Bernoulli **27**(1), 2021, 348–380
https://doi.org/10.3150/20-BEJ1242

On sub-geometric ergodicity of diffusion processes

PETRA LAZIĆ* and NIKOLA SANDRIƆ

Department of Mathematics, University of Zagreb, 10000 Zagreb, Croatia.
E-mail: *petralaz@math.hr; †nsandric@math.hr

In this article, we discuss ergodicity properties of a diffusion process given through an Itô stochastic differential equation. We identify conditions on the drift and diffusion coefficients which result in sub-geometric ergodicity of the corresponding semigroup with respect to the total variation distance. We also prove sub-geometric contractivity and ergodicity of the semigroup under a class of Wasserstein distances. Finally, we discuss sub-geometric ergodicity of two classes of Markov processes with jumps.

Keywords: asymptotic flatness; diffusion process; sub-geometric ergodicity; total variation distance; Wasserstein distance

1. Introduction

One of the classical directions in the analysis of Markov processes centers around their ergodicity properties. In this article, we focus on both qualitative and quantitative aspects of this problem. More precisely, we discuss sub-geometric ergodicity of a diffusion process given by

$$\mathrm{d}X_t^x = b\big(X_t^x\big)\,\mathrm{d}t + \sigma\big(X_t^x\big)\,\mathrm{d}B_t, \qquad X_0^x = x \in \mathbb{R}^d, \tag{1.1}$$

with respect to the total variation distance and/or a class of Wasserstein distances. Here, $\{B_t\}_{t\geq 0}$ stands for a standard n-dimensional Brownian motion (defined on a stochastic basis $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ satisfying the usual conditions), and the coefficients $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times n}$ satisfy:

(C1) for any $r > 0$,

$$\sup_{x\in B_r(0)} \big(|b(x)| + \|\sigma(x)\|_{\mathrm{HS}}\big) < \infty;$$

(C2) for any $r > 0$ there is $\Gamma_r > 0$ such that for all $x, y \in B_r(0)$,

$$2\big\langle x - y, b(x) - b(y)\big\rangle + \|\sigma(x) - \sigma(y)\|_{\mathrm{HS}}^2 \leq \Gamma_r |x - y|^2;$$

(C3) there is $\Gamma > 0$ such that for all $x \in \mathbb{R}^d$,

$$2\big\langle x, b(x)\big\rangle + \|\sigma(x)\|_{\mathrm{HS}}^2 \leq \Gamma\big(1 + |x|^2\big),$$

where $B_r(x)$ denotes the open ball with radius $r > 0$ around $x \in \mathbb{R}^d$, and $\|M\|_{\mathrm{HS}}^2 := \mathrm{Tr}\, MM^T$ is the Hilbert–Schmidt norm of a real matrix M.

1.1. Structural properties of the model

It is well known that under (C1)–(C3), for any $x \in \mathbb{R}^d$, the stochastic differential equation (SDE) in (1.1) admits a unique strong non-explosive solution $\{X_t^x\}_{t\geq 0}$ which is a strong Markov process with

1350-7265

continuous sample paths and transition kernel $p(t,x,\mathrm{d}y)=\mathbb{P}(X_t^x\in \mathrm{d}y)$, $t\geq 0$, $x\in\mathbb{R}^d$, (see [20], Theorems 5.4.1, 5.4.5 and 5.4.6, and [54], Theorem 3.1.1). In the context of Markov processes, it is natural that the underlying probability measure depends on the initial conditions of the process. Using standard arguments (Kolmogorov extension theorem), it is well known that for each $x\in\mathbb{R}^d$ the above defined transition kernel defines a unique probability measure $\mathbb{P}^x$ on the canonical (sample-path) space such that the projection process, denoted by $\{X_t\}_{t\geq0}$, is a strong Markov process (with respect to the completion of the corresponding natural filtration), it has continuous sample paths, and the same finite-dimensional distributions (with respect to $\mathbb{P}^x$) as $\{X_t^x\}_{t\geq0}$ (with respect to $\mathbb{P}$). Since we are interested in distributional properties of the solution to (1.1) only, in the sequel we rather deal with $\{X_t\}_{t\geq0}$ than with $\{X_t^x\}_{t\geq0}$. According to [43], Lemma 2.5, $\{X_t\}_{t\geq0}$ is also a C_b-Feller process, that is, the corresponding semigroup, defined by

$$P_tf(x):=\mathbb{E}^x\bigl[f(X_t)\bigr]=\int_{\mathbb{R}^d}f(y)p(t,x,\mathrm{d}y),\quad t\geq0,\ x\in\mathbb{R}^d,\ f\in B_b\bigl(\mathbb{R}^d\bigr),$$

satisfies $P_t(C_b(\mathbb{R}^d))\subseteq C_b(\mathbb{R}^d)$. Here, $B_b(\mathbb{R}^d)$ and $C_b(\mathbb{R}^d)$ denote the spaces of bounded Borel measurable functions and bounded continuous functions, respectively. Let us remark that in the above-mentioned lemma the author assumes that $b(x)$ is continuous, but the assertion of the lemma also holds true in the case when $b(x)$ is locally bounded (condition (C1)). In particular, this automatically implies that $\{X_t\}_{t\geq0}$ is a strong Markov process with respect to the right-continuous and completed version of the underlying natural filtration. Further, in [55], Theorem V.21.1, it is shown that

$$f(X_t)-f(X_0)-\int_0^t\mathcal{L}f(X_s)\,\mathrm{d}s,\quad t\geq0,$$

is a $\mathbb{P}^x$-local martingale for every $x\in\mathbb{R}^d$ and every $f\in C^2(\mathbb{R}^d)$, where

$$\mathcal{L}f(x):=\bigl\langle b(x),\nabla f(x)\bigr\rangle+\frac{1}{2}\operatorname{Tr}\sigma(x)\sigma(x)^T\nabla^2f(x).$$

If $b(x)$ and $\sigma(x)$ are continuous, then the infinitesimal generator $(\mathcal{A},\mathcal{D}_\mathcal{A})$ of $\{X_t\}_{t\geq0}$ (with respect to the Banach space $(B_b(\mathbb{R}^d),\|\cdot\|_\infty)$) satisfies $C_c^2(\mathbb{R}^d)\subseteq\mathcal{D}_\mathcal{A}$ and $\mathcal{A}|_{\mathcal{D}_\mathcal{A}}=\mathcal{L}$. Here, $\|\cdot\|_\infty$ and $C_c^2(\mathbb{R}^d)$ denote the supremum norm and the space of twice continuously differentiable functions with compact support, respectively. Recall, the infinitesimal generator (with respect to $(\|\cdot\|_\infty, B_b(\mathbb{R}^d))$) of an $\mathbb{R}^d$-valued Markov process $\{M_t\}_{t\geq0}$ with semigroup $\{P_t\}_{t\geq0}$ (defined as above) is a linear operator $\mathcal{A}:\mathcal{D}_\mathcal{A}\to B_b(\mathbb{R}^d)$ defined by

$$\mathcal{A}f:=\lim_{t\to0}\frac{P_tf-f}{t},\quad f\in\mathcal{D}_\mathcal{A}:=\left\{f\in B_b\bigl(\mathbb{R}^d\bigr):\lim_{t\to0}\frac{P_tf-f}{t}\text{ exists in }\|\cdot\|_\infty\right\}.$$

If $b(x)$ and $\sigma(x)$ are Lipschitz continuous then $\{X_t\}_{t\geq0}$ is a C_∞-Feller process, that is, $P_t(C_\infty(\mathbb{R}^d))\subseteq C_\infty(\mathbb{R}^d)$ for all $t\geq0$ (see [55], page 164), where $C_\infty(\mathbb{R}^d)$ stands for the space of continuous functions vanishing at infinity.

1.2. Notation and preliminaries

We first recall some definitions and general results from the ergodic theory of Markov processes. Our main references are [49] and [64]. Let $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geq0},\{\theta_t\}_{t\geq0},\{M_t\}_{t\geq0},\{\mathbb{P}^x\}_{x\in\mathbb{R}^d})$, denoted by $\{M_t\}_{t\geq0}$ in the sequel, be a Markov process with càdlàg sample paths and state space $(\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d))$ (see

[7]). We let $p(t,x,\mathrm{d}y):=\mathbb{P}^x(M_t\in \mathrm{d}y)$, $t\geq 0$, $x\in\mathbb{R}^d$, denote the corresponding transition kernel. For $t\geq 0$ and a (not necessarily finite) measure μ on $\mathcal{B}(\mathbb{R}^d)$, μP_t stands for $\int_{\mathbb{R}^d} p(t,x,\mathrm{d}y)\mu(\mathrm{d}x)$. Also, assume that $p(t,x,\mathrm{d}y)$ is a probability measure, that is, $\{M_t\}_{t\geq0}$ does not admit a cemetery point in the sense of [7]. Observe that this is not a restriction since, as we have already commented, $\{X_t\}_{t\geq0}$ is non-explosive. The process $\{M_t\}_{t\geq0}$ is called

(i) ϕ-irreducible if there exists a σ-finite measure ϕ on $\mathcal{B}(\mathbb{R}^d)$ such that whenever $\phi(B)>0$ we have $\int_0^\infty p(t,x,B)\,\mathrm{d}t>0$ for all $x\in\mathbb{R}^d$;
(ii) transient if it is ϕ-irreducible, and if there exists a countable covering of $\mathbb{R}^d$ with sets $\{B_j\}_{j\in\mathbb{N}}\subseteq\mathcal{B}(\mathbb{R}^d)$, and for each $j\in\mathbb{N}$ there exists a finite constant $\gamma_j\geq0$ such that $\int_0^\infty p(t,x,B_j)\,\mathrm{d}t\leq\gamma_j$ holds for all $x\in\mathbb{R}^d$;
(iii) recurrent if it is ϕ-irreducible, and $\phi(B)>0$ implies $\int_0^\infty p(t,x,B)\,\mathrm{d}t=\infty$ for all $x\in\mathbb{R}^d$.

Let us remark that if $\{M_t\}_{t\geq0}$ is a ϕ-irreducible Markov process, then the irreducibility measure ϕ can be maximized. This means that there exists a unique "maximal" irreducibility measure ψ such that for any measure $\bar\phi$, $\{M_t\}_{t\geq0}$ is $\bar\phi$-irreducible if and only if $\bar\phi$ is absolutely continuous with respect to ψ (see [64], Theorem 2.1). In view to this, when we refer to an irreducibility measure we actually refer to the maximal irreducibility measure. It is also well known that every ψ-irreducible Markov process is either transient or recurrent (see [64], Theorem 2.3). Further, recall that a Markov process $\{M_t\}_{t\geq0}$ is called

(i) open-set irreducible if the support of its maximal irreducibility measure ψ,

$$\operatorname{supp}\psi=\left\{x\in\mathbb{R}^d:\psi(O)>0 \text{ for every open neighborhood } O \text{ of } x\right\},$$

has a non-empty interior;
(ii) aperiodic if it admits an irreducible skeleton chain, that is, there exist $t_0>0$ and a σ-finite measure ϕ on $\mathcal{B}(\mathbb{R}^d)$, such that $\phi(B)>0$ implies $\sum_{n=0}^\infty p(nt_0,x,B)>0$ for all $x\in\mathbb{R}^d$.

A (not necessarily finite) measure π on $\mathcal{B}(\mathbb{R}^d)$ is called invariant for $\{M_t\}_{t\geq0}$ if $\pi P_t=\pi$ for all $t\geq0$. It is well known that if $\{M_t\}_{t\geq0}$ is recurrent, then it possesses a unique (up to constant multiples) invariant measure π (see [64], Theorem 2.6). If the invariant measure is finite, then it may be normalized to a probability measure. If $\{M_t\}_{t\geq0}$ is recurrent with finite invariant measure, then $\{M_t\}_{t\geq0}$ is called positive recurrent; otherwise it is called null recurrent. Note that a transient Markov process cannot have a finite invariant measure. Indeed, assume that $\{M_t\}_{t\geq0}$ is transient and that it admits a finite invariant measure π, and fix some $t>0$. Then, for each $j\in\mathbb{N}$, with γ_j and B_j as above, we have

$$t\pi(B_j)=\int_0^t\pi P_s(B_j)\,\mathrm{d}s\leq\gamma_j\pi\left(\mathbb{R}^d\right).$$

Now, by letting $t\to\infty$ we obtain $\pi(B_j)=0$ for all $j\in\mathbb{N}$, which is impossible. A Markov process $\{M_t\}_{t\geq0}$ is called ergodic if it possesses an invariant probability measure π and there exists a nondecreasing function $r:[0,\infty)\to[1,\infty)$ such that

$$\lim_{t\to\infty}r(t)\left\|p(t,x,\mathrm{d}y)-\pi(\mathrm{d}y)\right\|_{\mathrm{TV}}=0,\quad x\in\mathbb{R}^d,$$

where $\|\mu\|_{\mathrm{TV}}:=\sup_{B\in\mathcal{B}(\mathbb{R}^d)}|\mu(B)|$ is the total variation norm of a signed measure μ (on $\mathcal{B}(\mathbb{R}^d)$). We say that $\{M_t\}_{t\geq0}$ is sub-geometrically ergodic if it is ergodic and $\lim_{t\to\infty}\ln r(t)/t=0$, and that it is geometrically ergodic if it is ergodic and $r(t)=\mathrm{e}^{\kappa t}$ for some $\kappa>0$. Let us remark that (under the assumptions of C_b-Feller property, open-set irreducibility and aperiodicity) ergodicity is equivalent to positive recurrence (see [49], Theorem 6.1, and [64], Theorems 4.1, 4.2 and 7.1).

We now recall the notion and some general facts about Wasserstein distances (on $\mathbb{R}^d$). Let ρ be a metric on $\mathbb{R}^d$. Denote by $\mathbb{R}^d_\rho$ the topology induced by ρ, and let $\mathcal{B}(\mathbb{R}^d_\rho)$ be the corresponding Borel σ-algebra. For $p \geq 0$ denote by $\mathcal{P}_{\rho,p}$ the space of all probability measures μ on $\mathcal{B}(\mathbb{R}^d_\rho)$ having finite p-th moment, that is, $\int_{\mathbb{R}^d} \rho(x_0, x)^p \mu(\mathrm{d}x) < \infty$ for some (and then any) $x_0 \in \mathbb{R}^d$. Also, $\mathcal{P}_{\rho,0}$ is denoted by $\mathcal{P}_\rho$. If ρ is the standard d-dimensional Euclidean metric, then $\mathcal{P}_{\rho,p}$ and $\mathcal{P}_\rho$ are denoted by $\mathcal{P}_p$ and $\mathcal{P}$, respectively. For $p \geq 1$ and $\mu, \nu \in \mathcal{P}$, the $\mathcal{L}^p$-Wasserstein distance between μ and ν is defined as

$$\mathcal{W}_{\rho,p}(\mu,\nu) := \inf_{\Pi \in \mathcal{C}(\mu,\nu)} \left(\int_{\mathbb{R}^d \times \mathbb{R}^d} \rho(x,y)^p \Pi(\mathrm{d}x, \mathrm{d}y) \right)^{1/p},$$

where $\mathcal{C}(\mu,\nu)$ is the family of couplings of μ and ν, that is, $\Pi \in \mathcal{C}(\mu,\nu)$ if and only if Π is a probability measure on $\mathbb{R}^d \times \mathbb{R}^d$ having μ and ν as its marginals. It is not hard to see that $\mathcal{W}_{\rho,p}$ satisfies the axioms of a (not necessarily finite) distance on $\mathcal{P}_\rho$. The restriction of $\mathcal{W}_{\rho,p}$ to $\mathcal{P}_{\rho,p}$ defines a finite distance. If $(\mathbb{R}^d, \rho)$ is a Polish space, then it is well known that $(\mathcal{P}_{\rho,p}, \mathcal{W}_{\rho,p})$ is also a Polish space (see [67], Theorem 6.18). Of our special interest will be the situation when ρ takes the form $\rho(x,y) = f(|x-y|)$, where $f : [0,\infty) \to [0,\infty)$ is a non-decreasing concave function satisfying $f(t) = 0$ if and only if $t = 0$. In this situation, the corresponding Wasserstein space is denoted by $(\mathcal{P}_{f,p}, \mathcal{W}_{f,p})$ (which does not have to be a Polish space). Observe that if $f(t) = \mathbb{1}_{(0,\infty)}(t)$, then $\mathcal{W}_{f,p}(\mu,\nu) = \|\mu - \nu\|_{\mathrm{TV}}$ for all $p \geq 1$. In the case when $f(t) = t$, the corresponding Wasserstein space is denoted just by $(\mathcal{P}_p, \mathcal{W}_p)$ (which is always a Polish space). For more on Wasserstein distances, we refer the readers to [67].

1.3. Main results

The main goal of this article is to obtain (sharp) conditions for sub-geometric ergodicity of $\{X_t\}_{t\geq 0}$ with respect to the total variation distance and/or a class of Wasserstein distances. Before stating the main results, we introduce some notation we need in the sequel. Fix $x_0 \in \mathbb{R}^d$ and $r_0 \geq 0$, and put

$$\begin{aligned}
c(x) &:= \sigma(x)\sigma(x)^T, \\
A(x) &:= \frac{1}{2} \operatorname{Tr} c(x), \quad x \in \mathbb{R}^d, \\
B_{x_0}(x) &:= \langle x - x_0, b(x) \rangle, \quad x \in \mathbb{R}^d, \\
C_{x_0}(x) &:= \frac{\langle x - x_0, c(x)(x - x_0) \rangle}{|x - x_0|^2}, \quad x \in \mathbb{R}^d \setminus \{x_0\}, \\
\gamma_{x_0}(r) &:= \inf_{|x - x_0| = r} C_{x_0}(x), \quad r > 0, \\
\iota_{x_0}(r) &:= \sup_{|x - x_0| = r} \frac{2A(x) - C_{x_0}(x) + 2B_{x_0}(x)}{C_{x_0}(x)}, \quad r > 0, \\
I_{x_0}(r) &:= \int_{r_0}^r \frac{\iota_{x_0}(s)}{s} \, \mathrm{d}s, \quad r \geq r_0.
\end{aligned}$$

Theorem 1.1. *Assume* (C1)–(C3), *and assume that* $\{X_t\}_{t\geq 0}$ *is open-set irreducible and aperiodic. Further, let* $\varphi : [1,\infty) \to (0,\infty)$ *be a non-decreasing, differentiable and concave function satisfying* $\lim_{t\to\infty} \varphi'(t) = 0$ *and*

$$\Lambda := \int_{r_0}^\infty \varphi\left(\int_{r_0}^u \mathrm{e}^{-I_{x_0}(v)} \, \mathrm{d}v + 1 \right) \frac{\mathrm{e}^{I_{x_0}(u)}}{\gamma_{x_0}(u)} \, \mathrm{d}u < \infty \tag{1.2}$$

for some $x_0 \in \mathbb{R}^d$ *and* $r_0 \geq 0$, *and assume that* $c(x)$ *is positive definite for all* $x \in \mathbb{R}^d$, $|x - x_0| \geq r_0$ (*hence, the above functions and the relation in* (1.2) *are well defined*). *Then,* $\{X_t\}_{t\geq 0}$ *admits a unique invariant* $\pi \in \mathcal{P}$ *satisfying*

$$\lim_{t\to\infty} \varphi\big(\Phi^{-1}(t)\big)\|\delta_x P_t - \pi\|_{\mathrm{TV}} = 0, \quad x \in \mathbb{R}^d,$$

where

$$\Phi(t) := \int_1^t \frac{\mathrm{d}s}{\varphi(s)}, \quad t \geq 1.$$

The proof of Theorem 1.1 is based on the Foster-Lyapunov method for sub-geometric ergodicity of Markov processes developed in [15]. The method itself consists of finding an appropriate recurrent set $\mathcal{C} \in \mathcal{B}(\mathbb{R}^d)$, and constructing an appropriate function $\mathcal{V} : \mathbb{R}^d \to [1, \infty)$ (the so-called Lyapunov (energy) function) contained in the domain of the extended generator $\mathcal{A}^e$ of the underlying Markov process $\{M_t\}_{t\geq 0}$ (see [50], Section 1, for details), such that the Lyapunov equation

$$\mathcal{A}^e\mathcal{V}(x) \leq -\varphi\big(\mathcal{V}(x)\big) + \beta\mathbb{1}_{\mathcal{C}}(x), \quad x \in \mathbb{R}^d, \tag{1.3}$$

holds for some $\beta \in \mathbb{R}$ (see [15], Theorem 3.4). The equation in (1.3) implies that for any $\delta > 0$ the $\varphi \circ \Phi^{-1}$-moment of the δ-shifted hitting time $\tau_{\mathcal{C}}^{\delta} := \inf\{t \geq \delta : M_t \in \mathcal{C}\}$ of $\{M_t\}_{t\geq 0}$ on $\mathcal{C}$ (with respect to $\mathbb{P}^x$) is finite and controlled by $\mathcal{V}(x)$ (see [15], Theorem 4.1). However, this property in general does not immediately imply ergodicity of $\{M_t\}_{t\geq 0}$. Namely, we also need to ensure that a similar property holds for any other "reasonable" set. If $\{M_t\}_{t\geq 0}$ is ψ-irreducible and $\mathcal{C}$ is a petite set, then indeed for any $\delta > 0$ the $\varphi \circ \Phi^{-1}$-moment of τ_B^{δ}, for any $B \in \mathcal{B}(\mathbb{R}^d)$ with $\psi(B) > 0$, is again finite and controlled by $\mathcal{V}(x)$ (see [15], the discussion after Theorem 4.1). Recall, a set $C \in \mathcal{B}(\mathbb{R}^d)$ is said to be petite if it satisfies a Harris-type minorization condition: there are a probability measure η_C on $\mathcal{B}((0,\infty))$ (the standard Borel σ-algebra on $(0,\infty)$) and a non-trivial measure ν_C on $\mathcal{B}(\mathbb{R}^d)$, such that $\int_0^\infty p(t,x,B)\eta_C(\mathrm{d}t) \geq \nu_C(B)$ for all $x \in C$ and $B \in \mathcal{B}(\mathbb{R}^d)$. Recall also that ψ-irreducibility implies that the state space (in this case $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ can be covered by a countable union of petite sets (see [49], Propositio 4.1. Also, C_b-Feller property and open-set irreducibility of $\{M_t\}_{t\geq 0}$ ensure that every compact set is petite (see [64], Theorems 5.1 and 7.1). Intuitively, petite sets take a role of singletons for Markov processes on non-discrete state spaces (see [49], Section 4, and [48], Chapter 5, for details). However, as in the discrete setting, $\{M_t\}_{t\geq 0}$ can also show certain cyclic behavior which causes ergodicity not to hold (see [49], Section 5, and [48], Chapter 5). By assuming aperiodicity (which excludes this type of behavior), the sub-geometric ergodicity of $\{M_t\}_{t\geq 0}$ follows from [25], Theorem 1, which states that finiteness of the $\varphi \circ \Phi^{-1}$-moment of $\tau_{\mathcal{C}}^{\delta}$ implies sub-geometric ergodicity of $\{M_t\}_{t\geq 0}$ with rate $r(t) = \varphi(\Phi^{-1}(t))$. Let us remark that, in the context of the process $\{X_t\}_{t\geq 0}$, the relation in (1.2) is crucial in the construction of (actually it appears as a part of) the appropriate Lyapunov function (see the proof of Theorem 1.1). Thus, through this relation we control the $\varphi \circ \Phi^{-1}$-moment of $\tau_{\mathcal{C}}^{\delta}$ with $\mathcal{C}$ being a closed ball around the origin with large enough radius. We also remark that using an analogous approach as above in [32], Chapter 4, positive recurrence of the process $\{X_t\}_{t\geq 0}$ with globally Lipschitz coefficients and with $c(x)$ being positive definite (hence, according to Theorem 2.3, $\{X_t\}_{t\geq 0}$ is open-set irreducible and aperiodic) has been discussed. Based on this result, and analyzing polynomial moments of hitting times of compact sets, in [65], Theorem 6, polynomial ergodicity of $\{X_t\}_{t\geq 0}$ has been obtained. In the follow up work, by using analogous techniques the same author established polynomial ergodicity of $\{X_t\}_{t\geq 0}$ without directly assuming ψ-irreducibility and aperiodicity of the process, but basing on a local irreducibility condition which we discuss below (see [66], Theorem 6).

An alternative and, in a certain sense, more general approach to this problem is based on a local irreducibility condition. In this approach, instead of (1.3), we assume a slightly more general form of the Lyapunov equation:

$$\mathcal{A}^e\mathcal{V}(x) \le -\varphi\big(\mathcal{V}(x)\big) + \beta, \quad x \in \mathbb{R}^d, \tag{1.4}$$

for some $\beta \in \mathbb{R}$, and instead of assuming ψ-irreducibility and aperiodicity of $\{M_t\}_{t\ge0}$, we assume the so-called (local) Dobrushin condition (also known as Markov-Dobrushin condition): the Lyapunov function $\mathcal{V}(x)$ has precompact sub-level sets, and for every $\gamma > 0$ there is $t_\gamma > 0$ such that

$$\sup_{(x,y)\in\{(u,v):\mathcal{V}(u)+\mathcal{V}(v)\le\gamma\}} \big\| p(t_\gamma, x, \mathrm{d}z) - p(t_\gamma, y, \mathrm{d}z)\big\|_{\mathrm{TV}} < 1, \tag{1.5}$$

see [27], Theorem 4.1, (see also [38], Chapter 1.4, and [39], Chapter 3). Observe that this condition actually means that for each $(x, y) \in \{(u, v) : \mathcal{V}(u) + \mathcal{V}(v) \le \gamma\}$ the probability measures $p(t_\gamma, x, \mathrm{d}z)$ and $p(t_\gamma, y, \mathrm{d}z)$ are not mutually singular. Intuitively, the Dobrushin condition encodes ψ-irreducibility and aperiodicity of $\{M_t\}_{t\ge0}$, and petiteness of sub-level sets of $\mathcal{V}(x)$. By using a coupling approach with an appropriately chosen Markov coupling of $\{M_t\}_{t\ge0}$, say $\{M^c_t\}_{t\ge0}$, the Lyapunov equation and Dobrushin condition, analogously as before, imply that the hitting (that is, coupling) time $\tau_c := \inf\{t \ge 0 : M^c_t \in \mathrm{diag}\}$ of $\{M^c_t\}_{t\ge0}$ on $\mathrm{diag} := \{(x, x) : x \in \mathbb{R}^d\}$ is a.s. finite (with respect to the probability measure corresponding to $\{M^c_t\}_{t\ge0}$ with any initial position $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$). Moreover, it follows that the Φ^{-1}-moment of τ_c is finite and controlled by $\mathcal{V}(x) + \mathcal{V}(y)$. Then from the coupling inequality it follows that $\{M_t\}_{t\ge0}$ admits a unique invariant $\pi \in \mathcal{P}$, and

$$\sup_{t\ge0} \varphi\big(\Phi^{-1}(t)\big)\big\| p(t, x, \mathrm{d}y) - \pi(\mathrm{d}y)\big\|_{\mathrm{TV}} < \infty, \quad x \in \mathbb{R}^d,$$

(see [27], Theorem 4.1, or [38], Chapter 1.4, and [39], Chapter 3, for the skeleton chain approach).

Observe that (1.4) follows from (1.3). Also, ψ-irreducibility and aperiodicity (together with (1.3)) imply that the Dobrushin condition holds on the Cartesian product of any petite set with itself. Namely, according to [49], Proposition 6.1, for any petite set C there is $t_C > 0$ such that for the measure η_C (in the definition of petiteness) the Dirac measure in t_C can be taken (with some, possibly different, non-trivial measure ν_C). Thus, $p(t_C, x, B) \ge \nu_C(B)$ for any $x \in C$ and $B \in \mathcal{B}(\mathbb{R}^d)$, which implies

$$\sup_{(x,y)\in C\times C} \big\| p(t_C, x, \mathrm{d}z) - p(t_C, y, \mathrm{d}z)\big\|_{\mathrm{TV}} < 1. \tag{1.6}$$

If in addition $\{M_t\}_{t\ge0}$ is C_b-Feller and open-set irreducible, as we have already commented, every compact set is petite so the above relation holds for any bounded set C, showing that, at least in this particular situation, the approach based on the Dobrushin condition is more general than the approach based on ψ-irreducibility and aperiodicity. Situations where it shows a clear advantage are discussed in [40] and [1]. In the first reference, the author considers a Markov process obtained as a solution to a Lévy-driven SDE with highly irregular coefficients and noise term, while in the second a diffusion process with highly irregular (discontinuous) drift function and uniformly elliptic diffusion coefficient has been considered. In these concrete situations, it is not clear whether one can obtain ψ-irreducibility and aperiodicity of the processes, whereas the authors obtain (1.6) for any compact set C (see [40], Theorem 1.3, and [1], Lemma 3). For more on ergodic properties of Markov processes based on the Dobrushin condition, we refer the readers to [27,38] and [39].

In the case of the process $\{X_t\}_{t\ge0}$, open-set irreducibility and aperiodicity will be satisfied if the coefficient $c(x)$ is Lipschitz continuous and uniformly elliptic (see the discussion after Proposition 2.2). In Theorem 2.3 we show that $\{X_t\}_{t\ge0}$ will be open-set irreducible and aperiodic if $b(x)$ and $c(x)$ are

Hölder continuous, and $c(x)$ is uniformly elliptic on an open ball only. Let us also remark that, without further regularity assumptions on $b(x)$ and $c(x)$, it is not clear how to check the Dobrushin condition in these two situations.

The problem of sub-geometric ergodicity of diffusion processes (with respect to the total variation distance) has already been considered in the literature (see [15,25,38,39,56,65] and [66]). In these works, it has been shown that $\{X_t\}_{t\geq 0}$ will be sub-geometrically ergodic with rate $t^{\alpha/(1-\alpha)}$ (that is, $\varphi(t) = t^{\alpha}$), $\alpha \in (0,1)$, if there exist $\gamma > 0$, $\Gamma > 0$ and $r_0 \geq 0$, such that

$$A(x) - \left(1 - \frac{\gamma}{2}\right) C_0(x) + B_0(x) \leq -\Gamma |x|^{\gamma\alpha - \gamma + 2}, \quad |x| \geq r_0. \tag{1.7}$$

However, this result is far for being sharp (optimal). Namely, in Proposition 2.6 we show that (1.7) implies (1.2), and in Example 2.5 we give an example of a diffusion process satisfying conditions from Theorem 1.1, but not the condition in (1.7).

On the other hand, in the case when $c(x)$ is not regular enough, the topology induced by the total variation distance becomes too "rough", that is, it cannot completely capture the singular behavior of $\{X_t\}_{t\geq 0}$. In other words, $p(t, x, \mathrm{d}y)$ cannot converge to the underlying invariant probability measure (if it exists) in this topology, but in a weaker sense (see [58] and the references therein). Therefore, in this situation, we naturally resort to Wasserstein distances which, in a certain sense, induce a finer topology, that is, convergence with respect to a Wasserstein distance implies the weak convergence of probability measures (see [67], Theorems 6.9 and 6.15).

Theorem 1.2. *Let $\sigma(x) \equiv \sigma$ be an arbitrary $d \times n$ matrix, and assume* (C1)–(C3). *Further, let $p \geq 1$ and let $f, \psi : [0,\infty) \to [0,\infty)$ be such that*

(i) *$f(t)$ is concave, non-decreasing, absolutely continuous on $[t_0, t_1]$ for any $0 < t_0 < t_1 < \infty$, and $f(t) = 0$ if and only if $t = 0$;*
(ii) *$\psi(t)$ is convex and $\psi(t) = 0$ if and only if $t = 0$;*
(iii) *there are $\gamma > 0$, $\Gamma > 0$ and $t_0 > 0$, such that $f(t_0) \leq \gamma$ and*

$$f'\big(|x-y|\big)\big\langle x-y, b(x)-b(y)\big\rangle \leq \begin{cases} -\Gamma|x-y|\psi\big(f\big(|x-y|\big)\big), & f\big(|x-y|\big) \leq \gamma, \\ 0, & f\big(|x-y|\big) > \gamma, \end{cases} \tag{1.8}$$

a.e. on $\mathbb{R}^d$.

Then,

(a) *for all $x, y \in \mathbb{R}^d$, $f(|x-y|) \leq \gamma$, it holds that*

$$\mathcal{W}_{f,p}(\delta_x P_t, \delta_y P_t) \leq \Psi^{-1}_{f(|x-y|)}(\Gamma t), \quad t \geq 0, \tag{1.9}$$

where $\Psi_\kappa(t) := \int_t^\kappa \frac{\mathrm{d}s}{\psi(s)}$ for $\kappa > 0$ and $t \in (0, \kappa]$.
(b) *for all $x, y \in \mathbb{R}^d$, $f(|x-y|) \leq \gamma$, and all $\kappa \geq \gamma$ it holds that*

$$\mathcal{W}_{f,p}(\delta_x P_t, \delta_y P_t) \leq \Psi^{-1}_{\kappa}(\Gamma t), \quad t \geq 0. \tag{1.10}$$

In addition, if $\Psi_\infty(t) := \int_t^\infty \frac{\mathrm{d}s}{\psi(s)} < \infty$ for $t \in (0,\infty)$, then

$$\mathcal{W}_{f,p}(\delta_x P_t, \delta_y P_t) \leq \Psi^{-1}_{\infty}(\Gamma t), \quad t \geq 0. \tag{1.11}$$

(c) *for any* $x, y \in \mathbb{R}^d$ *it holds that*

$$\mathcal{W}_{f,p}(\delta_x P_t, \delta_y P_t) \le \lceil \delta |x-y| \rceil \Psi_\gamma^{-1}(\Gamma t), \quad t \ge 0, \tag{1.12}$$

where $\delta := \inf\{t > 0 : f(t^{-1}) \le \gamma\}$ *and* $\lceil u \rceil$ *denotes the least integer greater than or equal to* $u \in \mathbb{R}$. *Also, according to (b),* $\Psi_\gamma^{-1}(\Gamma t)$ *in (1.12) can be replaced by* $\Psi_\kappa^{-1}(\Gamma t)$ *for any* $\kappa \ge \gamma$, *and by* $\Psi_\infty^{-1}(\Gamma t)$ *if* $\Psi_\infty(t) < \infty$ *for* $t \in (0,\infty)$.

Observe that $f(t)$ is $\mathcal{B}((0,\infty))$-measurable, implying that the relation in (1.9) is well defined. The proof of Theorem 1.2 is based on the so-called synchronous coupling method (see [12], Example 2.16, for details) and the asymptotic flatness condition given in (1.8). Let us remark that in a special case when $p = 2$ and $f(t) = \psi(t) = t$ in [68] it has been shown that the relation in (1.9) (observe that in this case $\Psi_{f(|x-y|)}^{-1}(\Gamma t) = |x-y|\mathrm{e}^{-\Gamma t}$) is equivalent to the asymptotic flatness condition (in the sense of [3])

$$\langle x-y, b(x)-b(y) \rangle \le -\Gamma |x-y|^2, \quad x, y \in \mathbb{R}^d. \tag{1.13}$$

Even though at first sight the condition in (1.8) seems to be less restrictive than the condition in (1.13), they are actually equivalent. This can be easily observed by taking an equidistant subdivision of the line segment connecting x and y, such that the distance between consecutive points is strictly less than γ, and then applying triangle inequality. On the other hand, in the case when $\psi(t)$ is not the identity function this does not hold in general. Namely, $\psi(t)$ is not sub-additive, but super-additive. A typical example of a drift function (in dimension $d = 1$) satisfying (1.8) (and (1.14)), but not (1.13), is $b(x) = -\operatorname{sgn}(x)|x|^p$, $p > 1$, together with $f(t) = t$ and $\psi(t) = |t|^p$ (see Example 3.3). More generally, no drift function that is sub-linear near the origin can satisfy (1.13), but it might satisfy (1.8).

Finally, as a consequence of Theorem 1.2 we conclude the following.

Theorem 1.3. *In addition to the assumptions of Theorem 1.2 with* $f(t) = t$, *assume*

$$\langle x-y, b(x)-b(y) \rangle \le -\Gamma |x-y| \psi(|x-y|), \quad x, y \in \mathbb{R}^d. \tag{1.14}$$

Then, the process $\{X_t\}_{t\ge0}$ *admits a unique invariant* $\pi \in \bigcap_{p\ge1} \mathcal{P}_p$, *and for any* $\kappa > 0$, $p \ge 1$ *and* $\mu \in \mathcal{P}_p$,

$$\mathcal{W}_p(\mu P_t, \pi) \le \left(\frac{\mathcal{W}_p(\mu, \pi)}{\kappa} + 1 \right) \Psi_\kappa^{-1}(\Gamma t), \quad t \ge 0. \tag{1.15}$$

Let us also remark that if $\sigma(x) \equiv \sigma$ is quadratic and non-singular matrix, and $b(x)$ satisfies the following asymptotic flatness condition

$$\langle x-y, b(x)-b(y) \rangle \le \begin{cases} \Gamma_1 |x-y|^2, & |x-y| \le \Delta, \\ -\Gamma_2 |x-y|^2, & |x-y| \ge \Delta, \end{cases} \quad x, y \in \mathbb{R}^d, \tag{1.16}$$

for some $\Gamma_1 > 0$, $\Gamma_2 > 0$ and $\Delta > 0$, by using the so-called coupling by reflection method (see [12], Example 2.16, for details), in [21] (see also [22] and [42]) it has been shown that there is a concave function $f(t)$ (given explicitly in terms of the constants Γ_1, Γ_2 and Δ, and coefficients σ and $b(x)$) defining a metric $\rho(x, y) = f(|x-y|)$ on $\mathbb{R}^d$ under which $\{X_t\}_{t\ge0}$ satisfies contraction property of the type (1.12) with geometric rate of convergence, and geometric ergodicity property of the type (1.15). As we have already commented, $b(x) = -\operatorname{sgn}(x)|x|^p$, $p > 1$, satisfies (1.8) and (1.14), but clearly it also satisfies (1.16). However, in the later case, in order to conclude contractivity or ergodicity it is necessary to assume non-singularity of σ, while in the former case we can allow σ to be singular. Let us also remark that in the case when σ is non-singular, by taking $y = 0$ in (1.16), one can easily see that $\{X_t\}_{t\geq 0}$ is geometrically ergodic with respect to the total variation distance (see Proposition 2.2).

1.4. Literature review

Our work relates to the active research on ergodicity properties of Markov processes, and the vast literature on SDEs. In [3,6,38,39,61] and [65] ergodicity properties with respect to the total variation distance of diffusion processes are established using the Foster-Lyapunov(-type) method. In this article, we generalize the ideas from [6] (see also [26], Chapter 9, and [29], Supplement) and obtain sharp conditions which ensure ergodicity properties with sub-geometric rates of convergence of this class of processes. Furthermore, we adapt these results and discuss also ergodicity properties of a class of diffusion processes with jumps and a class of Markov processes obtained through the Bochner's subordination. These results are related to [2,4,13–15,18,25,31,40,46,47,49,50,56,69–71] and [73] where the ergodicity properties of general Markov processes are established using the Foster-Lyapunov method again.

The studies on ergodicity properties with respect to the total variation distance assume that the Markov processes are irreducible and aperiodic. This is satisfied if the process does not show a singular behavior in its motion, that is, its diffusion part is non-singular and/or its jump part shows enough jump activity. For Markov processes that do not converge in total variation, ergodic properties under Wasserstein distances are studied since they may converge weakly under certain conditions, see [8, 10,21,22,28,42,44,68] and [72]. In [8] and [68], the coupling approach and the asymptotic flatness property in (1.13) are employed to establish geometric contractivity and ergodicity of the semigroup of a diffusion process with possibly singular diffusion coefficient, with respect to a Wasserstein distance. However, in many situations the condition in (1.13) is too restrictive. For example, as we have already commented, drift functions which are sub-linear near the origin do not satisfy (1.13). The first step in relaxing this condition has been recently done in [21] (see also [22] and [42]) where (1.13) is replaced by the asymptotic flatness property in (1.16), but at the price of assuming that the diffusion coefficient is non-singular. Under these assumptions geometric contractivity and ergodicity of the semigroup of a diffusion process with respect to a Wasserstein distance are again established. In this article, we relax (1.13) to the asymptotic flatness conditions in (1.8) and (1.14), and obtain sub-geometric contractivity and sub-geometric ergodicity of the semigroup of a diffusion process, with possibly singular diffusion coefficient, with respect to a Wasserstein distance. At the end, we again discuss ergodicity properties, but with respect to Wasserstein distances, of a class of diffusion processes with jumps and a class of Markov processes obtained through the Bochner's subordination.

At the end, we remark that an analogous results, with respect to the total variation distance and Wasserstein distances, have also been obtained in the discrete-time setting, see [16,17,19,24,38,39,48, 63,65,66] and the references therein.

1.5. Organization of the article

In the next section, we prove Theorem 1.1, and discuss open-set irreducibility and aperiodicity of diffusion processes. Also, we discuss sub-geometric ergodicity of two classes of Markov processes with jumps. In Section 3, we prove Theorems 1.2 and 1.3, and again discuss sub-geometric ergodicity of Markov processes with jumps, but with respect to Wasserstein distances.

2. Ergodicity with respect to the total variation distance

In this section, we first prove Theorem 1.1. Then, we discuss open-set irreducibility and aperiodicity of diffusion processes. Finally, at the end, we discuss sub-geometric ergodicity of two classes of Markov processes with jumps.

2.1. Ergodicity of diffusion processes

We start with the proof of Theorem 1.1.

Proof of Theorem 1.1. Set $\varphi_\Lambda(t)=\varphi(t)/\Lambda$, where Λ is given in (1.2), and observe that $\varphi_\Lambda(t)$ has the same properties as $\varphi(t)$. Next, define

$$\bar{\mathcal{V}}(r):=\int_{r_0}^{r}\mathrm{e}^{-I_{x_0}(u)}\int_u^\infty\varphi_\Lambda\left(\int_{r_0}^{v}\mathrm{e}^{-I_{x_0}(w)}\,\mathrm{d}w+1\right)\frac{\mathrm{e}^{I_{x_0}(v)}}{\gamma_{x_0}(v)}\,\mathrm{d}v\,\mathrm{d}u,\qquad r\ge r_0.$$

Clearly, for $r\ge r_0$ it holds that

$$\bar{\mathcal{V}}(r)\le\int_{r_0}^{r}\mathrm{e}^{-I_{x_0}(u)}\,\mathrm{d}u, \tag{2.1}$$

and

$$\bar{\mathcal{V}}'(r)=\mathrm{e}^{-I_{x_0}(r)}\int_r^\infty\varphi_\Lambda\left(\int_{r_0}^{u}\mathrm{e}^{-I_{x_0}(v)}\,\mathrm{d}v+1\right)\frac{\mathrm{e}^{I_{x_0}(u)}}{\gamma_{x_0}(u)}\,\mathrm{d}u,$$

$$\bar{\mathcal{V}}''(r)=-\frac{\iota_{x_0}(r)}{r}\mathrm{e}^{-I_{x_0}(r)}\int_r^\infty\varphi_\Lambda\left(\int_{r_0}^{u}\mathrm{e}^{-I_{x_0}(v)}\,\mathrm{d}v+1\right)\frac{\mathrm{e}^{I_{x_0}(u)}}{\gamma_{x_0}(u)}\,\mathrm{d}u-\frac{\varphi_\Lambda(\int_{r_0}^{r}\mathrm{e}^{-I_{x_0}(u)}\,\mathrm{d}u+1)}{\gamma_{x_0}(r)}.$$

Further, fix $r_1>r_0$ and let $\mathcal{V}:\mathbb{R}^d\to[0,\infty)$, $\mathcal{V}\in C^2(\mathbb{R}^d)$, be such that $\mathcal{V}(x)=\bar{\mathcal{V}}(|x-x_0|)+1$ for $x\in\mathbb{R}^d$, $|x-x_0|\ge r_1$. Now, for $x\in\mathbb{R}^d$, $|x-x_0|\ge r_1$, we have

$$\begin{aligned}\mathcal{L}\mathcal{V}(x)&=\frac{1}{2}C_{x_0}(x)\bar{\mathcal{V}}''\big(|x-x_0|\big)+\frac{\bar{\mathcal{V}}'(|x-x_0|)}{2|x-x_0|}\big(2A(x)-C_{x_0}(x)+2B_{x_0}(x)\big)\\&\le-\frac{1}{2}\varphi_\Lambda\left(\int_{r_0}^{|x-x_0|}\mathrm{e}^{-I_{x_0}(u)}\,\mathrm{d}u+1\right)\\&\le-\frac{1}{2}\varphi_\Lambda\big(\mathcal{V}(x)\big),\end{aligned}$$

where in the final step we employed the fact that $\varphi(t)$ (that is, $\varphi_\Lambda(t)$) is non-decreasing and (2.1). Thus, we have obtained the relation in (3.11) in [15], Theorem 3.4(i), with $\phi(t)=\varphi_\Lambda(t)$, $C=\bar{B}_{r_1}(x_0)$ (the topological closure of the open ball $B_{r_1}(x_0)$), and $b=\sup_{x\in C}|\mathcal{L}V(x)|$. Now, [64], Theorems 5.1 and 7.1, together with open-set irreducibility, aperiodicity and C_b-Feller property of $\{X_t\}_{t\ge0}$, imply that $\{X_t\}_{t\ge0}$ meets the conditions of [15], Theorem 3.2, with $\Psi_1(t)=t$ and $\Psi_2(t)=1$, which concludes the proof. □

As a direct consequence of Theorem 1.1, we conclude the following.

Corollary 2.1. *If in Theorem 1.1, we take $\varphi(t) = t^\alpha$ with $\alpha \in (0,1)$, then $\{X_t\}_{t\ge0}$ is sub-geometrically ergodic with rate $t^{\alpha/(1-\alpha)}$.*

If $\varphi(t)$ is bounded, then the condition in (1.2) reduces to

$$\int_{r_0}^\infty \frac{\mathrm{e}^{I_{x_0}(u)}}{\gamma_{x_0}(u)}\,\mathrm{d}u < \infty,$$

which is exactly the condition for ergodicity obtained in [6], Theorem 3.5, (see also [70], Theorem 1.2, and [45], Chapter IV, for the one-dimensional case). By taking $\varphi(t) = t$, one expects to obtain geometric ergodicity of $\{X_t\}_{t\ge0}$. However, we cannot apply Theorem 1.1 directly since $\lim_{t\to\infty}\varphi'(t) \neq 0$. By employing analogous ideas as in Theorem 1.1, in [70], Theorem 1.3, the author proves geometric ergodicity of $\{X_t\}_{t\ge0}$ under (1.2) (with $\varphi(t) = t$) in the one-dimensional case. In what follows, we give a multi-dimensional version of this result.

Proposition 2.2. *If in Theorem 1.1 $\liminf_{t\to\infty}\varphi'(t) > 0$, then $\{X_t\}_{t\ge0}$ is geometrically ergodic.*

Proof. First, observe that since $\varphi(t)$ is differentiable and concave, $t \mapsto \varphi'(t)$ is non-increasing. Thus, since $\varphi(t)$ is also non-decreasing, there are constants $\Gamma \ge \gamma > 0$ such that

$$\gamma t - \gamma + \varphi(1) \le \varphi(t) \le \Gamma t - \Gamma + \varphi(1), \quad t \ge 1.$$

Consequently, the condition in (1.2) is equivalent to

$$\int_{r_0}^\infty \left(\int_{r_0}^u \mathrm{e}^{-I_{x_0}(v)}\,\mathrm{d}v + 1\right)\frac{\mathrm{e}^{I_{x_0}(u)}}{\gamma_{x_0}(u)}\,\mathrm{d}u < \infty$$

(recall that $\varphi(1) > 0$). Denote this constant again by Λ. Analogously as in the proof of Theorem 1.1, let

$$\bar{\mathcal{V}}(r) := \frac{1}{\Lambda}\int_{r_0}^r \mathrm{e}^{-I_{x_0}(u)}\int_u^\infty\left(\int_{r_0}^v \mathrm{e}^{-I_{x_0}(w)}\,\mathrm{d}w + 1\right)\frac{\mathrm{e}^{I_{x_0}(v)}}{\gamma_{x_0}(v)}\,\mathrm{d}v\,\mathrm{d}u, \quad r \ge r_0,$$

and, for arbitrary but fixed $r_1 > r_0$, let $\mathcal{V} : \mathbb{R}^d \to [0,\infty)$, $\mathcal{V} \in C^2(\mathbb{R}^d)$, be such that $\mathcal{V}(x) = \bar{\mathcal{V}}(|x - x_0|) + 1$ for $x \in \mathbb{R}^d$, $|x - x_0| \ge r_1$. Then, for all $x \in \mathbb{R}^d$, $|x - x_0| \ge r_1$, it holds that

$$\mathcal{L}\mathcal{V}(x) \le -\frac{1}{2\Lambda}\mathcal{V}(x), \tag{2.2}$$

which is exactly the Lyapunov equation on [50], page 529, with $c = 1/2\Lambda$, $f(x) = \mathcal{V}(x)$, $C = \bar{B}_{r_1}(x_0)$ and $b = \sup_{x\in C}|\mathcal{L}V(x)|$. The fact that C is a petite set follows from [64], Theorems 5.1 and 7.1, together with open-set irreducibility and C_b-Feller property of $\{X_t\}_{t\ge0}$. Next, from [49], Proposition 6.1, [50], Theorem 4.2, and aperiodicity it follows now that the are a petite set $\mathcal{C} \in \mathcal{B}(\mathbb{R}^d)$, $T > 0$ and a non-trivial measure $\nu_{\mathcal{C}}$ on $\mathcal{B}(\mathbb{R}^d)$, such that $\nu_{\mathcal{C}}(\mathcal{C}) > 0$ and

$$p(t,x,B) \ge \nu_{\mathcal{C}}(B), \quad x \in \mathcal{C},\ t \ge T,\ B \in \mathcal{B}(\mathbb{R}^d).$$

In particular,

$$p(t,x,\mathcal{C}) > 0, \quad x \in \mathcal{C},\ t \ge T,$$

which is exactly the definition of aperiodicity used on [18], page 1675. Finally, observe that (2.2) is also the Lyapunov equation used on [18], page 1679, with $c = 1/2\Lambda$, $C = \bar{B}_{r_1}(x_0)$ and $b = \sup_{x\in C}|\mathcal{L}V(x)|$. The assertion now follows from [18], Theorem 5.2. □

Observe that in the proof of Theorem 1.1 we did not use the fact that $\{X_t\}_{t\geq 0}$ is a unique strong solution to (1.1). All that we needed is that the martingale problem for (b, c) is well posed, which is equivalent to that (1.1) admits a unique (in distribution) weak solution (see [55], Theorem V.20.1). According to [20], Theorem 7.3.8, and [55], Theorem V.24.1, the conclusion of Theorem 1.1 remains true if, in addition to (C1)–(C3), $c(x)$ is Lipschitz continuous and there are $\Gamma > 0$ and $\gamma \geq 1$ such that

$$\gamma^{-1}|y|^2 \leq \langle y, c(x)y\rangle \leq \gamma|y|^2 \quad \text{and} \quad \big|b(x)\big|^2 + \big\|c(x)\big\|_{\mathrm{HS}}^2 \leq \Gamma\big(1+|x|^2\big), \quad x, y \in \mathbb{R}^d. \tag{2.3}$$

Moreover, under the above assumptions, [55], Theorem V.24.1, states that $\{X_t\}_{t\geq 0}$ is a Feller and strong Feller process. Recall, strong Feller property means that the corresponding semigroup maps $B_b(\mathbb{R}^d)$ to $C_b(\mathbb{R}^d)$. Also, (2.3), together with (C1)–(C3) and Lipschitz continuity of $c(x)$, implies open-set irreducibility and aperiodicity of $\{X_t\}_{t\geq 0}$ (see [61], Remark 4.3).

In the following theorem we discuss open-set irreducibility and aperiodicity of $\{X_t\}_{t\geq 0}$ in the situation when $c(x)$ is not necessarily Lipschitz continuous and uniformly elliptic.

Theorem 2.3. *Assume* (C1)–(C3). *Further, assume that there are $x_0 \in \mathbb{R}^d$ and $r_0 > 0$, such that*

(i) *there are $\delta, \Gamma, \gamma > 0$, such that for all $x, y \in B_{r_0}(x_0)$ we have that*

$$\big|b(x) - b(y)\big| + \big\|c(x) - c(y)\big\|_{\mathrm{HS}} \leq \Gamma|x-y|^\delta \quad \textit{and} \quad \big\langle y, c(x)y\big\rangle \geq \gamma|y|^2;$$

(ii) $\mathbb{P}^x(\tau_{B_{r_0}(x_0)} < \infty) > 0$ *for all $x \in \mathbb{R}^d$, where $\tau_B := \inf\{t \geq 0 : X_t \in B\}$ is the first hitting time of a set $B \subseteq \mathbb{R}^d$.*

Then, $\{X_t\}_{t\geq 0}$ is open-set irreducible and aperiodic.

Proof. Due to [20], Theorems 7.3.6 and 7.3.7, there is a strictly positive function $q(t, x, y)$ on $(0, \infty) \times \bar{B}_{r_0}(x_0) \times \bar{B}_{r_0}(x_0)$, jointly continuous in t, x and y, and twice continuously differentiable in x on $B_{r_0}(x_0)$, satisfying

$$\mathbb{E}^x\big(f(X_t), \tau_{\bar{B}^c_{r_0}(x_0)} > t\big) = \int_{B_{r_0}(x_0)} q(t, x, y) f(y)\,\mathrm{d}y, \quad t > 0,\ x \in B_{r_0}(x_0),\ f \in C_b\big(\mathbb{R}^d\big),$$

where $\tau_{\bar{B}^c_{r_0}(x_0)} := \inf\{t \geq 0 : X_t \in \bar{B}^c_{r_0}(x_0)\}$. Clearly, by employing dominated convergence theorem, the above relation holds also for $\mathbb{1}_O$, for any open set $O \subseteq B_{r_0}(x_0)$. Denote by $\mathcal{D}$ the class of all $B \in \mathcal{B}(B_{r_0}(x_0))$ (the Borel σ-algebra on $B_{r_0}(x_0)$) such that

$$\mathbb{P}^x(X_t \in B, \tau_{\bar{B}^c_{r_0}(x_0)} > t) = \int_B q(t, x, y)\,\mathrm{d}y, \quad t > 0,\ x \in B_{r_0}(x_0).$$

Clearly, $\mathcal{D}$ contains the π-system of open rectangles in $B_{r_0}(x_0)$, and forms a λ-system. Hence, by employing Dynkin's π-λ theorem we conclude that $\mathcal{D} = \mathcal{B}(B_{r_0}(x_0))$. Consequently, for any $t > 0$, $x \in B_{r_0}(x_0)$ and $B \in \mathcal{B}(\mathbb{R}^d)$ we have that

$$p(t, x, B) \geq \int_{B\cap B_{r_0}(x_0)} q(t, x, y)\,\mathrm{d}y.$$

Set now $\phi(\cdot) := \lambda(\cdot \cap B_{r_0}(x_0))$, where λ stands for the Lebesgue measure on $\mathbb{R}^d$. Then, ϕ is a σ-finite measure whose support has a non-empty interior.

Let us now show that $\{X_t\}_{t\geq 0}$ is ϕ-irreducible. Let $x \in B^c_{r_0}(x_0)$ (for $x \in B_{r_0}(x_0)$ the assertion is obvious) and $B \in \mathcal{B}(\mathbb{R}^d)$, $\phi(B) > 0$, be arbitrary. For all $s > 0$ we have

$$\begin{aligned}\int_0^\infty p(t,x,B)\,\mathrm{d}t &\geq \int_s^\infty p(t,x,B)\,\mathrm{d}t\\ &= \int_s^\infty \int_{\mathbb{R}^d} p(t-s,x,\mathrm{d}y)p(s,y,B)\,\mathrm{d}t\\ &\geq \int_s^\infty \int_{B_{r_0}(x_0)} p(t-s,x,\mathrm{d}y)p(s,y,B)\,\mathrm{d}t\\ &= \int_{B_{r_0}(x_0)} p(s,y,B) \int_s^\infty p(t-s,x,\mathrm{d}y)\,\mathrm{d}t.\end{aligned}$$

The assertion now follows from the fact that $p(s,y,B) > 0$ for $y \in B_{r_0}(x_0)$, and

$$\int_s^\infty p\big(t-s,x,B_{r_0}(x_0)\big)\,\mathrm{d}t = \int_0^\infty p\big(t,x,B_{r_0}(x_0)\big)\,\mathrm{d}t = \mathbb{E}^x\left[\int_0^\infty \mathbb{1}_{\{X_t \in B_{r_0}(x_0)\}}\,\mathrm{d}t\right] > 0,$$

since $\{X_t\}_{t\geq 0}$ has continuous sample paths, $B_{r_0}(x_0)$ is an open set and, by assumption, $\mathbb{P}^x(\tau_{B_{r_0}(x_0)} < \infty) > 0$ for every $x \in \mathbb{R}^d$.

Finally, let us prove that $\{X_t\}_{t\geq 0}$ is aperiodic. We show that

$$\sum_{n=1}^\infty p(n,x,B) > 0, \quad x \in \mathbb{R}^d,$$

whenever $\phi(B) > 0$, $B \in \mathcal{B}(\mathbb{R}^d)$. Again, for $x \in B_{r_0}(x_0)$ the relation obviously holds. For $x \in B^c_{r_0}(x_0)$ and $B \in \mathcal{B}(\mathbb{R}^d)$, $\phi(B) > 0$, we have that

$$\sum_{n=1}^\infty p(n,x,B) \geq \int_{B_{r_0}(x_0)} \sum_{n=1}^\infty p(n-t,x,\mathrm{d}y)p(t,y,B), \quad t \in (0,1).$$

Since $p(t,y,B) > 0$ for $y \in B_{r_0}(x_0)$, it suffices to show that

$$\sum_{n=1}^\infty p\big(n-t,x,B_{r_0}(x_0)\big) \geq \mathbb{P}^x\left(\bigcup_{n=1}^\infty \big\{X_{n-t} \in B_{r_0}(x_0)\big\}\right) > 0$$

for some $t \in (0,1)$. Assume this is not the case, that is,

$$\mathbb{P}^x\left(\bigcup_{n=1}^\infty \big\{X_{n-t} \in B_{r_0}(x_0)\big\}\right) = 0, \quad t \in (0,1).$$

This, in particular, implies that

$$\mathbb{P}^x\Bigg(\bigcup_{q\in\mathbb{Q}_+\setminus\mathbb{Z}_+} \big\{X_q \in B_{r_0}(x_0)\big\}\Bigg) = 0,$$

which is impossible since $\{X_t\}_{t\geq 0}$ has continuous sample paths, $B_{r_0}(x_0)$ is an open set and $\mathbb{P}^x(\tau_{B_{r_0}(x_0)} < \infty) > 0$ for every $x \in \mathbb{R}^d$. Thus,

$$\sum_{n=1}^{\infty} p(n, x, B) > 0, \quad x \in \mathbb{R}^d,$$

whenever $\phi(B) > 0$, which concludes the proof. □

In the following proposition, we give a sufficient condition for the second assumption in Theorem 2.3 to hold.

Proposition 2.4. *Assume* (C1)–(C3). *Then for any* $x_0 \in \mathbb{R}^d$ *and* $r_0 > 0$, *provided that* $c(x)$ *is positive definite for all* $x \in \mathbb{R}^d$, $|x - x_0| \geq r_0$, *it holds that*

$$\mathbb{P}^x(\tau_{B_{r_0}(x_0)} < \infty) > 0, \quad x \in \mathbb{R}^d.$$

Proof. Let $0 < \varepsilon < r_0$, and let

$$\bar{\mathcal{V}}(r) := \int_{r_0-\varepsilon}^{r} \mathrm{e}^{-I_{x_0}(u)}\,\mathrm{d}u, \quad r \geq r_0 - \varepsilon.$$

Then, for $r > r_0 - \varepsilon$ we have

$$\bar{\mathcal{V}}'(r) = \mathrm{e}^{-I_{x_0}(r)} > 0 \quad \text{and} \quad \bar{\mathcal{V}}''(r) = -\frac{\bar{\mathcal{V}}'(r)}{r}\iota_{x_0}(r).$$

Further, let $\mathcal{V} : \mathbb{R}^d \to [0, \infty)$, $\mathcal{V} \in C^2(\mathbb{R}^d)$, be such that $\mathcal{V}(x) = \bar{\mathcal{V}}(|x - x_0|)$ for $x \in \mathbb{R}^d$, $|x - x_0| \geq r_0$. Now, for $x \in \mathbb{R}^d$, $|x - x_0| \geq r_0$, we have

$$\begin{aligned} 2\mathcal{L}\mathcal{V}(x) &= C_{x_0}(x)\bar{\mathcal{V}}''\big(|x - x_0|\big) + \frac{\bar{\mathcal{V}}'(|x - x_0|)}{|x - x_0|}\big(2A(x) - C_{x_0}(x) + 2B_{x_0}(x)\big) \\ &= \frac{\bar{\mathcal{V}}'(|x - x_0|)}{|x - x_0|}\big(2A(x) - C_{x_0}(x) + 2B_{x_0}(x) - C_{x_0}(x)\iota\big(|x - x_0|\big)\big) \\ &\leq 0. \end{aligned}$$

Further, as we have already discussed, for every $x \in \mathbb{R}^d$ the process

$$\mathcal{V}(X_t) - \mathcal{V}(X_0) - \int_0^t \mathcal{L}\mathcal{V}(X_s)\,\mathrm{d}s, \quad t \geq 0,$$

is a local $\mathbb{P}^x$-martingale. For $n \in \mathbb{N}$, define $\tau_n := \tau_{B_n^c(x_0)}$. Clearly, τ_n, $n \in \mathbb{N}$, are stopping times such that (due to non-explosivity of $\{X_t\}_{t\geq 0}$) $\tau_n \to \infty$ $\mathbb{P}^x$-a.s. as $n \to \infty$ for all $x \in \mathbb{R}^d$. Hence, the processes

$$\mathcal{V}(X_{t\wedge\tau_n}) - \mathcal{V}(X_0) - \int_0^{t\wedge\tau_n} \mathcal{L}\mathcal{V}(X_s)\,\mathrm{d}s, \quad t \geq 0,\ n \in \mathbb{N},$$

are $\mathbb{P}^x$-martingales. Now, for $x \in \mathbb{R}^d$, $|x - x_0| \geq r_0$, we have

$$\begin{aligned} 2\mathbb{E}^x\big[\bar{\mathcal{V}}\big(|X_{t\wedge\tau_n\wedge\tau_{B_{r_0}(x_0)}} - x_0|\big)\big] - 2\bar{\mathcal{V}}\big(|x - x_0|\big) &= 2\mathbb{E}^x\big[\mathcal{V}(X_{t\wedge\tau_n\wedge\tau_{B_{r_0}(x_0)}})\big] - 2\mathbb{E}^x\big[\mathcal{V}(X_0)\big] \\ &= \mathbb{E}^x \int_0^{t\wedge\tau_n\wedge\tau_{B_{r_0}(x_0)}} 2\mathcal{L}\mathcal{V}(X_s)\,\mathrm{d}s \\ &\leq 0, \end{aligned}$$

that is,

$$\mathbb{E}^x\big[\bar{\mathcal{V}}\big(|X_{t\wedge\tau_n\wedge\tau_{B_{r_0}(x_0)}} - x_0|\big)\big] \leq \bar{\mathcal{V}}\big(|x - x_0|\big).$$

Thus,

$$\mathbb{E}^x\big[\bar{\mathcal{V}}\big(|X_{t\wedge\tau_n} - x_0|\big)\mathbb{1}_{\{\tau_{B_{r_0}(x_0)}>\tau_n\}}\big] \leq \bar{\mathcal{V}}\big(|x - x_0|\big), \quad x \in \mathbb{R}^d,\ |x - x_0| \geq r_0.$$

By letting $t \to \infty$ Fatou's lemma implies

$$\bar{\mathcal{V}}(n)\mathbb{P}^x(\tau_{B_{r_0}(x_0)} > \tau_n) \leq \bar{\mathcal{V}}\big(|x - x_0|\big), \quad x \in \mathbb{R}^d,\ |x - x_0| \geq r_0.$$

Consequently, by letting $n \to \infty$, we conclude

$$\mathbb{P}^x(\tau_{B_{r_0}(x_0)} = \infty) \leq \frac{\bar{\mathcal{V}}(|x - x_0|)}{\bar{\mathcal{V}}(\infty)} < 1, \quad x \in \mathbb{R}^d,\ |x - x_0| \geq r_0,$$

that is, $\mathbb{P}^x(\tau_{B_{r_0}(x_0)} < \infty) > 0$ for all $x \in \mathbb{R}^d$. □

As we have already commented, in [15], Theorem 5.4, [25], page 1581, [38], Theorem 1.30, [39], Theorem 3.3.6, [56], Theorem 3.3(iv), [65], Theorem 6, and [66], Theorem 6, it has been shown that a diffusion process $\{X_t\}_{t\geq 0}$ (satisfying the assumptions from Corollary 2.1) is sub-geometrically ergodic with rate $t^{\alpha/(1-\alpha)}$, $0 < \alpha < 1$, if there are $\gamma > 0$, $\Gamma > 0$ and $r_0 \geq 0$, such that (1.7) holds true. A simple example which satisfies the relation in (1.2) but not the one in (1.7) is the following.

Example 2.5. Let $\sigma(x) \equiv 1$, and let $b(x)$ be locally Lipschitz continuous and such that $b(x) = -\operatorname{sgn}(x)(\cos x + 1)$ for all $|x|$ large enough, where

$$\operatorname{sgn}(x) := \begin{cases} 1, & x \geq 0, \\ -1, & x < 0. \end{cases}$$

Clearly, $b(x)$ and $\sigma(x)$ satisfy (C1)–(C3) and define, through (1.1), an open-set irreducible and aperiodic diffusion process $\{X_t\}_{t\geq 0}$. The condition in (1.2) now reduces to showing that there is $r_0 \geq 0$ such that

$$\int_{r_0}^{\infty}\left(\int_{r_0}^{u} \mathrm{e}^{2\sin v + 2v} + 1\right)^{\alpha} \mathrm{e}^{-2\sin u - 2u}\,\mathrm{d}u < \infty,$$

which can be obviously obtained for any $0 < \alpha < 1$. On the other hand, the condition in (1.7) is equivalent to showing that there are $\gamma > 0$, $\Gamma > 0$ and $r_0 \geq 0$, such that

$$\frac{\gamma - 1}{2} - x\operatorname{sgn}(x)(\cos x + 1) \leq -\Gamma|x|^{\gamma\alpha-\gamma+2}, \quad |x| \geq r_0.$$

However, observe that in the points of the form $x = (2k+1)\pi$, $k \in \mathbb{Z}$, the second term on the left-hand side in the above inequality vanishes. Thus, we conclude that it is necessary that $0 < \gamma < 1$ and $\gamma\alpha - \gamma + 2 < 0$, which is impossible. Note also that if we take $b(x)$ to be locally Lipschitz continuous and such that $b(x) = -\operatorname{sgn}(x)(\cos x + \varrho)$ for all $|x|$ large enough, where $\varrho > 0$, then we again easily conclude that (1.2) holds for any $0 < \alpha < 1$. On the other hand, by the same reasoning as above, (1.7) can never hold. Observe that for $0 < \varrho < 1$ the drift function generates a region in which the process is "pushed towards infinity" (set of points for which $\operatorname{sgn}(x)b(x) > 0$). The condition in (1.2) says that this region is small compared to the region in which the process is "pushed towards the center of the state space" (set of points for which $\operatorname{sgn}(x)b(x) < 0$) and which is responsible for the ergodic behavior.

Proposition 2.6. *Assume* (C1)–(C3). *Further, assume that* $\gamma < 2/(1-\alpha)$ *and there are* $r_0 \geq 0$ *and* $\Delta \geq 1$, *such that* $\Delta^{-1} \leq C_0(x) \leq \Delta$ *for all* $|x| \geq r_0$. *Then*, (1.2) (*with* $x_0 = 0$) *is a consequence of* (1.7).

Proof. We have that

$$\iota_0(r) = \sup_{|x|=r} \frac{2(A(x) - (1-\frac{\gamma}{2})C_0(x) + B_0(x)) + (1-\gamma)C_0(x)}{C_0(x)} \leq -\frac{2\Gamma}{\Delta} r^{\gamma\alpha-\gamma+2} + 1 - \gamma$$

for all $r \geq r_1$, for some $r_1 \geq r_0$ large enough. Thus, there are $\Gamma_1 > 0$ and $r_2 \geq r_1$, such that

$$\iota_0(r) \leq -\Gamma_1 r^{\gamma\alpha-\gamma+2}, \quad r \geq r_2.$$

This automatically implies that there are $\Gamma_2 > 0$ and $r_3 \geq r_2$, such that

$$I_0(r) \leq -\Gamma_2 r^{\gamma\alpha-\gamma+2}, \quad r \geq r_3.$$

Now, by employing L'Hospital's rule (here we use the assumption $\gamma < 2/(1-\alpha)$), we have that

$$\lim_{u\to\infty} \frac{(\int_{r_3}^u \mathrm{e}^{-I_0(v)}\,\mathrm{d}v + 1)}{\mathrm{e}^{-I_0(u)}} = 0.$$

Hence, there is $r_4 \geq r_3$ such that

$$\int_{r_3}^u \mathrm{e}^{-I_0(v)}\,\mathrm{d}v + 1 \leq \mathrm{e}^{-I_0(u)} u \geq r_4.$$

Finally, we conclude

$$\int_{r_4}^\infty \left(\int_{r_4}^u \mathrm{e}^{-I_0(v)}\,\mathrm{d}v + 1\right)^\alpha \mathrm{e}^{I_0(u)}\,\mathrm{d}u \leq \int_{r_4}^\infty \mathrm{e}^{(1-\alpha)I_0(u)}\,\mathrm{d}u < \infty,$$

which proves the assertion. □

In the following proposition, which generalizes [11], Lemma 1.2, to the sub-geometric case, we give sufficient conditions ensuring (1.2).

Proposition 2.7. *Let* $c \geq 0$, *and let* $\rho(t)$ *be a non-negative and non-decreasing differentiable function defined on* $[0,\infty)$. *Further, let* $f(r)$ *and* $g(r)$ *be non-negative Borel measurable functions, also defined* *on* $[0,\infty)$, *satisfying*

$$\Delta := \sup_{r\geq r_0}\rho\left(\int_{r_0}^{r} g(u)\,\mathrm{d}u + c\right)^{1+\beta}\int_{r}^{\infty} f(u)\,\mathrm{d}u < \infty \tag{2.4}$$

for some $r_0 \geq 0$ *and* $\beta \geq 0$. *Then,*

(i) *if* $\beta > 0$,

$$\int_{r}^{\infty}\rho\left(\int_{r_0}^{u} g(v)\,\mathrm{d}v + c\right) f(u)\,\mathrm{d}u \leq \frac{\Delta(1+\beta)}{\beta}\rho\left(\int_{r_0}^{r} g(u)\,\mathrm{d}u + c\right)^{-\beta}, \quad r \geq r_0.$$

(ii) *if* $\beta = 0$, *and* $\int_{r_0}^{\infty} g(r)\,\mathrm{d}r < \infty$ *or* $\rho(t)$ *is bounded,*

$$\int_{r}^{\infty}\rho\left(\int_{r_0}^{u} g(v)\,\mathrm{d}v + c\right) f(u)\,\mathrm{d}u \leq \Delta + \Delta\ln\frac{\rho(\int_{r_0}^{\infty} g(u)\,\mathrm{d}u + c)}{\rho(\int_{r_0}^{r} g(u)\,\mathrm{d}u + c)}, \quad r \geq r_0.$$

Proof. Set $F(r) = \int_{r}^{\infty} f(u)\,\mathrm{d}u$, $r \geq r_0$. Then, by assumption,

$$F(r) \leq \Delta\rho\left(\int_{r_0}^{r} g(u)\,\mathrm{d}u + c\right)^{-1-\beta}, \quad r \geq r_0.$$

Consequently, for $r \geq r_0$, we have that

$$\begin{aligned}
&\int_{r}^{\infty}\rho\left(\int_{r_0}^{u} g(v)\,\mathrm{d}v + c\right) f(u)\,\mathrm{d}u\\
&= -\int_{r}^{\infty}\rho\left(\int_{r_0}^{u} g(v)\,\mathrm{d}v + c\right)\mathrm{d}F(u)\\
&\leq \rho\left(\int_{r_0}^{r} g(u)\,\mathrm{d}u + c\right)F(r) + \int_{r}^{\infty}\rho'\left(\int_{r_0}^{u} g(v)\,\mathrm{d}v + c\right)g(u)F(u)\,\mathrm{d}u\\
&\leq \Delta\rho\left(\int_{r_0}^{r} g(u)\,\mathrm{d}u + c\right)^{-\beta} + \Delta\int_{r}^{\infty}\rho'\left(\int_{r_0}^{u} g(v)\,\mathrm{d}v + c\right)g(u)\rho\left(\int_{r_0}^{u} g(v)\,\mathrm{d}v + c\right)^{-1-\beta}\mathrm{d}u.
\end{aligned}$$

Now, under the assumption in (i) we have that

$$\begin{aligned}
&\int_{r}^{\infty}\rho\left(\int_{r_0}^{u} g(v)\,\mathrm{d}v + c\right) f(u)\,\mathrm{d}u\\
&\leq \Delta\rho\left(\int_{r_0}^{r} g(u)\,\mathrm{d}u + c\right)^{-\beta} - \frac{\Delta}{\beta}\int_{r}^{\infty}\mathrm{d}\rho\left(\int_{r_0}^{u} g(v)\,\mathrm{d}v + c\right)^{-\beta}\\
&\leq \Delta\rho\left(\int_{r_0}^{r} g(u)\,\mathrm{d}u + c\right)^{-\beta} + \frac{\Delta}{\beta}\rho\left(\int_{r_0}^{r} g(u)\,\mathrm{d}u + c\right)^{-\beta}\\
&= \frac{\Delta(1+\beta)}{\beta}\rho\left(\int_{r_0}^{r} g(u)\,\mathrm{d}u + c\right)^{-\beta},
\end{aligned}$$

where in the second step we employed integration by parts formula. On the other hand, under the assumptions in (ii),

$$\int_r^\infty \rho\Big(\int_{r_0}^u g(v)\,\mathrm{d}v + c\Big) f(u)\,\mathrm{d}u \le \Delta + \Delta \int_r^\infty d\ln\Big(\rho\Big(\int_{r_0}^u g(v)\,\mathrm{d}v + c\Big)\Big)$$
$$= \Delta + \Delta \ln \frac{\rho(\int_{r_0}^\infty g(u)\,\mathrm{d}u + c)}{\rho(\int_{r_0}^r g(u)\,\mathrm{d}u + c)},$$

which concludes the proof. □

As a direct consequence of the proposition we see that (1.2) holds true if

$$\sup_{r\ge r_0} \varphi\Big(\int_{r_0}^r \mathrm{e}^{-I_{x_0}(u)}\,\mathrm{d}u + 1\Big)^{1+\beta} \int_r^\infty \frac{\mathrm{e}^{I_{x_0}(u)}}{\gamma_{x_0}(u)}\,\mathrm{d}u < \infty$$

for some $\beta > 0$.

2.2. Ergodicity of Markov processes with jumps

In this subsection, as an application of Theorem 1.1, we discuss sub-geometric ergodicity of a class of Markov processes with jumps. First, we consider jump-diffusion processes generated by operator of the form

$$\mathcal{L}f(x) = \langle b(x), \nabla f(x)\rangle + \frac{1}{2}\operatorname{Tr} c(x)\nabla^2 f(x) + \int_{\mathbb{R}^d} \big(f(y+x) - f(x) - \langle y, \nabla f(x)\rangle \mathbb{1}_{B_1(0)}(y)\big)\nu(x,\mathrm{d}y), \tag{2.5}$$

where $b(x)$ is an $\mathbb{R}^d$-valued Borel measurable function, $c(x)$ is a symmetric non-negative definite $d\times d$ matrix-valued Borel measurable function, and $\nu(x,\mathrm{d}y)$ is a non-negative Borel kernel on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, called the Lévy kernel, satisfying

$$\nu\big(x,\{0\}\big) = 0, \quad \text{and} \quad \int_{\mathbb{R}^d} \big(1\wedge |y|^2\big)\nu(x,\mathrm{d}y) < \infty, \quad x\in\mathbb{R}^d.$$

Clearly, if $\nu(x,\mathrm{d}y)$ is a null-measure, then $\mathcal{L}$ becomes a diffusion operator. In the sequel, we assume that

(A1) there is a càdlàg Markov process $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\ge0}, \{\theta_t\}_{t\ge0}, \{X_t\}_{t\ge0}, \{\mathbb{P}^x\}_{x\in\mathbb{R}^d})$, denoted by $\{X_t\}_{t\ge0}$ in the sequel, which we call jump-diffusion process, such that for every $f\in C^2(\mathbb{R}^d)$ the process

$$f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s)\,\mathrm{d}s, \quad t\ge 0,$$

is a $\mathbb{P}^x$ local martingale for all $x\in\mathbb{R}^d$ under the natural filtration;

(A2) the process $\{X_t\}_{t\ge0}$ satisfies the C_b-Feller property;

(A3) the process $\{X_t\}_{t\ge0}$ is open-set irreducible and aperiodic.

Here, $C_b^2(\mathbb{R}^d)$ denotes the space of twice continuously differentiable functions with bounded derivatives. Let us remark that (A1) always holds for the infinitesimal generator $(\mathcal{A}, \mathcal{D}_{\mathcal{A}})$ of $\{X_t\}_{t\geq 0}$ (see [23], Theorem 2.2.13 and Proposition 4.1.7). We refer the readers to [9] for conditions, in terms of $b(x)$, $c(x)$ and $\nu(x, \mathrm{d}y)$, ensuring (A1) and (A2). Open-set irreducibility and aperiodicity of jump-diffusion processes is a very well-studied topic in the literature. In particular, we refer the readers to [36] and [37] for the case of so-called stable-like processes, to [34,35,41], [51], Remark 3.3, [57], Theorem 2.6, and [62] for the case of jump-diffusion processes with bounded coefficients, and to [4, 5,30,33,46,47] and [52,53] for the case of a class of jump-diffusion processes obtained as a solution to certain jump-type SDEs. According to [64], Theorem 3.2, $\{X_t\}_{t\geq 0}$ will be open-set irreducible and aperiodic if it is strong Feller (actually it suffices to assume that $\{X_t\}_{t\geq 0}$ is a T-model in the sense of [64], which is a certain weak version of the strong Feller property) and $\mathbb{P}^x(X_t \in O) > 0$ for every $t > 0$, $x \in \mathbb{R}^d$ and non-empty open set $O \subseteq \mathbb{R}^d$. If $b(x)$ is continuous and bounded, $c(x)$ continuous, bounded and positive definite, $x \mapsto \int_B (1 \wedge |y|^2)\nu(x, \mathrm{d}y)$ continuous and bounded for any $B \in \mathcal{B}(\mathbb{R}^d)$, and

$$(x, \xi) \mapsto i\langle \xi, b(x)\rangle + \frac{1}{2}\langle \xi, c(x)\xi\rangle + \int_{\mathbb{R}^d} \left(1 - e^{i\langle \xi, y\rangle} + i\langle \xi, y\rangle \mathbb{1}_{B_1(0)}(y)\right)\nu(x, \mathrm{d}y)$$

continuous, then

(i) there is a unique non-explosive strong Markov process $\{X_t\}_{t\geq 0}$ with infinitesimal generator $(\mathcal{A}, \mathcal{D}_{\mathcal{A}})$ such that $C_c^\infty(\mathbb{R}^d) \subseteq \mathcal{D}_{\mathcal{A}}$, and $\mathcal{A}|_{C_c^\infty(\mathbb{R}^d)}$ takes the form in (2.5), where $C_c^\infty(\mathbb{R}^d)$ stands for the space of smooth functions with compact support;
(ii) the operator $\mathcal{L} := \mathcal{A}|_{C_c^\infty(\mathbb{R}^d)}$ satisfies (A1);
(iii) the semigroup of $\{X_t\}_{t\geq 0}$ satisfies the Feller and strong Feller property,

(see [9], Theorems 2.37, 3.23, 3.24, 3.25, and [62], Remark after Theorem 4.3). Finally, we also assume

(A4) there is $\rho > 0$ such that $\nu(x, B^c_{|x|}(-x)) = 0$ and $\int_{B_1(0)} |y|\nu(x, \mathrm{d}y) < \infty$ for all $x \in \mathbb{R}^d$, $|x| \geq \rho$;
(A5) the functions $b(x)$, $c(x)$ and $x \mapsto \int_{B_1(0)} y\nu(x, \mathrm{d}y)$ are continuous on $B^c_\rho(0)$.

Assumption (A4) means that when $\{X_t\}_{t\geq 0}$ is far away from the center of the state space, it admits bounded jumps only, with maximal intensity equal twice the distance to the origin. Also, with each jump, it comes closer to the center of the state space.

In the following theorem we give sufficient conditions for sub-geometric ergodicity of a class of jump-diffusion processes satisfying (A1)–(A5). We use the same notation as in Theorem 1.1, with

$$B_{x_0}(x) := \left\langle x - x_0, b(x) - \int_{B_1(0)} y\nu(x, \mathrm{d}y)\right\rangle, \quad x \in \mathbb{R}^d.$$

Theorem 2.8. *Let $\{X_t\}_{t\geq 0}$ be an open-set irreducible and aperiodic jump-diffusion process with coefficients $b(x)$, $c(x)$ and $\nu(x, \mathrm{d}y)$, satisfying* (A1)–(A5). *Further, let $\varphi : [1, \infty) \longrightarrow (0, \infty)$ be a non-decreasing, differentiable and concave function satisfying $\lim_{t\to\infty} \varphi'(t) = 0$ and the relation in* (1.2) *for some $x_0 \in \mathbb{R}^d$ and $r_0 \geq \rho + |x_0|$, and assume that $c(x)$ is positive definite for all $x \in \mathbb{R}^d$, $|x - x_0| \geq r_0$. Then, $\{X_t\}_{t\geq 0}$ admits a unique invariant $\pi \in \mathcal{P}$ such that*

$$\lim_{t\to\infty} \varphi\left(\Phi^{-1}(t)\right)\|\delta_x P_t - \pi\|_{\mathrm{TV}} = 0, \quad x \in \mathbb{R}^d,$$

where $\Phi(t)$ is as in Theorem 1.1.

Proof. We proceed as in the proof of Theorem 1.1. Define

$$\bar{\mathcal{V}}(r) := \int_{r_0}^{r} \mathrm{e}^{-I_{x_0}(u)} \int_{u}^{\infty} \varphi_\Lambda\left(\int_{r_0}^{v} \mathrm{e}^{-I_{x_0}(w)}\,\mathrm{d}w + 1\right) \frac{\mathrm{e}^{I_{x_0}(v)}}{\gamma_{x_0}(v)}\,\mathrm{d}v\,\mathrm{d}u, \quad r \ge r_0,$$

where $\varphi_\Lambda(t) = \varphi(t)/\Lambda$. Clearly,

$$\bar{\mathcal{V}}(r) \le \int_{r_0}^{r} \mathrm{e}^{-I_{x_0}(u)}\,\mathrm{d}u, \quad r \ge r_0, \tag{2.6}$$

and, because of (A5), $\bar{\mathcal{V}}(r)$ is twice continuously differentiable on (r_0, ∞). Further, for arbitrary, but fixed, $r_1 > r_0$ let $\tilde{\mathcal{V}} : [0, \infty) \to [0, \infty)$ be non-decreasing on $[0, \infty)$, $\tilde{\mathcal{V}}(r) = \bar{\mathcal{V}}(r)$ on $[r_1, \infty)$, and such that $\mathcal{V}(x) := \tilde{\mathcal{V}}(|x - x_0|) + 1$ is twice continuously differentiable on $\mathbb{R}^d$. Now, because of (A1) and (A4), $\mathcal{L}\mathcal{V}(x)$ is well defined and the process

$$\mathcal{V}(X_t) - \mathcal{V}(X_0) - \int_0^t \mathcal{L}\mathcal{V}(X_s)\,\mathrm{d}s \quad t \ge 0,$$

is a local martingale. For $x \in \mathbb{R}^d$, $|x| \ge r_1$, we have that

$$\begin{aligned}
\mathcal{L}\mathcal{V}(x) &= \frac{1}{2} C_{x_0}(x) \bar{\mathcal{V}}''\big(|x - x_0|\big) + \frac{\bar{\mathcal{V}}'(|x - x_0|)}{2|x - x_0|}\big(2A(x) - C_{x_0}(x) + 2\langle x - x_0, b(x)\rangle\big) \\
&\quad + \int_{\mathbb{R}^d} \big(\mathcal{V}(y + x) - \mathcal{V}(x) - \langle y, \nabla\mathcal{V}(x)\rangle \mathbb{1}_{B_1(0)}(y)\big)\nu(x, \mathrm{d}y) \\
&\le \frac{1}{2} C_{x_0}(x) \bar{\mathcal{V}}''\big(|x - x_0|\big) + \frac{\bar{\mathcal{V}}'(|x - x_0|)}{2|x - x_0|}\big(2A(x) - C_{x_0}(x) + 2B_{x_0}(x)\big) \\
&\le -\frac{1}{2}\varphi_\Lambda\left(\int_{r_0}^{|x - x_0|} \mathrm{e}^{-I_{x_0}(u)}\,\mathrm{d}u + 1\right) \\
&\le -\frac{1}{2}\varphi_\Lambda\big(\mathcal{V}(x)\big),
\end{aligned}$$

where in the second step we used (A4) and properties of $\mathcal{V}(x)$ (that is, $\tilde{\mathcal{V}}(r)$), and the final step follows from (2.6). Finally, because of (A2) and (A5), as in the proof of Theorem 1.1, we are again in a position to apply [15], Theorems 3.2 and 3.4(i), and [64], Theorems 5.1 and 7.1, which concludes the proof. □

Let us now give several remarks.

Remark 2.9.

(a) If $2A(x) - C_{x_0}(x) + 2B_{x_0}(x) \le 0$ for all $x \in \mathbb{R}^d$, $|x - x_0| \ge r_0$, then we can replace $\gamma_{x_0}(r)$ and $\iota_{x_0}(r)$ by

$$\begin{aligned}
\gamma_{x_0}(r) &= \inf_{|x - x_0| = r} N_{x_0}(x), \quad r > 0, \\
\iota_{x_0}(r) &= \sup_{|x - x_0| = r} \frac{2A(x) - C_{x_0}(x) + 2B_{x_0}(x)}{N_{x_0}(x)}, \quad r > 0,
\end{aligned}$$

where

$$N_{x_0}(x) = \frac{\langle x - x_0, (c(x) + n(x))(x - x_0)\rangle}{|x - x_0|^2}, \quad x \in \mathbb{R}^d \setminus \{0\},$$

and $n(x) = (n_{ij}(x))_{i,j=1,\dots,d}$ with $n_{ij}(x) = \int_{B_1(0)} y_i y_j \nu(x, dy)$. Also, in this situation, the requirement in Theorem 2.8 that $c(x)$ is positive definite for all $x \in \mathbb{R}^d$, $|x - x_0| \geq r_0$, can be replaced by the requirement that $c(x) + n(x)$ is positive definite for all $x \in \mathbb{R}^d$, $|x - x_0| \geq r_0$.

(b) If $\varphi(t)$ is bounded, then (1.2) reads

$$\int_{r_0}^{\infty} \frac{e^{I_{x_0}(u)}}{\gamma_{x_0}(u)} du < \infty,$$

and gives a condition for ergodicity (see [70], Theorem 1.2, for the one-dimensional case).

(c) If in Theorem 2.8 $\liminf_{t\to\infty} \varphi'(t) > 0$ then, as in Proposition 2.2, we conclude that $\{X_t\}_{t\geq 0}$ is geometrically ergodic (see also [70], Theorem 1.3, for the one-dimensional case).

Let us now give an example satisfying conditions from Theorem 2.8.

Example 2.10 (Lévy-driven SDEs). Let $\{Y_t\}_{t\geq 0}$ be an n-dimensional Lévy process, and let $\Phi : \mathbb{R}^d \to \mathbb{R}^{d\times n}$ be bounded and locally Lipschitz continuous. Then, in [59], Theorems 3.1 and 3.5, and Corollary 3.3, (see also [9], Theorem 3.8) it has been shown that the SDE

$$dX_t = \Phi(X_{t-})\, dY_t, \quad X_0 = x \in \mathbb{R}^d, \tag{2.7}$$

admits a unique strong solution which is a non-explosive strong Markov process whose semigroup satisfies the Feller and C_b-Feller property (thus (A2) holds true). Also, it has been shown that $\{X_t\}_{t\geq 0}$ satisfies (A1) with certain coefficients $b(x)$, $c(x)$ and $\nu(x, dy)$, which in a special case we give below. Observe that the following SDE is a special case of (2.7),

$$dX_t = \Phi_1(X_{t-})\, dt + \Phi_2(X_{t-})\, dB_t + \Phi_3(X_{t-})\, dZ_t, \quad X_0 = x \in \mathbb{R}^d, \tag{2.8}$$

where $\Phi_1 : \mathbb{R}^d \to \mathbb{R}^d$, $\Phi_2 : \mathbb{R}^d \to \mathbb{R}^{d\times p}$ and $\Phi_3 : \mathbb{R}^d \to \mathbb{R}^{d\times q}$, with $p + q = n - 1$, are locally Lipschitz continuous and bounded, $\{B_t\}_{t\geq 0}$ is a p-dimensional Brownian motion, and $\{Z_t\}_{t\geq 0}$ is a q-dimensional pure-jump Lévy process (that is, a Lévy process determined by a Lévy triplet of the form $(0, 0, \nu_Z(dy)))$ independent of $\{B_t\}_{t\geq 0}$. Namely, set $\Phi(x) = (\Phi_1(x), \Phi_2(x), \Phi_3(x))$, and $Y_t = (t, B_t, Z_t)^T$, $t \geq 0$. Assume now that $d = p = q = 1$. Then, from [59], Theorem 3.1, we see that the corresponding coefficients read

$$b(x) = \begin{cases} \Phi_1(x), & \Phi_3(x) = 0, \\ \Phi_1(x) + \int_{\mathbb{R}} y\big(\mathbb{1}_{B_1(0)}(y) - \mathbb{1}_{B_{|\Phi_3(x)|}(0)}(y)\big)\nu_Z\Big(\frac{dy}{\Phi_3(x)}\Big), & \Phi_3(x) \neq 0, \end{cases}$$

$$c(x) = \Phi_2^2(x),$$

$$\nu(x, dy) = \begin{cases} 0, & \Phi_3(x) = 0, \\ \nu_Z\Big(\frac{dy}{\Phi_3(x)}\Big), & \Phi_3(x) \neq 0. \end{cases}$$

Take now, for simplicity,

$$\Phi_1(x) = \Phi_3(x) = \begin{cases} -1, & x \geq 1, \\ -x, & -1 \leq x \leq 1, \\ 1, & x \leq -1, \end{cases}$$

$\Phi_2(x) = 1$, and $\nu_Z(\mathrm{d}y) = f(y)\,\mathrm{d}y$ with $f(y)$ being the probability density function of the continuous uniform distribution on the segment [0, 1]. It is straightforward to see that $\{X_t\}_{t\geq 0}$ satisfies (A4) and (A5). Open-set irreducibility and aperiodicity of $\{X_t\}_{t\geq 0}$ have been considered on [46], page 43, (see also [41], Theorem 3.1). Finally, since

$$B_0(x) = \begin{cases} -\dfrac{1}{2}x, & x \geq 1, \\ \dfrac{1}{2}x, & x \leq -1, \end{cases}$$

it is elementary to check that $\{X_t\}_{t\geq 0}$ satisfies (1.2) with $x_0 = 0$, $r_0 = 1$ and $\varphi(t) = t^\alpha$, $\alpha \in (0, 1)$. Thus, $\{X_t\}_{t\geq 0}$ is sub-geometrically ergodic with rate $t^{\alpha/(1-\alpha)}$.

Observe that the same conclusion follows by employing a version of the relation in (1.7) including jumps (see [56], Theorem 3.3). However, if we take $\Phi_1(x) = -\operatorname{sgn}(x)(\cos x + 3/2)$ (analogously as in Example 2.5), then it is not hard to see that (1.7) does not hold. On the other hand, Theorem 2.8 (with $x_0 = 0$, $r_0 = 1$ and $\varphi(t) = t^\alpha$, $\alpha \in (0, 1)$) implies that $\{X_t\}_{t\geq 0}$ is again sub-geometrically ergodic with rate $t^{\alpha/(1-\alpha)}$.

An alternative approach in obtaining a class of Markov processes with jumps (from diffusion processes) is through the Bochner's subordination method. Recall, a subordinator $\{S_t\}_{t\geq 0}$ is a non-decreasing Lévy process on $[0, \infty)$ with Laplace transform

$$\mathbb{E}\big[\mathrm{e}^{-uS_t}\big] = \mathrm{e}^{-t\phi(u)}, \quad u > 0, t \geq 0.$$

The characteristic (Laplace) exponent $\phi : (0, \infty) \to (0, \infty)$ is a Bernstein function, that is, it is of class C^∞ and $(-1)^n \phi^{(n)}(u) \geq 0$ for all $n \in \mathbb{N}$. It is well known that every Bernstein function admits a unique (Lévy-Khintchine) representation

$$\phi(u) = bu + \int_{(0,\infty)} \big(1 - \mathrm{e}^{-uy}\big)\nu(\mathrm{d}y), \quad u > 0,$$

where $b \geq 0$ is the drift parameter and ν is a Lévy measure, that is, a measure on $\mathcal{B}((0, \infty))$ satisfying $\int_{(0,\infty)}(1 \wedge y)\nu(\mathrm{d}y) < \infty$. For more on subordinators and Bernstein functions, we refer the readers to the monograph [60]. Let now $\{M_t\}_{t\geq 0}$ be a Markov process with state space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ and transition kernel $p(t, x, \mathrm{d}y)$. Further, let $\{S_t\}_{t\geq 0}$ be a subordinator with characteristic exponent $\phi(u)$, independent of $\{M_t\}_{t\geq 0}$. The process $M_t^\phi := M_{S_t}$, $t \geq 0$, obtained from $\{M_t\}_{t\geq 0}$ by a random time change through $\{S_t\}_{t\geq 0}$, is referred to as the subordinate process $\{M_t\}_{t\geq 0}$ with subordinator $\{S_t\}_{t\geq 0}$ in the sense of Bochner. It is easy to see that $\{M_t^\phi\}_{t\geq 0}$ is again a Markov process with transition kernel

$$p^\phi(t, x, \mathrm{d}y) = \int_{[0,\infty)} p(s, x, \mathrm{d}y)\mu_t(\mathrm{d}s),$$

where $\mu_t(\cdot) = \mathbb{P}(S_t \in \cdot)$ is the transition probability of S_t, $t \geq 0$. Also, it is elementary to check that if π is an invariant probability measure for $\{M_t\}_{t\geq 0}$, then π is also invariant for the subordinate process

$\{M_t^\phi\}_{t\geq 0}$. In [13] it has been shown that if $\{M_t\}_{t\geq 0}$ is sub-geometrically ergodic with Borel measurable rate $r(t)$ (with respect to the total variation distance), then $\{M_t^\phi\}_{t\geq 0}$ is sub-geometrically ergodic with rate $r_\phi(t) = \mathbb{E}[r(S_t)]$. Therefore, as an direct application of Theorem 1.1, we obtain sub-geometric ergodicity results for a class of subordinate diffusion processes.

3. Ergodicity with respect to Wasserstein distances

In this section, we first prove Theorems 1.2 and 1.3. Then, we discuss sub-geometric ergodicity of two classes of Markov processes with jumps.

3.1. Proof of Theorems 1.2 and 1.3

In Theorem 1.1 we discussed sub-geometric ergodicity of a diffusion process $\{X_t\}_{t\geq 0}$ (given through (1.1)) with respect to the total variation distance. Crucial assumptions in this result were open-set irreducibility and aperiodicity of $\{X_t\}_{t\geq 0}$. In order to ensure these properties the discussion after Proposition 2.2 and Theorem 2.3 suggest that quite strong regularity and smoothness assumptions of the coefficient $c(x)$ are needed. By using a completely different approach to this problem, the so-called synchronous coupling method (see [12], Example 2.16, for details), we derive sub-geometric ergodicity for a class of diffusions with (possibly) singular diffusion coefficient.

We start with the following auxiliary result, which will be crucial in the proofs of Theorems 1.2 and 1.3, and which is a version of non-linear convex Gronwall's inequality.

Lemma 3.1. *Let* $\Gamma > 0$, *and let* $f : [0, T) \to [0, \infty)$, *with* $0 < T \leq \infty$, *and* $\psi : [0, \infty) \to [0, \infty)$ *be such that*

(i) $f(t)$ *is absolutely continuous on* $[t_0, t_1]$ *for any* $0 < t_0 < t_1 < T$;
(ii) $f'(t) \leq -\Gamma\psi(f(t))$ *a.e. on* $[0, T)$;
(iii) $\psi(f(t)) > 0$ *a.e. on* $[0, T)$, *and* $\Psi_{f(0)}(t) := \int_t^{f(0)} \frac{\mathrm{d}s}{\psi(s)} < \infty$ *for all* $t \in (0, f(0)]$.

Then,

$$f(t) \leq \Psi_{f(0)}^{-1}(\Gamma t), \quad 0 \leq t < \Gamma^{-1}\Psi_{f(0)}(0) \wedge T.$$

In addition, if there is $\kappa \in [f(0), \infty]$ *such that* $\Psi_\kappa(t) := \int_t^\kappa \frac{\mathrm{d}s}{\psi(s)} < \infty$ *for* $t \in (0, \kappa]$, *then*

$$f(t) \leq \Psi_\kappa^{-1}(\Gamma t), \quad 0 \leq t < \Gamma^{-1}\Psi_{f(0)}(0) \wedge T.$$

Also, if $\psi(t)$ *is convex and vanishes at zero, then* $\Psi_{f(0)}(0) = \infty$, *that is, the above relations hold for all* $t \in [0, T)$.

Proof. By assumption,

$$-\Psi_{f(0)}\big(f(t)\big) = \int_{f(0)}^{f(t)} \frac{\mathrm{d}s}{\psi(s)} = \int_0^t \frac{f'(s)\,\mathrm{d}s}{\psi(f(s))} \leq -\Gamma t, \quad t \in [0, T).$$

Now, the first assertion follows.

The second claim follows from the fact that $\Psi_{f(0)}(t) \leq \Psi_\kappa(t)$ for all $t \in (0, f(0)]$, while the last part follows from

$$\psi(t) = \psi\big(t + (1-t)0\big) \leq t\psi(1) + (1-t)\psi(0) = t\psi(1), \quad t \in [0, 1]. \qquad \square$$

Now, we are in position to prove Theorem 1.2.

Proof of Theorem 1.2. Fix $x, y \in \mathbb{R}^d$, $x \neq y$, and let $\{X_t\}_{t\geq 0}$ and $\{Y_t\}_{t\geq 0}$ be solutions to (1.1) starting from x and y, respectively. Further, define $\tau := \inf\{t > 0 : X_t = Y_t\}$ and

$$Z_t := \begin{cases} Y_t, & t < \tau, \\ X_t, & t \geq \tau, \end{cases} \quad t \geq 0.$$

By employing the strong Markov property it is easy to see that $\mathbb{P}^y(Z_t \in \cdot) = \mathbb{P}^y(Y_t \in \cdot)$ for all $t \geq 0$. Consequently,

$$\mathcal{W}_{f,p}(\delta_x P_t, \delta_y P_t) \leq \big(\mathbb{E}\big(f\big(|X_t - Z_t|\big)^p\big)\big)^{1/p}, \quad t \geq 0.$$

Next, since the mapping $t \mapsto |X_t - Z_t|$ is absolutely continuous on $[0, \tau)$, the function $t \mapsto f(|X_t - Z_t|)$ is differentiable a.e. on $[0, \tau)$ and we have that

$$\frac{\mathrm{d}}{\mathrm{d}t} f\big(|X_t - Z_t|\big) = \frac{f'(|X_t - Z_t|)}{|X_t - Z_t|}\big\langle X_t - Z_t, b(X_t) - b(Z_t)\big\rangle,$$

a.e. on $[0, \tau)$. Now, by assumption, we get

$$\frac{\mathrm{d}}{\mathrm{d}t} f\big(|X_t - Z_t|\big) \leq 0,$$

a.e. on $[0, \tau)$, which implies that the function $t \mapsto f(|X_t - Z_t|)$ is non-increasing on $[0, \infty)$. Take now $x, y \in \mathbb{R}^d$ such that $0 < f(|x - y|) \leq \gamma$ (which exist by (iii)). Thus, for such starting points, $f(|X_t - Z_t|) \leq \gamma$ on $[0, \infty)$. Now, by assumption,

$$\frac{\mathrm{d}}{\mathrm{d}t} f\big(|X_t - Z_t|\big) \leq -\Gamma\psi\big(f\big(|X_t - Z_t|\big)\big),$$

a.e. on $[0, \tau)$, which together with Lemma 3.1 gives

$$f\big(|X_t - Z_t|\big) \leq \Psi^{-1}_{f(|x-y|)}(\Gamma t), \quad t \geq 0.$$

For $t \geq \tau$ the term on the left-hand side vanishes, and the term on the right-hand side is well defined and strictly positive ($\psi(t)$ is convex and $\psi(t) = 0$ if and only if $t = 0$). Now, by taking the expectation and infimum we conclude

$$\mathcal{W}_{f,p}(\delta_x P_t, \delta_y P_t) \leq \Psi^{-1}_{f(|x-y|)}(\Gamma t), \quad t \geq 0,$$

which proves (a).

The relations in (b) now follow from (a) and Lemma 3.1.

Let us prove (c). If $f(|x - y|) \leq \gamma$ for all $x, y \in \mathbb{R}^d$, then the assertion follows from (a). Assume that there are $x, y \in \mathbb{R}^d$ such that $f(|x - y|) > \gamma$. Observe that, $\delta = 0$ if and only if $f(t) \leq \gamma$ for all $t \in [0, \infty)$. Thus, $\delta > 0$, and we have that

$$f\left(\frac{|x - y|}{\lceil \delta|x - y|\rceil}\right) \leq f\big(\delta^{-1}\big) \leq \gamma.$$

Take $z_0, \ldots, z_{\lceil \delta|x-y|\rceil} \in \mathbb{R}^d$, such that $z_0 = x$ and

$$z_{i+1} = z_i + \frac{y - x}{\lceil \delta|x - y|\rceil}, \quad i = 0, \ldots, \big\lceil \delta|x - y|\big\rceil - 1.$$

By construction, $f(|z_0 - z_1|) = \cdots = f(|z_{\lceil\delta|x-y|\rceil-1} - z_{\lceil\delta|x-y|\rceil}|) \le \gamma$. Thus, using (b) we conclude that for $x, y \in \mathbb{R}^d$ such that $f(|x-y|) > \gamma$,

$$\begin{aligned}\mathcal{W}_{f,p}(\delta_x P_t, \delta_y P_t) &\le \mathcal{W}_{f,p}(\delta_{z_0} P_t, \delta_{z_1} P_t) + \cdots + \mathcal{W}_{f,p}(\delta_{z_{\lceil\delta|x-y|\rceil-1}} P_t, \delta_{z_{\lceil\delta|x-y|\rceil}} P_t)\\ &\le \lceil \delta|x-y| \rceil \Psi_\gamma^{-1}(\Gamma t), \quad t \ge 0.\end{aligned}$$

Finally, observe that if $t > 0$ is such that $f(t) \le \gamma$, then $\delta \le 1/t$, that is, $\delta t \le 1$. Hence, for $x, y \in \mathbb{R}^d$ such that $f(|x-y|) \le \gamma$ we have $\lceil\delta|x-y|\rceil = 1$, which concludes the proof. □

Let us now give several remarks.

Remark 3.2.

(i) If the condition in (1.8) holds for some $\gamma > 0$, then it also holds for any $0 < \bar{\gamma} \le \gamma$.

(ii) By replacing the condition in (1.8) with

$$\begin{aligned}&f\big(|x-y|\big)^{p-1} f'\big(|x-y|\big)\big\langle x-y, b(x)-b(y)\big\rangle\\ &\le \begin{cases} -\frac{\Gamma}{p}|x-y|\psi\big(f^p\big(|x-y|\big)\big), & f^p\big(|x-y|\big) \le \gamma,\\ 0, & f^p\big(|x-y|\big) > \gamma,\end{cases}\end{aligned}$$

a.e. on $\mathbb{R}^d$ for $\gamma > 0$ and $\Gamma > 0$, leads to analogous results ($f(t)$ is replaced by $f^p(t)$ in every relation).

(iii) For any $\mu, \nu \in \mathcal{P}$ it holds that

$$\mathcal{W}_{f,p}(\mu P_t, \nu P_t) \le \big(\delta \mathcal{W}_p(\mu,\nu) + 1\big)\Psi_\gamma^{-1}(\Gamma t), \quad t \ge 0.$$

In particular, for $f(t) = t$ we have that

$$\mathcal{W}_p(\mu P_t, \nu P_t) \le \left(\frac{\mathcal{W}_p(\mu,\nu)}{\gamma} + 1\right)\Psi_\gamma^{-1}(\Gamma t), \quad t \ge 0.$$

(iv) By taking $\psi(t) = t$, we obtain geometric rate of convergence with $\Psi^{-1}_{f(|x-y|)}(\Gamma t) = f(|x-y|)\mathrm{e}^{-\Gamma t}$. This result can be also obtained in an alternative way (without Lemma 3.1, that is, Gronwall's inequality), by applying Itô's lemma to the processes $\{f(|X_t - Z_t|)\}_{t\ge0}$ and $\{\mathrm{e}^{\Gamma t} f(|X_t - Z_t|)\}_{t\ge0}$.

(v) In the case when $f(t) = \psi(t) = t$, according to (1.13), we get

$$W_p(\mu P_t, \nu P_t) \le \mathcal{W}_p(\mu,\nu)\mathrm{e}^{-\Gamma t}, \quad p \ge 1,\ \mu,\nu \in \mathcal{P},\ t \ge 0, \tag{3.1}$$

which is the same results as in [68] (for $p = 2$). Also, by an analogous approach as in the proof of Theorem 1.3, from (3.1) we see that $\{X_t\}_{t\ge0}$ admits a unique invariant $\pi \in \bigcap_{p\ge1}\mathcal{P}_p$ such that

$$\mathcal{W}_p(\mu P_t, \pi) \le \mathcal{W}_p(\mu,\pi)\mathrm{e}^{-\Gamma t}, \quad p \ge 1,\ \mu \in \mathcal{P}_p,\ t \ge 0.$$

(vi) From (3.1), we see that the mapping $\mathcal{P} \ni \mu \mapsto \mu P_t \in \mathcal{P}$ is a contraction for fixed $t > 0$, that is, the right-hand side in (3.1) is strictly smaller than $\mathcal{W}_p(\mu,\nu)$. On the other hand, in the general situation, this is not the case anymore (see (iii)). However, if

$$f'\big(|x-y|\big)\big\langle x-y, b(x)-b(y)\big\rangle \le -\Gamma|x-y|\psi\big(f\big(|x-y|\big)\big), \quad x, y \in \mathbb{R}^d,$$

then from (1.9) we have that for all $x, y \in \mathbb{R}^d$ and all $t \ge 0$,

$$\mathcal{W}_{f,p}(\delta_x P_t, \delta_y P_t) \le \Psi^{-1}_{f(|x-y|)}(\Gamma t) \le \Psi^{-1}_{f(|x-y|)}(0) = f\big(|x-y|\big),$$

that is,

$$\mathcal{W}_{f,p}(\mu P_t, \nu P_t) \le \mathcal{W}_{f,p}(\mu, \nu), \quad p \ge 1,\ \mu, \nu \in \mathcal{P},\ t \ge 0.$$

Thus, the mapping $\mathcal{P} \ni \mu \mapsto \mu P_t \in \mathcal{P}$ is contractive for any fixed $t \ge 0$.

We now prove Theorem 1.3.

Proof of Theorem 1.3. First, we prove that $\{X_t\}_{t\ge 0}$ admits an invariant probability measure. According to [49], Theorem 3.1, this will follow if we show that for each $x \in \mathbb{R}^d$ and $0 < \varepsilon < 1$ there is a compact set $C \subset \mathbb{R}^d$ (possibly depending on x and ε) such that

$$\liminf_{t \nearrow \infty} \frac{1}{t} \int_0^t p(s, x, C)\, \mathrm{d}s \ge 1 - \varepsilon.$$

By taking $y = 0$ in (1.14), we have that

$$\big\langle x, b(x)\big\rangle \le \big\langle x, b(0)\big\rangle - \Gamma |x| \psi\big(|x|\big) \le \big|b(0)\big||x| - \Gamma |x| \psi\big(|x|\big), \quad x \in \mathbb{R}^d.$$

In particular, for $\mathcal{V}(x) = |x|^2$ we have that

$$\mathcal{L}\mathcal{V}(x) = 2\big\langle x, b(x)\big\rangle + \operatorname{Tr}\sigma\sigma^T \le \operatorname{Tr}\sigma\sigma^T + 2\big|b(0)\big||x| - 2\Gamma|x|\psi\big(|x|\big), \quad x \in \mathbb{R}^d.$$

Now, since every super-additive convex function is necessarily non-decreasing and unbounded, we conclude that there is $r_0 > 0$ large enough such that

$$\operatorname{Tr}\sigma\sigma^T + 2\big|b(0)\big||x| \le \Gamma|x|\psi\big(|x|\big), \quad |x| \ge r_0,$$

that is,

$$\begin{aligned}
\mathcal{L}\mathcal{V}(x) &\le \big(\operatorname{Tr}\sigma\sigma^T + 2\big|b(0)\big||x| - 2\Gamma|x|\psi\big(|x|\big)\big)\mathbb{1}_{B_{r_0}}(x) \\
&\quad + \big(\operatorname{Tr}\sigma\sigma^T + 2\big|b(0)\big||x| - 2\Gamma|x|\psi\big(|x|\big)\big)\mathbb{1}_{B^c_{r_0}}(x) \\
&\le \big(\operatorname{Tr}\sigma\sigma^T + 2\big|b(0)\big||x| - 2\Gamma|x|\psi\big(|x|\big)\big)1_{B_{r_0}}(x) - \Gamma|x|\psi\big(|x|\big)\mathbb{1}_{B^c_{r_0}}(x) \\
&\le \big(\operatorname{Tr}\sigma\sigma^T + 2\big|b(0)\big|r_0 + \Gamma r_0\psi(r_0)\big)\mathbb{1}_{B_{r_0}}(x) - \Gamma r_0\psi(r_0), \quad |x| \ge r_0.
\end{aligned}$$

Clearly, the above relation holds for all $r \ge r_0$ also. Now, according to [50], Theorem 1.1, we conclude that for each $x \in \mathbb{R}^d$ and $r \ge r_0$ we have

$$\liminf_{t \nearrow \infty} \frac{1}{t} \int_0^t p\big(s, x, \bar{B}_r(0)\big)\, \mathrm{d}s \ge \frac{\Gamma r \psi(r)}{\operatorname{Tr}\sigma\sigma^T + 2|b(0)|r + \Gamma r\psi(r)}.$$

The assertion now follows by choosing r large enough.

Let us now show that any invariant $\pi \in \mathcal{P}$ of $\{X_t\}_{t\ge 0}$ has finite all moments. Fix $p \ge 2$ and let $\mathcal{V}_p(x) = |x|^p$. By the same reasoning as above, it is easy to see that there are $r_p > 0$, $\Gamma_{p,1} > 0$ and $\Gamma_{p,2} > 0$ such that

$$\mathcal{L}\mathcal{V}_p(x) \le \Gamma_{p,1}\mathbb{1}_{B_{r_p}(0)}(x) - \Gamma_{p,2}|x|^{p-1}\psi\big(|x|\big), \quad x \in \mathbb{R}^d.$$

Now, from [50], Theorem 4.3, it follows that

$$\int_{\mathbb{R}^d} |x|^{p-1}\psi\big(|x|\big)\pi(\mathrm{d}x) \le \frac{\Gamma_{p,1}}{\Gamma_{p,2}}$$

for any corresponding invariant $\pi \in \mathcal{P}$.

Finally, let us prove that $\{X_t\}_{t\ge0}$ admits a unique invariant probability measure which satisfies (1.15). Let $\pi, \bar\pi \in \mathcal{P}$ be two invariant probability measures of $\{X_t\}_{t\ge0}$. Then, for any $\kappa > 0$ and $p \ge 1$ Remark 3.2 implies that

$$\mathcal{W}_p(\pi, \bar\pi) = \mathcal{W}_p(\pi P_t, \bar\pi P_t) \le \left(\frac{\mathcal{W}_p(\pi, \bar\pi)}{\kappa} + 1\right)\Psi_\kappa^{-1}(\Gamma t), \quad t \ge 0.$$

Now, by letting $t \to \infty$ we see that $\mathcal{W}_p(\pi, \bar\pi) = 0$, that is, $\{X_t\}_{t\ge0}$ admits a unique invariant $\pi \in \mathcal{P}$. Finally, for any $\kappa > 0$, $p \ge 1$ and $\mu \in \mathcal{P}_p$, by employing Remark 3.2 again, we have that

$$\mathcal{W}_p(\pi, \mu P_t) = \mathcal{W}_p(\pi P_t, \mu P_t) \le \left(\frac{\mathcal{W}_p(\pi, \mu)}{\kappa} + 1\right)\Psi_\kappa^{-1}(\Gamma t), \quad t \ge 0,$$

which concludes the proof. □

Let us now give a simple example satisfying (1.8) and (1.14).

Example 3.3. Let $p > 1$, $b(x) = -\operatorname{sgn}(x)|x|^p$, $\sigma(x) \equiv \sigma \in \mathbb{R}$, $f(t) = t$, $\gamma > 0$ and $\psi(t) = t^p$. Now, it is easy to see that $b(x)$ cannot satisfy the relation in (1.13). On the other hand, an elementary computation shows that there is $\Gamma > 0$ such that (1.8) holds true. Thus, we have (1.12) with $\delta = \gamma^{-1}$. Also, $\lim_{t\to\infty} \sqrt[p-1]{t}\,\Psi_\kappa^{-1}(t) = 1/\sqrt[p-1]{p-1}$, $\kappa > 0$.

Let us also remark that one can show that the same result holds in the multidimensional case with $b(x_1, \dots, x_d) = (-\operatorname{sgn}(x_1)|x_1|^p, \dots, -\operatorname{sgn}(x_d)|x_d|^p)$.

3.2. Ergodicity of Markov processes with jumps

Let $\{Y_t\}_{t\ge0}$ be a d-dimensional Lévy process with Lévy triplet (β, γ, ν). Further, let $b : \mathbb{R}^d \to \mathbb{R}^d$ be continuous and such that

(J1) for any $r > 0$ there is $\Gamma_r > 0$ such that for all $x, y \in B_r(0)$,

$$\big\langle x - y, b(x) - b(y)\big\rangle \le \Gamma_r |x - y|^2;$$

(J2) there is $\Gamma > 0$ such that for all $x \in \mathbb{R}^d$,

$$\big\langle x, b(x)\big\rangle \le \Gamma\big(1 + |x|^2\big).$$

Then, according to [43], Theorem 1.1, and Lemmas 2.4 and 2.5, the SDE

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \mathrm{d}Y_t, \quad X_0 = x \in \mathbb{R}^d, \tag{3.2}$$

admits a unique strong non-explosive solution $\{X_t\}_{t\ge0}$ which is a strong Markov process and satisfies the C_b-Feller property.

Lemma 3.4. *Assume that* $\mathbb{E}[|Y_1|^p] < \infty$ *(or, equivalently,* $\int_{B_1^c(0)} |y|^p \nu(\mathrm{d}y) < \infty$*) for some* $p > 0$*. Then, there is a constant* $\Delta > 0$ *such that*

$$\mathbb{E}^x\big[|X_t|^p\big] \le \big(|x|^p + 1\big)\mathrm{e}^{\Delta t}, \quad t \ge 0, \ x \in \mathbb{R}^d.$$

Proof. Let $\chi \in C^2(\mathbb{R}^d)$ be such that $\chi(x) \ge 0$, $\chi(x) \le |x|^p$ and $\chi(x) = |x|^p$ for $x \in B_1^c(0)$. Further, for $n \in \mathbb{N}$, let $\chi_n \in C_b^2(\mathbb{R}^d)$ be such that $\chi_n(x) \ge 0$, $\chi_n(x) = \chi|_{B_{n+1}(0)}(x)$ and $\chi_n(x) \to \chi(x)$ as $n \to \infty$, and $\tau_n := \inf\{t \ge 0 : X_t \in B_n^c(0)\}$. Then, according to Itô's formula (see [2], Remark 2.2), we have that

$$\begin{aligned}\mathbb{E}^x\big[\chi_n(X_{t\wedge\tau_n})\big] &\le \chi_n(x) + \Delta_n(t\wedge\tau_n) + \Delta_n\mathbb{E}^x\left[\int_0^{t\wedge\tau_n}\chi_n(X_s)\,\mathrm{d}s\right] \\ &\le \chi_n(x) + \Delta_n t + \Delta_n\int_0^t \mathbb{E}^x\big[\chi_n(X_{s\wedge\tau_n})\big]\,\mathrm{d}s, \quad n \in \mathbb{N},\ t \ge 0,\ x \in \mathbb{R}^d,\end{aligned}$$

where the constants $\Delta_n > 0$ depend on p, β, γ, $b(x)$ and constants $\int_{B_1(0)} |y|^2\nu(\mathrm{d}y)$, $\nu(B_1^c(0))$, $\sup_{x\in B_R(0)} |\nabla\chi_n(x)|$ and $\sup_{x\in B_R(0)} |\nabla^2\chi_n(x)|$, for $R > 0$ large enough. Clearly, the functions $\chi_n(x)$ can be chosen such that $\Delta := \sup_{n\in\mathbb{N}} \Delta_n < \infty$. Now, since the function $t \mapsto \mathbb{E}^x[\chi_n(X_{t\wedge\tau_n})]$ is bounded and càdlàg, Gronwall's lemma implies that

$$\mathbb{E}^x\big[\chi_n(X_{t\wedge\tau_n})\big] \le \big(\chi_n(x) + 1\big)e^{\Delta t} - 1, \quad n \in \mathbb{N},\ t \ge 0,\ x \in \mathbb{R}^d.$$

By letting $n \to \infty$ monotone convergence theorem and non-explosivity of $\{X_t\}_{t\ge0}$ imply that

$$\mathbb{E}^x\big[\chi(X_t)\big] \le \big(\chi(x) + 1\big)e^{\Delta t} - 1, \quad t \ge 0,\ x \in \mathbb{R}^d.$$

Finally, we have that

$$\mathbb{E}^x\big[|X_t|^p\big] \le \mathbb{E}^x\big[\chi(X_t)\big] + 1 \le \big(\chi(x) + 1\big)\mathrm{e}^{\Delta t} \le \big(|x|^p + 1\big)\mathrm{e}^{\Delta t}, \quad t \ge 0,\ x \in \mathbb{R}^d. \qquad \square$$

Lemma 3.5. *Assume that* $\nu(\mathbb{R}^d) < \infty$*. Then, the sample paths of* $\{X_t\}_{t\ge0}$ *are piecewise continuous* $\mathbb{P}^x$*-a.s.*

Proof. Define $\tau_0 := 0$ and

$$\tau_n := \inf\big\{t \ge \tau_{n-1} : |X_t - X_{t-}| > 0\big\} = \inf\big\{t \ge \tau_{n-1} : |Y_t - Y_{t-}| > 0\big\}, \quad n \ge 1.$$

Clearly, $\{\tau_n\}_{n\in\mathbb{N}}$ are i.i.d. and $\mathbb{P}^x(\tau_1 > t) = \mathrm{e}^{-\nu(\mathbb{R}^d)t}$ (that is, τ_1 is exponentially distributed with parameter $\nu(\mathbb{R}^d)$) for any $x \in \mathbb{R}^d$. Hence, $\{X_t\}_{t\ge0}$ is continuous on $[\tau_n, \tau_{n+1})$, $n \ge 0$, $\mathbb{P}^x$-a.s. for all $x \in \mathbb{R}^d$. $\square$

Let now $\{X_t\}_{t\ge0}$ be a solution to (3.2) with $b(x)$ satisfying (J1) and (J2), and with $\{Y_t\}_{t\ge0}$ having finite p-th moment, $p \ge 1$, and finite Lévy measure. Then, according to Lemmas 3.4 and 3.5, if $b(x)$ satisfies (1.8) we conclude that $\{X_t\}_{t\ge0}$ satisfies (1.9), (1.10), (1.11) and (1.12). Further, according to [2] and [46], for any $f \in C^2(\mathbb{R}^d)$ such that $x \mapsto \int_{B_1^c(0)} f(x+y)\nu(\mathrm{d}y)$ is locally bounded,

$$f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s)\,\mathrm{d}s, \quad t \ge 0,$$

is a local $\mathbb{P}^x$-martingale, $x \in \mathbb{R}^d$, where

$$\mathcal{L}f(x) = \langle b(x), \nabla f(x)\rangle + \langle \beta, \nabla f(x)\rangle + \frac{1}{2}\operatorname{Tr}\gamma\nabla^2 f(x) + \int_{\mathbb{R}^d} \big(f(y+y) - f(x) - \langle y, \nabla f(x)\rangle \mathbb{1}_{B_1(0)}(y)\big)\nu(\mathrm{d}y).$$

Proposition 3.6. *Let $p \ge 1$. Assume that $b(x)$ satisfies* (J1), (J2) *and* (1.14), *and that $\{Y_t\}_{t\ge0}$ has finite p-th moment and finite Lévy measure. Then, $\{X_t\}_{t\ge0}$ admits a unique invariant $\pi \in \mathcal{P}_p$ such that for any $\kappa > 0$, $1 \le q \le p$ and $\mu \in \mathcal{P}_q$ it holds that*

$$\mathcal{W}_q(\pi, \mu P_t) \le \left(\frac{\mathcal{W}_q(\pi,\mu)}{\kappa} + 1\right)\Psi_\kappa^{-1}(\Gamma t), \quad t \ge 0. \tag{3.3}$$

Proof. First, observe that

$$\mathcal{L}f(x) = \langle b(x), \nabla f(x)\rangle + \left\langle \beta + \int_{B_1^c(0)} y\nu(\mathrm{d}y), \nabla f(x)\right\rangle + \frac{1}{2}\operatorname{Tr}\gamma\nabla^2 f(x) + \int_{\mathbb{R}^d} \big(f(y+y) - f(x) - \langle y, \nabla f(x)\rangle\big)\nu(\mathrm{d}y).$$

By taking a non-negative $\mathcal{V}_p \in C^2(\mathbb{R}^d)$ such that $\mathcal{V}_p(x) = |x|^p$ on $B_1^c(0)$ from [4], Lemma 5.1, we have that

$$\sup_{x\in\mathbb{R}^d}\left|\int_{\mathbb{R}^d} \big(\mathcal{V}_p(y+y) - \mathcal{V}_p(x) - \langle y, \nabla\mathcal{V}_p(x)\rangle\big)\nu(\mathrm{d}y)\right| < \infty.$$

Now, by completely the same approach as in the proof of Theorem 1.3 we conclude that $\{X_t\}_{t\ge0}$ admits a unique invariant $\pi \in \mathcal{P}$ such that $\int_{\mathbb{R}^d} |x|^{p-1}\psi(|x|)\pi(\mathrm{d}x) < \infty$. Thus, $\pi \in \mathcal{P}_p$, and the relation in (3.3) follows by the same reasoning as in the proof of Theorem 1.3. □

Analogously as in Section 2.2, in the following proposition we discuss ergodicity of a class of Markov processes with jumps, obtained through Bochner's subordination method, with respect to Wasserstein distances.

Proposition 3.7. *Let $\{M_t\}_{t\ge0}$ be a Markov process with state space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ and semigroup $\{P_t\}_{t\ge0}$. Let $\{S_t\}_{t\ge0}$ be a subordinator with characteristic exponent $\phi(u)$, independent of $\{M_t\}_{t\ge0}$. Further, let ρ be a metric on $\mathbb{R}^d$ such that $(\mathbb{R}^d, \rho)$ is a Polish space and $\mathcal{B}(\mathbb{R}^d_\rho) \subseteq \mathcal{B}(\mathbb{R}^d)$, that is, ρ induces a coarser topology than the standard d-dimensional Euclidean metric on $\mathbb{R}^d$. Assume, that $\{M_t\}_{t\ge0}$ admits an invariant $\pi \in \mathcal{P}$ such that $\mathcal{W}_{\rho,p}(\delta_x P_t, \pi) \le \Gamma(x) r(t)$, $t \ge 0$, $x \in \mathbb{R}^d$, where $r : [0,\infty) \to [1,\infty)$ is Borel measurable and $\Gamma(x) \ge 0$. Then, $\mathcal{W}_{\rho,p}(\delta_x P_t^\phi, \pi) \le \Gamma(x) r_\phi(t)$, $t \ge 0$, $x \in \mathbb{R}^d$, where $r_\phi(t) = (\mathbb{E}[r^p(S_t)])^{1/p}$.*

Proof. First, recall that if π is an invariant measure for $\{M_t\}_{t\ge0}$, then it is also invariant for $\{M_t^\phi\}_{t\ge0}$. Next, [67], Theorem 4.1, implies that for each $s \in [0,\infty)$ there is $\Pi_s \in \mathcal{C}(\delta_x P_s, \pi)$ such that

$\mathcal{W}_{\rho,p}(\delta_x P_s^\phi,\pi)=\int_{\mathbb{R}^d\times\mathbb{R}^d}\rho(y,z)\Pi_s(\mathrm{d}y,\mathrm{d}z)$. Now, we have that

$$\begin{aligned}
\mathcal{W}_{\rho,p}^p\big(\delta_x P_t^\phi,\pi\big) &= \inf_{\Pi\in\mathcal{C}(\delta_x P_t^\phi,\pi)}\int_{\mathbb{R}^d\times\mathbb{R}^d}\rho^p(y,z)\Pi(\mathrm{d}y,\mathrm{d}z)\\
&\le \int_{\mathbb{R}^d\times\mathbb{R}^d}\rho^p(y,z)\int_{[0,\infty)}\Pi_s(\mathrm{d}y,\mathrm{d}z)\mu_t(\mathrm{d}s)\\
&\le \int_{[0,\infty)}\mathcal{W}_{\rho,p}^p(\delta_x P_s,\pi)\mu_t(\mathrm{d}s)\\
&\le \Gamma^p(x)\int_{[0,\infty)}r^p(s)\mu_t(ds)\\
&= \Gamma^p(x)\mathbb{E}\big[r^p(S_t)\big],
\end{aligned}$$

which completes the proof. □

Acknowledgements

This research was supported by the Croatian Science Foundation (under Project 8958). We also thank two anonymous referees for the helpful comments that have led to significant improvements of the results in the article.

References

[1] Abourashchi, N. and Veretennikov, A.Y. (2010). On stochastic averaging and mixing. *Theory Stoch. Process.* **16** 111–129. MR2779833
[2] Albeverio, S., Brzeźniak, Z. and Wu, J.-L. (2010). Existence of global solutions and invariant measures for stochastic differential equations driven by Poisson type noise with non-Lipschitz coefficients. *J. Math. Anal. Appl.* **371** 309–322. MR2661009 https://doi.org/10.1016/j.jmaa.2010.05.039
[3] Arapostathis, A., Borkar, V.S. and Ghosh, M.K. (2012). *Ergodic Control of Diffusion Processes*. *Encyclopedia of Mathematics and Its Applications* **143**. Cambridge: Cambridge Univ. Press. MR2884272
[4] Arapostathis, A., Pang, G. and Sandrić, N. (2019). Ergodicity of a Lévy-driven SDE arising from multiclass many-server queues. *Ann. Appl. Probab.* **29** 1070–1126. MR3910024 https://doi.org/10.1214/18-AAP1430
[5] Bass, R.F. and Cranston, M. (1986). The Malliavin calculus for pure jump processes and applications to local time. *Ann. Probab.* **14** 490–532. MR0832021
[6] Bhattacharya, R.N. (1978). Criteria for recurrence and existence of invariant measures for multidimensional diffusions. *Ann. Probab.* **6** 541–553. MR0494525
[7] Blumenthal, R.M. and Getoor, R.K. (1968). *Markov Processes and Potential Theory*. *Pure and Applied Mathematics*, *Vol.* 29. New York: Academic Press. MR0264757
[8] Bolley, F., Gentil, I. and Guillin, A. (2012). Convergence to equilibrium in Wasserstein distance for Fokker–Planck equations. *J. Funct. Anal.* **263** 2430–2457. MR2964689 https://doi.org/10.1016/j.jfa.2012.07.007
[9] Böttcher, B., Schilling, R. and Wang, J. (2013). *Lévy Matters. III*. *Lecture Notes in Math.* **2099**. Cham: Springer. MR3156646 https://doi.org/10.1007/978-3-319-02684-8
[10] Butkovsky, O. (2014). Subgeometric rates of convergence of Markov processes in the Wasserstein metric. *Ann. Appl. Probab.* **24** 526–552. MR3178490 https://doi.org/10.1214/13-AAP922
[11] Chen, M. (2000). Explicit bounds of the first eigenvalue. *Sci. China Ser. A* **43** 1051–1059. MR1802148 https://doi.org/10.1007/BF02898239

[12] Chen, M.-F. (2005). *Eigenvalues, Inequalities, and Ergodic Theory. Probability and Its Applications (New York)*. London: Springer. MR2105651

[13] Deng, C.-S., Schilling, R.L. and Song, Y.-H. (2017). Subgeometric rates of convergence for Markov processes under subordination. *Adv. in Appl. Probab.* **49** 162–181. MR3631220 https://doi.org/10.1017/apr.2016.83

[14] Deng, C.-S., Schilling, R.L. and Song, Y.-H. (2018). Correction: "Subgeometric rates of convergence for Markov processes under subordination" [MR3631220]. *Adv. in Appl. Probab.* **50** 1005. MR3877262 https://doi.org/10.1017/apr.2018.44

[15] Douc, R., Fort, G. and Guillin, A. (2009). Subgeometric rates of convergence of f-ergodic strong Markov processes. *Stochastic Process. Appl.* **119** 897–923. MR2499863 https://doi.org/10.1016/j.spa.2008.03.007

[16] Douc, R., Fort, G., Moulines, E. and Soulier, P. (2004). Practical drift conditions for subgeometric rates of convergence. *Ann. Appl. Probab.* **14** 1353–1377. MR2071426 https://doi.org/10.1214/105051604000000323

[17] Douc, R., Moulines, E., Priouret, P. and Soulier, P. (2018). *Markov Chains. Springer Series in Operations Research and Financial Engineering*. Cham: Springer. MR3889011 https://doi.org/10.1007/978-3-319-97704-1

[18] Down, D., Meyn, S.P. and Tweedie, R.L. (1995). Exponential and uniform ergodicity of Markov processes. *Ann. Probab.* **23** 1671–1691. MR1379163

[19] Durmus, A., Fort, G. and Moulines, É. (2016). Subgeometric rates of convergence in Wasserstein distance for Markov chains. *Ann. Inst. Henri Poincaré Probab. Stat.* **52** 1799–1822. MR3573296 https://doi.org/10.1214/15-AIHP699

[20] Durrett, R. (1996). *Stochastic Calculus. Probability and Stochastics Series*. Boca Raton, FL: CRC Press. MR1398879

[21] Eberle, A. (2011). Reflection coupling and Wasserstein contractivity without convexity. *C. R. Math. Acad. Sci. Paris* **349** 1101–1104. MR2843007 https://doi.org/10.1016/j.crma.2011.09.003

[22] Eberle, A. (2016). Reflection couplings and contraction rates for diffusions. *Probab. Theory Related Fields* **166** 851–886. MR3568041 https://doi.org/10.1007/s00440-015-0673-1

[23] Ethier, S.N. and Kurtz, T.G. (1986). *Markov Processes. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. New York: Wiley. MR0838085 https://doi.org/10.1002/9780470316658

[24] Fort, G. and Moulines, E. (2003). Polynomial ergodicity of Markov transition kernels. *Stochastic Process. Appl.* **103** 57–99. MR1947960 https://doi.org/10.1016/S0304-4149(02)00182-5

[25] Fort, G. and Roberts, G.O. (2005). Subgeometric ergodicity of strong Markov processes. *Ann. Appl. Probab.* **15** 1565–1589. MR2134115 https://doi.org/10.1214/105051605000000115

[26] Friedman, A. (1975). *Stochastic Differential Equations and Applications. Vol.* 1. New York–London: Academic Press [Harcourt Brace Jovanovich, Publishers]. MR0494490

[27] Hairer, M. (2016). *Convergence of Markov Processes. Lecture Notes, University of Warwick*. Available at http://www.hairer.org/notes/Convergence.pdf.

[28] Hairer, M., Mattingly, J.C. and Scheutzow, M. (2011). Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations. *Probab. Theory Related Fields* **149** 223–259. MR2773030 https://doi.org/10.1007/s00440-009-0250-6

[29] Has'minskiĭ, R.Z. (1960). Ergodic properties of recurrent diffusion processes and stabilization of the solution of the Cauchy problem for parabolic equations. *Teor. Veroyatn. Primen.* **5** 196–214. MR0133871

[30] Ishikawa, Y. (2001). Density estimate in small time for jump processes with singular Lévy measures. *Tohoku Math. J.* (2) **53** 183–202. MR1829978 https://doi.org/10.2748/tmj/1178207478

[31] Kevei, P. (2018). Ergodic properties of generalized Ornstein–Uhlenbeck processes. *Stochastic Process. Appl.* **128** 156–181. MR3729534 https://doi.org/10.1016/j.spa.2017.04.010

[32] Khasminskii, R. (2012). *Stochastic Stability of Differential Equations*, 2nd ed. *Stochastic Modelling and Applied Probability* **66**. Heidelberg: Springer. MR2894052 https://doi.org/10.1007/978-3-642-23280-0

[33] Knopova, V. and Kulik, A. (2018). Parametrix construction of the transition probability density of the solution to an SDE driven by α-stable noise. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** 100–140. MR3765882 https://doi.org/10.1214/16-AIHP796

[34] Knopova, V. and Schilling, R.L. (2012). Transition density estimates for a class of Lévy and Lévy-type processes. *J. Theoret. Probab.* **25** 144–170. MR2886383 https://doi.org/10.1007/s10959-010-0300-0

[35] Knopova, V. and Schilling, R.L. (2013). A note on the existence of transition probability densities of Lévy processes. *Forum Math*. **25** 125–149. MR3010850 https://doi.org/10.1515/form.2011.108

[36] Kolokoltsov, V. (2000). Symmetric stable laws and stable-like jump-diffusions. *Proc. Lond. Math. Soc.* (3) **80** 725–768. MR1744782 https://doi.org/10.1112/S0024611500012314

[37] Kolokoltsov, V.N. (2011). *Markov Processes, Semigroups and Generators. De Gruyter Studies in Mathematics* **38**. Berlin: de Gruyter. MR2780345

[38] Kulik, A. (2015). *Introduction to Ergodic Rates for Markov Chains and Processes. Lectures in Pure and Applied Mathematics* **2**. Potsdam: Potsdam Univ. Press. MR3617049

[39] Kulik, A. (2018). *Ergodic Behavior of Markov Processes. De Gruyter Studies in Mathematics* **67**. Berlin: de Gruyter. MR3791835

[40] Kulik, A.M. (2009). Exponential ergodicity of the solutions to SDE's with a jump noise. *Stochastic Process. Appl*. **119** 602–632. MR2494006 https://doi.org/10.1016/j.spa.2008.02.006

[41] Kwon, Y. and Lee, C. (1999). Strong Feller property and irreducibility of diffusions with jumps. *Stoch. Stoch. Rep*. **67** 147–157. MR1717795 https://doi.org/10.1080/17442509908834206

[42] Luo, D. and Wang, J. (2016). Exponential convergence in L^p-Wasserstein distance for diffusion processes without uniformly dissipative drift. *Math. Nachr.* **289** 1909–1926. MR3563910 https://doi.org/10.1002/mana.201500351

[43] Majka, M.B. (2016). A note on existence of global solutions and invariant measures for jump SDEs with locally one-sided lipschitz drift. Preprint. Available at https://arxiv.org/pdf/1612.03824.pdf.

[44] Majka, M.B. (2017). Coupling and exponential ergodicity for stochastic differential equations driven by Lévy processes. *Stochastic Process. Appl*. **127** 4083–4125. MR3718107 https://doi.org/10.1016/j.spa.2017.03.020

[45] Mandl, P. (1968). *Analytical Treatment of One-Dimensional Markov Processes. Die Grundlehren der Mathematischen Wissenschaften, Band* 151. Prague: Academia Publishing House of the Czechoslovak Academy of Sciences; New York: Springer. MR0247667

[46] Masuda, H. (2007). Ergodicity and exponential β-mixing bounds for multidimensional diffusions with jumps. *Stochastic Process. Appl*. **117** 35–56. MR2287102 https://doi.org/10.1016/j.spa.2006.04.010

[47] Masuda, H. (2009). Erratum to: "Ergodicity and exponential β-mixing bound for multidimensional diffusions with jumps" [Stochastic Process. Appl. 117 (2007) 35–56] [MR2287102]. *Stochastic Process. Appl*. **119** 676–678. MR2494009 https://doi.org/10.1016/j.spa.2008.02.010

[48] Meyn, S. and Tweedie, R.L. (2009). *Markov Chains and Stochastic Stability*, 2nd ed. Cambridge: Cambridge Univ. Press. MR2509253 https://doi.org/10.1017/CBO9780511626630

[49] Meyn, S.P. and Tweedie, R.L. (1993). Stability of Markovian processes. II. Continuous-time processes and sampled chains. *Adv. in Appl. Probab*. **25** 487–517. MR1234294 https://doi.org/10.2307/1427521

[50] Meyn, S.P. and Tweedie, R.L. (1993). Stability of Markovian processes. III. Foster–Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab*. **25** 518–548. MR1234295 https://doi.org/10.2307/1427522

[51] Pang, G. and Sandrić, N. (2016). Ergodicity and fluctuations of a fluid particle driven by diffusions with jumps. *Commun. Math. Sci*. **14** 327–362. MR3436243 https://doi.org/10.4310/CMS.2016.v14.n2.a2

[52] Picard, J. (1996). On the existence of smooth densities for jump processes. *Probab. Theory Related Fields* **105** 481–511. MR1402654 https://doi.org/10.1007/BF01191910

[53] Picard, J. (2010). Erratum to: On the existence of smooth densities for jump processes [MR1402654]. *Probab. Theory Related Fields* **147** 711–713. MR2639720 https://doi.org/10.1007/s00440-010-0267-x

[54] Prévôt, C. and Röckner, M. (2007). *A Concise Course on Stochastic Partial Differential Equations. Lecture Notes in Math*. **1905**. Berlin: Springer. MR2329435

[55] Rogers, L.C.G. and Williams, D. (2000). *Diffusions, Markov Processes, and Martingales. Vol. 2. Cambridge Mathematical Library*. Cambridge: Cambridge Univ. Press. MR1780932 https://doi.org/10.1017/CBO9781107590120

[56] Sandrić, N. (2016). Ergodicity of Lévy-type processes. *ESAIM Probab. Stat*. **20** 154–177. MR3528622 https://doi.org/10.1051/ps/2016009

[57] Sandrić, N. (2016). Long-time behavior for a class of Feller processes. *Trans. Amer. Math. Soc*. **368** 1871–1910. MR3449227 https://doi.org/10.1090/tran/6371

[58] Sandrić, N. (2017). A note on the Birkhoff ergodic theorem. *Results Math*. **72** 715–730. MR3684455 https://doi.org/10.1007/s00025-017-0681-9

[59] Schilling, R.L. and Schnurr, A. (2010). The symbol associated with the solution of a stochastic differential equation. *Electron. J. Probab.* **15** 1369–1393. MR2721050 https://doi.org/10.1214/EJP.v15-807

[60] Schilling, R.L., Song, R. and Vondraček, Z. (2010). *Bernstein Functions*. *De Gruyter Studies in Mathematics* **37**. Berlin: de Gruyter. MR2598208

[61] Stramer, O. and Tweedie, R.L. (1997). Existence and stability of weak solutions to stochastic differential equations with non-smooth coefficients. *Statist. Sinica* **7** 577–593. MR1467449

[62] Stroock, D.W. (1975). Diffusion processes associated with Lévy generators. *Z. Wahrsch. Verw. Gebiete* **32** 209–244. MR0433614 https://doi.org/10.1007/BF00532614

[63] Tuominen, P. and Tweedie, R.L. (1994). Subgeometric rates of convergence of f-ergodic Markov chains. *Adv. in Appl. Probab.* **26** 775–798. MR1285459 https://doi.org/10.2307/1427820

[64] Tweedie, R.L. (1994). Topological conditions enabling use of Harris methods in discrete and continuous time. *Acta Appl. Math.* **34** 175–188. MR1273853 https://doi.org/10.1007/BF00994264

[65] Veretennikov, A.Y. (1997). On polynomial mixing bounds for stochastic differential equations. *Stochastic Process. Appl.* **70** 115–127. MR1472961 https://doi.org/10.1016/S0304-4149(97)00056-2

[66] Veretennikov, A.Y. (1999). On polynomial mixing and the rate of convergence for stochastic differential and difference equations. *Teor. Veroyatn. Primen.* **44** 312–327. MR1751475 https://doi.org/10.1137/S0040585X97977550

[67] Villani, C. (2009). *Optimal Transport*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **338**. Berlin: Springer. MR2459454 https://doi.org/10.1007/978-3-540-71050-9

[68] von Renesse, M.-K. and Sturm, K.-T. (2005). Transport inequalities, gradient estimates, entropy, and Ricci curvature. *Comm. Pure Appl. Math.* **58** 923–940. MR2142879 https://doi.org/10.1002/cpa.20060

[69] Wang, F.-Y. (2011). Coupling for Ornstein–Uhlenbeck processes with jumps. *Bernoulli* **17** 1136–1158. MR2854767 https://doi.org/10.3150/10-BEJ308

[70] Wang, J. (2008). Criteria for ergodicity of Lévy type operators in dimension one. *Stochastic Process. Appl.* **118** 1909–1928. MR2454470 https://doi.org/10.1016/j.spa.2007.11.003

[71] Wang, J. (2011). Stability of Markov processes generated by Lévy-type operators. *Chinese Ann. Math. Ser. A* **32** 33–50. MR2663819

[72] Wang, J. (2016). L^p-Wasserstein distance for stochastic differential equations driven by Lévy processes. *Bernoulli* **22** 1598–1616. MR3474827 https://doi.org/10.3150/15-BEJ705

[73] Wee, I.-S. (1999). Stability for multidimensional jump-diffusion processes. *Stochastic Process. Appl.* **80** 193–209. MR1682247 https://doi.org/10.1016/S0304-4149(98)00078-7

Received March 2019 and revised November 2019