robert, Interacting multi-class transmissions in large stochastic networks, Ann. Appl. Probab. 19 (6) (2009) 2334–2361.
[25] M. Hairer, Exponential mixing properties of stochastic PDEs through asymptotic coupling, Probab. Theory Related Fields 124 (3) (2002) 345–380.
[26] K. Horbacz, Invariant measures related with randomly connected Poisson driven differential equations, Ann. Polon. Math. 79 (1) (2002) 31–44.
[27] N. Ikeda, S. Wanatabe, Stochastic Differential Equations and Diffusion Processes, North-Holland, Amsterdam, 1981.
[28] K. Ito, F. Kappel, Evolution Equations and Approximations, in: Ser. Adv. Math. Appl. Sci., vol. 61, World Scientific, New Jersey, 2002.
[29] R. Kapica, M. Ślęczka, Random iteration with place dependent probabilities, Probab. Math. Statist. 40 (1) (2020) 119–137.
[30] J. Kazak, Piecewise-deterministic Markov processes, Ann. Polon. Math. 109 (3) (2013) 279–296.
[31] A. Lasota, From fractals to stochastic differential equations, in: P. Garbaczewski, M. Wolf, A. Weron (Eds.), Chaos – the Interplay Between Stochastic and Deterministic Behaviour (Proceedings of the XXXIst Winter School of Theoretical Physics, Karpacz, Poland, 13-24 1995), in: Lecture Notes in Physics, vol. 457, Springer Verlag, Berlin, 1995, pp. 235–255.

[32] A. Lasota, J. Traple, Invariant measures related with Poisson driven stochastic differential equation, Stoch. Proc. Appl. 106 (1) (2003) 81–93.
[33] A. Lasota, J. Yorke, Lower bound technique for Markov operators and iterated function systems, Rand. Comput. Dynam. 2 (1994) 41–77.
[34] M. Mackey, M. Tyran-Kamińska, R. Yvinec, Dynamic behavior of stochastic gene expression models in the presence of bursting, SIAM J. Appl. Math. 73 (5) (2013) 1830–1852.
[35] C. Marinelli, G. Ziglio, Ergodicity for nonlinear stochastic evolution equations with multiplicative Poisson noise, Dyn. Partial Differential Equations 7 (1) (2010) 1–23.
[36] S. Meyn, R. Tweedie, Criteria for stability of Markovian processes III: Foster–Lyapunov criteria for continuous time processes, with examples, Adv. Appl. Probab. 25 (3) (1993) 518–548.
[37] S. Meyn, R. Tweedie, Markov Chains and Stochastic Stability, Springer-Verlag, Berlin, 2003.
[38] D. Revuz, Markov Chains, North-Holland Elsevier, Amsterdam, 1975.
[39] R. Rudnicki, M. Tyran-Kamińska, Piecewise Deterministic Processes in Biological Models, Springer Briefs in Applied Sciences and Technology, Cham, 2017.
[40] R. Situ, Theory of Stochastic Differential Equations with Jumps and Applications, Springer US, 2005.
[41] M. Ślęczka, The rate of convergence for iterated function systems, Studia Math. 205 (3) (2011) 201–214.
[42] T. Szarek, Invariant measures for Markov operators with application to function systems, Studia Math. 154 (3) (2003) 207–222.
[43] H. Wojewódka, Exponential rate of convergence for some Markov operators, Statist. Probab. Lett. 83 (10) (2013) 2337–2347.