Title: On the connection between exponential ergodicity of a piecewise deterministic Markov process and the chain given by its post-jump locations

Author: Dawid Czapla, Katarzyna Horbacz, Hanna Wojewódka-Ściążko

Citation style: Czapla Dawid, Horbacz Katarzyna, Wojewódka-Ściążko Hanna. (2020). On the connection between exponential ergodicity of a piecewise deterministic Markov process and the chain given by its post-jump locations. "AIP Conference Proceedings" Vol. 2293, iss. 1 (2020), art. no 420056, doi 10.1063/5.0026519

On the Connection between Exponential Ergodicity of a Piecewise Deterministic Markov Process and the Chain Given by its Post-jump Locations

Dawid Czapla[a), Katarzyna Horbacz[b) and Hanna Wojewódka-Ściążko[c)

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

[a)Corresponding author: dawid.czapla@us.edu.pl
[b)horbacz@math.us.edu.pl
[c)hanna.wojewodka@us.edu.pl

Abstract. The aim of this paper is to derive the exponential ergodicity in the Wasserstein distance for a piecewise-deterministic Markov process (PDMP), being typically encountered in biological models, defined via interpolation of some discrete-time Markov chain. The key idea of the presented approach is to show that existence of an appropriate Markovian coupling between two instances of the chain implies that the transition semigroup associated with the continuous-time process is exponentially contracting.

INTRODUCTION

We shall consider a special case of the random dynamical system described in [1, 2, 3] (cf. also [4, 5]), which can serve as a framework for a stochastic description of single gene expression process in the presence of transcriptional bursting (see e.g. [6]). Such a system can be viewed as a PDMP that evolves through random jumps, occurring according to a homogeneous Poisson process, while the behaviour between jumps is governed by a continuous semiflow. Our goal is to show that the transition semigroup of this process is exponentially ergodic in the Wasserstein distance (cf. [7, 4]).

Let (X, ρ) be a Polish metric space endowed with the Borel σ-field $\mathcal{B}(X)$. In the analysis that follows, we will also confine ourselves to the case where X is bounded, and, for simplicity, we further assume that $\rho \leq 1$. The deterministic evolution of the aforementioned process will be governed by a continuous semiflow $S : \mathbb{R}_+ \times X \to X$ satisfying

$$\rho(S(t,x), S(t,y)) \leq L e^{\alpha t} \rho(x,y) \quad \text{for any} \quad x, y \in X,\ t \geq 0, \tag{1}$$

with some $\alpha < 0$ and some $L < \infty$. Further, we let $\{w_\theta : \theta \in \Theta\}$ be a collection of transformations from X to itself, such that $\Theta \times X \ni (\theta, x) \mapsto w_\theta(x)$ is continuous. The set of indexes Θ can be chosen as an arbitrary topological space equipped with a finite Borel measure ϑ. The transformations w_θ will be related to the post-jump locations of the process; more specifically, if the system is in the state x just before a jump, then its position directly after the jump should be $w_\theta(x)$ with some randomly selected $\theta \in \Theta$. The choice of θ depends on x and is determined by a probability density function $\theta \mapsto p(x,\theta)$, such that $p : X \times \Theta \to \mathbb{R}_+$ is continuous, and $\int_\Theta p(x,\theta)\,\vartheta(d\theta) = 1$ for any $x \in X$. The jump rate of the process will be denoted by $\lambda > 0$.

To provide that the Markov chain constituted by the post-jump positions enjoys certain suitable ergodic properties, we impose several additional restrictions on the functions p and $\{w_\theta : \theta \in \Theta\}$. Namely, we assume that there exist positive constants L_w, L_p and δ_p such that, for any $x, y \in X$, the following conditions hold:

$$L L_w + \frac{\alpha}{\lambda} < 1, \tag{2}$$

$$\int_\Theta \rho(w_\theta(x), w_\theta(y)) p(x,\theta)\,\vartheta(d\theta) \leq L_w \rho(x,y), \qquad \int_\Theta |p(x,\theta) - p(y,\theta)|\,\vartheta(d\theta) \leq L_p \rho(x,y), \tag{3}$$

$$\int_{\Theta(x,y)} \min\{p(x,\theta), p(y,\theta)\}\,\vartheta(d\theta) \geq \delta_p, \quad \text{where} \quad \Theta(x,y) := \{\theta \in \Theta : \rho(w_\theta(x), w_\theta(y)) \leq L_w \rho(x,y)\}. \tag{4}$$

For any given probability measure μ on X, we first introduce a discrete-time X-valued stochastic process $\{\Phi_n\}_{n\in\mathbb{N}_0}$ with initial distribution μ, defined on a suitable probability space endowed with a probability measure $\mathbb{P}_\mu$, so that

$$\Phi_n = w_{\theta_n}(S(\Delta\tau_n, \Phi_{n-1})) \quad \text{with} \quad \Delta\tau_n := \tau_n - \tau_{n-1} \quad \text{for any} \quad n \in \mathbb{N}, \tag{5}$$

where $\{\tau_n\}_{n\in\mathbb{N}_0}$ and $\{\theta_n\}_{n\in\mathbb{N}}$ are two sequences of random variables with values in $\mathbb{R}_+$ and Θ, respectively, constructed in such a way that $\tau_0 = 0$, $\tau_{n-1} < \tau_n$ for any $n \in \mathbb{N}$, $\tau_n \to \infty$ (as $n \to \infty$) $\mathbb{P}_\mu$-a.s, and, for every $n \in \mathbb{N}$,

$$\mathbb{P}_\mu(\Delta\tau_n \le t \,|\, \mathcal{G}_{n-1}) = 1 - e^{-\lambda t} \quad \text{whenever} \quad t \ge 0,$$
$$\mathbb{P}_\mu(\theta_n \in D \,|\, S(\Delta\tau_n, \Phi_{n-1}) = x;\, \mathcal{G}_{n-1}) = \int_D p(x,\theta)\, \vartheta(d\theta) \quad \text{for all} \quad D \in \mathcal{B}(\Theta),\ x \in X,$$

where $\mathcal{G}_{n-1}$ is the σ-field generated by the variables $\Phi_0, \tau_1, \ldots, \tau_{n-1}$ and $\theta_1, \ldots, \theta_{n-1}$.

Under the assumption that θ_n and $\Delta\tau_n$ are conditionally independent given $\mathcal{G}_{n-1}$ for any n, it is easy to check that $\{\Phi_n\}_{n\in\mathbb{N}_0}$ is a time-homogeneous Markov chain with transition probability kernel $\Pi : X \times \mathcal{B}(X) \to [0,1]$ of the form

$$\Pi(x,A) = \mathbb{P}_\mu(\Phi_{n+1} \in A \,|\, \Phi_n = x) = \int_0^\infty \int_\Theta \lambda e^{-\lambda t} \mathbb{1}_B(w_\theta(S(t,x))) p(S(t,x),\theta)\, \vartheta(d\theta)\, dt \quad \text{for} \quad x \in X,\ A \in \mathcal{B}(X). \tag{6}$$

On the same probability space, we can now define an interpolation of the chain $\{\Phi_n\}_{n\in\mathbb{N}_0}$ as follows:

$$\Phi(t) := S(t - \tau_n, \Phi_n) \quad \text{whenever} \quad t \in [\tau_n, \tau_{n+1}) \quad \text{for any} \quad n \in \mathbb{N}_0. \tag{7}$$

It is easily seen that $\{\Phi(t)\}_{t\ge0}$ is a time-homogeneous Markov process satisfying $\Phi(\tau_n) = \Phi_n$ for any $n \in \mathbb{N}_0$. By $\{\Pi_t\}_{t\ge0}$ we will denote the transition semigroup of this process, i.e.

$$\Pi_t(x,A) = \mathbb{P}_\mu(\Phi(s+t) \in A \,|\, \Phi(s) = x) \quad \text{for any} \quad x \in X,\ A \in \mathcal{B}(X),\ s,t \ge 0. \tag{8}$$

Letting $\mathcal{P}(X)$ denote the set of all probability Borel measures on X, we can introduce the Markov operators $P, P_t : \mathcal{P}(X) \to \mathcal{P}(X)$. $t \ge 0$, associated with the kernels (6) and (8), respectively, given by

$$P\mu(A) := \int_X \Pi(x,A)\, \mu(dx) \quad \text{and} \quad P_t\,\mu(A) := \int_X \Pi_t(x,A)\, \mu(dx) \quad \text{for any} \quad A \in \mathcal{B}(X),\ \mu \in \mathcal{P}(X).$$

Obviously, P and $\{P_t\}_{t\ge0}$ describe the evolution of the distributions $\mu_n := \mathbb{P}_\mu(\Phi_n \in \cdot)$ and $\mu(t) := \mathbb{P}_\mu(\Phi(t) \in \cdot)$, respectively, that is, $\mu_n = P\mu_{n-1}$ for any $n \in \mathbb{N}$, and $\mu(s+t) = P_t\,\mu(s)$ for any $s,t \ge 0$.

Since X is bounded, we can endow $\mathcal{P}(X)$ with the *Wasserstein metric*, defined by

$$d_{\mathcal{W}}(\mu,\nu) := \sup\left\{\left|\int_X f\, d(\mu-\nu)\right| :\ f \in BL(X),\ |f|_{Lip} \le 1\right\} \quad \text{for any} \quad \mu,\nu \in \mathcal{P}(X), \tag{9}$$

where $BL(X)$ stands for the space of all bounded, Lipschitz continuous functions $f : X \to \mathbb{R}$, and $|f|_{Lip}$ denotes the minimal Lipschitz constant of $f \in BL(X)$. In our framework, where $\operatorname{diam} X \le 1$, the Wasserstein metric coincides with the so-called *Fortet–Mourier metric* (a.k.a. *bounded Lipschitz distance*), defined in a similar fashion, but taking the supremum over all $f \in BL(X)$ such that $|f|_{Lip} \le 1$ and $\sup_{x\in X}|f(x)| \le 1$ (see [8, p. 234]). Consequently, the topology induced on $\mathcal{P}(X)$ by $d_{\mathcal{W}}$ equals to the topology of weak convergence of measures (see [8, Theorem 8.3.2]).

In what follows, we will show that, under hypotheses (1)-(4), the process $\{\Phi(t)\}_{t\ge0}$ admits a unique stationary distribution $\nu_* \in \mathcal{P}(X)$ such that, for any $\mu \in \mathcal{P}(X)$, $d_{\mathcal{W}}(P_t\,\mu, \nu_*) \to 0$ (as $t \to \infty$) at an exponential rate.

MAIN RESULTS

First of all, we need to consider a joint Markov chain $\Psi := \{(\overline{\Phi}_n, \overline{\tau}_n)\}_{n\in\mathbb{N}_0}$ with state space $X^2 \times \mathbb{R}_+$, wherein $\overline{\Phi} := \{(\Phi_n^{(1)}, \Phi_n^{(2)})\}_{n\in\mathbb{N}_0}$ is a *Markovian coupling* between two instances of the chain $\{\Phi_n\}_{n\in\mathbb{N}_0}$, and $\{\overline{\tau}_n\}_{n\in\mathbb{N}_0}$ is a copy of $\{\tau_n\}_{n\in\mathbb{N}_0}$. More precisely, the transition probability kernel of Ψ, say B, should satisfy

$$\begin{gathered} B((x,y,s),(A\times X)\times\mathbb{R}_+) = P(x,A), \quad B((x,y,s),(X\times A)\times\mathbb{R}_+) = P(y,A) \quad \text{for all} \quad A \in \mathcal{B}(X), \\ B\Big((x,y,s), X^2 \times T\Big) = \int_T \lambda e^{-\lambda(t-s)} \mathbb{1}_{(s,\infty)}(t)\, dt \quad \text{for every} \quad T \in \mathcal{B}(\mathbb{R}_+), \end{gathered} \tag{10}$$

with any $(x,y,s) \in X^2 \times \mathbb{R}_+$. In practise, it is convenient to assume that Ψ is the coordinate process defined on the path space $(\Omega, \mathcal{F}) := \left((X^2 \times \mathbb{R}_+)^{\mathbb{N}}, \mathcal{B}(X^2 \times \mathbb{R}_+)^{\otimes \mathbb{N}}\right)$, endowed with an appropriate collection $\{\mathbb{B}_{\mathbf{m}} : \mathbf{m} \in \mathcal{P}(X^2 \times \mathbb{R}_+)\}$ of probability measures. This means that $\Psi_n(\omega) = \omega_n$ for $\omega = (\omega_0, \omega_1, \ldots) \in \Omega$, and each $\mathbb{B}_{\mathbf{m}}$ is constructed so that Ψ is a Markov chain on $(\Omega, \mathcal{F}, \mathbb{B}_{\mathbf{m}})$ with transition law B and initial measure $\mathbf{m}$. As usual, we then write $\mathbb{E}_{\mathbf{m}}$ for the associated expectation operator. If $\mathbf{m}$ is the Dirac measure at some $(x,y,s) \in X^2 \times \mathbb{R}_+$, i.e. $\mathbf{m} = \delta_{(x,y,s)}$, then we write $\mathbb{E}_{(x,y,s)}$ instead. In the case where $\mathbf{m} = \pi \otimes \delta_0$, where $\pi \in \mathcal{P}(X^2)$, we will simply write $\mathbb{E}_\pi$ rather than $\mathbb{E}_{\pi \otimes \delta_0}$.

Within this framework, we can also consider two copies of the process $\{\Phi(t)\}_{t \geq 0}$, defined as follows:

$$\Phi^{(i)}(t) := S\left(t - \overline{\tau}_n, \Phi_n^{(i)}\right) \quad \text{whenever} \quad t \in [\overline{\tau}_n, \overline{\tau}_{n+1}) \quad \text{for any} \quad n \in \mathbb{N}_0,\ i \in \{1,2\}.$$

Keeping in mind that $\rho \leq 1$ and proceeding similarly as in the proof of [9, Lemmas 2.1 and 2.2] (cf. also [10]), under conditions (1)-(4), one can construct $\Psi = \{(\Phi_n^{(1)}, \Phi_n^{(2)}, \overline{\tau}_n)\}_{n \in \mathbb{N}_0}$ in such a way that its transition probability kernel B satisfies (10) and the following holds:

$$\mathbb{E}_\pi\left[\rho\left(\Phi_n^{(1)}, \Phi_n^{(2)}\right)\right] \leq Cq^n \quad \text{for all} \quad \pi \in \mathcal{P}(X^2),\ n \in \mathbb{N}, \tag{11}$$

with some $q \in (0,1)$ and some $C \in \mathbb{R}$ (being independent on π). In particular, from (9) and (11) it then follows that

$$d_{\mathcal{W}}\left(P^n\mu, P^n\nu\right) \leq Cq^n \quad \text{for any} \quad \mu, \nu \in \mathcal{P}(X),\ n \in \mathbb{N}. \tag{12}$$

We now aim to show that the semigroup $\{P_t\}_{t \geq 0}$ is *exponentially contracting* in the Wasserstein metric, i.e. there exists $\overline{L} \in \mathbb{R}$ such that

$$d_{\mathcal{W}}(P_t\mu, P_t\nu) \leq \overline{L}e^{\alpha t} d_{\mathcal{W}}(\mu, \nu) \quad \text{for any} \quad \mu, \nu \in \mathcal{P}(X) \quad \text{and any} \quad t \geq 0. \tag{13}$$

It is essential to stress that, although the framework adopted here essentially refers to a particular model, our proof of this statement uses only the existence of a coupling of $\{\Phi_n\}_{n \in \mathbb{N}_0}$ with property (11) and condition (1), imposed on the semiflow S. The crucial piece of the proof is the following lemma, which draws heavily on the coupling technique developed by Hairer in [7].

Lemma 1. *Suppose that condition* (11) *holds with certain constants* $C < \infty$ *and* $q \in (0,1)$ *for some properly constructed* Ψ *with transition law satisfying* (10)*. Moreover, assume that* (1) *is fulfilled with some* $L < \infty$ *and some* $\alpha < 0$*. Then, there exists* $t_0 > 0$ *such that*

$$\mathbb{E}_\pi\left[\rho\left(\Phi^{(1)}(t), \Phi^{(2)}(t)\right)\right] \leq Le^{\alpha t}\mathbb{E}_\pi\left[\rho\left(\Phi_0^{(1)}, \Phi_0^{(2)}\right)\right] \quad \textit{for all} \quad t \geq t_0 \quad \textit{and any} \quad \pi \in \mathcal{P}(X^2).$$

Sketch of the proof. Let $\pi \in \mathcal{P}(X^2)$. For each $t \geq 0$, we define two sequences $\{\xi_n^{(i)}(t)\}_{n \in \mathbb{N}_0}$ $(i = 1,2)$ of random variables on the space $(\Omega, \mathcal{F}, \mathbb{B}_\pi)$ (associated with Ψ) by setting $\xi_n^{(i)}(t) := \Phi^{(i)}(\overline{\tau}_n \wedge t)$ for $i \in \{1,2\}$. Letting $\{N(s)\}_{s \geq 0}$ denote the Poisson process with arrival times $\{\overline{\tau}_n\}_{n \in \mathbb{N}_0}$, i.e. $N(s) = \max\{n \in \mathbb{N}_0 : \overline{\tau}_n \leq s\}$, we see that $\xi_n^{(i)}(t)$ can be also written as

$$\xi_n^{(i)}(t) = \begin{cases} \Phi_n^{(i)} & \text{if } \overline{\tau}_n \leq t, \\ S\left(t - \overline{\tau}_{N(t)}, \Phi_{N(t)}^{(i)}\right) & \text{if } \overline{\tau}_n > t. \end{cases}$$

Obviously, $\xi_0^{(i)}(t) = \Phi_0^{(i)} = \Phi^{(i)}(0)$ and $\lim_{n \to \infty} \xi_n^{(i)}(t) = \Phi^{(i)}(t)$ for any $i = 1,2$. Moreover, $\eta(t) := \{(\xi_n^{(1)}(t), \xi_n^{(2)}(t), \overline{\tau}_n)\}_{n \in \mathbb{N}_0}$ is a time-homogeneous Markov chain with respect to its natural filtration, further denoted by $\{\mathcal{F}_n\}_{n \in \mathbb{N}_0}$. Let us fix $t \geq 0$, $n \in \mathbb{N}$, and write

$$\mathbb{E}_\pi\left[\rho\left(\xi_n^{(1)}(t), \xi_n^{(2)}(t)\right) | \mathcal{F}_{n-1}\right] = \mathbb{E}_\pi\left[\mathbb{1}_{\{\overline{\tau}_n \leq t\}}\, \rho\left(\Phi_n^{(1)}, \Phi_n^{(2)}\right) | \mathcal{F}_{n-1}\right] + \mathbb{E}_\pi\left[\mathbb{1}_{\{\overline{\tau}_n > t\}}\, \rho\left(\Phi^{(1)}(t), \Phi^{(2)}(t)\right) | \mathcal{F}_{n-1}\right]. \tag{14}$$

Applying sequentially the Markov property and (1) we obtain

$$\begin{aligned}\mathbb{E}_\pi\left[\mathbb{1}_{\{\overline{\tau}_n > t\}}\, \rho\left(\Phi^{(1)}(t), \Phi^{(2)}(t)\right) | \mathcal{F}_{n-1}\right] &= \mathbb{E}_{\eta_{n-1}(t)}\left[\mathbb{1}_{\{\overline{\tau}_1 > t\}}\, \rho\left(S(t, \Phi_0^{(1)}), S(t, \Phi_0^{(2)})\right)\right] \leq Le^{\alpha t}\, \mathbb{E}_{\eta_{n-1}(t)}\left[\rho\left(\Phi_0^{(1)}, \Phi_0^{(2)}\right)\right] \\ &= Le^{\alpha t}\mathbb{E}_{\eta_{n-1}(t)}\left[\rho\left(\xi_0^{(1)}(t), \xi_0^{(2)}(t)\right)\right] = Le^{\alpha t}\rho\left(\xi_{n-1}^{(1)}(t), \xi_{n-1}^{(2)}(t)\right).\end{aligned}$$

If we now combine the last estimate with (14) and take the expectation of both sides of the resulting inequality, then we see that

$$\mathbb{E}_\pi\left[\rho\left(\xi_n^{(1)}(t), \xi_n^{(2)}(t)\right)\right] \leq \mathbb{E}_\pi\left[\rho\left(\Phi_n^{(1)}, \Phi_n^{(2)}\right)\right] + Le^{\alpha t}\mathbb{E}_\pi\left[\rho\left(\xi_{n-1}^{(1)}(t), \xi_{n-1}^{(2)}(t)\right)\right].$$

Now, using condition (11) and proceeding inductively on n, one can deduce that, for sufficiently large $t \geq 0$,

$$\mathbb{E}_\pi \left[\rho\left(\xi_n^{(1)}(t), \xi_n^{(2)}(t)\right)\right] \leq C(Le^{\alpha t} + q)^n + Le^{\alpha t}\, \mathbb{E}_\pi \left[\rho\left(\Phi_0^{(1)}, \Phi_0^{(2)}\right)\right] \quad \text{for all} \quad n \in \mathbb{N}. \tag{15}$$

Finally, choosing $t_0 > 0$ so large that $Le^{\alpha t_0} + q < 1$ and passing to limit as $n \to \infty$ in (15) for any $t \geq t_0$, we obtain the assertion of the lemma. □

For any two probability measures $\mu, \nu \in \mathcal{P}(X)$, let us now define

$$\mathcal{R}(\mu,\nu) := \left\{\pi \in \mathcal{P}(X^2) : \ \pi(dx \times X) = \mu(dx) \ \text{ and } \ \pi(X \times dx) = \nu(dx)\right\} \quad \text{and} \quad \langle \rho, \pi \rangle := \int_{X^2} \rho\, d\pi \ \text{ for } \ \pi \in \mathcal{R}(\mu,\nu).$$

From Lemma 1 and the definition of $d_{\mathcal{W}}$, given in (9), it follows that, for any $\mu, \nu \in \mathcal{P}(X)$,

$$d_{\mathcal{W}}(P_t\,\mu, P_t\,\nu) \leq \mathbb{E}_\pi \left[\rho\left(\Phi^{(1)}(t), \Phi^{(2)}(t)\right)\right] \leq Le^{\alpha t}\, \langle \rho, \pi \rangle \quad \text{whenever} \quad \pi \in \mathcal{R}(\mu,\nu) \quad \text{for all} \quad t \geq t_0. \tag{16}$$

On the other hand, due to the celebrated Kantorovich–Rubinshtein theorem (see e.g. [8, Theorem 8.10.45]), we have

$$d_{\mathcal{W}}(\mu,\nu) = \inf\{\langle \rho, \pi \rangle : \ \pi \in \mathcal{R}(\mu,\nu)\} \quad \text{for any} \quad \mu, \nu \in \mathcal{P}(X). \tag{17}$$

Consequently, taking the infimum over all $\pi \in \mathcal{R}(\mu,\nu)$ in (16) and applying (17), we obtain the announced result:

Theorem 1. *Under hypotheses of Lemma 1, the transition semigroup* $\{P_t\}_{t \geq 0}$ *of the Markov process* (7) *is exponentially contracting in the Wasserstein metric, that is,* (13) *holds with some* $\overline{L} \geq L$*. Moreover,* $\overline{L}$ *depends only on* L *and* α*.*

Finally, let us note that (12) implies that $\{P^n\mu\}_{n \in \mathbb{N}_0}$ is a Cauchy sequence in $(\mathcal{P}(X), d_{\mathcal{W}})$ for any $\mu \in \mathcal{P}(X)$. Consequently, since the space $(\mathcal{P}(X), d_{\mathcal{W}})$ is complete (cf. [8, Theorem 8.9.4]), there must exist $\mu_* \in \mathcal{P}(X)$ such that $d_{\mathcal{W}}(P^n\mu, \mu_*) \to 0$ (as $n \to \infty$), independently of the choice of μ. If the operator P enjoys the *Feller property*, i.e. $x \mapsto \langle f, P\delta_x \rangle$ is continuous for any bounded continuous function $f : X \to \mathbb{R}$, which is the case if P is given by (6), then μ_* is a unique invariant probability measure for P, i.e. $P\mu_* = \mu_*$. Consequently, P is then *geometrically ergodic* in $d_{\mathcal{W}}$ (c.f. Theorem 4.1 in [1] or [2]), i.e.

$$d_{\mathcal{W}}(P^n\mu, \mu_*) \leq Cq^n \quad \text{for any} \quad \mu \in \mathcal{P}(X), \ n \in \mathbb{N}.$$

In the case where P is defined by (6), there is a one-to-one correspondence between invariant probability measures for P and those for the semigroup $\{P_t\}_{t \geq 0}$ (see [1, Theorem 4.4]). In particular there is a unique invariant measure $\nu_* \in \mathcal{P}(X)$ for $\{P_t\}_{t \geq 0}$, i.e. $P_t\nu_* = \nu_*$ for any $t \geq 0$. Hence, appealing to (13), we come to the following conclusion:

Corollary 1. *Suppose that the process* $\{\Phi_t\}_{t \geq 0}$*, given by* (7)*, interpolates the Markov chain with transition law* (6)*, and that conditions* (1)-(4) *are satisfied. Then, the transition semigroup* $\{P_t\}_{t \geq 0}$*, corresponding to* $\{\Phi_t\}_{t \geq 0}$*, admits a unique invariant measure* $\nu_* \in \mathcal{P}(X)$*, and it is exponentially ergodic in the Wasserstein metric, that is*

$$d_{\mathcal{W}}(P_t\,\mu, \nu_*) \leq \overline{L}e^{\alpha t} d_{\mathcal{W}}(\mu, \nu_*) \quad \textit{for any} \quad \mu \in \mathcal{P}(X) \quad \textit{and any} \quad t \geq 0.$$

ACKNOWLEDGEMENTS

Hanna Wojewódka-Ściążko is supported by the National Science Centre of Poland. Grant ID: 2018/02/X/ST1/01518.

REFERENCES

[1] D. Czapla, K. Horbacz, and H. Wojewódka, arXiv **1707.06489** (2017).
[2] D. Czapla and J. Kubieniec, Dynamical Systems **34**, 130–156 (2019).
[3] D. Czapla, H. Wojewódka, and K. Horbacz, AIP Conference Proceedings **1978**, 470008:1–4 (2018).
[4] M. Benaïm, S. Le Borgne, F. Malrieu, and P.-A. Zitt, Ann. Inst. Henri Poincaré Probab. **51**, 1040–1075 (2014).
[5] M. Benaïm, S. Le Borgne, F. Malrieu, and P.-A. Zitt, Electron. Commun. Probab. **17** (2012).
[6] M. Mackey, M. Tyran-Kamińska, and R. Yvinec, SIAM J. Appl. Math. **73**, 1830–1852 (2013).
[7] B. Cloez and M. Hairer, Bernoulli **21**, 505–536 (2015).
[8] V. Bogachev, *Measure Theory*, Vol. 2 (Springer-Verlag, Berlin, 2007).
[9] D. Czapla, K. Horbacz, and H. Wojewódka, arXiv **1804.09220** (2018).
[10] M. Hairer, Probab. Theory Related Fields **124**, 345–380 (2002).