Ergodic properties of some piecewise-deterministic Markov process with application to a gene expression model

Katarzyna Horbacz¹ Dawid Czapla² Hanna Wojewódka³

Institute of Mathematics, University of Silesia in Katowice, Bankowa 14, 40-007 Katowice, Poland

Abstract

A piecewise-deterministic Markov process, specified by switching flows and random jumps, as well as the associated chain of post-jump locations, are investigated in this paper. Asymptotic stability with exponential rate of convergence to the unique invariant measure and the strong law of large numbers are proven for the chain. Further, the relations between the aforementioned ergodic properties of the chain and of the corresponding process are studied. In particular, it is established that the existence and the uniqueness of an invariant measure for the continuous-time process, as well as the strong law of large numbers for this process, follow from the results obtained for the discrete-time model. An abstract random dynamical system, associated with the studied process, is inspired by a certain biological model for gene expression, which is also discussed within this paper.

Keywords: Piecewise-deterministic Markov process, invariant measure, asymptotic stability, exponential rate of convergence, strong law of large numbers, gene expression

MSC 2010: 60J05, 60J25, 37A30, 37A25.

Introduction

In this paper we study a class of piecewise-deterministic Markov processes (PDMPs), introduced by M. Davis [8] as a general family of non-diffusion Markov processes involving deterministic motion punctuated by random jumps. Due to its wide applications, especially (but not only) in life and human sciences (see e.g. molecular biology models [5,29], gene expression models [26,27] or models for erythroid production [24,34]), PDMPs have already been widely studied. The research is mainly focused on their long time behaviour and ergodic properties (see e.g. the recent results established by O.L.V. Costa and F. Dufour [7], B. Cloez and M. Hairer [6] or M. Benaim et al. [2,3]).
The paper considers PDMPs arising from dynamical systems governed by a specific jump mechanism. Roughly speaking, the continuous (and deterministic) component of the system evolves according to a finite collection of flows, which are randomly switched with time. However, the randomness of post-jump locations in our model stems not only from the switching flows (like in [2,3]), but also from jumps, which occur directly before choosing a new flow. Each of these jumps is obtained as a randomly chosen transformation of the current position, additionally perturbed by a random shift within an \( \varepsilon \)-neighbourhood. This makes the model more interesting, especially from the point of view of applications. The novelty of our approach is also the fact that the process evolves on a general phase space (which is not necessarily the subset of \( \mathbb{R}^d \), as it is in [2,3,7]). Therefore the applied methods are not conventional and require a certain subtlety.

The evolution of our dynamical system can be described, more precisely, as follows. An initial state and the number of the flow, which transforms it, are described by arbitrarily distributed random variables \( Y_0 \) and \( \xi_0 \), respectively. The process \( Y(t) \) (where \( Y(0) = Y_0 \)) is then driven by the flow \( S_{\xi_0} \) until some random moment \( \tau_1 \), at which it jumps to a random point in the \( \varepsilon \)-neighbourhood of \( \omega_{\theta_1} (S_{\xi_0} (\Delta \tau_1, Y_0)) \), where \( (y, \theta) \mapsto w_\theta (y) \) stands for a given continuous map, and \( \theta_1 \) is a random variable depending on \( Y(t) \). Let \( Y_1 = \overline{Y}(\tau_1) = \omega_{\theta_1} (S_{\xi_0} (\Delta \tau_1, Y_0)) + H_1 \) denote the position of the process directly after this jump. The number of the flow that \( Y(t) \) follows in the time period \( [\tau_1, \tau_2) \) is given by \( \xi_1 \), which depends on the current location \( Y_1 \) and also on \( \xi_0 \). The procedure then restarts for \( (Y_1, \xi_1) \) and is continued inductively. As a result, we obtain a piecewise-deterministic trajectory \( (\overline{Y}(t))_{t \geq 0} \) with jump times \( \tau_1, \tau_2, \ldots \) and post-jump locations \( Y_1, Y_2, \ldots \), as illustrated in the figure below.

![Diagram](image)

Figure 1: The scheme illustrates a certain trajectory of the process \( (\overline{Y}(t))_{t \geq 0} \).

In general, \( (\overline{Y}(t))_{t \geq 0} \) is not a Markov process. Therefore, in order to provide the possibility for analysis through the tools of Markov semigroup theory, we investigate the con-
pled Markov process \((\bar{Y}(t), \bar{\xi}(t))_{t \geq 0}\), defined via interpolation of the coupled Markov chain \((Y_n, \xi_n)_{n \in \mathbb{N}}\), which describes post-jump locations and numbers of the currently running flows.

The above-described model (whose detailed assumptions are summarized in Section 3) generalises, among others, the widely applied random dynamical systems considered in [17,20]. In addition to this, it includes as special cases: a simple cell cycle model due to J.J. Tyson and K.B. Hannsgen [33] (further examined by A. Lasota and M.C. Mackey [23]) and a stochastic model for an autoregulated gene, introduced by S.C. Hille et al. [15]. What is more, we show (in Section 6) how to apply our results to a specific model for gene expression in prokaryotes (which is motivated by [27]). Being more precise, we demonstrate that the concentration of proteins (encoded by genes within a single operon) in the presence of transcriptional bursting (see e.g. [1, Ch.8]) can be modeled by a PDMP consistent with our framework. It is worth stressing here that, within our quite natural assumptions, the existence of a unique invariant measure (not necessarily absolutely continuous) is guaranteed. In contrast, an invariant distribution in [27] can only be obtained by solving explicitly some differential equation (which is possible to do only in certain particular cases) and proving that the solution is a strictly positive probability density function. It should be mentioned, however, that the process of gene expression can be described with just a single flow. On the other hand, we are convinced of the usefulness of considering our abstract model in its full generality. An idea for the future research is to study the possibility of applying it to the so-called metabolic networks, whose mathematical description requires both: random jumps and switching flows.

Let us emphasise that we investigate both: the Markov chain, describing post-jump locations, and the corresponding PDMP. We first prove asymptotic stability (see e.g. [22]), with the exponential rate of convergence in Fortet–Mourier distance (the so-called spectral gap property), and the strong law of large numbers (SLLN) for the Markov chain \((Y_n, \xi_n)_{n \in \mathbb{N}}\). To obtain the spectral gap property, we apply the results of R. Kapica and M. Ślęczka [21], which in turn are based on the coupling methods introduced by M. Hairer [13] (also applied in [16,18,32,35]). The SLLN is established with the help of a theorem of Shirikyan [31]. Similar approach is followed e.g. in [19]. Having verified these properties for the chain \((Y_n, \xi_n)_{n \in \mathbb{N}}\), we further prove that they imply the existence and the uniqueness of an invariant measure, as well as the the SLLN, for the corresponding continuous-time process \((\bar{Y}(t), \bar{\xi}(t))_{t \geq 0}\). Our proofs require the use of several results from the theory of semigroups of linear operators in Banach spaces (see e.g. [?11]), as well as a martingale method (cf. [3,18]). Indicating direct relations between the properties of the PDMP and the associated chain is certainly a new contribution into the research. On the other hand, it is still an open question (which we have already started to investigate) whether asymptotic stability and the spectral gap property of the discrete-time model can imply the same properties for the associated continuous-time model.

The paper is organised as follows. In Section 1 we introduce notation and definitions used throughout the paper. They mainly relate to the theory of Markov operators, discussed more widely e.g. in [1,22,28,30,36]. Section 2 contains the statements of the two aforementioned results: theorem of Kapica and Ślęczka [21] (with a sketch of the proof given in the Appendix) and a Shirikyan’s version of SLLN for a class of Markov chains [31]. Section 3 deals with the structure and assumptions of the model. Main results and their proofs,
shortly summarized above, are given in Sections 4 and 5 for the discrete- and the corresponding continuous-time model, respectively. Finally, Section 6 provides the example on the biological application of the results.

1 Preliminaries

Let us begin with introducing a piece of notation. Given a metric space \((E, \rho)\), endowed with a \(\sigma\)-field \(\mathcal{B}(E)\), we define

\[
B_b(E) = \text{the space of all bounded, Borel, real valued functions defined on } E, \text{ endowed with the supremum norm: } \|f\|_\infty = \sup_{x \in E} |f(x)|, \ f \in B_b(E).
\]

\[
C_b(E) = \text{the subspace of } B_b(E) \text{ consisting of continuous functions},
\]

\[
Lip_b(E) = \text{the subspace of } B_b(E) \text{ consisting of Lipschitz continuous functions},
\]

\[
B_{E}(x,r) = \{y \in E : \rho(x,y) < r\}, \ r > 0,
\]

\[
\mathcal{M}_s(E) = \text{the space of all finite, countably additive functions (signed measures) on } \mathcal{B}(E), \text{ endowed with the total variation norm: }
\]

\[
\|\mu\|_{TV} = \sup \{ |\langle f, \mu \rangle| : f \in B_b(X), \|f\|_\infty \leq 1\}, \ \mu \in \mathcal{M}_s(X),
\]

where \(\langle f, \mu \rangle\) denotes \(\int_X f \, d\mu\),

\[
\mathcal{M}(E) = \text{the subset of } \mathcal{M}_s(E) \text{ consisting of non-negative measures},
\]

\[
\mathcal{M}_1(E) = \text{the subspace of } \mathcal{M}(E) \text{ consisting of probability measures},
\]

\[
\mathcal{M}_1^1(E) = \text{the set of all } \mu \in \mathcal{M}_1(E) \text{ additionally satisfying } \int\rho(x,x^*) \mu(dx) < \infty,
\]

where \(x^*\) is an arbitrary (and fixed) point of \(E\).

The set \(\mathcal{M}_1(E)\) will be considered with the topology induced by the so-called Fortet–Mourier distance (see e.g. [22]), defined as follows

\[
d_{FM}(\mu_1, \mu_2) = \sup \{ |\langle f, \mu_1 - \mu_2 \rangle| : f \in \mathcal{R}_{FM}(E)\}, \ \mu_1, \mu_2 \in \mathcal{M}_1(E),
\]

and

\[
\mathcal{R}_{FM}(E) = \{ f \in B_b(E) : |f| \leq 1, |f(x) - f(y)| \leq \rho(x,y) \text{ for } x,y \in E\}.
\]

It is well-known that whenever \(E\) is a Polish space, i.e. a complete separable metric space, then the convergence in \(d_{FM}\) is equivalent to the weak convergence of probability measures (cf. [9]). Let us recall that a sequence \(\mu_n \in \mathcal{M}(E), \ n \in \mathbb{N}\), is weakly convergent to \(\mu \in \mathcal{M}(E)\) (which is denoted by \(\mu_n \overset{w}{\rightharpoonup} \mu\)) whenever \(\langle f, \mu_n \rangle \to \langle f, \mu \rangle\) for all \(f \in C_b(E)\).

A function \(V : E \to [0, \infty)\) is called a Lyapunov function whenever it is continuous, bounded on bounded sets, and (if \(E\) is unbounded) \(V(x) \to \infty\) as \(\rho(x, x_0) \to \infty\) for some \(x_0 \in E\).
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 measurable map on $E$, and for each $x \in E$, $A \mapsto P(x,A)$ is a (sub)probability Borel measure on $\mathcal{B}(E)$. For an arbitrary (sub)stochastic kernel $P$ we consider two operators:

$$\mu P(A) = \int_E P(x,A) \mu(dx) \text{ for } \mu \in \mathcal{M}(E), \ A \in \mathcal{B}(E), \quad (1.1)$$

and

$$P f(x) = \int_E f(y) P(x,dy) \text{ for } x \in E, \ f \in B_b(E). \quad (1.2)$$

If the kernel $P$ is stochastic, then $\langle \cdot \rangle P : \mathcal{M}(E) \to \mathcal{M}(E)$ given by $(1.1)$ is called a regular Markov operator, and $P(\cdot) : B_b(E) \to B_b(E)$ defined by $(1.2)$ is said to be its dual operator (see [22]). It is easy to check that

$$\langle f, \mu P \rangle = \langle Pf, \mu \rangle \text{ for } f \in B(E), \ \mu \in \mathcal{M}_1(E). \quad (1.3)$$

Let us note that $P(\cdot)$, given by $(1.2)$, can be extended in the usual way to a linear operator $\mathcal{P}$ on the space of all bounded below Borel functions $\mathcal{B}_b(X)$ so that $(1.3)$ holds for all $f \in \mathcal{B}_b(X)$. For notational simplicity, in our further analysis we shall use the same symbol for the extension as for the original operator on $B_b(X)$.

A regular Markov operator $P$ is said to be Feller if $P f \in C_b(E)$ for every $f \in C_b(E)$. We say that $P$ is asymptotically stable whenever it has an invariant probability measure $\mu^* \in \mathcal{M}_1(E)$ and

$$d_{FM}(\mu P^n, \mu^*) \to 0 \text{ for any } \mu \in \mathcal{M}_1(E).$$

Suppose we are given a time-homogeneous Markov chain $(\Phi_n)_{n \in \mathbb{N}_0}$ with state space $E$, defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. The transition law of this chain is defined by

$$P(x,A) = \mathbb{P}(\Phi_{n+1} \in A | \Phi_n = x) \text{ for } x \in E, \ A \in \mathcal{B}(E), \ n \in \mathbb{N}_0. \quad (1.4)$$

Let $(\cdot)P$ denote the Markov operator corresponding to the kernel $(1.4)$. Assuming that $\mu_n$ stands for the distribution of $\Phi_n$, we see that

$$\mu_{n+1} = \mu_n P \text{ for } n \in \mathbb{N}_0.$$

A regular Markov semigroup $(P^t)_{t \geq 0}$ is a family of regular Markov operators $P^t : \mathcal{M}(E) \to \mathcal{M}(E)$, $t \geq 0$, which form a semigroup (under composition) with the identity transformation $P^0$ as the unity element. The semigroup $(P^t)_{t \geq 0}$ is called Feller whenever each $P^t$, $t \geq 0$, is Feller. A measure $\mu^* \in \mathcal{M}(E)$ is said to be invariant for the Markov semigroup $(P^t)_{t \geq 0}$ if $\mu^* P^t = \mu^*$ for all $t \geq 0$.

Let $(\Psi_t)_{t \geq 0}$ be an $E$-valued time-homogeneous Markov process with continuous time parameter $t \in \mathbb{R}_+$. The transition law of $(\Psi_t)_{t \geq 0}$ is defined by the collection of stochastic kernels of the form

$$P^t(x,A) = \mathbb{P}(\psi_{s+t} \in A | \psi_s = x), \text{ for } x \in E, \ A \in \mathcal{B}(E), \ s,t \geq 0. \quad (1.5)$$
Due to the Chapman-Kolmogorov equation, the family \((P^t)_{t \geq 0}\) of Markov operators corresponding to the kernels given by (1.5) is then a regular Markov semigroup.

In our further considerations we will use the symbol \(\mathbb{P}_x\) for the probability measure \(\mathbb{P}(\cdot | \Phi_0 = x)\) (or, depending on the context, for \(\mathbb{P}(\cdot | \Psi_0 = x)\)) and \(\mathbb{E}_x\) for the expectation with respect to \(\mathbb{P}_x\).

For a given stochastic kernel \(P : E \times \mathcal{B}(E) \to [0, 1]\), a time-homogeneous Markov chain evolving on a space \(E^2\) (endowed with the product topology) is said to be a Markovian coupling of \(P\) (see e.g. [25]) whenever its transition law \(B : E^2 \times \mathcal{B}(E^2) \to [0, 1]\) satisfies

\[
B(x, y, A \times E) = P(x, A) \quad \text{and} \quad B(x, y, E \times A) = P(y, A), \quad x, y \in E, \ A \in \mathcal{B}(E).
\]

Note that if \(Q : E^2 \times \mathcal{B}(E^2) \to [0, 1]\) is a substochastic kernel satisfying

\[
Q(x, y, A \times E) \leq P(x, A) \quad \text{and} \quad Q(x, y, E \times A) \leq P(y, A), \quad x, y \in E, \ A \in \mathcal{B}(E), \quad (1.6)
\]

then we can always construct a Markovian coupling of \(P\) whose transition function \(B\) satisfies \(Q \leq B\). Indeed, it suffices to define the family \(\{R(x, y, \cdot) : x, y \in E\}\) of measures on \(\mathcal{B}(E^2)\), which on rectangles \(A \times B \in \mathcal{B}(E^2)\) are given by

\[
R(x, y, A \times E) = \frac{1}{1 - Q(x, y, E^2)} \left[ P(x, A) - Q(x, y, A \times E) \right] \left[ P(y, B) - Q(x, y, E \times B) \right]
\]

when \(Q(x, y, E^2) < 1\), and \(R(x, y, A \times B) = 0\) otherwise. It is then easy to see that \(B := Q + R\) is a stochastic kernel satisfying \(Q \leq B\), and that the Markov chain with transition function \(B\) is a Markovian coupling of \(P\).

2 Some auxiliary results

In this section we formulate (without proofs) two known theorems concerning general Markov chains.

Firstly, let us quote [21, Theorem 2.1] of Kapica and Ślęczka. It provides sufficient conditions for a Markov–Feller operator to be asymptotically stable with an exponential rate of convergence. To explain a bit more the essential idea underlying this theorem (which is based on the coupling techniques [13]), we provide a brief sketch of its proof in the Appendix. Note that the results outlined in [32,35] were established in the same spirit (although in some particular cases only).

**Theorem 2.1.** Suppose we are given a regular Markov operator \(P : \mathcal{M}(E) \to \mathcal{M}(E)\) with the Feller property, and that there exists a substochastic kernel \(Q : E^2 \times \mathcal{B}(E^2) \to [0, 1]\) satisfying (1.6). Furthermore, assume that the following conditions hold:

\[(B1)\] There exist a Lyapunov function \(V : E \to [0, \infty)\) and constants \(a \in (0, 1)\) and \(b > 0\) satisfying

\[
PV(x) \leq aV(x) + b \quad \text{for} \quad x \in E.
\]

\[(B2)\] For some \(F \in \mathcal{B}(E^2)\) and some \(R > 0\) the following two properties are satisfied:
• $\text{supp } Q(x, y, \cdot) \subset F$ for $(x, y) \in F$;
• There exists a Markovian coupling $(\Phi^1_n, \Phi^2_n)_{n \in \mathbb{N}_0}$ of $P$ with transition function $B$, satisfying $Q \leq B$, such that for

$$K := \{(x, y) \in F : V(x) + V(y) < R\} \quad (2.1)$$

and $\kappa := \inf\{n \in \mathbb{N} : (\Phi^1_n, \Phi^2_n) \in K\}$ we can choose constants $\zeta \in (0, 1)$ and $\bar{C} > 0$, for which

$$E_{(x, y)}(\zeta^{-\kappa}) \leq \bar{C} \quad \text{whenever } V(x) + V(y) < \frac{4b}{1-a}. \quad (2.2)$$

(B3) There exists a constant $q \in (0, 1)$ such that

$$\int_{E^2} \rho(u, v) Q(x, y, du, dv) \leq q \rho(x, y) \quad \text{for } (x, y) \in F.$$

(B4) Letting $U(r) = \{(x, y) : \rho(x, y) \leq r\}$ for $r > 0$, we have

$$\inf_{(x, y) \in F} Q(x, y, U(q \rho(x, y))) > 0.$$

(B5) There exist constants $l > 0$ and $\nu \in (0, 1]$ such that

$$Q(x, y, E^2) \geq 1 - l \rho(x, y)^\nu \quad \text{for } (x, y) \in F.$$

Then, the operator $P$ possesses a unique invariant distribution $\mu^*$. Moreover, there exist constants $C \in \mathbb{R}$ and $\beta \in [0, 1)$ such that

$$d_{FM}(\mu^*, \mu^*) \leq C \beta^n((V, \mu + \mu^*) + 1) \quad (2.3)$$

for all $n \in \mathbb{N}$ and every $\mu \in M_1(E)$ satisfying $\langle V, \mu \rangle < \infty$. In particular, $P$ is then asymptotically stable.

Secondly, in Section 4.2 we need a modified version of SLLN from Shirikyan [31]. This result is originally stated for Markov chains evolving on a Hilbert space. However, a simple analysis of its proof shows that it can be easily reformulated to the following version, which remains valid in the case of Polish spaces.

**Theorem 2.2.** Let $(\Phi_n)_{n \in \mathbb{N}_0}$ be a time-homogeneous Markov chain, taking values in $E$, and let $P$ be the transition function of this chain. Suppose that the following conditions hold:

(C1) $P$ has a unique invariant measure $\mu^* \in M_1(E)$.

(C2) There exist a continuous function $\varphi : E \to \mathbb{R}_+$ and a sequence $(\gamma_k)_{k \in \mathbb{N}_0}$ of positive numbers satisfying $\sum_{k=0}^{\infty} \gamma_k < \infty$, such that for every $f \in \text{Lip}_b(E)$ we have

$$|P^n f(x) - \langle f, \mu^* \rangle| \leq \gamma_n \varphi(x) \|f\|_{\text{Lip}} \quad \text{for } x \in E, n \in \mathbb{N},$$
where
\[ \|f\|_{\text{Lip}} = \|f\|_{\infty} + \sup \left\{ \frac{|f(x) - f(y)|}{\rho(x,y)} : x, y \in E, \ x \neq y \right\}. \] (2.4)

(C3) there exists a continuous function \( h : E \to \mathbb{R}_+ \) such that
\[ \mathbb{E}_x \varphi(\Phi_n) \leq h(x) \text{ for } n \in \mathbb{N}_0, x \in E. \]

Then for every \( f \in \text{Lip}_b(E) \) and each initial state \( x \in E \)
\[ \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(\Phi_k) = \langle f, \mu^* \rangle \mathbb{P}_x - \text{almost surely (a.s.).} \]

3 Structure and assumptions of the model

Let \((H, \|\cdot\|)\) be a separable Banach space, and let \( Y \) be a closed subset of \( H \). Further, assume that \((\Theta, B(\Theta), \Delta)\) is a topological measure space with a finite Borel measure \( \Delta \), and define \( I = \{1, \ldots, N\} \) endowed with the discrete metric:
\[ \delta_{ij} = \begin{cases} 1 & \text{for } i \neq j, \\ 0 & \text{for } i = j. \end{cases} \] (3.1)

Consider a collection \( \{S_i : i \in I\} \) of flows \( S_i : \mathbb{R}_+ \times Y \to Y \), where \( \mathbb{R}_+ := [0, \infty) \), which are continuous with respect to each variable. The flow property means, as usual, that
\[ S_i(0, y) = y \text{ and } S_i(s + t, y) = S_i(s, S_i(t, y)) \text{ for } y \in Y, s, t \geq 0. \]
The flows \( S_i \) will being switched according to a matrix of continuous functions (probabilities) \( \pi_{ij} : Y \to [0, 1], i, j \in I \), satisfying
\[ \sum_{j \in I} \pi_{ij}(y) = 1 \text{ for } y \in Y, i \in I. \]

Further, assume that we are given a family \( \{w_\theta : \theta \in \Theta\} \) of transformations from \( Y \) to itself, which will be related to the post-jump locations of our dynamical system. We will require that the map \((y, \theta) \mapsto w_\theta(y)\) is continuous, and that there exists \( \varepsilon^* \in (0, \infty) \) such that
\[ w_\theta(y) + h \in Y \text{ whenever } h \in B_H(0, \varepsilon^*), \ \theta \in \Theta, \ y \in Y. \]

Let \( p : Y \times \Theta \to [0, \infty) \) be a continuous map such that \( \int_{\Theta} p(y, \theta) \, d\kappa(\theta) = 1 \) for any \( y \in Y \). For simplicity, in what follows, we will write \( d\theta \) for \( \Delta(d\theta) \). The place-dependent probability density function \( \theta \mapsto p(y, \theta) \) will capture the likelihood of occurrence of \( w_\theta \) at any jump time.

Now fix \( \varepsilon \in (0, \varepsilon^*] \), and assume that \( \nu^\varepsilon \in \mathcal{M}_1(H) \) is an arbitrary measure supported on \( B_H(0, \varepsilon) \). On a suitable probability space, say \((\Omega, \mathcal{F}, \mathbb{P})\), we define a sequence of random
variables $(Y_n)_{n \in \mathbb{N}_0}$, taking values in $Y$, in such a way that

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

(3.2)

where

- $Y_0 : \Omega \to Y$ and $\xi_0 : \Omega \to I$ are random variables with arbitrary distributions;

- $\tau_n : \Omega \to [0, \infty)$, $n \in \mathbb{N}_0$, form a strictly increasing sequence of random variables with $\tau_0 = 0$ and $\tau_n \to \infty$, whose increments $\Delta \tau_{n+1} = \tau_{n+1} - \tau_n$ are mutually independent and have common exponential distribution with parameter $\lambda > 0$;

- $H_n : \Omega \to Y$, $n \in \mathbb{N}$, are identically distributed random variables with distribution $\nu^\varepsilon$;

- $\theta_n : \Omega \to \Theta$ and $\xi_n : \Omega \to I$, $n \in \mathbb{N}$, are random variables defined (inductively) by

$$\mathbb{P}(\theta_{n+1} \in D \mid S_{\xi_n}(\Delta \tau_{n+1}, Y_n) = y) = \int_D p(y, \theta) d\theta \quad \text{for} \quad D \in \mathcal{B}(\Theta), \ y \in Y, \ n \in \mathbb{N}_0,$$

(3.3)

$$\mathbb{P}(\xi_{n+1} = j \mid Y_{n+1} = y, \xi_n = i) = \pi_{ij}(y) \quad \text{for} \quad i, j \in I, \ y \in Y, \ n \in \mathbb{N}_0,$$

so that $\theta_{n+1}$ is conditionally independent of $\mathcal{G}_n$ given $S_{\xi_n}(\Delta \tau_{n+1}, Y_n) = y$, and $\xi_{n+1}$ is conditionally independent of $\mathcal{G}_n$ given $(Y_{n+1}, \xi_n) = (y,i)$, where

$$\mathcal{G}_0 = \sigma(Y_0, \xi_0) \quad \text{and} \quad \mathcal{G}_n := \sigma(\mathcal{G}_0 \cup \{H_i, \tau_i, \theta_i, \xi_i : 1 \leq i \leq n\}) \quad \text{for} \quad n \in \mathbb{N}.$$ Simultaneously, we require that, for any $n \in \mathbb{N}_0$, $\Delta \tau_{n+1}$, $H_{n+1}$, $\theta_{n+1}$ and $\xi_{n+1}$ are (mutually) conditionally independent given $\mathcal{G}_n$, and that $\Delta \tau_{n+1}$ and $H_{n+1}$ are independent of $\mathcal{G}_n$.

In our further considerations we shall extensively use the following assumptions:

(A1) There exists $y^* \in Y$ such that

$$\sup_{y \in Y} \int_0^\infty e^{-\lambda t} \int_{\Theta} \|w_\theta(S_i(t, y^*)) - y^*\| p(S_i(t,y), \theta) d\theta dt < \infty \quad \text{for} \quad i \in I;$$

(A2) There exist $\alpha \in (-\infty, \lambda)$, $L > 0$ and a non-decreasing continuous function $\mathcal{L} : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$\|S_i(t, y_1) - S_j(t, y_2)\| \leq Le^{\alpha t} \|y_1 - y_2\| + t \mathcal{L}(\|y_2\|) \delta_{ij} \quad \text{for} \quad t \geq 0, \ y_1, y_2 \in Y, \ i, j \in I,$$

where $\delta_{ij}$ is given by (3.1); 

(A3) There exists $L_w > 0$ such that

$$\int_{\Theta} \|w_\theta(y_1) - w_\theta(y_2)\| p(y_1, \theta) d\theta \leq L_w \|y_1 - y_2\| \quad \text{for} \quad y_1, y_2 \in Y;$$

(A4) There exist $L_\pi > 0$ and $L_\rho > 0$ such that

$$\sum_{j \in I} |\pi_{ij}(y_1) - \pi_{ij}(y_2)| \leq L_\pi \|y_1 - y_2\| \quad \text{for} \quad y_1, y_2 \in Y, \ i \in I,$$
and
\[ \int_{\Theta} |p(y_1, \theta) - p(y_2, \theta)| \, d\theta \leq L_p \|y_1 - y_2\| \quad \text{for } y_1, y_2 \in Y; \]

(A5) There exist \( \delta_\pi > 0 \) and \( \delta_p > 0 \) such that
\[ \sum_{j \in I} \min\{\pi_{i_1,j}(y_1), \pi_{i_2,j}(y_2)\} \geq \delta_\pi \quad \text{for } t \geq 0, \ i_1, i_2 \in I, \ y_1, y_2 \in Y, \]
and
\[ \int_{\Theta(y_1,y_2)} \min\{p(y_1, \theta), p(y_2, \theta)\} \, d\theta \geq \delta_p \quad \text{for } y_1, y_2 \in Y, \]
where
\[ \Theta(y_1,y_2) := \{\theta \in \Theta : \|w_\theta(y_1) - w_\theta(y_2)\| \leq L_w \|y_1 - y_2\|\}. \tag{3.4} \]

Remark 3.1. Suppose that
\[ \max_{i \in I} \|S_i(t,y) - y\| \leq t\tilde{L}([y]) \quad \text{and} \quad \|S_i(t,y_1) - S_i(t,y_2)\| \leq L\alpha t \|y_1 - y_2\|, \tag{3.5} \]
where \( \tilde{L} : \mathbb{R}_+ \to \mathbb{R}_+ \) is a non-decreasing continuous function, \( \alpha < \lambda \) and \( L > 0 \). Then condition (A2) holds.

Remark 3.2. Suppose that (A3) holds, and that \( p \) does not depend on \( y \), i.e. there exists a continuous probability density function \( p : \Theta \to [0, \infty) \) such that \( p(y, \theta) = p(\theta) \) for \( \theta \in \Theta \). Furthermore, assume that
\[ \max_{i \in I} \int_0^\infty e^{-\lambda t} \|S_i(t,y^*) - y^*\| \, dt < \infty \quad \text{and} \quad \int_{\Theta} \|w_\theta(y^*) - y^*\| \, d\theta < \infty. \tag{3.6} \]
Then condition (A1) is satisfied. To see this, it suffices to observe that, for \( t \geq 0 \) and \( i \in I, \)
\[ \int_{\Theta} \|w_\theta(S_i(t,y^*)) - y^*\| \, d\theta \leq L_w \|S_i(t,y^*) - y^*\| + \int_{\Theta} \|w_\theta(y^*) - y^*\| \, d\theta. \]

It is easily seen that \( (Y_n)_{n \in \mathbb{N}} \) may not be a Markov chain. In order to use the theory of Markov operators we construct a time-homogeneous Markov chain \( (Y_n, \xi_n)_{n \in \mathbb{N}_0} \) evolving on the space \( X := Y \times I \). We assume that \( X \) is equipped with the following metric:
\[ \rho_c((y_1,i), (y_2,j)) = \|y_1 - y_2\| + c \, \delta_{ij}, \quad (y_1,i), (y_2,j) \in X, \tag{3.7} \]
where \( \delta_{ij} \) is defined by (3.1), and \( c \) is a positive constant, which will be specified at the end of this section.

Let \( P : X \times \mathcal{B}(X) \to [0,1] \) be the transition law of the chain \( (Y_n, \xi_n)_{n \in \mathbb{N}_0} \), determined by (3.2) and (3.3). An easy computation shows, that with the notation
\[ \Delta_h f(y,i) := \int_{\Theta} \left[ \sum_{j \in I} f(w_\theta(S_i(t,y)) + h, j) \, \pi_{ij}(w_\theta(S_i(t,y)) + h) \right] p(S_i(t,y), \theta) \, d\theta \]
and \( \Delta_h^A((y,i),A) := \Delta_h 1_A(y,i) \) for \( h \in B_H(0,\varepsilon), \ t \geq 0, \ (y,i) \in X, \ f \in B_b(X), \ A \in \mathcal{B}(X), \)
we have
\[ P_\varepsilon((y,i),A) = \int_{B_H(0,\varepsilon)} \int_0^\infty \lambda e^{-\lambda t} \Delta^t_N((y,i),A) dt \nu^\varepsilon(dh) \quad \text{for} \quad (y,i) \in X, A \in \mathcal{B}(X). \tag{3.8} \]

Now define \((\bar{Y}(t),\bar{\xi}(t))_{t \geq 0}\) via interpolation by setting
\[ \bar{Y}(t) = S_{\xi_n}(t - \tau_n, Y_n), \quad \bar{\xi}(t) = \xi_n \quad \text{for} \quad t \in [\tau_n, \tau_{n+1}), \quad n \in \mathbb{N}_0. \tag{3.9} \]
It is easy to check that \((\bar{Y}(t),\bar{\xi}(t))_{t \geq 0}\) is a time-homogeneous Markov process and
\((\bar{Y}(\tau_n),\bar{\xi}(\tau_n)) = (X_n, \xi_n)\) for \(n \in \mathbb{N}_0\). By \((P^t_\varepsilon)_{t \geq 0}\) we shall denote the Markov semigroup associated with the process \((\bar{Y}(t),\bar{\xi}(t))_{t \geq 0}\). The dual operator of \((P_\varepsilon f(y,i) = \mathbb{E}(f(\bar{Y}(t),\bar{\xi}(t)) | Y_0 = y, \xi_0 = i) \quad \text{for} \quad f \in B_b(X), (y,i) \in X. \tag{3.10} \)

Let us now go back to the definition of \(\rho_c\), that is, the metric in \(X\) given by (3.7). All the results of this paper work under the assumption that \(c\) is sufficiently large. The choice of this constant depends of the parameters appearing in conditions (A1)-(A3). Namely, we require that
\[ c \geq \frac{\mathcal{L}(M)(\lambda - \alpha)}{\lambda L} \max \left\{ \frac{1}{\lambda} e^{\sup T}, \frac{\lambda}{\lambda - \alpha} \right\} + \frac{2(\lambda - \alpha)}{L} \tag{3.11} \]
where \(T \subset [0,\infty)\) is a fixed bounded set with positive measure such that
\[ e^{\alpha t} \leq \frac{\lambda}{\lambda - \alpha} \quad \text{for all} \quad t \in T, \tag{3.12} \]
and \(M := 4b/(1 - a) + \|y^*\|\) with
\[ a := (\lambda LL_w)(\lambda - \alpha)^{-1}, \]
\[ b := \lambda \max_{i \in I} \sup_{y \in Y} \int_0^\infty e^{-\lambda t} \int_\Theta \|w_\theta(S_i(t,y^*)) - y^*\| p(S_i(t,y),\theta) d\theta dt + \varepsilon. \tag{3.13} \]

4 The Markov chain with post-jump Locations

In this part of the paper we provide a proof of the existence of a unique invariant probability measure for the chain \((Y_n,\xi_n)_{n \in \mathbb{N}}\) determined by (3.2), (3.3). Moreover, we show that the Markov operator \(P_\varepsilon\) induced by (3.8), is asymptotically stable, and that the rate of convergence of \((\mu P_\varepsilon^n)_{n \in \mathbb{N}}\) is exponential for every \(\mu \in \mathcal{M}_1(X)\). This will allow us to state the SLLN for the discrete-time model.
4.1 Asymptotic stability for the dicrete-time model

For any \( x_1 = (y_1, i_1), x_2 = (y_2, i_2) \in X, t \geq 0, \theta \in \Theta \) and \( h \in B_H(0, \varepsilon) \) we set

\[
\mathbf{p}(x_1, x_2, t, \theta) := p(S_{i_1}(t, y_1), \theta) \land p(S_{i_2}(t, y_2), \theta),
\]

\[
\pi_j(x_1, x_2, t, \theta, h) := \pi_{i_1,j}(w_\theta(S_{i_1}(t, y_1)) + h) \land \pi_{i_2,j}(w_\theta(S_{i_2}(t, y_2)) + h),
\]

\[
w_j(x_1, x_2, t, \theta, h) := ((w_\theta(S_{i_1}(t, y_1)) + h, j), (w_\theta(S_{i_2}(t, y_2)) + h, j)),
\]

where \( \land \) denotes minimum, and

\[
\Lambda_h^t f(x_1, x_2) := \int_{\Theta} \left[ \sum_{j \in I} f(w_j(x_1, x_2, t, \theta, h)) \pi_j(x_1, x_2, t, \theta, h) \right] \mathbf{p}(x_1, x_2, t, \theta) d\theta,
\]

\[
\Lambda_h^t((x_1, x_2), A) := \Lambda_h^t \mathbb{1}_A(x_1, x_2)
\]

for \( f \in B_b(X) \) and \( A \in \mathcal{B}(X) \).

We intend to apply Theorem 2.1 with \( Q = Q_\varepsilon \) defined by

\[
Q_\varepsilon(x_1, x_2, A) := \int_{B_H(0, \varepsilon)} \int_0^\infty \lambda e^{-\lambda t} \Lambda_h^t((x_1, x_2), A) dt \nu^\varepsilon(dh)
\]

for \( (x_1, x_2) \in X^2 \) and \( A \in \mathcal{B}(X^2) \). It is easily seen that \( Q_\varepsilon : X^2 \times \mathcal{B}(X^2) \rightarrow [0, 1] \) is then a substochastic kernel, satisfying (1.6) for \( P = P_\varepsilon \). Note that

\[
Q_\varepsilon f(x_1, x_2) = \int_{X^2} f(u_1, u_2) Q_\varepsilon(x_1, x_2, du_1, du_2) = \int_{B_H(0, \varepsilon)} \int_0^\infty \lambda e^{-\lambda t} \Lambda_h^t f(x_1, x_2) dt \nu^\varepsilon(dh).
\]

for any \( x_1, x_2 \in X \) and \( f \in B_b(X) \).

We assume that \( X^2 \) is equipped with the following metric

\[
\bar{d}_c((x_1, x_2), (z_1, z_2)) = \rho_c(x_1, z_1) + \rho_c(x_2, z_2), \quad (x_1, x_2), (z_1, z_2) \in X^2.
\]

**Theorem 4.1.** Suppose that conditions [(A1)-(A5)] hold, and that

\[
LL_w + \frac{\alpha}{\lambda} < 1.
\]

Then the Markov operator \( P_\varepsilon \) generated by [3.8] has a unique invariant distribution \( \mu^* \). Moreover, there exist \( x^* \in X \) and constants \( C \in \mathbb{R} \) and \( \beta \in [0, 1) \) such that

\[
d_{FM}(\mu P_\varepsilon^n \mu^* \mu^*) \leq C\beta^n \left( \int_X \rho_c(x^*, x) (\mu + \mu^*)(dx) + 1 \right),
\]

for any \( n \in \mathbb{N} \) and any \( \mu \in \mathcal{M}_1^1(X) \). In particular, \( P_\varepsilon \) is then asymptotically stable.

**Proof.** To establish the proof, it suffices to verify the hypotheses of Theorem 2.1 for \( P = P_\varepsilon \) and \( Q = Q_\varepsilon \) given by [4.12]. First of all, let us note that \( P_\varepsilon \) is Feller, which follows immediately from the continuity of functions \( \pi_{i,j} : y \mapsto S_i(t, y), y \mapsto p(y, \theta) \) and \( w_\theta \). Our further reasoning falls naturally into five parts.
Step 1. Our goal in this step is to show that condition \( (B1) \) is satisfied for \( V : X \to [0, \infty) \) defined by

\[
V(y, j) := \|y - y^*\| \quad \text{for} \quad (y, j) \in X,
\]

and for constants \( a \) and \( b \) given by \((3.13)\). Clearly \( V \) is a Lyapunov function, and taking \( x^* = (y^*, i^*) \) (for an arbitrary \( i^* \in I \)) gives \( V(x) \leq \rho_c(x^*, x) \) for \( x \in X \). Further, we note that \( a \in (0, 1) \), due to \((4.4)\). For brevity, let us define

\[
B_j(t, y) = \int_\Omega \|w_\theta(S_j(t, y^*)) - y^*\| p(S_j(t, y), \theta) d\theta, \quad j \in I, \ t \geq 0, \ z \in Y.
\]

From \((A1)\) we know that \( b < \infty \), which, in particular, implies that \( B_j(t, y) < \infty \) for almost all \( t \geq 0 \) and each \( y \in Y \).

Let \( (y, i) \in X \). By conditions \((A3)\) and \((A2)\) we see that for \( t \geq 0 \) and \( h \in B_H(0, \varepsilon) \),

\[
\Delta_h^t V(y, i) = \int_\Theta \|w_\theta(S_i(t, y)) + h - y^*\| \left[ \sum_{j \in I} \pi_{ij}(w_\theta(S_i(t, y)) + h) \right] p(S_i(t, y), \theta) d\theta
\]

\[
\leq \int_\Theta \|w_\theta(S_i(t, y)) - w_\theta(S_i(t, y^*))\| p(S_i(t, y), \theta) d\theta
\]

\[
+ \int_\Theta \|w_\theta(S_i(t, y^*)) - y^*\| p(S_i(t, y), \theta) d\theta + \|h\|
\]

\[
\leq L_w \|S_i(t, y) - S_i(t, y^*)\| + B_i(t, y) + \varepsilon \leq LL_w e^{\alpha t} \|y - y^*\| + B_i(t, y) + \varepsilon.
\]

Hence,

\[
P_\varepsilon V(y, i) = \int_{B_H(0, \varepsilon)} \int_0^\infty \lambda e^{-\lambda t} \Delta_h^t V(y, i) dt \nu^\varepsilon(dh)
\]

\[
\leq \lambda LL_w \int_0^\infty e^{(\alpha - \lambda) t} dt \|y - y^*\| + \lambda \int_0^\infty e^{-\lambda t} B_i(t, y) dt + \varepsilon
\]

\[
\leq \frac{\lambda LL_w}{\lambda - \alpha} \|y - y^*\| + b = a V(y, i) + b.
\]

Step 2. Let us define

\[
R := \frac{4b}{1 - a} \quad \text{and} \quad F := F_1 \cup F_2,
\]

where \( F_1, F_2 \subset X^2 \) are given by

\[
F_1 := \{((y_1, i_1), (y_2, i_2)) : i_1 = i_2\}, \quad F_2 := \{((y_1, i_1), (y_2, i_2)) : V(y_1, i_1) + V(y_2, i_2) < R\}.
\]

We will show that condition \((B2)\) is satisfied for them.

First of all, we observe that \( \supp Q_\varepsilon(x_1, x_2, \cdot) \subseteq F \) for every \((x_1, x_2) \in X^2\). To see this, let \((x_1, x_2) := ((y_1, i_1), (y_2, i_2)) \in X^2 \) and \((z_1, z_2) := ((u_1, k_1), (u_2, k_2)) \in X^2 \setminus F \). Then, in particular, \((z_1, z_2) \notin F_1\), that is \( k_1 \neq k_2 \). We obtain

\[
\overline{\rho}_c(w_j(x_1, x_2, t, \theta, h), (z_1, z_2)) \geq c(\delta_{jk_1} + \delta_{jk_2}) \geq c \quad \text{for} \quad j \in I, \ t \geq 0, \ h \in B_H(0, \varepsilon),
\]

where \( c > 0 \).
where \( w_j \) and \( \rho_c \) were introduced in \([4.1], [4.3]\), respectively. Hence, taking \( \eta \in (0, c) \), we see that \( w_j(x_1, x_2, t, \theta, h) \notin B_{X^2}(z_1, z_2, \eta) \) for all \( j, t, h \), which yields \( Q_\varepsilon(x_1, x_2, B_{X^2}(z_1, z_2, \eta)) = 0 \) and thus \( (z_1, z_2) \in X^2(\text{supp} Q_\varepsilon(x_1, x_2, \cdot)) \).

Let \((X_n^1, X_n^2)_{n \geq 0}\) be an arbitrary Markovian coupling of \( P_\varepsilon \) with transition function \( B \) such that \( Q_\varepsilon \leq B \) (it is justified at the end of Section 1 that such a coupling exists). Further, put \( \kappa := \inf\{n \in \mathbb{N} : (X_n^1, X_n^2) \in K\} \), where \( K \subset X^2 \) is given by \([2.1]\), and define \( \mathcal{V} : X^2 \to [0, \infty) \) by

\[
\mathcal{V}(x_1, x_2) := V(x_1) + V(x_2) \quad \text{for} \quad (x_1, x_2) \in X^2.
\]

Since \( \{(x_1, x_2) \in X^2 : \mathcal{V}(x_1, x_2) < R\} = F_2 \subset F \), we obtain

\[
\kappa = \inf \{n \in \mathbb{N}_0 : \mathcal{V}(X_n^1, X_n^2) < R\}.
\]

Furthermore, \( \mathcal{V} \) is a Lyapunov function satisfying

\[
B \mathcal{V}(x_1, x_2) \leq a \mathcal{V}(x_1, x_2) + 2b \quad \text{for} \quad (x_1, x_2) \in X^2,
\]

which follows from \([4.7]\). From \([21\text{ Lemma 2.2}]\) it now follows that \([2.2]\) holds.

**Step 3.** We shall prove that condition \([B3]\) holds with \( q := a \in (0, 1) \). Let \((x_1, x_2) := ((y_1, i_1), (y_2, i_2)) \in F \). By conditions \([A3]\) and \([A2]\) we obtain that, for \( t \geq 0 \) and \( h \in B_H(0, \varepsilon) \),

\[
\Lambda_h^i \rho_c(x_1, x_2) = \int_{\Theta} \| w_y(S_i(t, y_1)) - w_y(S_i(t, y_2)) \| \left[ \sum_j \pi_j(x_1, x_2, t, \theta, h) \right] p(x_1, x_2, t, \theta) \, d\theta
\]

\[
\leq \int_{\Theta} \| w_y(S_i(t, y_1)) - w_y(S_i(t, y_2)) \| p(S_i(t, y_1), \theta) \, d\theta
\]

\[
\leq L_w \| S_i(t, y_1) - S_i(t, y_2) \| \leq L e^{\alpha t} \| y_1 - y_2 \| + t \mathcal{L}(\| y_2 \|) \delta_{i, i_2}.
\]

Let \( M = R + \| y^* \| \). In the case where \( i_1 \neq i_2 \), that is \((x_1, x_2) \in F_2 \), from the definition of \( V \) and \( F_2 \), it follows that \( \mathcal{L}(\| y_2 \|) \leq \mathcal{L}(V(y_2) + \| y^* \|) \leq \mathcal{L}(M) \). This, combined with \([4.8]\), gives

\[
\| S_i(t, y_1) - S_i(t, y_2) \| \leq L e^{\alpha t} \| y_1 - y_2 \| + t \mathcal{L}(M) \delta_{i, i_2}.
\]

Finally, applying \([4.8], [4.9] \) and \([3.11]\), we have

\[
Q_\varepsilon \rho_c(x_1, x_2) = \int_{B_H(0, \varepsilon)} \int_0^\infty \lambda e^{-\lambda t} \Lambda_h^i \rho_c(x_1, x_2) \, dt \, \nu^\varepsilon(dh)
\]

\[
\leq \lambda L_w \int_0^\infty e^{-\lambda t} \| S_i(t, y_1) - S_i(t, y_2) \| \, dt
\]

\[
\leq \lambda LL_w \left( \int_0^\infty e^{-\lambda t} \, dt \| y_1 - y_2 \| + \frac{\mathcal{L}(M)}{L} \int_0^\infty t e^{-\lambda t} \, dt \delta_{i, i_2} \right)
\]

\[
\leq \frac{\lambda LL_w}{\lambda - \alpha} \left( \| y_1 - y_2 \| + \frac{(\lambda - \alpha) \mathcal{L}(M)}{\lambda^2 L} \delta_{i, i_2} \right) \leq q \rho_c(x_1, x_2).
\]

**Step 4.** We now proceed to prove condition \([B4]\). For this purpose, let \( T \subset [0, \infty) \) be the bounded set with positive measure such that \([3.12]\) holds. Clearly, due to \([3.12]\) we
obtain
\[ LL_w e^{\alpha t} \leq q \quad \text{for} \quad t \in T \quad (\text{where} \quad q = \lambda LL_w (\lambda - \alpha)^{-1}). \] (4.10)

Define \( \delta = \delta \pi \delta_p \int_T e^{-\lambda t} dt \). Letting \( (x_1, x_2) := ((y_1, i_1), (y_2, i_2)) \in F \) and
\[ U := \{ (u_1, u_2) \in X^2 : \rho_c(u_1, u_2) \leq q \rho_c(x_1, x_2) \}, \]
we shall establish that \( Q_\varepsilon(x_1, x_2, U) \geq \delta \).

Recall the definition of \( \Theta(\cdot, \cdot) \) introduced in (3.4) and consider the following sets:
\[ \mathcal{R}_1(t) := \Theta(S_{i_1}(t, y_1), S_{i_2}(t, y_2)), \]
\[ \mathcal{R}_2(t) := \{ \theta \in \Theta : \| w_\theta(S_{i_1}(t, y_1)) - w_\theta(S_{i_2}(t, y_2)) \| \leq q \rho_c(x_1, x_2) \}. \]

Now, applying (4.9), (4.10) and (3.11), we obtain the following estimations:
\[ \| w_\theta(S_{i_1}(t, y_1)) - w_\theta(S_{i_2}(t, y_2)) \| \leq L_w \| S_{i_1}(t, y_1) - S_{i_2}(t, y_2) \| \]
\[ \leq L_w L e^{\alpha t} \| y_1 - y_2 \| + L_w \varepsilon \sup_T (M) \delta_{i_1, i_2} \]
\[ \leq q \| y_1 - y_2 \| + cq \delta_{i_1, i_2} = q \rho_c(x_1, x_2) \]
for \( t \in T \) and \( \theta \in \mathcal{R}_1(t) \). This obviously implies that \( \mathcal{R}_1(t) \subset \mathcal{R}_2(t) \). Furthermore, appealing to the notation introduced in (4.11), we can write \( \mathcal{R}_2(t) = \{ \theta \in \Theta : w_j(x_1, x_2, t, \theta, h) \in U \} \) for \( t \in T, \theta \in \Theta, h \in B_H(0, \varepsilon) \) and \( j \in I \), which gives
\[ \mathbb{1}_U(w_j(x_1, x_2, t, \theta, h)) = \mathbb{1}_{\mathcal{R}_2(t)}(\theta) \geq \mathbb{1}_{\mathcal{R}_1(t)}(\theta). \]

From (A5) it then follows that
\[ \Lambda^t_h(x_1, x_2, U) = \int_\Theta \left[ \sum_{j \in I} \mathbb{1}_U(w_j(x_1, x_2, t, \theta, h)) \pi_j(x_1, x_2, t, \theta, h) \right] p(x_1, x_2, t, \theta) d\theta \]
\[ \geq \delta \pi \int_{\mathcal{R}_1(t)} p(S_{i_1}(t, y_1), \theta) \land p(S_{i_2}(t, y_2), \theta) d\theta \geq \delta \pi \delta_p, \quad t \in T, \quad h \in B_H(0, \varepsilon), \]
and we finally obtain
\[ Q_\varepsilon(x_1, x_2, U) \geq \int_{B_H(0, \varepsilon)} \int_T \lambda e^{-\lambda t} \Lambda^t_h(x_1, x_2, U) dt \nu^x (dh) \geq \delta \pi \delta_p \int_T \lambda e^{-\lambda t} dt = \delta. \]

**Step 5.** To complete the proof, it remains to establish condition (B5). Let \( (x_1, x_2) := ((y_1, i_1), (y_2, i_2)) \in F \). Applying the inequality
\[ (s_1 \land s_2)(t_1 \land t_2) \geq s_1 t_1 - s_1 |t_1 - t_2| - |s_1 - s_2| t_1, \quad s_i, t_i \in \mathbb{R}, \quad i = 1, 2, \]
and setting
\[ A^t_h(x_1, x_2) = \int_\Theta \left[ \sum_{j \in I} \pi_{i_1 j}(w_\theta(S_{i_1}(t, y_1)) + h) \right] p(S_{i_1}(t, y_1), \theta) d\theta, \]

15
for all

and, following (A4) and (A3), sequentially, we also observe that

we infer (recalling the notation introduced in (4.1) that

Clearly, \( A_h^t(x_1, x_2) = 1 \). Further, due to (A4) we have

and, following (A4) and (A3) sequentially, we also observe that

Now, reconsidering (4.9) we conclude that

for all \( t \geq 0 \) and \( h \in B_H(0, \varepsilon) \). Finally, from (3.11) it follows that

\[
Q(\varepsilon, t, x_2) = \int_{\Omega} \int_{\Theta} \lambda e^{-\lambda t} A_h^t(x_1, x_2, X^2) dt \nu^\delta(dh) \\
\geq 1 - \lambda L(L_p + L_w L_{\pi}) \int_0^\infty e^{(\alpha - \lambda) t} \|y_1 - y_2\| + \frac{1}{L}(\|L(M) t + 2\|) \delta_{i_1, i_2} dt \\
= 1 - \lambda L(L_p + L_w L_{\pi} + 1) \left[ \|y_1 - y_2\| + \frac{\lambda - \alpha}{L} \left( \frac{L(M)}{\lambda^2} + 2\right) \delta_{i_1, i_2} \right] \\
\geq 1 - \frac{\lambda L(L_p + L_w L_{\pi} + 1)}{\lambda - \alpha} \rho_c(x_1, x_2). 
\]
Summarizing, we have shown that all the hypotheses of Theorem 2.1 hold, and thus the proof is now complete.

As a straightforward consequence of Theorem 4.1 we obtain a stability for the sequence $(Y_n)_{n \in \mathbb{N}}$, which describes post-jump locations of $(Y(t))_{t \geq 0}$. Let $\mu_n$ denote the distribution of $(Y_n, \xi_n)$, $n \in \mathbb{N}$, and let $\tilde{\mu}_n$ stand for the distribution of $Y_n$, $n \in \mathbb{N}_0$. Clearly, we then have

$$\tilde{\mu}_n(B) = \mu_n(B \times I) \quad \text{for} \quad B \in \mathcal{B}(Y), \ n \geq 1.$$ 

It should be noted that results similar to these presented below have been proven by Horbacz in [19].

**Corollary 4.2.** Suppose that the hypotheses of Theorem 4.1 hold, and let $\mu^* \in \mathcal{M}_1(X)$ be the unique invariant distribution for $P_\varepsilon$. Define

$$\tilde{\mu}^*(B) = \mu^*(B \times I) \quad \text{for} \quad B \in \mathcal{B}(Y).$$

Then

1. If $Y_0$ has the distribution $\tilde{\mu}_0 = \tilde{\mu}^*$ and $\mathbb{P}(\xi_0 = i | Y_0 = y) = \pi_i(y)$, $i \in I$, $y \in Y$, where $\pi_i$ is the Radon–Nikodym derivative of $\mu^*(\cdot \times \{i\})$ with respect to $\tilde{\mu}^*$, then

   $$\tilde{\mu}_n = \tilde{\mu}^* \quad \text{for} \quad n \in \mathbb{N}.$$

2. There exists $\beta \in [0, 1)$ with the property that for every distribution $\tilde{\mu}_0 \in \mathcal{M}_1(Y)$ of $Y_0$ we may find a constant $\tilde{C}(\tilde{\mu}_0) \in \mathbb{R}$ such that

   $$d_{FM}(\tilde{\mu}_n, \tilde{\mu}^*) \leq \tilde{C}(\tilde{\mu}_0) \beta^n \quad \text{for} \quad n \in \mathbb{N}.$$ 

**Proof.** In order to show (1) let $\tilde{\mu}_0 = \tilde{\mu}^*$, and choose $\pi_i : Y \to [0, \infty)$ so that

$$\mu^*(B \times \{i\}) = \int_B \pi_i(y) \tilde{\mu}^*(dy), \ B \in \mathcal{B}(Y), \ i \in I.$$ 

Since $\sum_{i \in I} \pi_i = 1$ almost everywhere with respect to $\tilde{\mu}_0$, we can take

$$\mathbb{P}(\xi_0 = i | Y_0 = y) = \pi_i(y) \quad \text{for} \quad y \in Y, \ i \in I.$$ 

Then

$$\mu_0(B \times J) = \mathbb{P}(Y_0 \in B, \ \xi_0 \in J) = \sum_{j \in J} \int_{\{Y_0 \in B\}} \mathbb{P}(\xi_0 = j | Y_0) \ d\mathbb{P} = \sum_{j \in J} \int_B \pi_j(y) \tilde{\mu}^*(dy)$$

$$= \sum_{j \in J} \mu^*(B \times \{j\}) = \mu^*(B \times J) \quad \text{for} \quad B \in \mathcal{B}(Y), \ J \subset I,$$

and thus $\mu_0 = \mu^*$. Consequently, we obtain

$$\tilde{\mu}_n(B) = \mu_n(B \times I) = \mu_0 P^n(B \times I) = \mu^* P^n(B \times I) = \mu^*(B \times I) = \tilde{\mu}^*(B), \ B \in \mathcal{B}(Y).$$
For the proof of assertion (2), it suffices to observe that \(d_{FM}(\tilde{\mu}_n, \tilde{\mu}^*) \leq d_{FM}(\mu_n, \mu^*)\) for \(n \in \mathbb{N}\). It follows from the fact that for every \(f \in R_{FM}(Y)\),

\[
|\langle f, \tilde{\mu}_n - \tilde{\mu}^* \rangle| = \left| \int_Y f(y) (\tilde{\mu}_n - \tilde{\mu}^*)(dy) \right| = \left| \int_{Y \times I} f(y) (\mu_n - \mu^*)(dy, di) \right| \leq d_{FM}(\mu_n, \mu^*),
\]

where the last inequality holds since \([(y, i) \mapsto f(y)] \in R_{FM}(Y \times I)\). Then (2) is ensured by (4.5).

4.2 SLLN for the discrete-time model

We will now show that under the hypotheses of Theorem 4.1 the sequence \((f(X_n))_{n \in \mathbb{N}}\) obeys the SLLN for any bounded Lipschitz function \(f\) and any initial distribution of \(X_0\). To do this, we shall use Theorem 2.2.

**Theorem 4.3.** Suppose that conditions (A1)-(A5) and (4.4) hold. Then for every \(f \in Lip_b(X)\) and each initial state \(x \in X\) we have

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(X_k) = \langle f, \mu^* \rangle \quad \mathbb{P}_x - a.s.,
\]

where \(\mu^*\) is the unique invariant distribution for the Markov operator \(P_\varepsilon\) (which exists by Theorem 4.1).

**Proof.** It suffices to verify conditions of Theorem 2.2. By virtue of Theorem 4.1 there exists a unique invariant distribution \(\mu^*\), whence (C1) holds, and we can choose \(x^* = (y^*, i^*) \in X\), \(C \in \mathbb{R}\) and \(\beta \in [0, 1)\) such that (4.5) is satisfied. Recall that \(V\) is given by (4.6) and define

\[
D := \int_X \rho_c(x^*, u) \mu^*(du) + c + 1 \quad \text{and} \quad \varphi(y, i) := V(y, i) + D, \quad (y, i) \in X.
\]

Then, applying (4.5) for Dirac measures \(\mu = \delta_{(y, i)}\), we obtain

\[
|P_\varepsilon^n f(y, i) - \langle f, \mu^* \rangle| \leq C \beta^n \varphi(y, i) \quad \text{for} \quad (y, i) \in X \quad \text{and} \quad f \in R_{FM}(X).
\]

Letting \(f \in Lip_b(X)\) and using the latter inequality for \(f/ \|f\|_{Lip} \in R_{FM}(X)\) (when \(f \neq 0\)), we observe that

\[
|P_\varepsilon^n f(y, i) - \langle f, \mu^* \rangle| \leq C \beta^n \varphi(y, i) \|f\|_{Lip} \quad \text{for} \quad (y, i) \in X \quad \text{and} \quad f \in Lip_b(X),
\]

which ensures (C2).

We have seen in (4.7) that there exists \(a \in (0, 1)\) and \(b > 0\) such that \(P_\varepsilon V(y, i) \leq aV(y, i) + b\) for \((y, i) \in X\). This gives

\[
P_\varepsilon^n V(y, i) \leq a^n V(y, i) + b \sum_{k=0}^{n-1} a^k \leq V(y, i) + \frac{b}{1-a}, \quad (y, i) \in X, \quad n \in \mathbb{N},
\]

18
and therefore, we can conclude that

\[ \mathbb{E}(y,i)\varphi(X_n) = P^n \varphi(y,i) = P^n V(y,i) + D \leq V(y,i) + D + \frac{b}{1-a} \quad \text{for} \quad (y,i) \in X. \]

Hence, (C3) holds with \( h(y,i) := V(y,i) + D + b(1-a)^{-1} \), \( (y,i) \in X \), and this completes the proof. \( \square \)

5 The continuous-time model

This section will be devoted to the study of the Markov process \((\overline{Y}(t), \overline{\xi}(t))_{t \geq 0}\) defined by (3.9). We give an explicit correspondence between invariant measures for this continuous-time model and the chain \((Y_n, \xi_n)_{n \in \mathbb{N}}\). This result enables us to provide the existence and uniqueness of an invariant distribution for the process \((\overline{Y}(t), \overline{\xi}(t))_{t \geq 0}\), which in turn leads us to the SLLN. Let us recall that \((P^n_t)_{t \geq 0}\) stands for the transition semigroup of \((\overline{Y}(t), \overline{\xi}(t))_{t \geq 0}\).

5.1 Existence of an invariant measure for the continuous time-model

We begin our analysis by showing that \((P^n_t)_{t \geq 0}\) has the Feller property. For brevity, let us write \( j_n, s_n, \theta_n \) and \( h_n \) for the sequences \((j_1, \ldots, j_n), (s_1, \ldots, s_n), (\theta_1, \ldots, \theta_n) \) and \((h_1, \ldots, h_n)\), respectively. By \( d\theta_n \) and \( \nu_n(d\theta_n) \) we shall denote the product measures \( \Delta \otimes \nu_n(d\theta_1, \ldots, d\theta_n) \) and \( \nu^{\otimes n}(dh_1, \ldots, dh_n) \), respectively. Moreover, for any \( n \in \mathbb{N}, y \in Y, i \in I, j_n \in I^n, s_{n+1} \in [0, \infty)^{n+1}, \theta_n \in \Theta^n \) and \( h_n \in B_H(0, \varepsilon)^n \), we define

\[
\mathcal{W}_1(y,i, s_1, \theta_1, h_1) = w_1(S_1(s_1, y)) + h_1,
\]

\[
\mathcal{W}_n(y, i, j_{n-1}, s_n, \theta_n, h_n) = w_\theta_n(S_{j_{n-1}}(s_n, \mathcal{W}_{n-1}(y, i, j_{n-2}, s_{n-1}, \theta_{n-1}, h_{n-1}))) + h_n;
\]

\[
\mathcal{H}_1(y, i, j_1, s_1, s_2, \theta_1, h_1) = S_{j_1}(s_2, \mathcal{W}_1(y, i, s_1, \theta_1, h_1)),
\]

\[
\mathcal{H}_n(y, i, j_n, s_{n+1}, \theta_n, h_n) = S_{j_n}(s_{n+1}, \mathcal{W}_n(y, i, j_{n-1}, s_n, \theta_n, h_n));
\]

\[
\Pi_1(y, i, j_1, s_1, \theta_1, h_1) = \pi_{i,j_1}(\mathcal{W}_1(y, i, s_1, \theta_1, h_1)),
\]

\[
\Pi_n(y, i, j_n, s_n, \theta_n, h_n) = \Pi_{n-1}(y, i, j_{n-1}, s_{n-1}, \theta_{n-1}, h_{n-1}) \times \pi_{j_{n-1}j_n}(\mathcal{W}_n(y, i, j_{n-1}, s_n, \theta_n, h_n));
\]

\[
\mathcal{P}_1(y, i, s_1, \theta_1) = p(S_1(s_1, y), \theta_1),
\]

\[
\mathcal{P}_n(y, i, j_{n-1}, s_n, \theta_n, h_{n-1}) = \mathcal{P}_{n-1}(y, i, j_{n-2}, s_{n-1}, \theta_{n-1}, h_{n-2}) \times p(\mathcal{H}_{n-1}(y, i, j_{n-1}, s_n, \theta_{n-1}, h_{n-1}), \theta_n).
\]

Furthermore, we consider the Poisson counting process

\[ N_t := \max\{ n \in \mathbb{N}_0 : \tau_n \leq t \}, \quad t \geq 0, \]
Lemma 5.1. The semigroup \((P^t_\varepsilon)_{t \geq 0}\) is Feller.

Proof. Let \(t \geq 0\) and \(f \in C_b(X)\). According to the definition of \(N_t\) we have
\[
\{N_t = n\} = \{\omega \in \Omega : \tau_n(\omega) \leq t < \tau_{n+1}(\omega)\}, \quad n \in \mathbb{N}_0.
\]
For \((y, i) \in X\) and \(n \in \mathbb{N}_0\) define
\[
R^t_n f(y, i) := \int_{\{N_t = n\}} f(\bar{Y}(t), \bar{\xi}(t)) \, dP_{(y, i)} = \int_{\{N_t = n\}} f(S_{\xi_n}(t - \tau_n, Y_n), \xi_n) \, dP_{(y, i)}. \tag{5.1}
\]
This together with (3.10) allows us to write
\[
P^t_\varepsilon f(y, i) = E_{(y, i)} f(\bar{Y}(t), \bar{\xi}(t)) = \sum_{n=0}^{\infty} R^t_n f(y, i) \quad \text{for} \quad (y, i) \in X. \tag{5.2}
\]
We then see that
\[
R^t_0 f(y, i) = P(\tau_1 \geq t) f(S_i(t, y), i) = e^{-\lambda t} f(S_i(t, y), i), \tag{5.3}
\]
and
\[
R^t_1 f(y, i) = \int_{\{N_t = 1\}} f(S_{\xi_1}(t - \tau_1, w_{\theta_1}(S_i(\tau_1, y)) + H_1), \xi_1) \, dP_{(y, i)}
\]
\[
= \int_{\{N_t = 1\}} \sum_{j \in \Gamma} \int_{B_{\Delta t}(0, \varepsilon)} \int_{\Theta} f(\mathcal{H}_1(y, i, j, (\tau_1, t - \tau_1), \theta, h)) \Pi_1(y, i, j, \tau_1, \theta, h)
\]
\[
\times \mathcal{P}_1(y, i, \tau_1, \theta) \, d\theta \, \nu^\varepsilon(dh) \, dP.
\]
In the general case, by setting
\[
\Psi_n f(y, i, s_n, t) := \sum_{j_n \in \Gamma^n} \left[ \int_{B_{\Delta t}(0, \varepsilon)^n} \int_{\Theta^n} f(\mathcal{H}_n(y, i, j_n, (s_n, t - s_n), \theta_n, h_n), j_n)
\right.
\]
\[
\times \Pi_n(y, i, j_n, s_n, \theta_n, h_n) \mathcal{P}_n(y, i, j_{n-1}, s_n, \theta_n, h_{n-1}) \, d\theta_n \, \nu^\varepsilon(dh_n) \tag{5.4}
\]
for \((y, i) \in X, s_n \in \mathbb{R}^n_+, \text{ and } \Delta \tau_n := (\Delta \tau_1, \ldots, \Delta \tau_n)\), we obtain
\[
R^t_n f(y, i) = \int_{\{N_t = n\}} \Psi_n f(y, i, \Delta \tau_n, t) \, dP \quad \text{for} \quad (y, i) \in X, \quad n \geq 1. \tag{5.5}
\]
Clearly, \(\mathcal{H}_n, \Pi_n\) and \(\mathcal{P}_n\) are continuous with respect to \((y, i)\) since \(S_j(\cdot, s), w_\theta, p(\theta, \cdot)\) and \(\pi_{ij}\) are continuous. Therefore, and by the Lebesgue dominated convergence theorem, all the functions \(R^t_n\) are continuous. Noting that \(\|R^t_n f(y, i)\| \leq \|f\|_{\infty} \mathbb{P}(N_t = 1)\), we can again apply the Lebesgue theorem (in the discrete version) to deduce that \(P^t_\varepsilon f\) is continuous. \(\square\)

Using the decomposition (5.2) of \(P^t_\varepsilon\) appearing in the above proof, we can obtain an ap-
proximation of $P^t\varepsilon f$ (cf. [17]), which will be fundamental in our later analysis.

**Lemma 5.2.** For each $f \in B_b(X)$ there exists a function $u_f : X \times (0, \infty) \to \mathbb{R}$ such that

$$
\lim_{t \to 0} \|u_f(\cdot, t)\|_\infty = 0
$$

and

$$
P^t\varepsilon f(y, i) = e^{-\lambda t} f(S_i(t, y), i) + \lambda e^{-\lambda t} \int_0^t \Psi_1 f(y, i, s, t) \, ds + u_f(y, i, t) \quad \text{for } t > 0,
$$

where $\Psi_1$ is defined by (5.4).

**Proof.** According to (5.2), we may write $P^t\varepsilon f = R^t_0 f + R^t_1 f + \sum_{n=2}^\infty R^t_n f$, where $R^t_n f$ are defined by (5.1). As we have already seen, $R^t_0 f$ can be expressed as (5.3). Since the distribution of $\tau_1$ conditional on $\{N_t = 1\}$ is uniform over $(0, t)$, it follows that

$$
R^t_1 f(y, i) = \int_{\{N_t = 1\}} \Psi_1 f(y, i, \Delta \tau_1, t) \, d\mathbb{P} = \mathbb{P}(N_t = 1) \int_{\Omega} \Psi_1 f(y, i, \Delta \tau_1, t) \, d\mathbb{P}(\cdot | N_t = 1)
= \lambda t e^{-\lambda t} \int_0^t \frac{1}{t} \Psi_1 f(y, i, s, t) \, ds = \lambda e^{-\lambda t} \int_0^t \Psi_1 f(y, i, s, t) \, ds.
$$

Define $u_f(y, i, t) := \sum_{n=2}^\infty R^t_n f(y, i)$ for $(y, i) \in X$ and $t > 0$. By (5.5) we obtain

$$
\frac{|u_f(y, i, t)|}{t} \leq \|f\|_\infty \frac{1}{t} \sum_{n=2}^\infty \mathbb{P}(N_t = n) = \|f\|_\infty \frac{1}{t} e^{-\lambda t} \sum_{n=2}^\infty \frac{(\lambda t)^n}{n!} = \|f\|_\infty \frac{1}{t} e^{-\lambda t} (e^{\lambda t} - 1 - \lambda t)
= \|f\|_\infty \left(1 - \frac{e^{-\lambda t}}{t} - \lambda e^{-\lambda t}\right)
$$

for all $t > 0$, $(y, i) \in X$, and thus $\lim_{t \to 0} (\|u_f(\cdot, t)\|_\infty / t) = 0$, as claimed.

Having established this, we immediately obtain the following:

**Lemma 5.3.** The semigroup $(P^t\varepsilon)_{t \geq 0}$ is stochastically continuous, i.e. $\lim_{t \to 0} P^t\varepsilon f(x) = f(x)$ for all $x \in X$ and $f \in C_b(X)$.

In order to prepare for the proof of our main result in this section, we need a few facts from the theory of semigroups of linear operators on Banach spaces, adapted from [7, 11].

Consider the Banach space $(\mathcal{M}_s(X), \|\cdot\|_{TV})$, and let $\mathcal{M}_s^*(X)$ denotes its dual space. For any function $f \in B_b(X)$ we define the functional $\ell_f : \mathcal{M}_s(X) \to \mathbb{R}$ by

$$
\ell_f(\mu) := \langle f, \mu \rangle, \quad \mu \in \mathcal{M}_s(X).
$$

Clearly $\ell_f \in \mathcal{M}_s^*(X)$, and $f \mapsto \ell_f$ is an isometric embedding of $B_b(X)$ in $\mathcal{M}_s^*(X)$, i.e. an injective linear map satisfying $\|\ell_f\| = \|f\|_\infty$ for all $f \in B_b(X)$. Therefore, $B_b(X)$ can be regarded as a subspace of $\mathcal{M}_s^*(X)$, and consequently, it can be endowed with the weak
star (\(w^\ast\)-) topology inherited from \(\mathcal{M}_s^\ast(X)\). Moreover, one can easily show that \(B_b(X)\) is \(w^\ast\)-closed in \(\mathcal{M}_s^\ast(X)\).

We say that a sequence \(f_n \in B_b(X), n \in \mathbb{N}\), converges \(*\)-weakly to \(f \in B_b(X)\) (and we write \(w^\ast\)-lim\(n \to \infty\) \(f_n = f\)) whenever \((\ell_{f_n})_{n \in \mathbb{N}}\) converges \(*\)-weakly to \(\ell_f\) in \(\mathcal{M}_s^\ast(X)\), that is

\[
\lim_{n \to \infty} \langle f_n, \mu \rangle = \langle f, \mu \rangle \quad \text{for all } \mu \in \mathcal{M}_s(X). \tag{5.7}
\]

It is not hard to show that \(w^\ast\)-lim\(n \to \infty\) \(f_n = f\) is equivalent to the requirement that \(f_n(x) \to f(x)\) for \(x \in X\), and the sequence \((\|f_n\|_\infty)_{n \in \mathbb{N}}\) is bounded.

Suppose we are given a subspace \(L_0 \leq \mathcal{M}_s^\ast(X)\) of bounded linear operators \(T^t : L \to L, t \geq 0\). Let

\[
L_0(T) := \{f \in L : w^\ast\lim_{t \to 0} T^t f = f\}. \tag{5.8}
\]

We can consider the so-called weak infinitesimal operator \([11] \text{ Ch.1 §6}\) of the semigroup \((T^t)_{t \geq 0}\). It is the function \(A : D(A) \to L_0(T)\) defined by

\[
Af = w^\ast\lim_{t \to 0} \frac{T^t f - f}{t} \quad \text{for } f \in D(A),
\]

where

\[
D(A) := \left\{ f \in L : w^\ast\lim_{t \to 0} \frac{T^t f - f}{t} \text{ exists and belongs to } L_0(T) \right\}.
\]

Clearly, \(D(A) \subset L_0(T)\). We now give several properties of such an operator, which turn out to be useful for our further considerations.

**Remark 5.4.** Under the assumptions made above the following statements hold (see \([11] \text{ p. } 40\) or \([?], \text{ pp. } 437-448\) for the proofs):

(i) \(w^\ast\)-cl \(D(A) = w^\ast\)-cl \(L_0(T)\), where \(w^\ast\)-cl denotes the \(*\)-weak closure in \(B_b(X)\).

(ii) For every \(f \in D(A)\) the derivative

\[
t \mapsto \frac{d^+ T^t f}{dt} := w^\ast\lim_{h \to 0^+} \frac{T^{t+h} f - T^t f}{h}
\]

exists, is \(*\)-weak continuous from the right, and

\[
\frac{d^+ T^t f}{dt} = AT^t f = T^t Af \quad \text{and} \quad T^t f - f = \int_0^t T^s Af \, ds \quad \text{for all } t \geq 0.
\]

(iii) For any \(\beta > 0\), the operator \(\beta \text{id} - A : D(A) \to L_0(T)\) is invertible and the inverse operator \(R_\beta := (\beta \text{id} - A)^{-1} : L_0(T) \to D(A)\) (called the resolvent of \(A\)) is given by

\[
R_\beta f(x) = \int_0^\infty e^{-\beta t} T^t f(x) \, dt, \quad x \in X, \ f \in L_0(t).
\]

We are now in a position to give a relationship between invariant measures of discrete
and continuous-time dynamical systems. Let us define the stochastic kernels \( G, W : X \times B(X) \rightarrow [0, 1] \) by

\[
G((y, i), A) = \int_0^\infty \lambda e^{-\lambda t} 1_A(S_i(t, y), i) \, dt
\]

\[
W((y, i), A) = \sum_{j \in I} \int_{B_H(0, \varepsilon)} \int_\Theta 1_A(w_\theta(y) + h, j) \pi_{ij}(w_\theta(y) + h)p(y, \theta) \, d\theta \nu^\varepsilon(dh).
\]

By the convention, we use the same symbols to denote the corresponding Markov operators. It is easily seen that both \( G \) and \( W \) are Feller.

Our proof of the following theorem uses similar techniques to those developed in \cite{17} Theorem 5.3.1 and \cite{3} Proposition 2.1 and 2.4.

**Theorem 5.5.** Let \( P_\varepsilon \) and \((P^t_\varepsilon)_{t \geq 0} \) denote the Markov operator and the Markov semigroup corresponding to \((3.8)\) and \((3.10)\), respectively.

1. If the Markov operator \( P_\varepsilon \) has an invariant probability measure \( \mu^* \), then \( \nu^* := \mu^* G \) is an invariant measure for the Markov semigroup \((P^t_\varepsilon)_{t \geq 0} \), and \( \nu^* W = \mu^* \).

2. If the Markov semigroup \((P^t_\varepsilon)_{t \geq 0} \) has an invariant probability measure \( \nu^* \), then \( \mu^* := \nu^* W \) is an invariant measure for the Markov operator \( P_\varepsilon \), and \( \mu^* G = \nu^* \).

**Proof.** According to Lemma \(5.1\) we may consider the contraction semigroup \( \tilde{P}^t : C_b(X) \rightarrow C_b(X) \), \( t \geq 0 \), given by

\[
\tilde{P}^t f = P^t_\varepsilon f \quad \text{for} \quad f \in C_b(X).
\]

From Lemma \(5.3\) we know that \( L_0(P) = C_b(X) \), and we can define the weak infinitesimal generator \( B : D(B) \rightarrow C_b(X) \) of the semigroup \((\tilde{P}^t)_{t \geq 0} \).

Define the maps \( Q^t : C_b(X) \rightarrow C_b(X), t \geq 0 \), by

\[
Q^t f(y, i) = f(S_t(y, i)) \quad \text{for} \quad (y, i) \in X.
\]

Obviously, \((Q^t)_{t \geq 0} \) forms a contraction semigroup of linear operators and \( L_0(Q) = C_b(X) \). Let \( A : D(A) \rightarrow C_b(X) \) denote the weak infinitesimal generator of this semigroup.

We shall prove that \( D(A) = D(B) \) and

\[
B f = \lambda W f + Af - \lambda f \quad \text{for} \quad f \in D(B).
\]

To do this, fix \( f \in D(A) \), and let \( \Psi_1 \) be the function determined by \((5.4)\) for \( n = 1 \). It now follows from Lemma \(5.2\) that there exists \( u_f : X \times (0, \infty) \rightarrow \mathbb{R} \) such that

\[
\frac{1}{t} \left( \tilde{P}^t f(y, i) - f(y, i) \right) = \lambda e^{-\lambda t} \int_0^t \Psi_1 f(y, i, s, t) \, ds + e^{-\lambda t} \frac{1}{t} \left( f(S_t(y, i), i) - f(y, i) \right)
\]

\[
- \lambda \frac{1 - e^{-\lambda t}}{\lambda t} f(y, i) + u_f(y, i, t) \quad \text{for} \quad (y, i) \in X, t > 0,
\]

and \( w^* \lim_{t \rightarrow 0} u_f(\cdot, t)/t = 0 \). From the continuity of \( f, S_t(\cdot, y), w_\theta \) and the boundedness of \( f \), we infer that the maps \( (s, t) \mapsto \Psi_1 f(y, i, s, t) \) are continuous. Therefore, and by the
mean value theorem, for each \( t > 0 \) and \((y, i) \in X\), we can choose \( s_{(y,i)}(t) \in (0, t)\) such that
\[
\frac{1}{t} \int_0^t \Psi_1 f(y, i, s, t) \, ds = \Psi_1 f(y, i, s_{(y,i)}(t), t) \quad \text{for} \quad (y, i) \in X, \; t > 0.
\]
Keeping in mind the fact that \( |\Psi_1| \leq \|f\|_\infty \), we see that
\[
w^* - \lim_{t \to 0} e^{-\lambda t} \frac{1}{t} \int_0^t \Psi_1 f(\cdot, s, t) \, ds = \Psi_1 f(\cdot, 0, 0) = Wf,
\]
since the expression under the limit sign is bounded for all \((y, i) \in X\) and all \( t \) in some neighborhood of zero. Moreover, by the definition of \((Q^t)_{t \geq 0}\), we have
\[
\frac{1}{t} (f(S_i(t, y), i) - f(y, i)) = \frac{1}{t} (Q^t f(y, i) - f(y, i)) \quad \text{for} \quad (y, i) \in X, \; t > 0,
\]
and \( w^*-\lim_{t \to 0} e^{-\lambda t} (Q^t f - f) / t = Af \) since \( f \in D(A) \). Summarizing, we obtain
\[
w^* - \lim_{t \to 0} \frac{1}{t} (\mathcal{P}^t f - f) = \lambda Wf + Af - \lambda f \in C_b(X),
\]
which gives \( (5.10) \) and \( D(A) \subset D(B) \). Clearly, letting \( f \in D(B) \), we conclude analogously that \( D(B) \subset D(A) \).

Let us further observe that the operator \( P \) can be expressed as
\[
P = GW \tag{5.11}
\]
since
\[
GWf(y, i) = \int_0^\infty \lambda e^{-\lambda t} Wf(S_i(t, y), i) \, dt = \sum_{j \in I} \int_0^\infty \int_{B_H(0,c)} \int_{\Theta} \lambda e^{-\lambda t} f(w_\theta(S_i(t, y)) + h, j) \pi_{ij}(w_\theta(S_i(t, y)) + h) p(S_i(t, y), \theta) \, d\theta \, d\nu^x(dh) \, dt = P f(y, i) \quad \text{for} \quad f \in B_b(X), \; (y, i) \in X.
\]

Now, consider the resolvent \( R_\lambda : C_b(X) \to D(A) \) of the operator \( A \). As we pointed out in Remark 5.4(iii), we have
\[
R_\lambda f(y, i) = \int_0^\infty e^{-\lambda t} Q^t f(y, i) \, dt, \quad f \in C_b(X), \; (y, i) \in X,
\]
which implies \( G|_{C_b(X)} = \lambda R_\lambda \). Hence, in particular, \( G(C_b(X)) \subset D(A) \) and
\[
G(\lambda \text{id} - A) f = (\lambda \text{id} - A) G f = \lambda f \quad \text{for} \quad f \in D(A). \tag{5.12}
\]

We further proceed to show assertion \([1]\). For this purpose, suppose that \( P_\varepsilon \) has an invariant measure \( \mu^* \in \mathcal{M}_1(X) \), and let \( \nu^* := \mu^* G \). Applying \( (5.11) \), we obtain
\[
\nu^* W = \mu^* GW = \mu^* P = \mu^*.
\]
\[ \langle f, \nu^* \rangle = \langle Gf, \mu^* \rangle = \langle PGf, \mu^* \rangle = \langle GWGf, \mu^* \rangle = \langle Wf, \nu^* \rangle, \quad f \in B_b(X). \]  

(5.13)

Letting \( f \in D(A) = D(B) \), we can apply (5.13) with \( (\lambda \text{id} - A)f \) in place of \( f \), and use the identity (5.12) to obtain

\[ \langle (\lambda \text{id} - A)f, \nu^* \rangle = \langle WG(\lambda \text{id} - A)f, \nu^* \rangle = \langle \lambda Wf, \nu^* \rangle. \]

Consequently, by (5.10) we then see that

\[ \langle Bf, \nu^* \rangle = \langle \lambda Wf + Af - \lambda f, \nu^* \rangle = 0 \quad \text{for} \quad f \in D(B). \]  

(5.14)

From Remark 5.4 (ii) we know that, for \( f \in D(B) \), the map \( s \mapsto \tilde{P}_sBf \) is \( * \)-weak continuous from the right and

\[ \int_0^t \tilde{P}_sBf \, ds = \tilde{P}_t f - f. \]

Therefore, by (5.14), we obtain for \( f \in D(B) \) and \( t \geq 0 \),

\[ \left\langle \tilde{P}_t f - f, \nu^* \right\rangle = \left\langle \int_0^t B\tilde{P}_s f \, ds, \nu^* \right\rangle = \int_0^t \left\langle B\tilde{P}_s f, \nu^* \right\rangle \, ds = 0, \]

and thus

\[ \left\langle f, \nu^* \tilde{P}_t \right\rangle = \left\langle f, \nu^* \right\rangle. \]

Since Remark 5.4 (i) provides \( C_b(X) \subset w^* - \text{cl} D(B) \), by using (5.7), we can conclude that this equality holds for all \( f \in C_b(X) \). It then follows that \( \nu^* \) is invariant for the semigroup \((P^t_\varepsilon)_ {t \geq 0}\), which completes the proof of (1).

For the proof of statement (2), let \( \nu^* \in \mathcal{M}_1(X) \) be an invariant measure for the semigroup \((P^t_\varepsilon)_{t \geq 0}\), and set \( \mu^* := \nu^*W \).

Letting \( h \in D(B) \) and differentiating at \( t = 0 \) the identity \( \left\langle \tilde{P}^t h, \nu^* \right\rangle = \left\langle h, \nu^* \right\rangle \), we obtain \( \langle Bh, \nu^* \rangle = 0 \). According to (5.10) and the fact that \( D(B) = D(A) \), this implies \( \langle \lambda Wh, \nu^* \rangle = \langle (\lambda \text{id} - A)h, \nu^* \rangle \). If we now let \( f \in D(A) \) and take \( h = Gf \) then, in view of (5.12), we infer that

\[ \langle f, \nu^* \rangle = \frac{1}{\lambda} \langle (\lambda \text{id} - A)Gf, \nu^* \rangle = \langle WGf, \nu^* \rangle. \]

Since \( C_b(X) \subset w^* - \text{cl} D(A) \), it is clear that the latter equality holds for any \( f \in C_b(X) \), and therefore \( \nu^* = \nu^*WG \). Finally, using this and (5.11), we obtain \( \mu^*G = \nu^*WG = \nu^* \), as well as

\[ \mu^* = \nu^*W = (\nu^*WG)W = (\nu^*W)(GW) = \mu^*P, \]

which completes the proof.

\[ \square \]

**Corollary 5.6.** Suppose that conditions (A1)-(A5) and (4.4) hold. Then \((P^t_\varepsilon)_{t \geq 0}\), determined by (3.10), has a unique invariant distribution.
Lemma 5.8. For every arguments similar to those in the proof of [3, Lemma 2.5] and [18, Theorem 4], it follows from Theorem 5.5. In order to show that the above-mentioned limit holds, we use them vanishes at \( n \) and \( G \) operator that the function \( f \) appearing in (5.9) preserves the Lipschitz continuity, which is necessary for our proof method to work.

We now proceed with the proof of the SLLN for the process \((\overline{Y}(t), \overline{\xi}(t))_{t \geq 0}\). We will require that the function \( \mathcal{L} \) appearing in (A2) is constant. This will allow us to guarantee that the operator \( G \), corresponding to (5.9), preserves the Lipschitz continuity, which is necessary for our proof method to work.

The main idea of our approach is based on comparison of the averages \( t^{-1} \int_0^t f(\overline{Y}(s), \overline{\xi}(s)) \, ds \) and \( n^{-1} \sum_{k=0}^{N_t-1} Gf(Y_k, \xi_k) \), where \( f \in \text{Lip}_b(X) \). We aim to show that the difference between them vanishes at \( t \to \infty \), which enables the application of Theorem 4.3. The rest then follows from Theorem 5.5. In order to show that the above-mentioned limit holds, we use arguments similar to those in the proof of [3, Lemma 2.5] and [18, Theorem 4].

**Lemma 5.8.** For every \( f \in B_b(X) \) and \( G \) given by (5.9), we have

\[
\lim_{t \to \infty} \left( \frac{1}{t} \int_0^t f(\overline{Y}(s), \overline{\xi}(s)) \, ds - \frac{1}{N_t} \sum_{k=0}^{N_t-1} Gf(Y_k, \xi_k) \right) = 0 \text{ a.s.}
\]

**Proof.** Let \( f \in B_b(X) \), and introduce

\[
M_n = \sum_{k=0}^{n-1} \left( \int_{\tau_k}^{\tau_{k+1}} f(\overline{Y}(s), \overline{\xi}(s)) \, ds - \frac{1}{\lambda} Gf(Y_k, \xi_k) \right), \quad n \in \mathbb{N}.
\]

We will first prove that \((M_n)_{n \in \mathbb{N}}\) is a martingale in the filtration \( \mathcal{F}_n = \sigma\{Y_k, \xi_k, \tau_k : 0 \leq k \leq n\} \), \( n \in \mathbb{N} \), with uniformly bounded increments.

Let us define \( F : X \times \mathbb{R}_+ \to \mathbb{R} \) by

\[
F(y, i, t) = \int_0^t f(S_i(s, y), i) \, ds.
\]
Then $F$ is measurable, and by (3.9) we may write
\[
\int_{\tau_k}^{\tau_{k+1}} f(\bar{Y}(s), \xi(s)) \, ds = \int_{\tau_k}^{\tau_{k+1}} f(S_\xi_k(s - \tau_k, Y_k), \xi_k) \, ds = F(Y_k, \xi_k, \Delta \tau_{k+1}) \tag{5.15}
\]
for each $k \in \mathbb{N}_0$.

We now observe that, for $k \in \mathbb{N}_0$ and $(y, i, u) \in X \times \mathbb{R}_+$,
\[
\mathbb{E}[F(Y_k, \xi_k, \Delta \tau_{k+1}) \mid Y_k = y, \xi_k = i, \Delta \tau_k = u] = \int_0^\infty F(y, i, \Delta \tau_{k+1}) \mathbb{P}(\Delta \tau_{k+1} \in dt) \\
= \int_0^\infty \lambda e^{-\lambda t} \left( \int_0^t f(S_i(s, y), i) \, ds \right) \, dt = \int_0^\infty \left( \int_s^\infty \lambda e^{-\lambda t} \, dt \right) f(S_i(s, y), i) \, ds \\
= \int_0^\infty e^{-\lambda s} f(S_i(s, y), i) \, ds = \frac{1}{\lambda} Gf(Y_k, \xi_k).
\]

Consequently, using the Markov property we can conclude that
\[
\mathbb{E}[F(Y_k, \xi_k, \Delta \tau_{k+1}) \mid F_k] = \mathbb{E}[F(Y_k, \xi_k, \Delta \tau_{k+1}) \mid Y_k, \xi_k, \Delta \tau_k] = \frac{1}{\lambda} Gf(Y_k, \xi_k). \tag{5.16}
\]

From (5.15) and (5.16) it now follows that
\[
\mathbb{E}[M_{n+1} \mid F_n] = \sum_{k=0}^{n-1} \left( F(Y_k, \xi_k, \Delta \tau_{k+1}) - \frac{1}{\lambda} Gf(Y_k, \xi_k) \right) + \mathbb{E}[F(Y_n, \xi_n, \Delta \tau_{n+1}) \mid F_n] \\
- \frac{1}{\lambda} Gf(Y_n, \xi_n) = M_n \quad \text{for} \quad n \in \mathbb{N},
\]
so $(M_n)_{n \in \mathbb{N}}$ is a martingale. Finally, using the fact that $|F(\cdot, t)| \leq t \|f\|_\infty$ for $t \geq 0$, and $|Gf| \leq \|f\|_\infty$, we obtain
\[
\mathbb{E} \left[ (M_{n+1} - M_n)^2 \right] = \mathbb{E} \left[ \left( F(Y_n, \xi_n, \Delta \tau_{n+1}) - \frac{1}{\lambda} Gf(Y_n, \xi_n) \right)^2 \right] \\
\leq 2 \mathbb{E}[F(Y_n, \xi_n, \Delta \tau_{n+1})^2] + 2 \mathbb{E} \left[ \frac{1}{\lambda^2} Gf(Y_n, \xi_n)^2 \right] \\
\leq 2 \|f\|_\infty^2 \left( \mathbb{E}[(\Delta \tau_{n+1})^2] + \frac{1}{\lambda^2} \right) = \frac{6 \|f\|_\infty^2}{\lambda^2},
\]
which simply means that the martingale increments of $(M_n)_{n \in \mathbb{N}}$ are uniformly bounded.

Let us now define
\[
r_t = \frac{1}{N_t} \int_{\tau_N t}^{t} f(\bar{Y}(s), \xi(s)) \, ds \quad \text{when} \quad \tau_1 \leq t, \quad \text{and} \quad r_t = 0 \quad \text{otherwise}.
\]

Clearly
\[
|r_t| \leq \|f\|_\infty \frac{\Delta \tau_{N_t+1}}{N_t} \quad \text{whenever} \quad \tau_1 \leq t, \tag{5.17}
\]
and, for \( \tau_1 \leq t \), we can write

\[
\frac{1}{t} \int_0^t f(\bar{Y}(s), \bar{x}(s)) \, ds = \frac{N_t}{t} \left( \frac{1}{N_t} \sum_{k=0}^{N_t-1} \int_{t_k}^{t_{k+1}} f(\bar{Y}(s), \bar{x}(s)) \, ds + r_t \right). \tag{5.18}
\]

We then obtain what follows

\[
\begin{align*}
\frac{1}{t} \int_0^t f(\bar{Y}(s), \bar{x}(s)) \, ds &- \frac{1}{N_t} \sum_{k=0}^{N_t-1} Gf(Y_k, \xi_k) \\
&= \frac{N_t}{t} \left[ \frac{1}{N_t} \sum_{k=0}^{N_t-1} \left( \int_{t_k}^{t_{k+1}} f(\bar{Y}(s), \bar{x}(s)) \, ds - \frac{1}{\lambda} Gf(Y_k, \xi_k) \right) \right] \\
&\quad + \left( \lambda - \frac{N_t}{t} \right) \left( \frac{1}{N_t} \sum_{k=0}^{N_t-1} \frac{1}{\lambda} Gf(Y_k, \xi_k) + \frac{N_t}{t} r_t \right) \\
&= \frac{N_t}{t} \frac{M_{N_t}}{N_t} - \frac{1}{\lambda} \left( \lambda - \frac{N_t}{t} \right) \left( \frac{1}{N_t} \sum_{k=0}^{N_t-1} Gf(Y_k, \xi_k) + \frac{N_t}{t} r_t \right).
\end{align*}
\]

We now only need to show that the right-hand side of the above equality tends to 0 a.s., as \( t \to \infty \). For this purpose, we first note that \( n^{-1} \sum_{k=0}^{n-1} Gf(X_k) \) is bounded (by \( \|f\|_\infty \)) for all \( n \). Next, from the Elementary Renewal Theorem we know that \( N_t/t \to \lambda \), and thus \( N_t \to \infty \). From the Borel–Cantelli Lemma it then follows that \( \Delta \tau_{N_t+1}/N_t \to 0 \), since 

\[
\sum_{n=1}^{\infty} \mathbb{P}(\Delta \tau_{n+1}/n \geq \varepsilon) = \sum_{n=1}^{\infty} e^{-\lambda n \varepsilon} < \infty \quad \text{for any } \varepsilon > 0.
\]

This together with (5.17) ensures that \( r_t \to 0 \). Finally, by the SLLN for martingales (see Theorem 2.18) we have \( M_k/k \to 0 \), as \( k \to \infty \), and thus \( M_{N_t}/N_t \to 0 \). This yields the desired conclusion and completes the proof.

We are now ready to state the announced theorem.

**Theorem 5.9.** Suppose that conditions [A1] [A3] hold with \( \mathcal{L} = \text{const} \), and that (4.4) is satisfied. Then, for any \( f \in \text{Lip}_b(X) \) and any initial state \((y, i) \in X \), we have

\[
\lim_{t \to \infty} \frac{1}{t} \int_0^t f(\bar{Y}(s), \bar{x}(s)) \, ds = \langle f, \nu^* \rangle \quad \mathbb{P}_{(y, i)} - \text{a.s.}
\]

where \( \nu^* \) stands for the unique invariant distribution of \((\bar{Y}(t), \bar{x}(t))_{t \geq 0} \) (which exists by Corollary 5.6).

**Proof.** Fix \( f \in \text{Lip}_b(X) \), \((y, i) \in X \), and let \( \mu^* \) denote the unique invariant distribution of \((Y_n, \xi_n)_{n \in \mathbb{N}_0} \) (which exists by Theorem 4.1). It follows from Theorem 5.5 that \( \mu^* G = \nu^* \). Therefore, in view of Theorem 4.3 it suffices to show that \( Gf \in \text{Lip}_b(X) \), since then

\[
\lim_{t \to \infty} \frac{1}{N_t} \sum_{k=0}^{N_t-1} Gf(Y_k, \xi_k) = \langle Gf, \mu^* \rangle = \langle f, \nu^* \rangle \quad \mathbb{P}_{(y, i)} - \text{a.s.,}
\]

which, by Lemma 5.8 gives us the desired thesis.
By assumption, $\mathcal{L}$ is constant, say $\mathcal{L}(t) = \tilde{L}$ for $t \geq 0$. Let $(y, i), (z, l) \in X$, and let $L_f$ denote a Lipschitz constant of $f$. By condition [A2] we then have

\[
|Gf(y, i) - Gf(z, l)| \leq \int_0^\infty \lambda e^{-\lambda t} |f(S_i(t, y), i) - f(S_l(t, z), l)| ds
\]
\[
\leq \lambda L_f \int_0^\infty e^{-\lambda t} (\|S_i(t, y) - S_l(t, z)\| + c \delta_{il}) dt
\]
\[
\leq \lambda L_f \left[ L \int_0^\infty e^{(\lambda - \alpha) t} dt \|y - z\| + \mathcal{L}(\|z\|) \int_0^\infty \left( te^{-\lambda t} dt + c \int_0^\infty e^{-\lambda t} \right) \delta_{il} dt \right]
\]
\[
= \lambda L_f \left( \frac{L}{\lambda - \alpha} \|y - z\| + \left( \frac{L}{\lambda^2} + \frac{c}{\lambda} \right) \delta_{il} \right).
\]

Since $\tilde{L} = \mathcal{L}(M) \leq cL$ by (3.11), we see that

\[
|Gf(y, i) - Gf(z, l)| \leq \lambda L_f \left( \frac{L}{\lambda - \alpha} + \frac{L}{\lambda^2} + \frac{1}{\lambda} \right) \rho_c((y, i), (z, l)).
\]

The proof is now complete. \qed

6 A model for gene expression

We shall describe the dynamical system which occurs in a model for gene expression in the presence of transcriptional bursting (cf. [27]; for biological aspects, see [1, Ch.8] or [12, Ch.3]). To be more precise, we focus on the prokaryotic (bacterial) gene expression. Genes in prokaryotes are frequently organised in the so-called operons, that is, small groups of related structural genes, which are transcribed (at the same time) as a unit, into a single polycistronic mRNA (which encodes more than one protein). Typically, the proteins encoded by genes within the same operon interact in some way; for instance, the \textit{lac} operon in bacterium Escherichia coli has three genes involved in the uptake and breakdown of lactose.

Consider a prokaryotic cell and a single operon containing $d$ structural genes. Let $t \geq 0$ denote the age of the cell, and suppose that

\[
\mathbf{Y}(t) = (y_1(t), \ldots, y_d(t)) \in \mathbb{R}_+^d, \quad \text{where} \quad \mathbb{R}_+^d = [0, \infty)^d,
\]

describes the concentration of $d$ different protein types encoded by the genes within the operon.

The protein molecules undergo degradation, whose rate depends on the current amount of the gene product. We assume that this rate is determined by a Lipschitz vector field $\mathcal{D} : \mathbb{R}_+^d \to \mathbb{R}_+^d$. It follows from the Cauchy-Lipschitz theorem that there exists a unique solution of the initial value problem for the autonomous system of the form

\[
\frac{d}{dt} \mathbf{Y}(t) = -\mathcal{D}(\mathbf{Y}(t)),
\]
\[
\mathbf{Y}(0) = y.
\]

We shall denote this solution by $t \mapsto S(t, y) \in \mathbb{R}_+^d$. Then, $S(t, y)$ describes the amount of
the gene product (from the operon) at time $t$, assuming that $y$ determines its amount at the
time of the cell birth. Moreover, $(t, y) \mapsto S(t, y)$ is a flow, as the unique solution of (6.1).
In addition to the Lipschitz continuity of $D$, we require that $-D$ is dissipative, that is

$$
\langle y_1 - y_2, D(y_1) - D(y_2) \rangle \geq \alpha \| y_1 - y_2 \|^2 \quad \text{for} \quad y_1, y_2 \in \mathbb{R}_+^d
$$

(6.2)

for some $\alpha > 0$. Note that the simplest $D$ which satisfies condition (6.2) is just a linear
map $D(y_1, \ldots, y_d) = (a_1y_1, \ldots, a_ny_n)$ with positive $a_1, \ldots, a_d$ denoting different degradation
rates for each protein.

The degradation process is interrupted by transcription occurring in the so-called bursts. For all protein types encoded by the genes in the operon, the bursts will appear simulta-

nously at the same random times, say $0 =: \tau_0 < \tau_1 < \tau_2 < \ldots$. We assume that these times
occur at exponentially distributed intervals with a constant intensity $\lambda$. Since a prokaryotic
mRNA can be efficiently transcribed and translated at the same time (because of the lack
of nucleus), $\tau_0, \tau_1, \ldots$ determine the moments of production at once. Clearly, the number of
new translations (and so the produced proteins) can be different for each of the $d$ genes in the
operon.

Let $\kappa_n$ be an $\mathbb{R}_+^d$-valued random variable, which describes the amount of proteins pro-
duced (by genes in the operon) at the time $\tau_n$ . Then, the process $(\mathbf{Y}(t))_{t \geq 0}$ changes from
$\mathbf{Y}(\tau_k-) = \mathbf{Y}(\tau_k-) + \kappa_k$ for $k \in \mathbb{N}$. We assume that $\kappa_n$ depends only on the current
amount of the gene product $Y(\tau_n-)$ and is disturbed by a non-negative fluctuation. More
precisely, we require that it has the form $\kappa_n = \theta_n + H_n$, where $(H_n)_{n \in \mathbb{N}}$ forms a sequence
of $\mathbb{R}_+^d$-valued, identically distributed random variables satisfying $H_n \leq \varepsilon$, $n \in \mathbb{N}$, for some
$\varepsilon = (\varepsilon, \ldots, \varepsilon) \in \mathbb{R}_+^d$, and $\theta_n$ is a random variable, taking values in $\Theta := [0, \Delta]^d$, determined by

$$
P(\theta_n \in E | \mathbf{Y}(\tau_n-) = y) = \int_E p(y, \theta) d\theta, \quad E \in \mathcal{B}(\Theta),
$$

(6.3)

where $p : Y \times \Theta \to [0, \infty)$ is a continuous function satisfying $\int_\Theta p(y, \theta) d\theta = 1$ for $y \in \mathbb{R}_+^d$.

It is quite natural to require that all the variables defining the model, i.e. $\tau_n, \theta_n$ and $H_n$, satisfy
the independence conditions detailed in Section 3.

Suppose that the initial amount of the gene product (from the operon) is described by
a random variable $\mathbf{Y}(0)$ with an arbitrary (and fixed) distribution. Then, letting

$$
w_p(y) = y + \theta \quad \text{for} \quad \theta \in \Theta, \quad y \in \mathbb{R}_+^d,
$$

(6.4)

we see that for each $n \in \mathbb{N}$ and given $\mathbf{Y}(\tau_n-)$, the process $(\mathbf{Y}(t))_{t \geq 0}$ evolves as

$$
\mathbf{Y}(t) = \begin{cases}
S(t - \tau_n-, \mathbf{Y}(\tau_n-)) & \text{for} \quad t \in [\tau_n-, \tau_n), \\
w_{\theta_n} \left( S(\Delta \tau_n, \mathbf{Y}(\tau_n-)) + H_n \right) & \text{for} \quad t = \tau_n.
\end{cases}
$$

(6.5)

Such a dynamical system has the same form as $(\mathbf{Y}(t))_{t \geq 0}$ defined in Section 3.

We shall establish that the model described above satisfies conditions (A1) (A3) and
inequality (4.4). Clearly, (A3) is trivially satisfied with $L_w = 1$. Using condition (6.2) we
infer that, for any \( y_1, y_2 \in \mathbb{R}_+^d \),
\[
\frac{d}{dt} \| S(t, y_1) - S(t, y_2) \|^2 = 2 \left( S(t, y_1) - S(t, y_2), \frac{d}{dt} S(t, y_1) - \frac{d}{dt} S(t, y_2) \right) \\
= 2 \left( S(t, y_1) - S(t, y_2), D(S(t, y_2)) - D(S(t, y_1)) \right) \\
\leq -2\alpha \| S(t, y_1) - S(t, y_2) \|^2 .
\]

It then follows from Grönwall’s inequality that
\[
\| S(t, y_1) - S(t, y_2) \|^2 \leq e^{-2\alpha t} \| y_1 - y_2 \|^2 \quad \text{for} \quad t \geq 0,
\]
which ensures continuity of \( S \), and yields that \([A2]\) holds with \( \alpha := -\alpha < 0 \), \( L = 1 \) and \( \mathcal{L} = 0 \). We can observe that for these constants condition \([4.4] \) is also satisfied. The task is now to show \([A1] \). For this purpose, we first notice that \( S(t, 0) = 0 \) for \( t \geq 0 \). To see this, observe that for every \( y \in \mathbb{R}_+^d \),
\[
\langle y, D(y) \rangle \geq \langle y, D(y) \rangle - \langle y, D(0) \rangle = \langle y - 0, D(y) - D(0) \rangle \geq \alpha \| y \|^2
\]
due to \((6.2) \). In particular, this implies that
\[
\frac{d}{dt} \| S(t, 0) \|^2 = 2 \left( S(t, 0), \frac{d}{dt} S(t, 0) \right) = -2 \langle S(t, 0), D(S(t, 0)) \rangle \leq -2\alpha \| S(t, 0) \|^2
\]
for \( t \geq 0 \). Applying Grönwall’s inequality again, we obtain
\[
0 \leq \| S(t, 0) \|^2 \leq e^{-2\alpha t} \| S(0, 0) \|^2 = 0 \quad \text{for} \quad t \geq 0,
\]
which gives
\[
\| w_\theta(S(t, 0)) \| = \| w_\theta(0) \| = \| \theta \| \leq \sqrt{d} \Delta \quad \text{for} \quad t \geq 0, \ \theta \in \Theta.
\]
and provides that \([A1] \) is satisfied with \( y^* = 0 \).

Since only one flow \( S_1 := S \) is considered (i.e. \( I = \{ 1 \} \)), the auxiliary process \((\overline{Y}(t), \overline{\xi}(t))_{t \geq 0} \), determined by \((3.9) \), takes the form \((\overline{Y}(t), 1) \) for \( t \geq 0 \). Consequently, \((\overline{Y}(t))_{t \geq 0} \) is a Markov process and it can be identified with \((\overline{Y}(t), \overline{\xi}(t))_{t \geq 0} \). Using Corollary \( 5.6 \) and Theorem \( 5.9 \) we can now provide a criterion of the SLLN for this process.

**Proposition 6.1.** Let \((\overline{Y}(t))_{t \geq 0} \) be the process determined by \((6.1)-(6.5) \), and suppose that \([A4] \) and \([A5] \) hold for \( p \) appearing in \((6.3) \). Then \((\overline{Y}(t))_{t \geq 0} \) has a unique invariant distribution \( \nu^* \in \mathcal{M}_1(\mathbb{R}_+^d) \). Moreover, for any \( f \in \text{Lip}_b(\mathbb{R}_+^d) \) and any \( x \in \mathbb{R}_+^d \) we have
\[
\lim_{t \to \infty} \frac{1}{t} \int_0^t f(\overline{Y}(s)) \, ds = \langle f, \nu^* \rangle \quad \mathbb{P}_{x^*} \text{-a.s.}
\]

**Acknowledgement**

We acknowledge useful discussions with S.C. Hille. Hanna Wojewódka is supported by the Foundation for Polish Science (FNP).


References

[1] B. Alberts, D. Bray, K. Hopkin, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. *Essential Cell Biology, 4th edition*. Garland Science, New York, 2013.

[2] M. Benaïm, S. Le Borgne, F. Malrieu, and P.-A. Zitt. Quantitative ergodicity for some switched dynamical systems. *Electron. Commun. Probab.*, 17(56, 14), 2012.

[3] M. Benaïm, S. Le Borgne, F. Malrieu, and P.-A. Zitt. Qualitative properties of certain piecewise deterministic Markov processes. *Ann. Inst. Henri Poincar Probab.*, 51(3):1040–1075, 2014.

[4] P. Billingsley. *Convergence of Probability Measures*. Wiley, New York, 1999.

[5] E. Buckwar and M. G. Riedler. An exact stochastic hybrid model of excitable membranes including spatio-temporal evolution. *J. Math. Biol.*, 63(6):1051–1093, 2011.

[6] B. Cloez and M. Hairer. Exponential ergodicity for Markov processes with random switching. *Bernoulli*, 21(1):505–536, 2015.

[7] O.L.V. Costa and F. Dufour. Stability and ergodicity of piecewise deterministic Markov processes. *SIAM J. Control Optim.*, 47(2):1053–1077, 2008.

[8] M.H.A. Davis. Piecewise-deterministic Markov processes: a general class of nondiffusion stochastic models. *J. Roy. Statist. Soc. Ser. B*, 46(3):353–388, 1984.

[9] R.M. Dudley. Convergence of baire measures. *Studia Math.*, 27:251–268, 1966.

[10] E.B. Dynkin. Markov processes I. Springer-Verlag, Berlin, 1965.

[11] E.B. Dynkin. *Selected papers of E.B. Dynkin with commentary*. American Mathematical Society, RI; International Press, Cambridge, MA, 2000.

[12] M. Fitzerald-Hayes and F. Reichsman. *DNA and Biotechnology, 3rd edition*. Academic Press/Elsevier Inc., Burlington, 2010.

[13] M. Hairer. Exponential mixing properties of stochastic PDEs through asymptotic coupling. *Probab. Theory Related Fields*, 124(3):345–380, 2002.

[14] P. Hall and C.C. Heyde. *Martingale limits theory and its applications*. Academic Press, New York, 1980.

[15] S.C. Hille, K. Horbacz, and T. Szarek. Existence of a unique invariant measure for a class of equicontinuous Markov operators with application to a stochastic model for an autoregulated gene. *Ann. Math. Blaise Pascal*, 23(2):171–217, 2016.

[16] S.C. Hille, K. Horbacz, T. Szarek, and H. Wojewódka. Limit theorems for some Markov chains. *J. Math. Anal. Appl.*, 443(1):385–408, 2016.

[17] K. Horbacz. Invariant measures for random dynamical systems. *Dissertationes Math.*, 451, 2008.
[18] K. Horbacz. The central limit theorem for random dynamical systems. *J. Stat. Phys.*, 164(6):1261–1291, 2016.

[19] K. Horbacz. Strong law of large numbers for continuous random dynamical systems. *Statist. Probab. Lett.*, 118:70–79, 2016.

[20] K. Horbacz and M. Ślęczka. Law of large numbers for random dynamical systems. *J. Stat. Phys.*, 162(3):671–684, 2016.

[21] R. Kapica and M. Ślęczka. Random iterations with place dependent probabilities. *arXiv, 1107.0707v2*, 2012.

[22] A. Lasota. From fractals to stochastic differential equations, in: Chaos-the interplay between stochastic and deterministic behaviour. *Lecture Notes in Phys. (Springer Verlag)*, 457:235–255, 1995.

[23] A. Lasota and M.C. Mackey. Cell division and the stability of cellular populations. *J. Math. Biol.*, 38:241–261, 1999.

[24] A. Lasota, M.C. Mackey, and M. Ważewska-Czyżewska. Minimizing therapeutically induced anemia. *J. Math. Biol.*, 13(2):149–158, 1981.

[25] D.A. Levin, Y. Peres, and E.L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, 2009.

[26] T. Lipniacki, P. Paszek, A. Marciniak-Czochra, A.R. Brasier, and M. Kimmel. Transcriptional stochasticity in gene expression. *J. Theoret. Biol.*, 238(2):348–367, 2006.

[27] M.C. Mackey, M. Tyran-Kamińska, and R. Yvinec. Dynamic behavior of stochastic gene expression models in the presence of bursting. *SIAM J. Appl. Math.*, 73(5):1830–1852, 2013.

[28] S.P. Meyn and R.L. Tweedie. *Markov chains and stochastic stability*. Springer-Verlag, London, 1993.

[29] O. Radulescu, A. Muller, and A. Crudu. Théorèmes limites pour les processus de Markov à sauts. Synthèse des résultats et applications en biologie moléculaire. *Technique et Science Informatiques*, 26(3-4):443–469, 2007.

[30] D. Revuz. *Markov chains*. North-Holland Elsevier, Amsterdam, 1975.

[31] A. Shirikyan. A version of the law of large numbers and applications. In: *Probabilistic Methods in Fluids, Proceedings of the Swansea Workshop held on 14 - 19 April 2002, World Scientific, New Jersey*, pages 263–271, 2003.

[32] M. Ślęczka. Exponential convergence for Markov systems. *Ann. Math. Sil.*, 29:139–149, 2015.

[33] J.J. Tyson and K.B. Hannsgen. Cell growth and division: a deterministic/probabilistic model of the cell cycle. *J. Math. Biol.*, 23:231–246, 1986.
[34] M. Ważewska-Czyżewska and A. Lasota. Matematyczne problemy dynamiki układu krwinek czerwonych. *Mat. Stos.*, 6(3):23–40, 1976.

[35] H. Wojewódka. Exponential rate of convergence for some Markov operators. *Statist. Probab. Lett.*, 83(10):2337–2347, 2013.

[36] R. Zaharopol. *Invariant probabilities of Markov-Feller operators and their supports*. Frontiers in Mathematics. Birkhäuser Verlag, Basel, 2005.