

# Subexponential upper and lower bounds in Wasserstein distance for Markov processes

Nikola Sandrić<sup>1</sup>, Ari Arapostathis<sup>2</sup> and Guodong Pang<sup>3</sup>

<sup>1</sup>*Department of Mathematics, University of Zagreb  
Bijenička cesta 30, 10000 Zagreb, Croatia  
e-mail: nsandric@math.hr*

<sup>2</sup>*Department of Electrical and Computer Engineering  
University of Texas at Austin, 2501 Speedway, EER 7.824, Austin, TX 78712  
e-mail: ari@utexas.edu*

<sup>3</sup>*Department of Computational and Applied Mathematics  
Rice University, Houston, TX 77005  
e-mail: gdpang@rice.edu*

**Abstract:** In this article, relying on Foster-Lyapunov drift conditions, we establish subexponential upper and lower bounds on the rate of convergence in the  $L^p$ -Wasserstein distance for a class of irreducible and aperiodic Markov processes. We further discuss these results in the context of Markov Lévy-type processes. In the lack of irreducibility and/or aperiodicity properties, we obtain exponential ergodicity in the  $L^p$ -Wasserstein distance for a class of Itô processes under an asymptotic flatness (uniform dissipativity) assumption. Lastly, applications of these results to specific processes are presented, including Langevin tempered diffusion processes, piecewise Ornstein–Uhlenbeck processes with jumps under constant and stationary Markov controls, and backward recurrence time chains, for which we provide a sharp characterization of the rate of convergence via matching upper and lower bounds.

**MSC 2010 subject classifications:** Primary 60J05; 60J25; secondary 60H10; 60J75.

**Keywords and phrases:** exponential and subexponential ergodicity, Wasserstein distance, Itô process, Foster–Lyapunov condition, asymptotic flatness (uniform dissipativity), Langevin diffusion process, Ornstein–Uhlenbeck process.

## 1. Introduction

One of the classical directions in the analysis of Markov processes centers around their ergodic properties. In this article, we focus on both qualitative and quantitative aspects of this problem. Let  $\mathbb{X}$  be a locally compact Polish space, i.e. a locally compact separable completely metrizable topological space. Denote the corresponding metric by  $d$ , and let  $\mathbb{T} = \mathbb{R}_+$  or  $\mathbb{Z}_+$  be the time parameter set. We endow  $(\mathbb{X}, d)$  with its Borel  $\sigma$ -algebra  $\mathfrak{B}(\mathbb{X})$ . Further, let  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in \mathbb{T}}, \{\theta_t\}_{t \in \mathbb{T}}, \{X(t)\}_{t \in \mathbb{T}}, \{\mathbb{P}_x\}_{x \in \mathbb{X}})$ , denoted by  $\{X(t)\}_{t \in \mathbb{T}}$  in the sequel, be a time-homogeneous conservative strong Markov process with càdlàg sample paths (when  $\mathbb{T} = \mathbb{R}_+$ ) and state space  $(\mathbb{X}, \mathfrak{B}(\mathbb{X}))$ , in the sense of [10]. Here,  $(\Omega, \mathcal{F}, \mathbb{P}_x)_{x \in \mathbb{X}}$  is a family of probability spaces and  $\{X(t)\}_{t \in \mathbb{T}}$  satisfies  $\mathbb{P}_x(X(0) = x) = 1$ ,  $\{\mathcal{F}_t\}_{t \in \mathbb{T}}$  is a filtration on  $(\Omega, \mathcal{F})$  (non-decreasing family of sub- $\sigma$ -algebras of  $\mathcal{F}$ ) and  $\{\theta_t\}_{t \in \mathbb{T}}$  is a family of shift operators on  $\Omega$  satisfying  $X(t) \circ \theta_s = X(t + s)$  for all  $s, t \in \mathbb{T}$ . Recall,  $\{X(t)\}_{t \in \mathbb{T}}$  is said to be conservative if  $\mathbb{P}_x(X(t) \in \mathbb{X}) = 1$  for all  $t \in \mathbb{T}$  and  $x \in \mathbb{X}$ . In the present article, we present (sharp) sufficient conditions under which  $\{X(t)\}_{t \in \mathbb{T}}$  admits a unique invariant probability measure  $\pi(dx)$ , and which ensure that the marginals of  $\{X(t)\}_{t \in \mathbb{T}}$  converge to  $\pi(dx)$ , as  $t \rightarrow \infty$ , in the  $L^p$ -Wasserstein distance at exponential and subexponential rates.

### 1.1. Summary of the results

Before stating the main results of this article, we introduce some notation we need in the sequel. Denote by  $p(t, x, dy) := \mathbb{P}_x(X(t) \in dy)$  for  $t \in \mathbb{T}$  and  $x \in \mathbb{X}$ , the transition kernel of  $\{X(t)\}_{t \in \mathbb{T}}$ . We endow  $\mathbb{T}$  with the standard (Euclidean Borel in the case when  $\mathbb{T} = \mathbb{R}_+$ , and discrete when  $\mathbb{T} = \mathbb{Z}_+$ )  $\sigma$ -algebra. The process  $\{X(t)\}_{t \in \mathbb{T}}$  is called

- (i) irreducible if there exists a  $\sigma$ -finite measure  $\varphi(dx)$  on  $\mathfrak{B}(\mathbb{X})$  such that whenever  $\varphi(B) > 0$  we have  $\int_{\mathbb{T}} p(t, x, B) \tau(dt) > 0$  for all  $x \in \mathbb{X}$ , where  $\tau(dt)$  stands for the Lebesgue measure on  $\mathbb{T}$  when  $\mathbb{T} = \mathbb{R}_+$ , and the counting measure when  $\mathbb{T} = \mathbb{Z}_+$ ;

- (ii) transient if it is irreducible, and there exist  $\{b_k\}_{k \in \mathbb{N}} \subset [0, \infty)$  and a covering  $\{B_k\}_{k \in \mathbb{N}} \subseteq \mathfrak{B}(\mathbb{X})$  of  $\mathbb{X}$ , such that  $\int_{\mathbb{T}} p(t, x, B_k) \tau(dt) \leq b_k$  for all  $x \in \mathbb{X}$  and  $k \in \mathbb{N}$ ;
- (iii) recurrent if it is irreducible, and  $\varphi(B) > 0$  implies that  $\int_{\mathbb{T}} p(t, x, B) \tau(dt) = \infty$  for all  $x \in \mathbb{X}$ ;
- (iv) aperiodic if there exists  $t_0 > 0$  such that  $\{X_{kt_0}\}_{k \in \mathbb{Z}_+}$  is irreducible, in the case when  $\mathbb{T} = \mathbb{R}_+$ ; and there does not exist a partition  $\{B_1, \dots, B_k\} \subseteq \mathfrak{B}(\mathbb{X})$  with  $k \geq 2$  of  $\mathbb{X}$  such that  $p(1, x, B_{i+1}) = 1$  for all  $x \in B_i$  and all  $1 \leq i \leq k-1$ , and  $p(1, x, B_1) = 1$  for all  $x \in B_k$ , in the case when  $\mathbb{T} = \mathbb{Z}_+$ .

Let us remark that if  $\{X(t)\}_{t \in \mathbb{T}}$  is irreducible, then it is either transient or recurrent (see [79, Theorem 2.3]). A Borel measure  $\pi(dx)$  on  $\mathbb{X}$  is called invariant for  $\{X(t)\}_{t \in \mathbb{T}}$  if  $\int_{\mathbb{X}} p(t, x, dy) \pi(dx) = \pi(dy)$  for all  $t \in \mathbb{T}$ . It is well known that if  $\{X(t)\}_{t \in \mathbb{T}}$  is recurrent, then it possesses a unique (up to constant multiples) invariant measure (see [79, Theorem 2.6]). If the invariant measure is finite, then it may be normalized to a probability measure. If  $\{X(t)\}_{t \in \mathbb{T}}$  is recurrent with finite invariant measure, then it is called positive recurrent; otherwise it is called null recurrent. Note that a transient Markov process cannot have a finite invariant measure. A set  $C \in \mathfrak{B}(\mathbb{X})$  is called petite for  $\{X(t)\}_{t \in \mathbb{T}}$  if there exist a probability measure  $\chi(dt)$  on  $\mathbb{T}$  and a non-trivial Borel measure  $\nu_\chi(dx)$  on  $\mathbb{X}$ , such that

$$\int_{\mathbb{T}} p(t, x, B) \chi(dt) \geq \nu_\chi(B)$$

for all  $x \in C$  and  $B \in \mathfrak{B}(\mathbb{X})$ . Recall that petite sets play a role of singletons for Markov processes on general state spaces (see [64, Chapter 5] for a detailed discussion). Denote by  $\mathcal{P}(\mathbb{X})$  the class of all Borel probability measures on  $\mathbb{X}$ , and for  $f \in \mathcal{B}(\mathbb{X})$  (the space of real-valued Borel measurable functions on  $\mathbb{X}$ ) let  $\mathcal{P}_f(\mathbb{X})$  denote the class of all  $\mu \in \mathcal{P}(\mathbb{X})$  with the property that  $\int_{\mathbb{X}} |f(x)| \mu(dx) < \infty$ . When  $f(x) = (\mathbf{d}(x_0, x))^p$  for some  $p > 0$  and  $x_0 \in \mathbb{X}$ , we denote this as  $\mathcal{P}_p(\mathbb{X})$ . We adopt the usual notation

$$\mu P_t(dy) = \int_{\mathbb{X}} p(t, x, dy) \mu(dx), \quad \text{and} \quad \mu(f) = \int_{\mathbb{X}} f(x) \mu(dx)$$

for  $t \in \mathbb{T}$ ,  $x \in \mathbb{X}$ ,  $\mu \in \mathcal{P}(\mathbb{X})$  and  $f \in \mathcal{B}(\mathbb{X})$ . Therefore, with  $\delta_x$  denoting the Dirac measure concentrated at  $x \in \mathbb{X}$ , we have  $\delta_x P_t(dy) = p(t, x, dy)$ . Finally, recall that the  $L^p$ -Wasserstein distance on  $\mathcal{P}_p(\mathbb{X})$  with  $p \geq 1$  is defined by

$$\mathcal{W}_p(\mu_1, \mu_2) := \inf_{\Pi \in \mathcal{C}(\mu_1, \mu_2)} \left( \int_{\mathbb{X} \times \mathbb{X}} (\mathbf{d}(x, y))^p \Pi(dx, dy) \right)^{1/p},$$

where  $\mathcal{C}(\mu_1, \mu_2)$  is the family of couplings of  $\mu_1(dx)$  and  $\mu_2(dx)$ , i.e.  $\Pi \in \mathcal{C}(\mu_1, \mu_2)$  if, and only if,  $\Pi(dx, dy)$  is a probability measure on  $\mathbb{X} \times \mathbb{X}$  having  $\mu_1(dx)$  and  $\mu_2(dx)$  as its marginals. It is well known that  $\mathcal{P}_p(\mathbb{X})$  is a complete separable metric space under the metric  $\mathcal{W}_p$  [82, Theorem 6.18]. The topology generated by  $\mathcal{W}_p$  on  $\mathcal{P}_p(\mathbb{X})$  is finer than the Prokhorov topology, i.e. the topology of weak convergence.

We now state the main results of this article.

**Theorem 1.1.** *Suppose that  $\{X(t)\}_{t \in \mathbb{T}}$  is irreducible and aperiodic, and there exist a continuous  $\mathcal{V}: \mathbb{X} \rightarrow [1, \infty)$ , a constant  $b > 0$ , a nondecreasing differentiable concave function  $\phi: [1, \infty) \rightarrow (0, \infty)$ , and a (topologically) closed petite set  $C \subseteq \mathbb{X}$  such that*

$$\mathbb{E}_x[\mathcal{V}(X(t))] - \mathcal{V}(x) \leq b \int_{[0, t]} \mathbb{E}_x[\mathbf{1}_C(X(s))] \tau(ds) - \int_{[0, t]} \mathbb{E}_x[\phi \circ \mathcal{V}(X(s))] \tau(ds) \quad (1.1)$$

for all  $(t, x) \in T \times \mathbb{X}$ . Assume further that  $\sup_{x \in C} \mathcal{V}(x) < \infty$ , and

$$c := \inf_{x \in \mathbb{X}} \frac{\phi \circ \mathcal{V}(x)}{(1 + \mathbf{d}(x, x_0))^\eta} > 0 \quad (1.2)$$

for some  $\eta \geq 1$  and some (and therefore any)  $x_0 \in \mathbb{X}$ . Then  $\{X(t)\}_{t \in \mathbb{T}}$  admits a unique invariant  $\pi \in \mathcal{P}_{\phi \circ \mathcal{V}}(\mathbb{X})$ . In addition, with  $\Phi(t) := \int_1^t \frac{ds}{\phi(s)}$  and  $r(t) := \phi \circ \Phi^{-1}(t)$ , the following hold.

- (i) If  $\lim_{t \rightarrow \infty} \phi'(t) = 0$ , then for some  $\bar{c} > 0$  we have

$$\left(1 \vee (r(t))^{(\eta-1)/\eta}\right) \mathcal{W}_1(\delta_x P_t, \pi) \leq \bar{c} \mathcal{V}(x) \quad \forall (t, x) \in \mathbb{T} \times \mathbb{X}, \quad (1.3)$$

and

$$\int_{\mathbb{T}} \left(1 \vee (r(t))^{(n-1)/\eta}\right) \mathcal{W}_1(\delta_x P_t, \delta_y P_t) \tau(dt) \leq \bar{c} (\mathcal{V}(x) + \mathcal{V}(y)) \quad \forall x, y \in \mathbb{X}. \quad (1.4)$$

(ii) If  $\lim_{t \rightarrow \infty} \phi'(t) = 0$ , then for any  $p \in [1, \eta]$  there exists  $\tilde{c} > 0$  such that

$$\left(1 \vee \left(t^{(\eta-p)/p} \wedge t^{(1-p)/p}\right) (r(t))^{(n-1)/p\eta}\right) \mathcal{W}_p(\delta_x P_t, \pi) \leq \tilde{c} (\mathcal{V}(x) + \bar{m}_\eta) \quad (1.5)$$

for all  $(t, x) \in \mathbb{T} \times \mathbb{X}$ , where  $\bar{m}_\eta = \pi(\mathbf{d}(x_0, \cdot)^\eta)$ .

(iii) If  $\phi(t) = \hat{c}t$  for some  $\hat{c} > 0$ , then there exist  $\check{c} > 0$  and  $\gamma > 0$ , such that

$$e^{\gamma t} \mathcal{W}_1(\delta_x P_t, \pi) \leq \check{c} \mathcal{V}(x) \quad \forall (t, x) \in \mathbb{T} \times \mathbb{X}. \quad (1.6)$$

In addition, for any  $p \in [1, \eta]$  there exists  $\check{c} > 0$  such that

$$\left(1 \vee t^{\eta/p-1}\right) \mathcal{W}_p(\delta_x P_t, \pi) \leq \check{c} (\mathcal{V}(x) + \bar{m}_\eta)^{1/p} \quad \forall (t, x) \in \mathbb{T} \times \mathbb{X}. \quad (1.7)$$

The results in [Theorem 1.1](#) should be compared to equations (2.3) and (2.5) in [[12](#), Theorems 2.1 and 2.4] (see also [[28](#), Theorem 3 (ii)] and [[53](#), Chapter 4]). The underlying metric  $\mathbf{d}$  is assumed to be bounded in [[12](#)]. The starting point is a Foster-Lyapunov condition of the form in (1.1), and the irreducibility and aperiodicity assumptions are replaced by a closely related structural property: the metric  $\mathbf{d}$  is contracting, and the sublevel sets of  $(x, y) \mapsto \mathcal{V}(x) + \mathcal{V}(y)$  are  $\mathbf{d}$ -small (see (3) and (4) in [[12](#), Theorems 2.1 and 2.4]). Then an analogous estimate to (1.3) holds for the corresponding  $\mathcal{W}_1$ -distance. Observe that when  $\mathbf{d}$  is bounded, the relation in (1.1) trivially holds for any  $\eta \geq 0$ . Provided  $\{X(t)\}_{t \in \mathbb{T}}$  is irreducible and aperiodic, this gives an analogous result to the one obtained in [[12](#), Theorems 2.1 and 2.4] (in  $\mathcal{W}_1$ -distance) without assuming either contraction properties of  $\mathbf{d}$  or  $\mathbf{d}$ -smallness of the sublevel sets of  $(x, y) \mapsto \mathcal{V}(x) + \mathcal{V}(y)$ . The proof of [Theorem 1.1](#) relies on [[24](#), Theorem 3.2] and [[25](#), Theorem 2.8], where, under the assumptions of [Theorem 1.1](#), the authors show ergodicity of  $\{X(t)\}_{t \in \mathbb{T}}$  in the  $f$ -norm with rate  $\Psi_1 \circ r(t)$  and  $f(x) = \Psi_2 \circ \phi \circ \mathcal{V}(x) \vee 1$ , for any pair  $(\Psi_1^{-1}, \Psi_2^{-1})$  of Young's functions. Recall, for a signed Borel measure  $\mu(dx)$  on  $\mathbb{X}$  and a function  $f: \mathbb{X} \rightarrow [1, \infty]$  the so-called  $f$ -norm of  $\mu(dx)$  is defined as

$$\|\mu\|_f := \sup_{g \in \mathcal{B}(\mathbb{X}), |g| \leq f} |\mu(g)|, \quad (1.8)$$

generalizing the usual total variation norm  $\|\mu\|_{\text{TV}} := \sup_{g \in \mathcal{B}(\mathbb{X}), |g| \leq 1} |\mu(g)|$ . We remark here that convergence in the  $f$ -norm does not in general imply convergence in the  $\mathcal{W}_p$ -distance, and *vice versa* (see [Section 3](#) for examples of such Markov processes).

In the following theorem we establish a lower bound for  $\mathcal{W}_p$ -convergence, which matches the upper bounds obtained in (1.3) and (1.5). For  $\gamma \in C([0, 1]; \mathbb{X})$  (the space of continuous mappings from  $[0, 1]$  to  $\mathbb{X}$ ) let

$$\Lambda(\gamma) := \sup_{k \in \mathbb{N}} \sup_{0=u_0 < u_1 < \dots < u_{k-1} < u_k=1} \left( \mathbf{d}(\gamma(u_0), \gamma(u_1)) + \dots + \mathbf{d}(\gamma(u_{k-1}), \gamma(u_k)) \right).$$

The space  $\mathbb{X}$  is called a length space if

$$\mathbf{d}(x, y) = \inf_{\gamma \in C([0, 1]; \mathbb{X})} \{\Lambda(\gamma) : \gamma(0) = x, \gamma(1) = y\} \quad \forall x, y \in \mathbb{X}.$$

**Theorem 1.2.** *Assume that  $\mathbb{X}$  is a length space,  $\{X(t)\}_{t \in \mathbb{T}}$  satisfies (1.1), and there exist a Lipschitz continuous function  $L: \mathbb{X} \rightarrow [0, \infty)$  and constants  $\theta > \vartheta \geq 1$  and  $c > 0$ , such that*

$$\mathcal{V}(x) \geq c(L(x))^\theta, \quad \text{and} \quad \phi \circ \mathcal{V}(x) \geq c(L(x))^\vartheta \quad \forall x \in \mathbb{X}.$$

*In addition, suppose that  $\{X(t)\}_{t \in \mathbb{T}}$  admits an invariant  $\pi \in \mathcal{P}(\mathbb{X})$  such that  $\int_{\mathbb{X}} (L(x))^{\vartheta+\varepsilon} \pi(dx) = \infty$  for some  $\varepsilon \in (0, \theta - \vartheta)$ . Then, for each  $p \in [1, \vartheta]$ ,  $\iota \in (0, \theta - \vartheta - \varepsilon)$  and  $x \in \mathbb{X}$ , there exist a constant  $\bar{c} > 0$  and a diverging increasing sequence  $\{t_n\}_{n \in \mathbb{N}} \subseteq \mathbb{T}$ , depending on these parameters, such that*

$$\mathcal{W}_p(\delta_x P_{t_n}, \pi) \geq \bar{c} (t_n + \mathcal{V}(x))^{-\frac{\vartheta-p+\varepsilon+\iota}{(\theta-\vartheta-\varepsilon-\iota)p}} \quad \forall n \in \mathbb{N}. \quad (1.9)$$

Note that the parameters  $\theta$ ,  $\vartheta$ ,  $\varepsilon$ ,  $p$  and  $\iota$  are such that the exponent in the above expression is always strictly negative. Obtaining lower bound for the convergence in the total variation norm is discussed in [[38](#), Theorem 5.1 and Corollary 5.2]. Applications of [Theorem 1.2](#) are discussed in [Section 3](#).

### 1.2. Ergodicity of a class of Lévy-type processes

Here, we discuss ergodic properties of a class of Markov processes on the Euclidean space  $\mathbb{R}^n$  (endowed with the standard Euclidean metric) generated by a (Lévy-type) operator  $\mathcal{L}: \mathcal{D}(\mathcal{L}) \subseteq \mathcal{B}(\mathbb{R}^n) \rightarrow \mathcal{B}(\mathbb{R}^n)$  given by

$$\mathcal{L}f(x) = \langle b(x), \nabla f(x) \rangle + \frac{1}{2} \text{Tr}(a(x) \nabla^2 f(x)) + \int_{\mathbb{R}^n} \mathfrak{d}_1 f(x; y) \nu(x, dy), \quad x \in \mathbb{R}^n. \quad (1.10)$$

Here,  $b = (b_i)_{i=1, \dots, n}: \mathbb{R}^n \rightarrow \mathbb{R}^n$  is Borel measurable,  $a = (a_{ij})_{1 \leq i, j \leq n}: \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$  is a symmetric non-negative definite  $n \times n$  matrix-valued Borel measurable function,  $\nu(x, dy)$  is a nonnegative Borel kernel on  $\mathbb{R}^n \times \mathfrak{B}(\mathbb{R}^n)$ , called the Lévy kernel, satisfying

$$\nu(x, \{0\}) = 0, \quad \text{and} \quad \int_{\mathbb{R}^n} (1 \wedge |y|^2) \nu(x, dy) < \infty \quad \forall x \in \mathbb{R}^n,$$

and

$$\mathfrak{d}_1 f(x; y) := f(x + y) - f(x) - \mathbf{1}_{\mathfrak{B}}(y) \langle y, \nabla f(x) \rangle, \quad x, y \in \mathbb{R}^n, \quad f \in C^1(\mathbb{R}^n).$$

The symbol  $\mathcal{D}(\mathcal{L})$  stands for the domain of  $\mathcal{L}$ , i.e. the set of functions  $f \in \mathcal{B}(\mathbb{R}^n)$  for which (1.10) is well defined,  $\langle \cdot, \cdot \rangle$  and  $|\cdot|$  denote the standard inner product and the corresponding Euclidean norm on  $\mathbb{R}^n$ ,  $\text{Tr } M$  stands for the trace of a square matrix  $M$ , and  $\nabla^2 f(x)$  denotes the Hessian of  $f \in C^2(\mathbb{R}^n)$ . An open (resp. closed) ball of radius  $r > 0$  centered at  $x$  is denoted by  $\mathfrak{B}_r(x)$  (resp.  $\overline{\mathfrak{B}}_r(x)$ ). If  $x = 0$ , we write  $\mathfrak{B}_r$  (resp.  $\overline{\mathfrak{B}}_r$ ), and the unit open (resp. closed) ball centered at 0 is denoted by  $\mathfrak{B}$  (resp.  $\overline{\mathfrak{B}}$ ). Observe that  $C_b^k(\mathbb{R}^n) \subseteq \mathcal{D}(\mathcal{L})$ , where  $C_b^k(\mathbb{R}^n)$ ,  $k \geq 0$ , denotes the space of  $k$  times differentiable functions such that all derivatives up to order  $k$  are bounded. We also denote by  $\|M\| := (\text{Tr } MM')^{1/2}$  the Hilbert–Schmidt norm of a matrix  $M$ , where  $M'$  stands for the transpose of  $M$ .

We introduce the following assumption:

**(MP)** There exists a conservative strong Markov process  $\{X(t)\}_{t \geq 0}$  with càdlàg sample paths such that

$$M_f(t) := f(X(t)) - f(X(0)) - \int_0^t \mathcal{L}f(X(s)) \, ds, \quad t \geq 0, \quad (1.11)$$

is a  $\mathbb{P}_x$ -martingale (with respect to  $\{\mathcal{F}_t\}_{t \geq 0}$ ) for any  $f \in C_c^\infty(\mathbb{R}^n)$  (the space of smooth functions with compact support).

Define

$$q(x, \xi) := -i \langle \xi, b(x) \rangle + \frac{1}{2} \langle \xi, a(x) \xi \rangle + \int_{\mathbb{R}^n} (1 - e^{i \langle \xi, y \rangle} + i \langle \xi, y \rangle \mathbf{1}_{\mathfrak{B}}(y)) \nu(x, dy), \quad x, \xi \in \mathbb{R}^n,$$

and observe that

$$\mathcal{L}f(x) = - \int_{\mathbb{R}^n} e^{i \langle \xi, x \rangle} q(x, \xi) \hat{f}(\xi) \, d\xi$$

for all  $x \in \mathbb{R}^n$  and  $f \in C_c^\infty(\mathbb{R}^n)$ , where  $\hat{f}(\xi) := (2\pi)^{-n} \int_{\mathbb{R}^n} e^{-i \langle \xi, x \rangle} f(x) \, dx$  denotes the Fourier transform of  $f(x)$ . In other words,  $\mathcal{L}$  is a pseudo-differential operator with symbol  $q(x, \xi)$ . According to [52, Theorem 1.1], **(MP)** is satisfied if

**(LB)** The functions  $b(x)$ ,  $a(x)$ , and  $x \mapsto \int_{\mathbb{R}^n} (1 \wedge |y|^2) \nu(x, dy)$  are locally bounded.

**(SG)**  $x \mapsto q(x, \xi)$  is continuous for all  $\xi \in \mathbb{R}^n$ , and  $q(x, \xi)$  is locally uniformly continuous at  $\xi = 0$ , i.e.

$$\lim_{\rho \rightarrow \infty} \sup_{x \in \mathfrak{B}_\rho} \sup_{\xi \in \mathfrak{B}_{1/\rho}} |q(x, \xi)| = 0.$$

Observe that the second condition in **(SG)** essentially means that the coefficients  $b(x)$ ,  $a(x)$ , and  $\nu(x, dy)$  have a sublinear growth. Namely, it is satisfied if

$$\lim_{\rho \rightarrow \infty} \left( \frac{\sup_{x \in \mathfrak{B}_\rho} |b(x)|}{\rho} + \frac{\sup_{x \in \mathfrak{B}_\rho} |a(x)|}{\rho^2} + \frac{\sup_{x \in \mathfrak{B}_\rho} \int_{\mathfrak{B}} |y|^2 \nu(x, dy)}{\rho^2} + \sup_{x \in \mathfrak{B}_\rho} \sup_{\xi \in \mathfrak{B}_{1/\rho}} \int_{\mathfrak{B}^c} (1 - e^{i \langle \xi, y \rangle}) \nu(x, dy) \right) = 0.$$

In order to allow linear growth of the coefficients, we replace (LB) and (SG) by (LG)  $\mathcal{L}(C_c^\infty(\mathbb{R}^n)) \subseteq C_\infty(\mathbb{R}^n)$ ,  $x \mapsto q(x, \xi)$  is continuous for all  $\xi \in \mathbb{R}^n$ , and

$$\limsup_{|x| \rightarrow \infty} \sup_{\xi \in \mathcal{B}_{1/|x|}} |q(x, \xi)| < \infty$$

(see [50, Corollary 3.2]).

Here,  $C_\infty(\mathbb{R}^n)$  stands for the space of continuous functions vanishing at infinity. Clearly, the last condition in (LG) follows from

$$\limsup_{|x| \rightarrow \infty} \left( \frac{|b(x)|}{|x|} + \frac{\|a(x)\|}{|x|^2} + \frac{\int_{\mathcal{B}} |y|^2 \nu(x, dy)}{|x|^2} + \nu(x, \mathcal{B}^c) \right) < \infty. \quad (1.12)$$

Let us also remark that due to [51, Theorem A1] the map  $x \mapsto q(x, \xi)$  is continuous for all  $\xi \in \mathbb{R}^n$  if  $b(x)$  and  $a(x)$  are continuous, and for any  $r > 0$ ,  $x \in \mathbb{R}^n$  and  $f \in C_c(\mathbb{R}^n \setminus \{0\})$ ,

$$\lim_{\rho \rightarrow \infty} \sup_{y \in \mathcal{B}_r} \nu(y, \mathcal{B}_\rho^c) = 0, \quad \lim_{\rho \rightarrow 0} \sup_{y \in \mathcal{B}_r} \int_{\mathcal{B}_\rho} |z|^2 \nu(y, dz) = 0,$$

and

$$\lim_{y \rightarrow x} \int_{\mathbb{R}^n} f(z) \nu(y, dz) = \int_{\mathbb{R}^n} f(z) \nu(x, dz).$$

Furthermore, under the continuity of  $x \mapsto q(x, \xi)$  (for all  $\xi \in \mathbb{R}^n$ ) in the same reference it has been shown that  $\mathcal{L}(C_c^\infty(\mathbb{R}^n)) \subseteq C_b(\mathbb{R}^n)$ . In addition, if

$$\lim_{|x| \rightarrow \infty} \nu(x, \mathcal{B}_r(-x)) = 0 \quad \forall r > 0,$$

we easily see that  $\mathcal{L}(C_c^\infty(\mathbb{R}^n)) \subseteq C_\infty(\mathbb{R}^n)$ .

**Definition 1.1.** Let  $\mathcal{M}_+$  denote the class of positive definite matrices in  $\mathbb{R}^{n \times n}$ . For  $Q \in \mathcal{M}_+$ , let  $|x|_Q := \langle x, Qx \rangle^{1/2}$  for  $x \in \mathbb{R}^n$ , and  $\chi_Q \in C^\infty(\mathbb{R}^n)$  be some nonnegative, symmetric convex function such that  $\chi_Q(x) = |x|_Q$  for  $x \in \mathcal{B}^c$ . For  $Q \in \mathcal{M}_+$  and  $\zeta > 0$ , we define

$$\mathcal{V}_{Q, \zeta}(x) := (\chi_Q(x))^\zeta, \quad \text{and} \quad \tilde{\mathcal{V}}_{Q, \zeta}(x) := e^{\zeta \chi_Q(x)}, \quad x \in \mathbb{R}^n.$$

Further, let

$$\Theta_\nu := \left\{ \theta \geq 0 : \sup_{x \in \mathbb{R}^n} \int_{\mathbb{R}^n} (|y|^2 \mathbb{1}_{\mathcal{B}}(y) + |y|^\theta \mathbb{1}_{\mathcal{B}^c}(y)) \nu(x, dy) < \infty \right\},$$

and when  $\Theta_\nu \neq \emptyset$ , let  $\theta_\nu := \sup \Theta_\nu$ .

We now discuss ergodic properties of the Lévy-type process  $\{X(t)\}_{t \geq 0}$ .

**Theorem 1.3.** Assume (LB) and (MP), and suppose that  $\{X(t)\}_{t \geq 0}$  is irreducible and aperiodic, and that every compact set is petite for  $\{X(t)\}_{t \geq 0}$ . Then the following hold.

(i) If  $\theta_\nu > 0$ ,

$$\lim_{r \rightarrow \infty} \sup_{x \in \mathbb{R}^n} \int_{\mathcal{B}_r^c} |y|^\theta \nu(x, dy) = 0 \quad (1.13)$$

for some  $\theta \in (0, \theta_\nu] \cap \Theta_\nu$ , and there exist  $Q \in \mathcal{M}_+$  and  $\vartheta \in [0 \vee (2 - \theta), 2)$  such that

$$\limsup_{|x| \rightarrow \infty} \frac{\|a(x)\|}{|x|^\vartheta} = 0, \quad \text{and} \quad \limsup_{|x| \rightarrow \infty} \frac{\langle b(x) + \mathbb{1}_{[1, \infty)}(\theta) \int_{\mathcal{B}^c} y \nu(x, dy), Qx \rangle}{|x|^\vartheta} < 0,$$

then  $\{X(t)\}_{t \geq 0}$  admits a unique invariant  $\pi \in \mathcal{P}_{\theta-2+\vartheta}(\mathbb{R}^n)$ . In addition, if  $\theta - 3 + \vartheta \geq 0$ , then Theorem 1.1 (i) and (ii) hold with  $\mathcal{V}(x) = \mathcal{V}_{Q, \theta}(x) + 1$ ,  $\phi(t) = t^{(\theta-2+\vartheta)/\theta}$  and  $\eta = \theta - 2 + \vartheta$ .

(ii) If  $\theta_\nu > 0$ , (1.13) holds for some  $\theta \in (0, \theta_\nu] \cap \Theta_\nu$ , and there exists  $Q \in \mathcal{M}_+$  such that

$$\limsup_{|x| \rightarrow \infty} \frac{\|a(x)\|}{|x|^2} = 0, \quad \text{and} \quad \limsup_{|x| \rightarrow \infty} \frac{\langle b(x) + \mathbf{1}_{[1, \infty)}(\theta) \int_{\mathbb{B}^c} y \nu(x, dy), Qx \rangle}{|x|^2} < 0,$$

then  $\{X(t)\}_{t \geq 0}$  admits a unique invariant  $\pi \in \mathcal{P}_\theta(\mathbb{R}^n)$ . In addition, if  $\theta \geq 1$ , then the conclusion of Theorem 1.1 (iii) holds with  $\mathcal{V}(x) = \mathcal{V}_{Q, \theta}(x) + 1$  and  $\eta = \theta$ .

(iii) Suppose that  $a(x)$  is bounded, and there exist  $\theta > 0$  and  $Q \in \mathcal{M}_+$ , such that

$$\sup_{x \in \mathbb{R}^n} \int_{\mathbb{R}^n} (|y|^2 \mathbf{1}_{\mathbb{B}}(y) + e^{\theta|y|} \mathbf{1}_{\mathbb{B}^c}(y)) \nu(x, dy) < \infty, \quad (1.14)$$

and

$$\limsup_{|x| \rightarrow \infty} \frac{\langle b(x) + \int_{\mathbb{B}^c} y \nu(x, dy), Qx \rangle}{|x|} < 0.$$

Then the conclusion of Theorem 1.1 (iii) holds with  $\mathcal{V}(x) = \tilde{\mathcal{V}}_{Q, \zeta}(x)$  for any  $\zeta > 0$  sufficiently small and any  $\eta \geq 1$ .

Irreducibility and aperiodicity are crucial structural properties of the underlying process in Theorems 1.1 to 1.3. Roughly speaking, they ensure that the process does not show singular behavior in its motion, and together with the Foster-Lyapunov condition in (1.1) (which ensures controllability of the  $\phi \circ \Phi^{-1}$ -modulated moment of return-times to the petite set  $C$ , see [24, Theorem 4.1]) they lead to the ergodic properties stated.

Under an asymptotic flatness (uniform dissipativity) property (see (1.16)), we use a completely different approach to this problem, the so-called synchronous coupling method (see [14, Example 2.16] for details), to obtain ergodic properties for a class of Itô processes which are not necessarily irreducible and aperiodic. Recall that an Itô process is a solution to a stochastic differential equation (SDE) of the following form

$$\begin{aligned} X(t) &= x + \int_0^t b(X(s)) ds + \int_0^t \sigma(X(s)) dB_s \\ &\quad + \int_0^t \int_{\{w: |k(X(s-), w)| < 1\}} k(X(s-), v) (\nu_p(dv, ds) - \nu(dv) ds) \\ &\quad + \int_0^t \int_{\{w: |k(X(s-), w)| \geq 1\}} k(X(s-), v) \nu_p(dv, ds), \quad (t, x) \in [0, \infty) \times \mathbb{R}^n, \end{aligned} \quad (1.15)$$

where  $b: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $\sigma: \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$  and  $k: \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$  are Borel measurable,  $\{B(t)\}_{t \geq 0}$  is a standard  $n$ -dimensional Brownian motion, and  $\nu_p(dv, ds)$  is a Poisson random measure on  $\mathfrak{B}(\mathbb{R}) \otimes \mathfrak{B}([0, \infty))$ , with intensity measure  $\nu(dv) ds$  (a  $\sigma$ -finite measure on  $\mathfrak{B}(\mathbb{R}) \otimes \mathfrak{B}(\mathbb{R})$ ). According to [13, Theorem 3.33], every Itô process is a semimartingale Hunt process. In particular, it is a conservative strong Markov process with càdlàg sample paths. Conversely, again by [13, Theorem 3.33], for every  $n$ -dimensional semimartingale Hunt process  $\{X(t)\}_{t \geq 0}$ , and every  $\sigma$ -finite nonfinite and nonatomic measure  $\nu(dv)$  on  $\mathfrak{B}(\mathbb{R})$ , there exist  $b(x)$ ,  $\sigma(x)$ ,  $k(x, v)$ ,  $\{B(t)\}_{t \geq 0}$ , and  $\nu_p(dv, ds)$  as above (possibly defined on an enlargement of the initial stochastic basis), such that  $\{X(t)\}_{t \geq 0}$  satisfies (1.15). By setting

$$\nu_p(dy, ds) = \nu_p(\{(v, u) \in \mathbb{R} \times [0, \infty): (k(X(u-), v), u) \in (dy, ds)\}),$$

and

$$\nu(x, dy) = \nu(\{u \in \mathbb{R}: k(x, u) \in dy\}),$$

(1.15) reads as

$$\begin{aligned} X(t) &= x + \int_0^t b(X(s)) ds + \int_0^t \sigma(X(s)) dB_s + \int_0^t \int_{\mathbb{B}} y (\nu_p(dy, ds) - \nu(X(s-), dy) ds) \\ &\quad + \int_0^t \int_{\mathbb{B}^c} y \nu_p(dy, ds), \quad (t, x) \in [0, \infty) \times \mathbb{R}^n. \end{aligned}$$

Set  $a(x) := \sigma(x)\sigma(x)'$ , and let  $\mathcal{L}$  be as in (1.10). According to [40, Theorem II.2.42] (with  $h(x) = x\mathbb{1}_{\mathfrak{B}}(x)$ ), for any  $f \in C_b^2(\mathbb{R}^n)$ , the process  $\{M_f(t)\}_{t \geq 0}$ , defined as in (1.11), is a  $\mathbb{P}_x$ -local martingale for every  $x \in \mathbb{R}^n$ . In addition, if (LB) holds true, then  $\{M_f(t)\}_{t \geq 0}$  is a  $\mathbb{P}_x$ -local martingale for every  $f \in C_c^\infty(\mathbb{R}^n)$  and every  $x \in \mathbb{R}^n$ , i.e. (MP) is satisfied.

For  $x, z \in \mathbb{R}^n$  define

$$\begin{aligned} \Delta_z b(x) &:= b(x+z) - b(x), \quad \Delta_z \sigma(x) := \sigma(x+z) - \sigma(x), \quad \Delta_z \nu(x, dy) := \nu(x+z, dy) - \nu(x, dy), \\ \Delta_z \tilde{b}(x) &:= \Delta_z b(x) + \int_{\mathfrak{B}^c} y \Delta_z \nu(x, dy), \quad \text{and} \quad \tilde{a}(x; z) := \Delta_z \sigma(x) \Delta_z \sigma(x)'. \end{aligned}$$

If  $b(x) \equiv b$  (resp.  $\sigma(x) \equiv \sigma$ , or  $\nu(x, dy) \equiv \nu(dy)$ ), then of course  $\Delta_z b(x)$  (resp.  $\Delta_z \sigma(x)$ , or  $\Delta_z \nu(x, dy)$ ) is equal to zero.

**Theorem 1.4.** *Assume that  $b(x)$  and  $a(x)$  are locally bounded and satisfy the linear growth condition in (1.12), and that  $\nu(x, dy)$  is such that  $2 \in \Theta_\nu$ . If for some  $p \in [2, \theta_\nu] \cap \Theta_\nu$  there exist  $Q \in \mathcal{M}_+$ , and a  $\sigma$ -finite nonfinite and nonatomic measure  $\nu(dv)$  on  $\mathfrak{B}(\mathbb{R})$  such that (1.15) admits a unique strong solution  $\{X(t)\}_{t \geq 0}$ , and*

$$\begin{aligned} 2 \langle \Delta_z \tilde{b}(x), Qz \rangle + \text{Tr}(\tilde{a}(x; z)Q) + (p-2) \|\sqrt{Q} \Delta_z \sigma(x)\|^2 \\ + 2^{p-3} (1 + (p-2) \|Q^{-1}\|) \int_{\mathbb{R}} |k(x+z, v) - k(x, v)|_Q^2 \nu(dv) \\ + \frac{2^{p-2}}{p(p-1)} (1 + (p-2) \|Q^{-1}\|) |z|_Q^{2-p} \int_{\mathbb{R}} |k(x+z, v) - k(x, v)|_Q^p \nu(dv) \leq -\frac{2c(p)}{p} |z|_Q^2 \end{aligned} \quad (1.16)$$

for some  $c(p) > 0$  and all  $x, z \in \mathbb{R}^n$ , where  $k: \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$  is given in (1.15), then

$$\mathcal{W}_p(\delta_x P_t, \delta_y P_t) \leq \left( \frac{\bar{\lambda}_Q}{\underline{\lambda}_Q} \right)^{1/2} |x - y| e^{-\frac{c(p)t}{p}} \quad (1.17)$$

for all  $t \geq 0$  and  $x, y \in \mathbb{R}^n$ , where  $\bar{\lambda}_Q$  ( $\underline{\lambda}_Q$ ) stands for the largest (smallest) eigenvalue of  $Q$ . Furthermore,  $\{X(t)\}_{t \geq 0}$  admits a unique invariant  $\pi \in \mathcal{P}_p(\mathbb{R}^n)$ , and

$$\mathcal{W}_p(\mu P_t, \pi) \leq \left( \frac{\bar{\lambda}_Q}{\underline{\lambda}_Q} \right)^{1/2} \mathcal{W}_p(\mu, \pi) e^{-\frac{c(p)t}{p}} \quad (1.18)$$

for all  $t \geq 0$  and  $\mu \in \mathcal{P}_p(\mathbb{R}^n)$ .

In addition, if  $\sigma(x) \equiv \sigma$ , a constant,  $\nu(x, dy) \equiv \nu(dy)$ ,  $1 \in \Theta_\nu$ , and (1.16) holds for some  $p \in [1, \theta_\nu] \cap \Theta_\nu$ , then (1.17) and (1.18) remain valid.

We remark that ergodic properties of a Markov process with respect to the  $\mathcal{W}_p$ -distance are invariant under the Bochner's random time-change method. Recall that a subordinator  $\{S(t)\}_{t \geq 0}$  is a nondecreasing Lévy process on  $[0, \infty)$  with Laplace transform  $\mathbb{E}[e^{-uS_t}] = e^{-t\psi(u)}$ ,  $u, t \geq 0$ . The characteristic (Laplace) exponent  $\psi: (0, \infty) \rightarrow (0, \infty)$  is a Bernstein function, i.e. it is of class  $C^\infty$  and  $(-1)^n \psi^{(n)}(u) \geq 0$  for all  $n \in \mathbb{N}$ . It is well known that every Bernstein function admits a unique (Lévy-Khintchine) representation

$$\psi(u) = b_S u + \int_{(0, \infty)} (1 - e^{-uy}) \nu_S(dy) \quad \forall u \geq 0,$$

where  $b_S \geq 0$  is the drift parameter and  $\nu_S(dy)$  is a Lévy measure, i.e. a Borel measure on  $\mathfrak{B}((0, \infty))$  satisfying  $\int_{(0, \infty)} (1 \wedge y) \nu(dy) < \infty$ . For additional reading on subordinators and Bernstein functions we refer the reader to the monograph [74]. Suppose  $\{X(t)\}_{t \geq 0}$  is a Markov process on  $(\mathbb{X}, \mathfrak{B}(\mathbb{X}))$  with transition kernel  $p(t, x, dy)$ , and let  $\{S(t)\}_{t \geq 0}$  be a subordinator with characteristic exponent  $\psi(u)$ , independent of  $\{X(t)\}_{t \geq 0}$ . The process  $X^\psi(t) := X(\bar{S}(t))$ ,  $t \geq 0$ , obtained from  $\{X(t)\}_{t \geq 0}$  by a random time change through  $\{S(t)\}_{t \geq 0}$ ,

is referred to as the subordinate process  $\{X(t)\}_{t \geq 0}$  with subordinator  $\{S(t)\}_{t \geq 0}$  in the sense of Bochner. It is easy to see that  $\{X^\psi(t)\}_{t \geq 0}$  is again a Markov process with transition kernel

$$p^\psi(t, x, dy) = \int_{[0, \infty)} p(s, x, dy) \mu_t(ds), \quad t \geq 0, \quad x \in \mathbb{R}^n,$$

where  $\mu_t(\cdot) = \mathbb{P}(S(t) \in \cdot)$ . It is also elementary to check that if  $\pi(dx)$  is an invariant measure for  $\{X(t)\}_{t \geq 0}$ , then it is also invariant for the subordinate process  $\{X^\psi(t)\}_{t \geq 0}$ .

**Proposition 1.1.** *Assume that  $\{X(t)\}_{t \geq 0}$  admits an invariant  $\pi \in \mathcal{P}(\mathbb{X})$  such that  $\mathcal{W}_p(\delta_x P_t, \pi) \leq c(x) r(t)$  for some  $p \geq 1$ , and all  $t \geq 0$  and  $x \in \mathbb{X}$ , where  $r: [0, \infty) \rightarrow [1, \infty)$  is Borel measurable, and  $c: \mathbb{X} \rightarrow [0, \infty)$ . Then,*

$$\mathcal{W}_p(\delta_x P_t^\psi, \pi) \leq c(x) r_\psi(t) \quad \forall (t, x) \in [0, \infty) \times \mathbb{X},$$

where  $r_\psi(t) := \left( \mathbb{E} \left[ (r(S(t)))^p \right] \right)^{1/p}$ .

Ergodic properties of Markov processes under subordination in the  $f$ -norm are discussed in [20–22].

### 1.3. Literature review

Our work contributes to the understanding of the ergodic properties of Markov processes. Most of the existing literature focuses on characterizing the exponential or subexponential ergodicity under the  $f$ -norm, and in particular the total variation norm, see [3, 19, 24–26, 32, 33, 35, 58, 64–66, 78] and the references therein. However, there have been some recent developments in understanding ergodic properties of Markov processes (both continuous and discrete time) under the Wasserstein distances; see [12, 28–30, 34, 53, 56, 57, 59, 60, 85]. As already mentioned, exponential and subexponential convergence rates in the  $\mathcal{W}_1$ -distance for general Markov processes that are (possibly) not irreducible or aperiodic are established in [12, 28, 53], under the Foster-Lyapunov condition in (1.1), contractivity of the underlying metric, and smallness of sublevel sets of the corresponding Lyapunov function. Using the coupling approach, the authors in [29, 30, 59] studied exponential ergodicity with respect to a class of Wasserstein distances for SDEs driven by an additive Brownian noise term and a drift term satisfying an asymptotic flatness property at infinity. Under the same assumption on the drift term, these results have been extended in [60, 85] to allow for more general additive Lévy noises. Subexponential ergodicity with respect to the  $\mathcal{W}_p$ -distance for stochastic differential equations driven by an additive Lévy noise term, with a drift term satisfying asymptotic flatness property at zero, has been studied in [56]. By combining the Foster-Lyapunov method with the coupling approach, exponential ergodicity with respect to a class of  $f$ -norms and Wasserstein distances (given in terms of the underlying Lyapunov function) is established in [57] for a class of McKean-Vlasov SDE with Lévy noise. Lastly, exponential ergodicity with respect to the  $\mathcal{W}_1$ -distance for one-dimensional positive-valued stochastic differential equations with jumps and the drift term satisfying asymptotic flatness property has been studied in [34].

Our results on both exponential and subexponential ergodicity under the  $\mathcal{W}_p$ -distance contribute to this active research topic. Of particular interest is the result obtained in Theorem 1.2 which seems to be completely new in the literature, and which, in some cases, allows one to conclude that the obtained upper bound on the rate of convergence is sharp.

As we have already remarked, irreducibility and aperiodicity are crucial structural properties of the underlying process used in Theorems 1.1 to 1.3. There is a vast literature on these, and related questions such as the strong Feller property and heat kernel estimates of Markov processes. In particular, we refer the readers to [8, 15–18, 36, 42, 43, 45–48, 55, 67, 71, 77] for the case of a class of Markov Lévy-type processes with bounded coefficients, and to [7, 9, 39, 44, 56, 62, 63, 68, 69, 72, 76, 86] for the case of a class of Itô processes.

Recall that the Foster-Lyapunov condition in (1.1) implies that for any  $\varepsilon > 0$  the  $\phi \circ \Phi^{-1}$ -modulated moment of the  $\varepsilon$ -shifted hitting time  $\tau_C^\varepsilon := \inf\{t \geq \delta: X(t) \in C\}$  of  $\{X(t)\}_{t \geq 0}$  of  $C$  (with respect to  $\mathbb{P}_x$ ) is finite and controlled by  $\mathcal{V}(x)$  (see [24, Theorem 4.1]). However, this property in general does not immediately imply ergodicity of  $\{X(t)\}_{t \geq 0}$ . Namely, we also need to ensure that a similar property holds for any other “reasonable” set. If  $\{X(t)\}_{t \geq 0}$  is irreducible with irreducibility measure  $\varphi(dx)$ , then indeed for any  $\varepsilon > 0$

the  $\phi \circ \Phi^{-1}$ -modulated moment of  $\tau_B^\varepsilon$ , for any  $B \in \mathfrak{B}(\mathbb{X})$  with  $\varphi(B) > 0$ , is again finite and controlled by  $\mathcal{V}(x)$  (see [24, the discussion after Theorem 4.1]). However,  $\{X(t)\}_{t \geq 0}$  can also show certain cyclic behavior which destroys ergodicity (see [65, Section 5] and [64, Chapter 5]). By assuming aperiodicity, which excludes this type of behavior, (sub)exponential ergodicity in the  $\mathcal{W}_p$ -distance of  $\{X(t)\}_{t \geq 0}$  follows as discussed in Theorem 1.1, and in the  $f$ -norm as discussed in [33, Theorem 1].

### 1.4. Organization of the article

In Section 2, we give the proofs of Theorems 1.1 to 1.4 and Proposition 1.1 together with some auxiliary lemmas. Applications of the main results to several classes of Markov processes, including Langevin tempered diffusion processes, Ornstein-Uhlenbeck processes with jumps, piecewise Ornstein-Uhlenbeck processes with jumps under constant and stationary Markov controls, state-space models, and backward recurrence time chains, are contained in Section 3.

## 2. Proofs of the main results

We start with the proof of Theorem 1.1.

*Proof of Theorem 1.1.* We consider the case when  $\mathbb{T} = \mathbb{R}_+$  only. The case when  $\mathbb{T} = \mathbb{Z}_+$  proceeds in an analogous way, by employing the results from [25, Theorem 2.8] and [64, Theorem 15.0.2].

First, under the assumptions of the theorem, it has been shown in [24, Proposition 3.1] and [66, Theorem 4.2] that  $\{X(t)\}_{t \geq 0}$  admits a unique invariant  $\pi \in \mathcal{P}_{\phi \circ \mathcal{V}}(\mathbb{X})$ . This, together with (1.2), implies that  $\pi \in \mathcal{P}_\eta(\mathbb{X})$ . We continue now with the proof of part (i). By the Kantorovich-Rubinstein theorem, we have

$$\mathcal{W}_1(\mu_1, \mu_2) = \sup_{\{f: \text{Lip}(f) \leq 1\}} \left| \int_{\mathbb{X}} f(x) (\mu_1(dx) - \mu_2(dx)) \right| \quad \forall \mu_1, \mu_2 \in \mathcal{P}_1(\mathbb{X}),$$

where the supremum is taken over all Lipschitz continuous functions  $f: \mathbb{X} \rightarrow \mathbb{R}$  with Lipschitz constant  $\text{Lip}(f) \leq 1$ . We apply [24, Theorem 3.2],

$$r_*(t) = \phi \circ \Phi^{-1}(t), \quad f_*(x) = \phi \circ \mathcal{V}(x), \quad \Psi_1(z) = z^{(\eta-1)/\eta}, \quad \text{and} \quad \Psi_2(z) = c^{-1/\eta} z^{1/\eta}.$$

Note that if  $f: \mathbb{X} \rightarrow \mathbb{R}$  is such that  $\text{Lip}(f) \leq 1$  and  $f(x_0) = 0$ , then  $|f(x)| \leq \mathbf{d}(x, x_0) \leq \Psi_2 \circ f_*(x)$ . Thus

$$\sup_{f: \text{Lip}(f) \leq 1} \left| \int_{\mathbb{X}} f(x) (\mu_1(dx) - \mu_2(dx)) \right| \leq \sup_{|f| \leq \Psi_2 \circ f_* \vee 1} \left| \int_{\mathbb{X}} f(x) (\mu_1(dx) - \mu_2(dx)) \right| = \|\mu_1 - \mu_2\|_{\Psi_2 \circ f_* \vee 1}$$

(recall the definition of the  $f$ -norm in (1.8)). Now, from [24, (3.5) and (3.6)] we have

$$(\Psi_1 \circ r_*(t) \vee 1) \mathcal{W}_1(\delta_x P_t, \pi) \leq (\Psi_1 \circ r_*(t) \vee 1) \|\delta_x P_t - \pi\|_{\Psi_2 \circ f_* \vee 1} \leq \bar{c} \mathcal{V}(x),$$

and

$$\begin{aligned} \int_0^\infty (\Psi_1 \circ r_*(s) \vee 1) \mathcal{W}_1(\delta_x P_s, \delta_y P_s) ds &\leq \int_0^\infty (\Psi_1 \circ r_*(s) \vee 1) \|\delta_x P_s - \pi\|_{\Psi_2(f_*) \vee 1} ds \\ &\leq \bar{c} (\mathcal{V}(x) + \mathcal{V}(y)), \end{aligned}$$

for some  $\bar{c} > 0$ , and all  $t \geq 0$  and  $x, y \in \mathbb{R}^n$ , which proves (1.3) and (1.4), respectively.

We next prove part (ii). Applying (1.3) and [24, (3.5)] with  $\Psi_1(z) = 1$ , and  $\Psi_2(z) = z$ , we obtain  $\mathbb{E}_x[\mathbf{d}(X(t), x_0)^\eta] \leq \bar{m}_\eta + \check{c} \mathcal{V}(x)$ , for some  $\check{c} > 0$ , and all  $t \geq 0$  and  $x \in \mathbb{X}$ . Hence

$$\mathbb{E}_x[\mathbf{d}(X(t), x_0)^p \mathbf{1}_{\mathcal{B}_t^c(x_0)}(X(t))] \leq t^{p-\eta} (\bar{m}_\eta + \check{c} \mathcal{V}(x)) \quad \forall (t, x) \in [0, \infty) \times \mathbb{X}. \quad (2.1)$$

Further, for  $t \geq 0$ ,  $z \in \mathbb{X}$ , and  $\Pi \in \mathcal{C}(\delta_z P_t, \pi)$ , we have

$$\begin{aligned} & \int_{\mathbb{X} \times \mathbb{X}} (\mathbf{d}(x, y))^p \Pi(\mathrm{d}x, \mathrm{d}y) \\ &= \int_{\mathcal{B}_t(x_0) \times \mathcal{B}_t(x_0)} (\mathbf{d}(x, y))^p \Pi(\mathrm{d}x, \mathrm{d}y) + \int_{(\mathcal{B}_t(x_0) \times \mathcal{B}_t(x_0))^c} (\mathbf{d}(x, y))^p \Pi(\mathrm{d}x, \mathrm{d}y) \\ &\leq (2t)^{p-1} \int_{\mathbb{X} \times \mathbb{X}} \mathbf{d}(x, y) \Pi(\mathrm{d}x, \mathrm{d}y) + 2^{p-1} \int_{\mathcal{B}_t^c(x_0)} (\mathbf{d}(x, x_0))^p [\delta_z P_t(\mathrm{d}x) + \pi(\mathrm{d}x)]. \end{aligned} \quad (2.2)$$

Using (2.1) and (2.2), and the bound  $\int_{\mathcal{B}_t^c(x_0)} (\mathbf{d}(x, x_0))^p \pi(\mathrm{d}x) \leq t^{p-\eta} \bar{m}_\eta$ , we have

$$\mathcal{W}_p^p(\delta_x P_t, \pi) \leq (2t)^{p-1} \mathcal{W}_1(\delta_x P_t, \pi) + 2^{p-1} t^{p-\eta} (2\bar{m}_\eta + \check{c} \mathcal{V}(x)) \quad \forall (t, x) \in [0, \infty) \times \mathbb{X},$$

and combining this with (1.3) we obtain

$$\left(1 \vee \left(t^{\eta-p} \wedge t^{1-p} (r_*(t))^{(\eta-1)/\eta}\right)\right) \mathcal{W}_p^p(\delta_x P_t, \pi) \leq 2^{p-1} \bar{c} \mathcal{V}(x) + 2^{p-1} (2\bar{m}_\eta + \check{c} \mathcal{V}(x))$$

for all  $t \geq 0$  and  $x \in \mathbb{X}$ , from which (1.5) follows with  $\tilde{c} = 2 \max\{1, \bar{c}, \check{c}\}^{1/p}$ .

Moving on to the proof of part (iii), note that according to [65, Proposition 6.1], [66, Theorem 4.2], and [26, Theorem 5.2], there exist constants  $\check{c} > 0$  and  $\gamma > 0$ , such that

$$\|\delta_x P_t - \pi\|_{\mathcal{V}} \leq \check{c} \mathcal{V}(x) e^{-\gamma t} \quad \forall (t, x) \in [0, \infty) \times \mathbb{X}. \quad (2.3)$$

Equation (1.6) now follows from the Kantorovich-Rubinstein theorem and (1.2). Let  $p \in [1, \eta]$ . First, from (2.3) we obtain  $\mathbb{E}_x[\mathbf{d}(X_t, x_0)^\eta] \leq \bar{m}_\eta + \check{c} \mathcal{V}(x)$ , for some  $\check{c} > 0$ , and all  $t \geq 0$  and  $x \in \mathbb{X}$ , which again implies (2.1). By (2.1) and (2.2), we have

$$\mathcal{W}_p^p(\delta_x P_t, \pi) \leq (2t)^{p-1} \mathcal{W}_1(\delta_x P_t, \pi) + 2^{p-1} t^{p-\eta} (2\bar{m}_\eta + \check{c} \mathcal{V}(x)) \quad \forall (t, x) \in [0, \infty) \times \mathbb{X},$$

and combining this with (1.6) we obtain

$$(1 \vee t^{\eta-p}) \mathcal{W}_p^p(\delta_x P_t, \pi) \leq 2^{p-1} \check{c} \mathcal{V}(x) + 2^{p-1} (2\bar{m}_\eta + \check{c} \mathcal{V}(x)) \quad \forall (t, x) \in [0, \infty) \times \mathbb{X},$$

from which (1.7) follows again with  $\check{c} = 2 \max\{1, \bar{c}, \check{c}\}^{1/p}$ . This completes the proof.  $\square$

We proceed with the proof of Theorem 1.2.

*Proof of Theorem 1.2.* We again consider the case when  $\mathbb{T} = \mathbb{R}_+$  only. The case when  $\mathbb{T} = \mathbb{Z}_+$  proceeds in a similar manner.

Fix some  $x_0 \in \mathbb{X}$ ,  $p \in [1, \vartheta]$  and  $\iota \in (0, \theta - \vartheta - \varepsilon)$ . For  $s > 0$ , define  $f_s: \mathbb{X} \rightarrow [0, \infty)$  by

$$f_s(x) := \begin{cases} 0, & \text{if } L(x) \leq \frac{s}{2}, \\ L(x) - \frac{s}{2}, & \text{if } L(x) > \frac{s}{2}. \end{cases}$$

We have

$$\int_{\mathbb{X}} (f_s(x))^p \pi(\mathrm{d}x) \geq \left(\frac{s}{2}\right)^p \pi(\{x: L(x) > s\}) \quad \forall s > 0. \quad (2.4)$$

Since, by assumption,  $\int_{\mathbb{X}} (L(x))^{\vartheta+\varepsilon} \pi(\mathrm{d}x) = \infty$ , there exists an increasing diverging sequence  $\{s_n\}_{n \in \mathbb{N}} \subset [0, \infty)$  such that

$$\left(\frac{s_n}{2}\right)^p \pi(\{x: L(x) > s_n\}) \geq 2^p s_n^{p-\vartheta-\varepsilon-\iota}. \quad (2.5)$$

Note also that  $(f_s(x))^p \leq 2^{\theta-p} s^{p-\theta} (L(x))^\theta \leq \frac{2^{\theta-p}}{c} s^{p-\theta} \mathcal{V}(x)$  for all  $s > 0$  and  $x \in \mathbb{X}$ . This follows from the facts that  $f_s(x) = 0$  for  $s > 0$  and  $x \in \mathbb{X}$  such that  $L(x) \leq s/2$ ,

$$0 \leq \frac{f_s(x)}{\frac{s}{2}} \leq \frac{L(x)}{\frac{s}{2}} \quad \forall (s, x) \in (0, \infty) \times \mathbb{X},$$

and  $\theta > p \geq 1$ . Thus, by the Foster-Lyapunov equation (1.1) (see [66, Theorem 1.1]), we obtain

$$\int_{\mathbb{X}} (f_s(x))^p \delta_{x_0} P_t(dx) \leq \frac{2^{\theta-p}}{c} s^{p-\theta} (bt + \mathcal{V}(x_0)) \quad \forall s, t > 0. \quad (2.6)$$

Select a sequence  $\{t_n\}_{n \in \mathbb{N}} \subset [0, \infty)$  such that

$$s_n^{\theta-\vartheta-\varepsilon-\iota} = \frac{2^{\theta-p}}{c} (bt_n + \mathcal{V}(x_0)). \quad (2.7)$$

Combining (2.4)–(2.7) above we have

$$\begin{aligned} \left( \int_{\mathbb{X}} (f_{s_n}(x))^p \pi(dx) \right)^{1/p} - \left( \int_{\mathbb{X}} (f_{s_n}(x))^p \delta_{x_0} P_{t_n}(dx) \right)^{1/p} &\geq (s_n)^{\frac{p-\vartheta-\varepsilon-\iota}{p}} \\ &\geq \left( \frac{2^{\theta-p}}{c} (bt_n + \mathcal{V}(x_0)) \right)^{-\frac{\vartheta-p+\varepsilon+\iota}{(\theta-\vartheta-\varepsilon-\iota)p}} \quad \forall n \in \mathbb{N}. \end{aligned}$$

The result then follows by [82, Proposition 7.29], which asserts that

$$\left| \left( \int_{\mathbb{X}} (f(x))^p \mu_1(dx) \right)^{1/p} - \left( \int_{\mathbb{X}} (f(x))^p \mu_2(dx) \right)^{1/p} \right| \leq \text{Lip}(f) \mathcal{W}_p(\mu_1, \mu_2)$$

for all  $\mu_1, \mu_2 \in \mathcal{P}_p(\mathbb{X})$  and Lipschitz  $f : \mathbb{X} \rightarrow \mathbb{R}$  with Lipschitz constant  $\text{Lip}(f)$ .  $\square$

For the proof of Theorem 1.3 we need two auxiliary results given in Lemmas 2.1 and 2.2 below. First, recall that  $\{X(t)\}_{t \geq 0}$  is said to be conservative if  $\mathbb{P}_x(X(t) \in \mathbb{R}^n) = 1$  for all  $t \geq 0$  and  $x \in \mathbb{R}^n$ , and note that this is equivalent to

$$\mathbb{P}_x \left( \lim_{k \rightarrow \infty} \tau_k = \infty \right) = 1 \quad \forall x \in \mathbb{R}^n,$$

where  $\tau_k := \inf\{t \geq 0 : X_t \in \mathcal{B}_k^c\}$  for  $k \in \mathbb{N}$  (here it is also essential that  $\{X(t)\}_{t \geq 0}$  has càdlàg sample paths). Namely, for  $t \geq 0$  and  $x \in \mathbb{R}^n$  it holds that

$$\mathbb{P}_x(X(t) \in \mathbb{R}^n) = \mathbb{P}_x \left( \lim_{k \rightarrow \infty} \tau_k > t \right) \geq \mathbb{P}_x \left( \lim_{k \rightarrow \infty} \tau_k = \infty \right) = \lim_{t \rightarrow \infty} \mathbb{P}_x \left( \lim_{k \rightarrow \infty} \tau_k > t \right) = \lim_{t \rightarrow \infty} \mathbb{P}_x(X(t) \in \mathbb{R}^n).$$

**Lemma 2.1.** *Assume (LB) and (MP). Then for any  $x \in \mathbb{R}^n$  and any nonnegative  $f \in C^\infty(\mathbb{R}^n)$  such that the map  $y \mapsto \int_{\mathcal{B}^c} f(y+z) \nu(y, dz)$  is locally bounded,  $\{M_f(t)\}_{t \geq 0}$  is a  $\mathbb{P}_x$ -local martingale (with respect to  $\{\mathcal{F}_t\}_{t \geq 0}$ ).*

*Proof.* For  $k \in \mathbb{N}$ , let  $\chi_k \in C_c^\infty(\mathbb{R}^n)$  be such that  $\mathbf{1}_{\mathcal{B}_k}(x) \leq \chi_k(x) \leq \mathbf{1}_{\mathcal{B}_{k+1}}(x)$  and  $\chi_k(x) \leq \chi_{k+1}(x)$  for  $x \in \mathbb{R}^n$ . Then, for any  $x \in \mathbb{R}^n$ ,  $k, j \in \mathbb{N}$  and  $s, t \geq 0$ ,  $s \leq t$ , [31, Theorem 2.2.13] implies that

$$\mathbb{E}_x [M_{f\chi_k}(t \wedge \tau_j) | \mathcal{F}_s] = M_{f\chi_k}(s \wedge \tau_j).$$

Next, by employing the monotone and dominated convergence theorems, we easily see that

$$\mathbb{E}_x \left[ \left| \int_0^t \mathcal{L}f(X(s \wedge \tau_j)) ds \right| \right] < \infty \quad \forall (x, j) \in \mathbb{R}^n \times \mathbb{N},$$

and

$$\begin{aligned} \mathbb{E}_x [f(X(t \wedge \tau_j))] &= \lim_{k \rightarrow \infty} \mathbb{E}_x [f(X(t \wedge \tau_j)) \chi_k(X(t \wedge \tau_j))] \\ &= f(x) + \lim_{k \rightarrow \infty} \mathbb{E}_x \left[ \int_0^t \mathcal{L}(f(X(s \wedge \tau_j)) \chi_k(X(s \wedge \tau_j))) ds \right] \\ &= f(x) + \mathbb{E}_x \left[ \int_0^t \mathcal{L}f(X(s \wedge \tau_j)) ds \right] \quad \forall (x, j) \in \mathbb{R}^n \times \mathbb{N}. \end{aligned}$$

Hence, for each  $x \in \mathbb{R}^n$ ,  $t \geq 0$  and  $j \in \mathbb{N}$ ,  $M_f(t \wedge \tau_j)$  is integrable. Also,

$$\lim_{k \rightarrow \infty} \mathbb{E}_x [M_{f\chi_k}(t \wedge \tau_j) | \mathcal{F}_s] = \mathbb{E}_x [M_f(t \wedge \tau_j) | \mathcal{F}_s], \quad \text{and} \quad \lim_{k \rightarrow \infty} M_{f\chi_k}(s \wedge \tau_j) = M_f(s \wedge \tau_j),$$

for all  $x \in \mathbb{R}^n$ ,  $t \geq s \geq 0$ , and  $j \in \mathbb{N}$ . The assertion now follows from the conservativeness of  $\{X(t)\}_{t \geq 0}$ .  $\square$

For  $f \in C^1(\mathbb{R}^n)$  we let

$$\mathfrak{d}f(x; y) := f(x + y) - f(x) - \langle y, \nabla f(x) \rangle, \quad x, y \in \mathbb{R}^n,$$

$$\mathfrak{J}_{1, \nu}[f](x) := \int_{\mathbb{R}^n} \mathfrak{d}_1 f(x; y) \nu(x, dy), \quad \text{and} \quad \mathfrak{J}_\nu[f](x) := \int_{\mathbb{R}^n} \mathfrak{d}f(x; y) \nu(x, dy), \quad x \in \mathbb{R}^n,$$

whenever the integrals are well defined.

**Lemma 2.2.** *Suppose that  $\theta_\nu > 0$ , and that (1.13) holds for some  $\theta \in (0, \theta_\nu] \cap \Theta_\nu$ . Then, we have the following:*

(i) *If  $\theta \in (0, 1)$ , and  $f \in C^2(\mathbb{R}^n)$  satisfies*

$$\sup_{x \in \mathbb{B}^c} |x|^{-\theta} \max(|f(x)|, |x| |\nabla f(x)|, |x|^2 \|\nabla^2 f(x)\|) < \infty,$$

*then  $\mathfrak{J}_{1, \nu}[f](x)$  vanishes at infinity.*

(ii) *If  $\theta \geq 1$ , and  $f \in C^2(\mathbb{R}^n)$  satisfies*

$$\sup_{x \in \mathbb{B}^c} |x|^{1-\theta} \max(|\nabla f(x)|, |x| \|\nabla^2 f(x)\|) < \infty,$$

*then  $\mathfrak{J}_\nu[f](x)$  vanishes at infinity when  $\theta \in [1, 2)$ , and the map  $x \mapsto (1 + |x|)^{2-\theta} \mathfrak{J}_\nu[f](x)$  is bounded when  $\theta \geq 2$ .*

(iii) *If (1.14) holds for some  $\theta > 0$ , then there exist  $c > 0$  and  $r = r(\zeta) > 0$ , such that for any  $\zeta \in (0, \frac{1}{2}\theta\|Q\|^{-1/2})$  we have*

$$\mathfrak{J}_\nu[\tilde{\mathfrak{V}}_{Q, \zeta}](x) \leq c \zeta^{3/2} \tilde{\mathfrak{V}}_{Q, \zeta}(x) \quad \forall x \in \mathbb{B}_r^c. \quad (2.8)$$

*Proof.* The proof of parts (i) and (ii) follows as a straightforward adaptation of [7, Lemma 5.1] by setting

$$C_0(\theta) := \sup_{x \in \mathbb{R}^n} \int_{\mathbb{R}^n} (|y|^2 \wedge |y|^\theta) \nu(x, dy), \quad \widehat{C}_0(\theta) := \sup_{x \in \mathbb{R}^n} \int_{\mathbb{B}} |y|^2 \nu(x, dy),$$

and

$$\check{C}_0(r; \theta) := \sup_{x \in \mathbb{R}^n} \int_{\mathbb{B}_r^c} |y|^\theta \nu(x, dy), \quad r > 0.$$

To prove part (iii), we use the identity

$$\int_{\mathbb{R}^n} \mathfrak{d} \tilde{\mathfrak{V}}_{Q, \zeta}(x; y) \nu(x, dy) = \int_{\mathbb{R}^n} \int_0^1 (1-t) \langle y, \nabla^2 \tilde{\mathfrak{V}}_{Q, \zeta}(x + ty) y \rangle dt \nu(x, dy) \quad (2.9)$$

Consider the set

$$A_x := \{(t, y) \in [0, 1] \times \mathbb{R}^n : |x + ty|_Q \leq \frac{1}{2}|x|_Q\}, \quad x \in \mathbb{R}^n.$$

On this set we have the bound

$$|\langle y, \nabla^2 \tilde{\mathfrak{V}}_{Q, \zeta}(x + ty) y \rangle| \leq \bar{c}(\zeta + \zeta^2) |y|^2 e^{\zeta \|Q\|^{1/2} |ty|} \tilde{\mathfrak{V}}_{Q, \zeta}(x) \quad (2.10)$$

for some  $\bar{c} \geq 1$ . Since  $\zeta \|Q\|^{1/2} < \theta$ , and  $|y|_Q \geq |ty|_Q \geq \frac{1}{2}|x|_Q$  on the set  $A_x$ , there exists  $\rho = \rho(\zeta) \geq 1$  such that

$$\zeta^{-1/2} (1 + \zeta) |y|^2 e^{\zeta \|Q\|^{1/2} |ty|} \leq e^{\theta |y|} \quad (2.11)$$

for all  $x \in \mathbb{B}_{2\rho}^c$  and  $(t, y) \in A_x$ . Hence, using (2.10) and (2.11) and Fubini's theorem, we have

$$\iint_{A_x} (1-t) \langle y, \nabla^2 \tilde{\mathfrak{V}}_{Q, \zeta}(x + ty) y \rangle dt \nu(x, dy) \leq 2\bar{c} \zeta^{3/2} \left( \int_{\mathbb{B}_\rho^c} e^{\theta |y|} \nu(x, dy) \right) \tilde{\mathfrak{V}}_{Q, \zeta}(x) \quad (2.12)$$

for all  $x \in \mathcal{B}_{2\rho}^c$ . Next, since  $|x + ty|_Q > \frac{1}{2}|x|_Q$  on the set  $A_x^c$ , we have a bound of the form

$$\langle y, \nabla^2 \tilde{\mathcal{V}}_{Q,\zeta}(x + ty)y \rangle \leq \bar{c} \left( \zeta^2 + \frac{\zeta}{|x|} \right) |y|^2 e^{\zeta \|Q\|^{1/2} |y|} \tilde{\mathcal{V}}_{Q,\zeta}(x) \quad (2.13)$$

for all  $x \in \mathcal{B}^c$  and  $(t, y) \in A_x^c$ , where, without loss of generality, we use the same constant  $\bar{c}$  as in (2.10). Since  $\theta > 2\zeta\sqrt{\|Q\|}$ , it is clear that there exists  $\tilde{c} > 0$ , independent of  $\zeta$ , such that

$$\left( \zeta^2 + \frac{\zeta}{|x|} \right) |y|^2 e^{\zeta \|Q\|^{1/2} |y|} \leq \tilde{c} \zeta^{3/2} (|y|^2 \mathbf{1}_{\mathcal{B}}(y) + e^{\theta|y|} \mathbf{1}_{\mathcal{B}^c}(y)) \quad \forall (x, y) \in \mathcal{B}_{1/\zeta}^c \times \mathbb{R}^n. \quad (2.14)$$

Thus, by (1.14), (2.13), and (2.14), there exists  $\hat{c} > 0$  such that

$$\iint_{A_x^c} (1-t) \langle y, \nabla^2 \tilde{\mathcal{V}}_{Q,\zeta}(x + ty)y \rangle dt \nu(x, dy) \leq \hat{c} \zeta^{3/2} \tilde{\mathcal{V}}_{Q,\zeta}(x) \quad \forall x \in \mathcal{B}_{1/\zeta}^c. \quad (2.15)$$

The estimate in (2.8) follows from (1.14), (2.9), (2.12), and (2.15). This completes the proof.  $\square$

We next prove [Theorem 1.3](#).

*Proof of Theorem 1.3.* In cases (i) and (ii), we take  $\mathcal{V}(x) = \mathcal{V}_{Q,\theta}(x) + 1$ , while in case (iii) we use  $\mathcal{V}(x) = \tilde{\mathcal{V}}_{Q,\zeta}(x)$  with  $\zeta > 0$  sufficiently small. Then, in view of [Lemma 2.2](#) it is straightforward to see that there exist constants  $\bar{c} > 0$ ,  $\tilde{c} > 0$ , and  $r > 0$ , such that

$$\mathcal{L}\mathcal{V}(x) \leq \bar{c} \mathbf{1}_{\mathcal{B}_r}(x) - \tilde{c} (\mathcal{V}(x))^{\frac{(\theta-2+\vartheta)}{\theta}} \quad \forall x \in \mathbb{R}^n,$$

in case (i), and

$$\mathcal{L}\mathcal{V}(x) \leq \bar{c} \mathbf{1}_{\mathcal{B}_r}(x) - \tilde{c} \mathcal{V}(x) \quad \forall x \in \mathbb{R}^n,$$

in cases (ii) and (iii). Observe that the above relations, together with [66, Theorem 2.1] and [Lemma 2.1](#), imply that  $\{X(t)\}_{t \geq 0}$  is conservative. Finally, according to [Lemma 2.1](#) and [24, Theorem 3.4] the process  $\{X(t)\}_{t \geq 0}$  satisfies (1.1) with  $\phi(t) = t^{(\theta-2+\vartheta)/\theta}$  in case (i), and  $\phi(t) = t$  in cases (ii) and (iii) (for some  $b > 0$  and closed petite set  $C$ ).  $\square$

The proof of [Theorem 1.4](#) is based on the following lemma.

**Lemma 2.3.** *Let  $\{X(t)\}_{t \geq 0}$  be an Itô process with locally bounded coefficients  $b(x)$  and  $a(x)$  and satisfying the linear growth condition in (1.12), and  $\nu(x, dy)$  such that  $\theta_\nu > 0$ . Then, for any  $\theta \in [0, \theta_\nu] \cap \Theta_\nu$ , there exists a constant  $c > 0$  such that*

$$\mathbb{E}_x[|X(t)|^\theta] \leq (1 + |x|^\theta) e^{ct} \quad \forall (t, x) \in [0, \infty) \times \mathbb{R}^n.$$

*Proof.* Let  $\varphi \in C^2(\mathbb{R}^n)$  be such that  $\varphi(x) \geq 0$  and  $\varphi(x) \leq |x|^\theta$  for  $x \in \mathbb{R}^n$ , and  $\varphi(x) = |x|^\theta$  for  $x \in \mathcal{B}^c$ . Further, for  $k \in \mathbb{N}$ , let  $\varphi_k \in C_b^2(\mathbb{R}^n)$  be such that  $\varphi_k(x) \geq 0$ ,  $\varphi_k(x) = \varphi|_{\mathcal{B}_{k+1}}(x)$ , and  $\varphi_k(x) \nearrow \varphi(x)$ , as  $k \rightarrow \infty$ , for every  $x \in \mathbb{R}^n$ . Then, according to Itô's formula and the conservativeness of  $\{X(t)\}_{t \geq 0}$  we have

$$\begin{aligned} \mathbb{E}_x[\varphi_k(X(t \wedge \tau_k))] &= \varphi_k(x) + \mathbb{E}_x \left[ \int_0^{t \wedge \tau_k} \mathcal{L}\varphi_k(X(s)) ds \right] \\ &\leq \varphi_k(x) + c_k(t \wedge \tau_k) + c_k \mathbb{E}_x \left[ \int_0^{t \wedge \tau_k} \varphi_k(X(s)) ds \right] \\ &\leq \varphi_k(x) + c_k t + c_k \int_0^t \mathbb{E}_x[\varphi_k(X(s \wedge \tau_k))] ds \end{aligned}$$

for all  $k \in \mathbb{N}$ ,  $t \geq 0$ , and  $x \in \mathbb{R}^n$ , where the constants  $c_k > 0$  depend on  $\theta$ ,  $b(x)$ ,  $a(x)$ , and the quantities

$$\sup_{x \in \mathbb{R}^n} \int_{\mathbb{R}^n} (|y|^2 \mathbf{1}_{\mathcal{B}}(y) + |y|^\theta \mathbf{1}_{\mathcal{B}^c}(y)) \nu(x, dy) \quad \text{and} \quad \sup_{x \in \mathcal{B}_r} (|\varphi_k(x)| + |\nabla \varphi_k(x)| + |\nabla^2 \varphi_k(x)|),$$

for  $r > 0$  large enough. Clearly, the functions  $\varphi_k(x)$  can be chosen such that  $c := \sup_{k \in \mathbb{N}} c_k < \infty$ . Now, since the function  $t \mapsto \mathbb{E}_x[\varphi_k(X(t \wedge \tau_k))]$  is bounded and càdlàg, Gronwall's lemma implies that

$$\mathbb{E}_x[\varphi_k(X(t \wedge \tau_k))] \leq (1 + \varphi_k(x)) e^{ct} - 1$$

for all  $k \in \mathbb{N}$ ,  $t \geq 0$ , and  $x \in \mathbb{R}^n$ . By letting  $k \rightarrow \infty$ , Fatou's lemma and the conservativeness of  $\{X(t)\}_{t \geq 0}$  imply that

$$\mathbb{E}_x[\varphi(X(t))] \leq (1 + \varphi(x)) e^{ct} - 1 \quad \forall (t, x) \in [0, \infty) \times \mathbb{R}^n.$$

Finally, we have that

$$\mathbb{E}_x[|X(t)|^\theta] \leq \mathbb{E}_x[\varphi(X(t))] + 1 \leq (1 + \varphi(x)) e^{ct} \leq (1 + |x|^\theta) e^{ct} \quad \forall (t, x) \in [0, \infty) \times \mathbb{R}^n.$$

This completes the proof.  $\square$

We next prove [Theorem 1.4](#).

*Proof of [Theorem 1.4](#).* For  $p \in [2, \theta_\nu] \cap \Theta_\nu$ , define  $\mathcal{V}_p(x) := |x|_Q^p$ ,  $x \in \mathbb{R}^n$ , and

$$\begin{aligned} \tilde{\mathcal{L}}f(x; z) &:= \sum_{i=1}^n \Delta_z \tilde{b}_i(x) \frac{\partial f}{\partial x_i}(z) + \frac{1}{2} \sum_{i,j=1}^n \tilde{a}_{ij}(x; z) \frac{\partial^2 f}{\partial x_i \partial x_j}(z) \\ &\quad + \int_{\mathbb{R}^n} \left( f(z+y) - f(y) - \sum_{i=1}^n y_i \frac{\partial f}{\partial x_i}(z) \right) \Delta_z \nu(x, dy), \quad x, z \in \mathbb{R}^n. \end{aligned}$$

Calculating  $\tilde{\mathcal{L}}\mathcal{V}_p(x; z)$ , using [\(1.16\)](#), we obtain

$$\begin{aligned} \tilde{\mathcal{L}}\mathcal{V}_p(x; z) &= \frac{p}{2} |z|_Q^{p-2} \left( 2 \langle \Delta_z \tilde{b}(x), Qz \rangle + \text{Tr}(\tilde{a}(x; z)Q) \right) + \frac{p(p-2)}{2} |z|_Q^{p-4} |\Delta_z \sigma'(x)Qz|^2 \\ &\quad + \int_{\mathbb{R}^n} \int_0^1 (1-t) \langle y, \nabla^2 \mathcal{V}_{\varepsilon, p}(z+ty)y \rangle dt \Delta_z \nu(x, dy) \\ &= \frac{p}{2} |z|_Q^{p-2} \left( 2 \langle \Delta_z \tilde{b}(x), Qz \rangle + \text{Tr}(\tilde{a}(x; z)Q) \right) + \frac{p(p-2)}{2} |z|_Q^{p-4} |\Delta_z \sigma'(x)Qz|^2 \\ &\quad + \int_{\mathbb{R}^n} \int_0^1 (1-t) \left( p|z+ty|_Q^{p-2} |y|_Q^2 + p(p-2)|z+ty|_Q^{p-4} |y'Q(z+ty)|^2 \right) dt \Delta_z \nu(x, dy) \\ &\leq \frac{p}{2} |z|_Q^{p-2} \left( 2 \langle \Delta_z \tilde{b}(x), Qz \rangle + \text{Tr}(\tilde{a}(x; z)Q) + (p-2) |\sqrt{Q} \Delta_z \sigma(x)|^2 \right) \\ &\quad + p(1+(p-2)\|Q^{-1}\|) \int_{\mathbb{R}^n} |y|_Q^2 \int_0^1 (1-t) |z+ty|_Q^{p-2} dt \Delta_z \nu(x, dy) \\ &\leq \frac{p}{2} |z|_Q^{p-2} \left( 2 \langle \Delta_z \tilde{b}(x), Qz \rangle + \text{Tr}(\tilde{a}(x; z)Q) + (p-2) |\sqrt{Q} \Delta_z \sigma(x)|^2 \right) \\ &\quad + p2^{p-4} (1+(p-2)\|Q^{-1}\|) |z|_Q^{p-2} \int_{\mathbb{R}^n} |y|_Q^2 \Delta_z \nu(x, dy) \\ &\quad + \frac{2^{p-3}}{p-1} (1+(p-2)\|Q^{-1}\|) \int_{\mathbb{R}^n} |y|_Q^p \Delta_z \nu(x, dy) \\ &= \frac{p}{2} |z|_Q^{p-2} \left( 2 \langle \Delta_z \tilde{b}(x), Qz \rangle + \text{Tr}(\tilde{a}(x; z)Q) + (p-2) |\sqrt{Q} \Delta_z \sigma(x)|^2 \right. \\ &\quad \left. + 2^{p-3} (1+(p-2)\|Q^{-1}\|) \int_{\mathbb{R}^n} |y|_Q^2 \Delta_z \nu(x, dy) \right. \\ &\quad \left. + \frac{2^{p-2}}{p(p-1)} (1+(p-2)\|Q^{-1}\|) |z|_Q^{2-p} \int_{\mathbb{R}^n} |y|_Q^p \Delta_z \nu(x, dy) \right) \end{aligned}$$

$$\begin{aligned}
 &= \frac{p}{2} |z|_Q^{p-2} \left( 2 \langle \Delta_z \tilde{b}(x), Qz \rangle + \text{Tr}(\tilde{a}(x; z)Q) + (p-2) |\sqrt{Q} \Delta_z \sigma(x)|^2 \right. \\
 &\quad + 2^{p-3} (1 + (p-2) \|Q^{-1}\|) \int_{\mathbb{R}} |k(x+z, v) - k(x, v)|_Q^2 \nu(dv) \\
 &\quad \left. + \frac{2^{p-2}}{p(p-1)} (1 + (p-2) \|Q^{-1}\|) |z|_Q^{2-p} \int_{\mathbb{R}} |k(x+z, v) - k(x, v)|_Q^p \nu(dv) \right) \\
 &\leq -c(p) \mathcal{V}_p(z)
 \end{aligned}$$

for all  $x, z \in \mathbb{R}^n$ . Next, for  $x, z \in \mathbb{R}^n$ , let  $\tau := \inf\{t \geq 0: X^{x+z}(t) = X^x(t)\}$  (possibly  $+\infty$ ), where  $\{X^x(t)\}_{t \geq 0}$  denotes the solution to (1.15) with  $X^x(0) = x$  for  $x \in \mathbb{R}^n$ . By Itô's formula and the conservativeness of  $\{X(t)\}_{t \geq 0}$  we obtain

$$\begin{aligned}
 &\mathbb{E}[\mathcal{V}_p(X^{x+z}(t \wedge \tau \wedge \tau_k) - X^x(t \wedge \tau \wedge \tau_k))] - \mathcal{V}_p(z) \\
 &= \mathbb{E}\left[\int_0^{t \wedge \tau \wedge \tau_k} \tilde{\mathcal{L}}\mathcal{V}_p(X^x(s); X^{x+z}(s) - X^z(s)) ds\right] \\
 &= \mathbb{E}\left[\int_0^t \tilde{\mathcal{L}}\mathcal{V}_p(X^x(s \wedge \tau \wedge \tau_k); X^{x+z}(s \wedge \tau) - X^z(s \wedge \tau \wedge \tau_k)) ds\right] \\
 &= \int_0^t \mathbb{E}\left[\tilde{\mathcal{L}}\mathcal{V}_p(X^x(s \wedge \tau \wedge \tau_k); X^{x+z}(s \wedge \tau \wedge \tau_k) - X^z(s \wedge \tau \wedge \tau_k))\right] ds
 \end{aligned}$$

for all  $t \geq 0$  and  $k \in \mathbb{N}$ , since, for  $t \geq \tau$ ,  $X^{x+z}(t) = X^x(t)$  a.s. by the pathwise uniqueness of the solution to (1.15). From this and Lemma 2.3 we conclude that the function  $t \mapsto \mathbb{E}[\mathcal{V}_p(X^{x+z}(t \wedge \tau \wedge \tau_k) - X^x(t \wedge \tau \wedge \tau_k))]$  is differentiable a.e. on  $(0, \infty)$ . Note that  $|\tilde{\mathcal{L}}\mathcal{V}_p(x; z)| \leq c|z|^p$  for some  $c > 0$  and all  $x, z \in \mathbb{R}^n$ . We conclude now that

$$\begin{aligned}
 &\frac{d}{dt} \mathbb{E}[\mathcal{V}_p(X^{x+z}(t \wedge \tau \wedge \tau_k) - X^x(t \wedge \tau \wedge \tau_k))] \\
 &= \mathbb{E}[\tilde{\mathcal{L}}\mathcal{V}_p(X^x(s \wedge \tau \wedge \tau_k); X^{x+z}(t \wedge \tau \wedge \tau_k) - X^z(t \wedge \tau \wedge \tau_k))] \\
 &\leq -c(p) \mathbb{E}[\mathcal{V}_p(X^{x+z}(t \wedge \tau \wedge \tau_k) - X^x(t \wedge \tau \wedge \tau_k))] \quad \text{a.e. on } (0, \infty)
 \end{aligned}$$

for all  $k \in \mathbb{N}$ . Thus by Gronwall's lemma, it follows that

$$\mathbb{E}[\mathcal{V}_p(X^{x+z}(t \wedge \tau_k) - X^x(t \wedge \tau_k))] = \mathbb{E}[\mathcal{V}_p(X^{x+z}(t \wedge \tau \wedge \tau_k) - X^x(t \wedge \tau \wedge \tau_k))] \leq \mathcal{V}_p(z) e^{-c(p)t},$$

and Fatou's lemma implies that

$$\mathbb{E}[\mathcal{V}_p(X^{x+z}(t) - X^x(t))] \leq \mathcal{V}_p(z) e^{-c(p)t}$$

for all  $t \geq 0$  and  $x, z \in \mathbb{R}^n$ . Next, from the bound  $\underline{\lambda}_Q |z|^2 \leq |z|_Q^2 \leq \bar{\lambda}_Q |z|^2$  we obtain

$$\begin{aligned}
 \mathbb{E}[|X^{x+z}(t) - X^x(t)|^p] &\leq (\underline{\lambda}_Q)^{-p/2} \mathbb{E}[|X^{x+z}(t) - X^x(t)|_Q^p] \\
 &\leq (\underline{\lambda}_Q)^{-p/2} (\bar{\lambda}_Q)^{p/2} |z|^p e^{-c(p)t}
 \end{aligned}$$

for all  $t \geq 0$  and  $x, z \in \mathbb{R}^n$ , thus establishing (1.17).

Finally, in order to establish (1.18), we follow the idea from [59, Proof of Corollary 1.8] or [49, Proof of Theorem 2.1]. Observe first that, according to Lemma 2.3, for any  $\mu \in \mathcal{P}_p(\mathbb{R}^n)$ ,  $\mu P_t \in \mathcal{P}_p(\mathbb{R}^n)$  for all  $t \geq 0$ . Next, let  $\mu_1, \mu_2 \in \mathcal{P}_p(\mathbb{R}^n)$  be arbitrary. According to (1.17), we have

$$\mathcal{W}_p(\mu_1 P_t, \mu_2 P_t) \leq \left(\frac{\bar{\lambda}_Q}{\underline{\lambda}_Q}\right)^{1/2} \mathcal{W}_p(\mu_1, \mu_2) e^{-\frac{c(p)t}{p}} \quad \forall t \geq 0.$$

Fix  $t_0 \geq 0$  such that

$$\left(\frac{\bar{\lambda}_Q}{\underline{\lambda}_Q}\right)^{1/2} e^{-\frac{c(p)t_0}{p}} < 1.$$

Then, the mapping  $\mu \mapsto \mu P_{t_0}$  is a contraction on  $\mathcal{P}_p(\mathbb{R}^n)$ . Thus, since  $(\mathcal{P}_p(\mathbb{R}^n), \mathcal{W}_p)$  is a complete metric space, the Banach fixed point theorem entails that there exists a unique  $\pi_0 \in \mathcal{P}_p(\mathbb{R}^n)$  such that  $\pi_0 P_{t_0}(dx) = \pi_0(dx)$ . By defining  $\pi(dx) := t_0^{-1} \int_0^{t_0} \pi_0 P_s(dx) ds$ , we can easily see that  $\pi P_t(dx) = \pi(dx)$  for all  $t \geq 0$ , i.e.  $\pi(dx)$  is an invariant probability measure for  $\{X(t)\}_{t \geq 0}$ . By employing [Lemma 2.1](#) again, we also see that  $\pi \in \mathcal{P}_p(\mathbb{R}^n)$ . Finally, for any  $\mu \in \mathcal{P}_p(\mathbb{R}^n)$  we have

$$\mathcal{W}_p(\mu P_t, \pi) = \mathcal{W}_p(\mu P_t, \pi P_t) \leq \left( \frac{\bar{\lambda}_Q}{\underline{\lambda}_Q} \right)^{1/2} \mathcal{W}_p(\mu, \pi) e^{-\frac{c(p)t}{p}} \quad \forall t \geq 0,$$

which also proves uniqueness of  $\pi(dx)$ .

To prove the second assertion, we adapt the proof of [\[4, Lemma 7.3.4\]](#), where an analogous result is shown for  $p = 1$ . Define

$$V_{\varepsilon,p}(x) := \frac{|x|_Q^{p+1}}{(\varepsilon + |x|_Q^2)^{1/2}}, \quad \varepsilon > 0, \quad x \in \mathbb{R}^n,$$

and observe that in this case  $\tilde{\mathcal{L}}$  reduces to

$$\tilde{\mathcal{L}}f(x; z) = \sum_{i=1}^d \Delta_z b^i(x) \frac{\partial f}{\partial x_i}(z) \quad \forall x, z \in \mathbb{R}^n.$$

Calculating  $\tilde{\mathcal{L}}V_{\varepsilon,p}(x; z)$ , using [\(1.16\)](#), we obtain

$$\begin{aligned} \tilde{\mathcal{L}}V_{\varepsilon,p}(x; z) &= \frac{(\varepsilon(p+1) + p|z|_Q^2) |z|_Q^{p-1}}{(\varepsilon + |z|_Q^2)^{3/2}} \langle Qz, \Delta_z b(x) \rangle \\ &\leq -c(p) \frac{\varepsilon^{p+1} + |z|_Q^2}{\varepsilon + |z|_Q^2} V_{\varepsilon,p}(z) \\ &\leq -c(p) V_{\varepsilon,p}(z) \quad \forall x, z \in \mathbb{R}^n. \end{aligned}$$

As before, by Itô's formula and the conservativeness of  $\{X(t)\}_{t \geq 0}$ , combined with the fact that the Lévy noise does not depend on the state, we obtain

$$\begin{aligned} &\mathbb{E}[V_{\varepsilon,p}(X^{x+z}(t \wedge \tau \wedge \tau_k) - X^x(t \wedge \tau \wedge \tau_k))] - V_{\varepsilon,p}(z) \\ &= \int_0^t \mathbb{E}[\tilde{\mathcal{L}}V_{\varepsilon,p}(X^x(s \wedge \tau \wedge \tau_k); X^{x+z}(s \wedge \tau \wedge \tau_k) - X^z(s \wedge \tau \wedge \tau_k))] ds \end{aligned}$$

for all  $t \geq 0$  and  $k \in \mathbb{N}$ , and

$$\begin{aligned} &\frac{d}{dt} \mathbb{E}[V_{\varepsilon,p}(X^{x+z}(t \wedge \tau \wedge \tau_k) - X^x(t \wedge \tau \wedge \tau_k))] \\ &= \mathbb{E}[\tilde{\mathcal{L}}V_{\varepsilon,p}(X^x(t \wedge \tau \wedge \tau_k); X^{x+z}(t \wedge \tau \wedge \tau_k) - X^z(t \wedge \tau \wedge \tau_k))] \\ &\leq -c(p) \mathbb{E}[V_{\varepsilon,p}(X^{x+z}(t \wedge \tau \wedge \tau_k) - X^x(t \wedge \tau \wedge \tau_k))] \quad \text{a.e. on } (0, \infty) \end{aligned}$$

for all  $k \in \mathbb{N}$ . Thus by Gronwall's and Fatou's lemmas it follows that

$$\mathbb{E}[V_{\varepsilon,p}(X^{x+z}(t) - X^x(t))] \leq V_{\varepsilon,p}(z) e^{-c_p t}$$

for all  $t \geq 0$  and  $x, z \in \mathbb{R}^n$ . Taking limits as  $\varepsilon \rightarrow 0$ , and using monotone convergence, the assertion follows.  $\square$

In what follows we give an alternative proof of [Theorem 1.4](#) in the case when  $\sigma(x) \equiv \sigma$  and  $\nu(x, dy) \equiv \nu(dy)$ . Let  $\bar{X}(t) := Q^{1/2} X(t)$  for  $t \geq 0$ . Clearly,  $\{\bar{X}(t)\}_{t \geq 0}$  is again an Itô process which satisfies

$$\bar{X}(t) = x + Q^{1/2} \int_0^t b(Q^{-1/2} \bar{X}(s)) ds + Q^{1/2} \sigma B(t) + Q^{1/2} L(t) \quad \forall t \geq 0,$$

where  $\{L(t)\}_{t \geq 0}$  is an  $n$ -dimensional pure-jump and zero-drift Lévy process determined by  $\nu(dy)$ . The corresponding transition probability satisfies

$$\begin{aligned} \bar{p}(t, x, dy) &= \bar{\mathbb{P}}^x(\bar{X}(t) \in dy) \\ &= \mathbb{P}^{Q^{-1/2}x}(X(t) \in Q^{-1/2}dy) = p(t, Q^{-1/2}x, Q^{-1/2}dy) \quad \forall (t, x) \in [0, \infty) \times \mathbb{R}^n. \end{aligned}$$

Thus, we have

$$\begin{aligned} \langle \Delta_z \bar{b}(x), Qz \rangle &= \langle Q^{1/2} \Delta_{Q^{-1/2}z} b(Q^{-1/2}x), Qz \rangle \\ &= \langle \Delta_{Q^{-1/2}z} b(Q^{-1/2}x), Q^{1/2}z \rangle \\ &\leq -\frac{c(p)}{p} |Q^{-1/2}z|_Q = -\frac{c(p)}{p} |z|_Q \quad \forall x, z \in \mathbb{R}^n. \end{aligned} \tag{2.16}$$

Now, in [11] it has been shown that (2.16) implies that

$$\mathcal{W}_p(\delta_x \bar{P}_t, \delta_y \bar{P}_t) \leq |x - y| e^{-\frac{c(p)t}{p}}$$

for all  $t \geq 0$  and  $x, y \in \mathbb{R}^n$ . Finally we get

$$\begin{aligned} \mathcal{W}_p(\delta_x P_t, \delta_y P_t) &= \mathcal{W}_p(\bar{p}(t, Q^{1/2}x, Q^{1/2}dz), \bar{p}(t, Q^{1/2}y, Q^{1/2}dz)) \\ &\leq (\lambda_Q)^{-1/2} |Q^{1/2}(x - y)| e^{-\frac{c(p)t}{p}} \\ &\leq \left( \frac{\bar{\lambda}_Q}{\lambda_Q} \right)^{1/2} |x - y| e^{-\frac{c(p)t}{p}} \end{aligned}$$

for all  $t \geq 0$  and  $x, y \in \mathbb{R}^n$ , which is (1.17).

Lastly, we prove Proposition 1.1.

*Proof of Proposition 1.1.* According to [82, Theorem 4.1], for each  $s \in [0, \infty)$  there exists  $\Pi_s \in \mathcal{C}(\delta_x P_s, \pi)$  such that  $\mathcal{W}_p(\delta_x P_s^\psi, \pi) = \int_{\mathbb{X} \times \mathbb{X}} \mathbf{d}(y, z) \Pi_s(dy, dz)$ . Now, we have that

$$\begin{aligned} \mathcal{W}_p^p(\delta_x P_t^\psi, \pi) &= \inf_{\Pi \in \mathcal{C}(\delta_x P_t^\psi, \pi)} \int_{\mathbb{X} \times \mathbb{X}} (\mathbf{d}(y, z))^p \Pi(dy, dz) \\ &\leq \int_{\mathbb{X} \times \mathbb{X}} (\mathbf{d}(y, z))^p \int_{[0, \infty)} \Pi_s(dy, dz) \mu_t(ds) \\ &\leq \int_{[0, \infty)} \mathcal{W}_p^p(\delta_x P_s, \pi) \mu_t(ds) \\ &\leq (c(x))^p \int_{[0, \infty)} (r^p(s))^p \mu_t(ds) = (c(x))^p \mathbb{E}[(r(S(t)))^p] \quad \forall (t, x) \in [0, \infty) \times \mathbb{X}, \end{aligned}$$

which completes the proof. □

### 3. Examples

In this section, we consider applications of the main results to several classes of Markov processes, including Langevin tempered diffusion processes, Ornstein-Uhlenbeck processes with jumps, piecewise Ornstein-Uhlenbeck processes with jumps under constant and stationary Markov controls, state-space models and backward recurrence time chains. Further examples can be found in [24, 25, 32, 33, 78].

#### 3.1. Langevin tempered diffusion processes

We first consider a class of Langevin tempered diffusion processes. Let  $\alpha \in (0, 1/n)$ , and let  $\pi \in C^2(\mathbb{R}^n)$  be strictly positive,  $\pi(x) = c|x|^{-1/\alpha}$  for some  $c > 0$  and all  $x \in \mathcal{B}^c$ , and  $\int_{\mathbb{R}^n} \pi(x) dx = 1$ . Further, for  $\beta \in [0, (1 + \alpha(2 - n))/2]$  and  $x \in \mathbb{R}^n$ , let

$$\sigma(x) := (\pi(x))^{-\beta} \mathbb{I}_n, \quad a(x) := \sigma(x)\sigma(x)' = (\pi(x))^{-2\beta} \mathbb{I}_n,$$

and

$$b(x) := \frac{1}{2} (a(x) \nabla \log(\pi(x)) + \nabla \cdot a(x)') = \frac{1 - 2\beta}{2} (\pi(x))^{-2\beta} \nabla \log(\pi(x)).$$

Then, in [33, Proposition 15] it has been shown that the SDE

$$X(t) = x + \int_0^t b(X(s)) ds + \int_0^t \sigma(X(s)) dB_s, \quad t \geq 0, \quad x \in \mathbb{R}^n,$$

admits a weak solution  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \{B(t)\}_{t \geq 0}, \{X(t)\}_{t \geq 0}, \mathbb{P})$ , which is a conservative strong Markov process with continuous sample paths. Moreover, it is irreducible, aperiodic, every compact set is petite, and  $\pi(dx) := \pi(x)dx$  is its unique invariant probability measure. Here,  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \{B(t)\}_{t \geq 0}, \mathbb{P})$  is a standard  $n$ -dimensional Brownian motion. Note also that according to Itô's formula  $\{X(t)\}_{t \geq 0}$  satisfies (MP) with

$$\mathcal{L}f(x) = \langle b(x), \nabla f(x) \rangle + \frac{1}{2} \text{Tr}(a(x) \nabla^2 f(x)), \quad x \in \mathbb{R}^n.$$

**Proposition 3.1.** (i) If  $\beta \in [\alpha, \frac{1}{2}(1 + \alpha(1 - n))]$ , then the assertions of Theorem 1.1 (iii) hold with  $\mathcal{V}(x) = 1 + \mathcal{V}_{\mathbb{I}_n, \frac{\gamma}{\alpha}}(x)$  and  $\eta = \frac{\gamma}{\alpha}$  for any  $\gamma \in [\alpha, 1 + \alpha(2 - n) - 2\beta]$ .

(ii) If  $\alpha \in (0, 1/(n + 1))$  and  $\beta \in [0, \alpha)$ , then the assertions of Theorem 1.1 (i) and (ii) hold with

$$\mathcal{V}(x) = 1 + \mathcal{V}_{\mathbb{I}_n, \frac{\gamma}{\alpha}}(x), \quad \phi(t) = t^{1 - \frac{2(\alpha - \beta)}{\gamma}}, \quad \text{and} \quad \eta = \alpha^{-1}(\gamma - 2(\alpha - \beta))$$

for any  $\gamma \in [3\alpha - 2\beta, 1 + \alpha(2 - n) - 2\beta]$ .

(iii) Under the assumptions of (ii),  $\pi \in \mathcal{P}_{\alpha^{-1}(1 - \alpha n) - \iota}(\mathbb{R}^n)$  for  $\iota \in (0, \alpha^{-1}(1 - \alpha(n + 1)))$ . Let  $\rho \in (0, (1 - (n + 1)\alpha) \wedge 2(\alpha - \beta))$  and  $\varepsilon \in [\alpha^{-1}\rho, 2\alpha^{-1}(\alpha - \beta)]$  be fixed. Then, for every  $p \in [1, \alpha^{-1}(1 - n\alpha - \rho)]$  and  $\iota \in (0, 2\alpha^{-1}(\alpha - \beta) - \varepsilon)$  there exist a positive constant  $\bar{c}$  and a diverging increasing sequence  $\{t_n\}_{n \in \mathbb{N}} \subset [0, \infty)$ , depending on the above parameters, such that (1.9) in Theorem 1.2 holds with  $\mathcal{V}(x)$  as above,  $\theta = \alpha^{-1}(1 + \alpha(2 - n) - 2\beta - \rho)$ , and  $\vartheta = \alpha^{-1}(1 - n\alpha - \rho)$ .

*Proof.* (i) In [33, Theorem 16 (i)] it has been shown that for  $\beta \in [\alpha, \frac{1}{2}(1 + \alpha(1 - n))]$  and  $\gamma \in (0, 1 + \alpha(2 - n) - 2\beta)$  the Foster-Lyapunov condition in (1.1) holds with  $\mathcal{V}(x)$  as above,  $\phi(t) = t$  and  $C = \bar{\mathcal{B}}_r$  for some  $r > 0$  large enough. Also, the relation in (1.2) easily follows from the form of  $\mathcal{V}(x)$  and  $\phi(t)$ , and the choice of  $\eta$ .

(ii) In [33, Theorem 16 (ii)] it has been shown that for  $\alpha \in (0, 1/n)$ ,  $\beta \in [0, \alpha)$  and  $\gamma \in (2(\alpha - \beta), 1 + \alpha(2 - n) - 2\beta)$ , the Foster-Lyapunov condition in (1.1) holds with  $\mathcal{V}(x)$  and  $\phi(t)$  as above and  $C = \bar{\mathcal{B}}_r$  for some  $r > 0$  large enough. The relation in (1.2) can again be easily verified due to the form of  $\mathcal{V}(x)$  and  $\phi(t)$ , and the choice of  $\eta$ .

(iii) Since  $\vartheta + \varepsilon - 1/\alpha + n - 1 \geq -1$ , we have  $\int_{\mathbb{R}^n} |x|^{\vartheta + \varepsilon} \pi(dx) = \infty$ . The assertion now follows from Theorem 1.2 by taking  $L(x) = |x|$ .

This completes the proof.  $\square$

*Remark 3.1.* Observe that the rates obtained in Proposition 3.1 (ii) and (iii) match. Also, in Proposition 3.1 (ii) we assume that  $\alpha \in (0, (n + 1)^{-1})$ . Namely, for  $\alpha \in [(n + 1)^{-1}, n^{-1})$  it holds that  $\int_{\mathbb{R}^n} |x| \pi(dx) = \infty$ , and hence convergence in the  $\mathcal{W}_p$ -distance cannot hold. On the other hand, in this case, [33, Theorem 16 (ii)] shows subexponential convergence in the  $f$ -norm. In the following subsections we give examples of Markov processes which are ergodic in the  $\mathcal{W}_p$ -distance but not in the  $f$ -norm. For additional results on ergodic properties of Langevin tempered diffusion processes with respect to the  $f$ -norm see [24] and [33].

### 3.2. Ornstein-Uhlenbeck processes with jumps

We next consider a class of Itô processes with linear drift. Let  $H$  be an  $n \times n$  matrix, and let  $\{L(t)\}_{t \geq 0}$  be an  $n$ -dimensional Lévy process determined by Lévy triplet  $(b_L, a_L, \nu_L(dy))$ . It is well known that the SDE

$$X(t) = x + H \int_0^t X(s) ds + L(t), \quad t \geq 0, \quad x \in \mathbb{R}^n,$$

admits a unique conservative strong solution  $\{X(t)\}_{t \geq 0}$  which is a strong Markov process with càdlàg sample paths (see e.g. [2, Theorem 3.1 and Proposition 4.2]). In particular,  $\{X(t)\}_{t \geq 0}$  is an Itô process satisfying (MP) with  $b(x) = b_L + Mx$ ,  $a(x) = a_L$ , and  $\nu(x, dy) = \nu_L(dy)$ . This process is known as an Ornstein-Uhlenbeck process with jumps. In the case when  $\{L(t)\}_{t \geq 0}$  is a standard Brownian motion,  $\{X(t)\}_{t \geq 0}$  is the classical Ornstein-Uhlenbeck process. If  $H$  is a Hurwitz matrix (a square matrix whose eigenvalues have all strictly negative real parts), it has been shown in [73, Theorems 4.1 and 4.2] that  $\{X(t)\}_{t \geq 0}$  admits a unique invariant  $\pi \in \mathcal{P}(\mathbb{R}^n)$  if, and only if,

$$\int_{\mathbb{B}^c} \log(|y|) \nu_L(dy) < \infty.$$

Moreover, if this is the case, then  $\lim_{t \rightarrow \infty} \delta_x P_t(f) = \pi(f)$  for all  $x \in \mathbb{R}^n$  and  $f \in C_b(\mathbb{R}^n)$ , i.e. for any  $x \in \mathbb{R}^n$ , the transition kernel  $p(t, x, dy)$  converges weakly, as  $t \rightarrow \infty$ , to  $\pi(dy)$ . However, this is not enough for  $\mathcal{W}_p$ -convergence of  $p(t, x, dy)$  to  $\pi(dy)$  (see [82, Theorem 6.9]). Assume additionally that  $1 \in \Theta_\nu$ , and let  $p \in [1, \theta_\nu] \cap \Theta_\nu$ . Since  $H$  is Hurwitz, there exists  $Q \in \mathcal{M}_+$  such that  $-(QH + H'Q) \in \mathcal{M}_+$  (see [27, Lemma 2.2]). The left-hand side of (1.16) then reads

$$2 \langle \Delta_z \tilde{b}(x), Qz \rangle = 2 \langle Hz, Qz \rangle = \langle (QH + H'Q)z, z \rangle \quad \forall x, z \in \mathbb{R}^n.$$

Now, by setting

$$c(p) := \inf_{z \in \mathbb{R}^n} \frac{-p \langle (QH + H'Q)z, z \rangle}{2|z|_Q^2},$$

the assertions of Theorem 1.4 follow. We remark here that this result does not necessarily imply ergodicity of  $\{X(t)\}_{t \geq 0}$  in the  $f$ -norm. Indeed, let  $n = 1$ , and take  $L_t \equiv 0$  for  $t \geq 0$ . Then it is easy to see that  $X_t = x e^{Ht}$  for  $t \geq 0$ . Thus,  $\pi(dx) = \delta_0(dx)$ , and  $\delta_x P_t$  converges to  $\pi(dx)$ , as  $t \rightarrow \infty$ , in  $\mathcal{W}_p$ -distance for any  $p \geq 1$ , but clearly this convergence cannot hold in the  $f$ -norm.

If  $\theta_\nu > 1$ , and  $\{X(t)\}_{t \geq 0}$  satisfies the assumptions in [7, Theorem 3.1] (which ensure that  $\{X(t)\}_{t \geq 0}$  is irreducible and aperiodic, and that the support of the corresponding irreducibility measure has nonempty interior), then according to [2, Proposition 4.3] and [79, Theorems 5.1 and 7.1] (which imply that every compact set is petite for  $\{X(t)\}_{t \geq 0}$ ) the conclusions of Theorem 1.3 (ii) hold true for any  $\theta \in (1, \theta_\nu] \cap \Theta_\nu$ . If  $\theta_\nu > 0$ , then under the same assumptions as above, [26, Theorem 5.2], [65, Proposition 6.1], and [66, Theorem 4.2] (and [2, Proposition 4.3], and [79, Theorems 5.1 and 7.1]) imply that for any  $\theta \in (0, \theta_\nu] \cap \Theta_\nu$  the process  $\{X(t)\}_{t \geq 0}$  is exponentially ergodic in the  $f$ -norm with  $f(x) = \mathcal{V}_{Q, \theta}(x) + 1$ . However, this does not necessarily imply ergodicity of  $\{X(t)\}_{t \geq 0}$  in the  $\mathcal{W}_p$ -distance. To see this take again  $n = 1$ , and let  $\{L(t)\}_{t \geq 0}$  be a one-dimensional symmetric  $\alpha$ -stable Lévy process with  $\alpha \in (0, 1)$  and symbol (characteristic exponent)  $q(\xi) = |\xi|^\alpha$ . Thus,  $1 \notin \Theta_\nu$ , and  $\theta_\nu = \alpha$ . We claim that  $\pi \notin \mathcal{P}_1(\mathbb{R})$ . Assume this is not the case. Then,

$$\int_{\mathbb{R}} \int_{\mathbb{R}} |y| p(t, x, dy) \pi(dx) = \int_{\mathbb{R}} |x| \pi(dx) < \infty \quad \forall t \geq 0.$$

In particular, for every  $t > 0$  it holds that  $\int_{\mathbb{R}} |y| p(t, x, dy) < \infty$ ,  $\pi$ -a.e. On the other hand, according to [73, Theorem 3.1], we have

$$P_t f(x) = \int_{\mathbb{R}} f(e^{Ht}x + y) \mu_t(dy)$$

for all  $t \geq 0$ ,  $x \in \mathbb{R}$  and  $f \in \mathcal{B}_b(\mathbb{R})$ , where  $\mu_t(dy)$  is a probability measure on  $\mathbb{R}$  with characteristic function

$$\hat{\mu}_t(\xi) = e^{-\int_0^t q(e^{Hs}\xi) ds} = e^{\frac{1-e^{\alpha Ht}}{\alpha H} |\xi|^\alpha}, \quad t \geq 0, \quad \xi \in \mathbb{R},$$

and  $\mathcal{B}_b(\mathbb{R})$  denotes the space of bounded functions in  $\mathcal{B}(\mathbb{R})$ . Hence,  $\mu_t(dy)$  is the law of a symmetric  $\alpha$ -stable random variable. Now, the monotone convergence theorem implies that

$$\int_{\mathbb{R}} |y| p(t, x, dy) = \int_{\mathbb{R}} |e^{Ht}x + y| \mu_t(dy) < \infty \quad \forall (t, x) \in [0, \infty) \times \mathbb{R},$$

which is impossible.

Let us mention that ergodic properties of Ornstein-Uhlenbeck processes with jumps in the  $f$ -norm, and in particular in the total variation norm, have been considered in [41, 61, 75, 83, 84].

### 3.3. Piecewise Ornstein-Uhlenbeck processes with jumps

We extend the results from the previous subsection to a class of Itô processes with a piecewise linear drift. Consider an  $n$ -dimensional SDE of the form

$$X(t) = x + \int_0^t \bar{b}(X(s)) \, ds + \int_0^t \sigma(X(s)) \, dB(s) + L(t), \quad t \geq 0, \quad x \in \mathbb{R}^n, \quad (3.1)$$

where

- (i) the function  $\bar{b}: \mathbb{R}^n \rightarrow \mathbb{R}^n$  is given by

$$\bar{b}(x) = l - M(x - \langle e, x \rangle^+ v) - \langle e, x \rangle^+ \Gamma v,$$

where  $l \in \mathbb{R}^n$ ,  $v \in \mathbb{R}^n$  has nonnegative components and satisfies  $\langle e, v \rangle = 1$  with  $e = (1, \dots, 1)' \in \mathbb{R}^n$ ,  $M \in \mathbb{R}^{n \times n}$  is a nonsingular M-matrix such that the vector  $e'M$  has nonnegative components, and  $\Gamma = \text{diag}(\gamma_1, \dots, \gamma_n)$  with  $\gamma_i \geq 0$  for  $i = 1, \dots, n$ ;

- (ii)  $\{B(t)\}_{t \geq 0}$  is a standard  $m$ -dimensional Brownian motion, and the covariance function  $\sigma: \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$  is locally Lipschitz continuous and satisfies, for some  $c > 0$ ,

$$\|\sigma(x)\|^2 \leq c(1 + |x|^2) \quad \forall x \in \mathbb{R}^n;$$

- (iii)  $\{L(t)\}_{t \geq 0}$  is a  $n$ -dimensional pure-jump Lévy process specified by a drift  $b_L \in \mathbb{R}^n$  and Lévy measure  $\nu_L(dy)$ .

Recall that a  $n \times n$  matrix  $M$  is called an M-matrix if it can be expressed as  $M = \mu \mathbb{I}_n - N$  for some  $\mu > 0$  and some nonnegative  $n \times n$  matrix  $N$  with the property that  $\rho(N) \leq \mu$ , where  $\mathbb{I}_n$  and  $\rho(N)$  denote the  $n \times n$  identity matrix and the spectral radius of  $N$ , respectively. Clearly, the matrix  $M$  is nonsingular if  $\rho(N) < \mu$ . It is well known that the SDE in (3.1) admits a unique conservative strong solution  $\{X(t)\}_{t \geq 0}$  which is a strong Markov process with càdlàg sample paths (see e.g. [2, Theorem 3.1 and Proposition 4.2]). In particular,  $\{X(t)\}_{t \geq 0}$  is an Itô process satisfying (MP) with  $b(x) = b_L + \bar{b}(x)$ ,  $a(x) = \sigma(x)\sigma(x)'$ , and  $\nu(x, dy) = \nu_L(dy)$ . This process is often called a piecewise Ornstein-Uhlenbeck process with jumps. It arises as a limit of the suitably scaled queueing processes of multiclass many-server queueing networks with heavy-tailed (bursty) arrivals and/or asymptotically negligible service interruptions. In these models, if the scheduling policy is based on a static priority assignment on the queues, then the vector  $v$  in the limiting diffusion (3.1) corresponds to a constant control. The process  $\{X(t)\}_{t \geq 0}$  also arises in many-server queues with phase-type service times, where the constant vector  $v$  corresponds to the probability distribution of the phases. For a multiclass queueing network with independent heavy-tailed arrivals, the process  $\{L(t)\}_{t \geq 0}$  is an anisotropic Lévy process consisting of independent one-dimensional symmetric  $\alpha$ -stable components. Under service interruptions,  $\{L(t)\}_{t \geq 0}$  is either a compound Poisson process, or an anisotropic Lévy process described above together with a compound Poisson component. More details on these queueing models can be found in [7, Section 4].

We first discuss the case when  $\Gamma v = 0$ . This corresponds to the case when the control gives lowest priority to queues whose abandonment rate is zero. When  $1 \in \Theta_\nu$ , we define

$$\tilde{l} := l + b_L + \int_{\mathbb{B}^c} y \nu_L(dy). \quad (3.2)$$

**Proposition 3.2.** *In addition to the assumptions of [7, Theorem 3.1] (which ensure that  $\{X(t)\}_{t \geq 0}$  is irreducible and aperiodic with irreducibility measure having support with nonempty interior), suppose that  $\Gamma v = 0$ ,  $2 \in \Theta_\nu$ , and  $\langle e, M^{-1} \tilde{l} \rangle < 0$ .*

- (i) If

$$\limsup_{|x| \rightarrow \infty} \frac{\|a(x)\|}{|x|} = 0, \quad (3.3)$$

then there exists  $Q \in \mathcal{M}_+$  such that the assertions of Theorem 1.3 (i) hold true with  $\vartheta = 1$ .

- (ii) If  $a(x)$  is bounded, and  $\int_{\mathbb{B}^c} e^{\theta|y|} \nu_L(dy) < \infty$  for some  $\theta > 0$ , then there exists  $Q \in \mathcal{M}_+$  such that the assertions of Theorem 1.3 (iii) hold.

*Proof.* (i) In [7, Theorem 3.2 (i)] it has been shown that there exist  $Q \in \mathcal{M}_+$ ,  $\bar{c} = \bar{c}(\theta) > 0$ , and  $\tilde{c} = \tilde{c}(\theta) > 0$ , such that for any  $\theta \in [1, \theta_v] \cap \Theta_v$ , we have

$$\mathcal{L}\mathcal{V}_{Q,\theta}(x) \leq \bar{c} - \tilde{c}\mathcal{V}_{Q,\theta-1}(x) \quad \forall x \in \mathbb{R}^n.$$

It is easy to see that the above relation implies that there exist  $r > 0$ ,  $\hat{c} > 0$ , and  $\check{c} > 0$ , such that

$$\mathcal{L}\mathcal{V}(x) \leq \hat{c}\mathbb{1}_{\bar{\mathcal{B}}_r}(x) - \check{c}(\mathcal{V}(x))^{(\theta-1)/\theta} \quad \forall x \in \mathbb{R}^n,$$

with  $\mathcal{V}(x) = \mathcal{V}_{Q,\theta}(x) + 1$ . The assertion now follows from [Theorem 1.3](#) (i), together with [\[2, Proposition 4.3\]](#), [\[24, Theorem 3.4\]](#), and [\[79, Theorems 5.1 and 7.1\]](#).

(ii) Let  $\tilde{b}(x) := \bar{b}(x) + \tilde{l} - l$ . As shown in the proof of [\[7, Theorem 3.2 \(ii\)\]](#), there exist  $Q \in \mathcal{M}_+$ ,  $\bar{c} = \bar{c}(\zeta) > 0$  and  $\tilde{c} = \tilde{c}(\zeta) > 0$ , such that for any  $\zeta \in (0, \theta\|Q\|^{-1/2})$ ,

$$\langle \tilde{b}(x), \nabla \tilde{V}_{Q,\zeta}(x) \rangle \leq \bar{c} - \tilde{c}\tilde{V}_{Q,\zeta}(x) \quad \forall x \in \mathbb{R}^n.$$

This together with [Lemma 2.2](#) (iii) imply that, for any  $\zeta > 0$  sufficiently small, there exist  $\hat{c} = \hat{c}(\zeta) > 0$  and  $\check{c} = \check{c}(\zeta) > 0$ , such that

$$\mathcal{L}\tilde{V}_{Q,\zeta}(x) \leq \hat{c} - \check{c}\tilde{V}_{Q,\zeta}(x) \quad \forall x \in \mathbb{R}^n.$$

Again, It is straightforward to see that the above relation implies that there exist  $r > 0$ ,  $\check{c} > 0$  and  $\hat{c} > 0$ , such that

$$\mathcal{L}\tilde{V}_{Q,\zeta}(x) \leq \check{c}\mathbb{1}_{\bar{\mathcal{B}}_r}(x) - \hat{c}\tilde{V}_{Q,\zeta}(x) \quad \forall x \in \mathbb{R}^n.$$

The assertion now follows from [Theorem 1.3](#) (iii), and the results from [\[2, 24, 79\]](#) cited in part (i).  $\square$

*Remark 3.2.* It has been shown in [\[7, Theorem 3.3 \(b\) and Lemma 5.7\]](#) that the assumptions  $1 \in \Theta_v$  and  $\langle e, M^{-1}\tilde{l} \rangle < 0$  are both necessary for the existence of an invariant probability measure of  $\{X(t)\}_{t \geq 0}$ . Using this, we can exhibit an example where we have ergodicity with respect to the  $f$ -norm but not with respect to  $\mathcal{W}_p$ -distance. Suppose that  $\Gamma v = 0$ ,  $\langle e, M^{-1}\tilde{l} \rangle < 0$ ,  $a(x)$  satisfies [\(3.3\)](#), and  $\{L(t)\}_{t \geq 0}$  is a rotationally invariant  $\alpha$ -stable process with  $\alpha \in (1, 2)$ . Then [\[7, Theorem 3.1 \(i\)\]](#) shows that  $\{X(t)\}_{t \geq 0}$  admits a unique invariant  $\pi \in \mathcal{P}_{\alpha-1-\iota}(\mathbb{R}^n)$  for  $\iota \in (0, \alpha - 1)$ , and

$$\lim_{t \rightarrow \infty} t^{\alpha-1-\iota} \|\delta_x P_t(\cdot) - \pi(\cdot)\|_{\text{TV}} = 0$$

for all  $x \in \mathbb{R}^n$  and  $\iota \in (0, \alpha - 1)$ . Here,  $\|\cdot\|_{\text{TV}}$  stands for the total variation norm, i.e. the  $f$ -norm with  $f(x) \equiv 1$ . However,  $\int_{\mathbb{R}^n} |x| \pi(dx) = \infty$  by [\[7, Theorem 3.4 \(b\)\]](#), so we cannot have convergence in  $\mathcal{W}_1$ -distance.

We next exhibit a lower bound on the polynomial rate of convergence in [Proposition 3.2](#) (i), which is analogous to [\[7, Theorem 3.4\]](#). We let

$$\tilde{\theta}_v := \sup \left\{ \theta \geq 0: \int_{\mathbb{R}^n} (\langle e, M^{-1}y \rangle^+)^{\theta} \nu_L(dy) < \infty \right\}. \quad (3.4)$$

Note that, in general,  $\tilde{\theta}_v \geq \theta_v$ . In [\[7\]](#) it is assumed that  $\{L(t)\}_{t \geq 0}$  is a compound Poisson process with drift  $b_L$ , and Lévy measure  $\nu_L(dy)$  which is supported on a half-line of the form  $\{\zeta w: \zeta \in [0, \infty)\}$  with  $\langle e, M^{-1}w \rangle > 0$ , and  $a(x)$  satisfies [\(3.3\)](#). This implies that  $\tilde{\theta}_v = \theta_v$ , and subsequently, this equality is used in the proof of [\[7, Lemma 5.7 \(b\)\]](#) to establish that, provided  $\Gamma v = 0$ ,  $\int_{\mathbb{R}^n} (\langle e, M^{-1}x \rangle^+)^{p-1} \pi(dx) < \infty$  implies  $p \in \Theta_v$  for  $p > 1$ . We use this fact, namely that the conclusions of [\[7, Lemma 5.7 \(b\)\]](#) hold under the weaker assumption that  $\tilde{\theta}_v = \theta_v$  in the proof of the following proposition.

**Proposition 3.3.** *In addition to the assumptions of [\[7, Theorem 3.1\]](#), assume that  $\Gamma v = 0$ ,  $\langle e, M^{-1}\tilde{l} \rangle < 0$ , and  $\tilde{\theta}_v = \theta_v \in (2, \infty)$ . Then, due to [Proposition 3.2](#) (i),  $\{X(t)\}_{t \geq 0}$  admits a unique invariant  $\pi \in \mathcal{P}_{\theta_v-1-\iota}(\mathbb{R}^n)$ ,  $\iota \in (0, \theta_v - 1)$ . Next, fix  $\rho \in (0, (\theta_v - 2) \wedge 1)$  and  $\varepsilon \in (\rho, 1)$ . Then, for any  $p \in [1, \theta_v - \rho - 1]$  and  $\iota \in (0, 1 - \varepsilon)$  there exist  $\bar{c} > 0$  and a diverging increasing sequence  $\{t_n\}_{n \in \mathbb{N}} \subset [0, \infty)$ , depending on these parameters, such that [\(1.9\)](#) holds with  $\theta = \theta_v - \rho$ ,  $\vartheta = \theta - 1$ , and  $\mathcal{V}(x) = \mathcal{V}_{Q,\theta}(x) + 1$ , where  $Q \in \mathcal{M}_+$  is given in [Proposition 3.2](#) (i).*

*Proof.* Observe first that  $\vartheta + \varepsilon > \theta_\nu - 1$ . Thus, according to [7, Lemma 5.7 (b)], we have

$$\int_{\mathbb{R}^n} (\langle e, M^{-1}x \rangle^+)^{\vartheta+\varepsilon} \pi(dx) = \infty.$$

The assertion now follows from the proof of Proposition 3.2 (i) (together with [2, Proposition 4.3], [24, Theorem 3.4], and [79, Theorems 5.1 and 7.1]), and Theorem 1.2 by setting  $L(x) = \langle e, M^{-1}x \rangle^+$  and  $\phi(t) = t^{(\vartheta-1)/\theta}$ .  $\square$

We now discuss the case when  $\Gamma v \neq 0$ . For  $x \in \mathbb{R}^n$ , we write  $x \geq 0$  ( $x \gneq 0$ ) to indicate that all components of  $x$  are nonnegative (nonnegative and at least one is strictly positive). Also, for  $x, y \in \mathbb{R}^n$  we write  $x \geq y$  if, and only if,  $x - y \geq 0$ .

**Proposition 3.4.** *In addition to the assumptions of [7, Theorem 3.1], suppose that  $\theta_\nu > 0$ ,*

$$\limsup_{|x| \rightarrow \infty} \frac{\|a(x)\|}{|x|^2} = 0, \tag{3.5}$$

and that one of the following holds:

- (i)  $Mv \geq \Gamma v \gneq 0$ ;
- (ii)  $M = \text{diag}(m_1, \dots, m_d)$  with  $m_i > 0$ ,  $i = 1, \dots, n$ , and  $\Gamma v \neq 0$ .

Then there exists  $Q \in \mathcal{M}_+$  such that the assertions of Theorem 1.3 (ii) hold true.

*Proof.* In [7, Theorem 3.5] it has been shown that there exist  $Q \in \mathcal{M}_+$ ,  $\bar{c} = \bar{c}(\theta) > 0$ , and  $\tilde{c} = \tilde{c}(\theta) > 0$ , such that for any  $\theta \in (0, \theta_\nu] \cap \Theta_\nu$ , we have

$$\mathcal{L}\mathcal{V}_{Q,\theta}(x) \leq \bar{c} - \tilde{c}\mathcal{V}_{Q,\theta}(x) \quad \forall x \in \mathbb{R}^n.$$

As in Proposition 3.2, it is easy to see that the above relation implies that there exist  $r > 0$ ,  $\hat{c} > 0$  and  $\check{c} > 0$ , such that

$$\mathcal{L}\mathcal{V}(x) \leq \hat{c}\mathbf{1}_{\mathbb{B}_r}(x) - \check{c}\mathcal{V}(x) \quad \forall x \in \mathbb{R}^n,$$

with  $\mathcal{V}(x) = \mathcal{V}_{Q,\theta}(x) + 1$ . The assertion now follows from Theorem 1.3 (ii), together with the results from [2, 24, 79] cited in the proof of Proposition 3.3.  $\square$

In the case when  $\Gamma v \neq 0$  (under (i) or (ii) in Proposition 3.4) the dynamics are contractive in the  $\mathcal{W}_p$ -distance. This is shown by establishing an asymptotic flatness (uniform dissipativity) property for  $\{X(t)\}_{t \geq 0}$ . As a consequence, we assert exponential ergodicity of  $\{X(t)\}_{t \geq 0}$  with respect to  $\mathcal{W}_p$ , without assuming irreducibility and aperiodicity, i.e. we allow the SDE in (3.1) to be degenerate.

**Proposition 3.5.** *Suppose that  $2 \in \Theta_\nu$ ,  $\sigma(x)$  is Lipschitz continuous, and either (i) or (ii) in Proposition 3.4 holds. Then there exists  $Q \in \mathcal{M}_+$  such that the matrices*

$$MQ + QM, \quad \text{and} \quad (M - ev'(M - \Gamma))Q + Q(M - (M - \Gamma)ve') \tag{3.6}$$

are in  $\mathcal{M}_+$ . Let  $\underline{\kappa}$  denote the smallest eigenvalue of the positive definite matrices in (3.6), and  $\bar{\lambda}_Q, \lambda_Q$  denote the largest, smallest eigenvalue of  $Q$ , respectively. For  $p \geq 1$ , let

$$c(p) := \frac{p}{2} \left( \frac{\underline{\kappa}}{\bar{\lambda}_Q} - \frac{(p-1) \text{Lip}^2(\sqrt{Q}\sigma)}{\lambda_Q} \right),$$

where  $\text{Lip}(\sqrt{Q}\sigma)$  is the Lipschitz constant of  $\sqrt{Q}\sigma(x)$  with respect to the Hilbert-Schmidt norm, and suppose that  $c(p) > 0$  for some  $p \in [2, \theta_\nu] \cap \Theta_\nu$ . Then the assertions of Theorem 1.4 hold true. If  $\sigma(x) \equiv \sigma$  and  $1 \in \Theta_\nu$ , the assertions of Theorem 1.4 hold true for any  $p \in [1, \theta_\nu] \cap \Theta_\nu$ .

*Proof.* Existence of the matrix  $Q$  has been proven in [7, Theorem 3.5]. We prove that (1.16) holds with  $c(p)$  defined above. First, clearly,

$$\begin{aligned} \text{Tr}(\tilde{a}(x; y-x)Q) + (p-2) \|\sqrt{Q} \Delta_{y-x} \sigma(x)\|^2 &\leq (p-1) \text{Lip}^2(\sqrt{Q} \sigma) |y-x|^2 \\ &\leq \frac{(p-1) \text{Lip}^2(\sqrt{Q} \sigma)}{\lambda_Q} |y-x|_Q^2 \end{aligned} \quad (3.7)$$

for all  $x, y \in \mathbb{R}^n$ . We next discuss the term  $\langle \Delta_{y-x} \tilde{b}(x), Q(y-x) \rangle$  for  $x, y \in \mathbb{R}^n$ . Clearly,  $\Delta_{y-x} \tilde{b}(x) = \Delta_{y-x} \bar{b}(x)$  for  $x, y \in \mathbb{R}^n$ . With  $\hat{v} = -M^{-1}(Mv - \Gamma v)$ , we have  $\bar{b}(x) = l - M(x + \langle e, x \rangle^+ \hat{v})$ . If both  $x$  and  $y$  are on the same half-space, i.e.  $\langle e, x \rangle \geq 0$  and  $\langle e, y \rangle \geq 0$ , or the opposite, then

$$\langle \Delta_{y-x} \tilde{b}(x), Q(y-x) \rangle \leq -\frac{\kappa}{2} |y-x|^2.$$

So suppose, without loss of generality, that  $\langle e, x \rangle \geq 0$  and  $\langle e, y \rangle \leq 0$ . Then we have

$$\langle y-x, Q \bar{b}(x) \rangle = \langle y-x, Ql \rangle - \langle y-x, QMx \rangle - \langle y-x, QM\hat{v}e'x \rangle \quad (3.8a)$$

$$\langle y-x, Q \bar{b}(y) \rangle = \langle y-x, Ql \rangle - \langle y-x, QMy \rangle. \quad (3.8b)$$

We distinguish two cases.

(i)  $\langle y-x, QM\hat{v}e'x \rangle \leq 0$ . Then of course subtracting (3.8a) from (3.8b), we obtain

$$\begin{aligned} \langle \Delta_{y-x} \tilde{b}(x), Q(y-x) \rangle &= -\langle y-x, QM(y-x) \rangle + \langle y-x, QM\hat{v}e'x \rangle \\ &\leq -\langle y-x, QM(y-x) \rangle \\ &\leq -\frac{\kappa}{2} |y-x|^2. \end{aligned}$$

(ii)  $\langle y-x, QM\hat{v}e'x \rangle > 0$ . Since  $\langle e, x \rangle \geq 0$ , we must have  $\langle y-x, QM\hat{v} \rangle > 0$ . This in turn implies, since  $\langle e, y \rangle \leq 0$ , that

$$\langle y-x, QM\hat{v}e'y \rangle \leq 0. \quad (3.9)$$

Adding (3.8a) and (3.9) and subtracting (3.8b) from the sum, we obtain

$$\begin{aligned} \langle \Delta_{y-x} \tilde{b}(x), Q(y-x) \rangle &\leq -\langle y-x, QM(y-x) \rangle - \langle y-x, QM\hat{v}e'(y-x) \rangle \\ &\leq -\langle y-x, QM(\mathbb{I}_n + \hat{v}e')(y-x) \rangle \leq -\frac{\kappa}{2} |y-x|^2. \end{aligned} \quad (3.10)$$

Finally, combining (3.7) and (3.10), we obtain

$$\begin{aligned} 2 \langle \Delta_{y-x} \tilde{b}(x), Q(y-x) \rangle + \text{Tr}(\tilde{a}(x; y-x)Q) + (p-2) \|\sqrt{Q} \Delta_{y-x} \sigma(x)\|^2 \\ \leq \left( -\frac{\kappa}{\lambda_Q} + \frac{(p-1) \text{Lip}^2(\sqrt{Q} \sigma)}{\lambda_Q} \right) |y-x|_Q^2 \\ = -\frac{2c(p)}{p} |y-x|_Q^2 \quad \forall x, y \in \mathbb{R}^n, \end{aligned}$$

thus completing the proof.  $\square$

The hypothesis in Proposition 3.5 that  $c(p) > 0$  is, of course, always true if  $\sigma(x) \equiv \sigma$ , in which case we have  $c(p) = p \frac{\kappa}{2\lambda_Q}$ . This is the scenario for multiclass queueing models with service interruptions described in [7, Section 4.2].

Some examples of degenerate SDEs of the form (3.1) for which Proposition 3.5 is applicable are the following.

(i)  $\{L(t)\}_{t \geq 0}$  is given by  $L(t) = R\tilde{L}(t)$  for  $t \geq 0$ , where  $R \in \mathbb{R}^{n \times r}$  has rank smaller than  $\min\{n, r\}$ , and  $\{\tilde{L}(t)\}_{t \geq 0}$  is a  $r$ -dimensional Lévy process. As a special case  $\{\tilde{L}(t)\}_{t \geq 0}$  may be composed of mutually independent  $\alpha$ -stable processes. This is the case in the queueing example described below.

(ii)  $\{L(t)\}_{t \geq 0}$  is a degenerate subordinate Brownian motion, as studied in [87].

The following is an example of a degenerate SDE that arises in applications for which Proposition 3.4 is applicable. Consider a two class  $GI/M/k + M$  queue with class-1 jobs having a Poisson process, and class-2 jobs having a heavy-tailed renewal arrival process. Service and patience times are exponentially distributed with rates  $m_i$  and  $\gamma_i$  for  $i = 1, 2$ , respectively. Assume that the arrival, service and abandonment processes are mutually independent, and that the number of servers is  $k$ . Consider a sequence of such models indexed by  $k$ , operating in the critically loaded asymptotic modified Halfin-Whitt regime as  $k \rightarrow \infty$ . Let  $\{A_i^k(t)\}_{t \geq 0}$  denote the arrival process for class  $i = 1, 2$ , with arrival rates  $\lambda_i^k$ . Assume that  $m_i$  and  $\gamma_i$  for  $i = 1, 2$  are independent of  $k$ , and that  $\frac{\lambda_i^k}{k} \rightarrow \lambda_i > 0$  as  $k \rightarrow \infty$ , for  $i = 1, 2$ . The arrival process  $\{A_1^k(t)\}_{t \geq 0}$  satisfies a functional central limit theorem (FCLT) with a Brownian motion limit  $\{\hat{A}_1(t)\}_{t \geq 0} = \{\sqrt{\lambda_1} B_1(t)\}_{t \geq 0}$ , where  $\{B_1(t)\}_{t \geq 0}$  is a standard Brownian motion, i.e.

$$\{\hat{A}_1^k(t)\}_{t \geq 0} = \{k^{-1/2}(A_1^k(t) - \lambda_1^k t)\}_{t \geq 0} \xrightarrow[k \rightarrow \infty]{J_1} \{\hat{A}_1(t)\}_{t \geq 0}.$$

Here,  $\xrightarrow{J_1}$  denotes the convergence in the space  $D = D([0, \infty), \mathbb{R})$  of càdlàg functions endowed with the Skorokhod  $J_1$  topology. We assume that the arrival process  $\{A_2^k(t)\}_{t \geq 0}$  satisfies a FCLT with a symmetric  $\alpha$ -stable Lévy process  $\{\hat{A}_2(t)\}_{t \geq 0}$ ,  $\alpha \in (1, 2)$ , in the limit, i.e.

$$\{\hat{A}_2^k(t)\}_{t \geq 0} = \{k^{-1/\alpha}(A_2^k(t) - \lambda_2^k t)\}_{t \geq 0} \xrightarrow[k \rightarrow \infty]{M_1} \{\hat{A}_2(t)\}_{t \geq 0}.$$

Here,  $\xrightarrow{M_1}$  denotes the convergence in the space  $D$  with the  $M_1$  topology. Let  $\rho_i^k = \frac{\lambda_i^k}{k m_i}$  and  $\rho_i = \frac{\lambda_i}{m_i}$  for  $i = 1, 2$ . The modified Halfin-Whitt regime requires the parameters satisfy

$$\lim_{k \rightarrow \infty} k^{1-1/\alpha} \left( 1 - \sum_{i=1}^2 \rho_i^k \right) = \hat{\rho} \in \mathbb{R}, \quad \text{and} \quad \sum_{i=1}^2 \rho_i = 1.$$

In addition, we assume that  $k^{-1/\alpha}(\lambda_i^k - k \lambda_i) \rightarrow l_i$  as  $k \rightarrow \infty$  for  $i = 1, 2$ . Next, let  $\{X_i^k(t)\}_{t \geq 0}$  denote the number of class- $i$  jobs in the system. Define the scaled processes  $\hat{X}_i^k(t) = k^{-1/\alpha}(X_i^k(t) - k \rho_i t)$  for  $t \geq 0$ . Let  $\{U_i^k(t)\}_{t \geq 0}$  be the scheduling control process, representing allocations of service capacity to class  $i$ . Let  $\hat{X}^k(t) = (\hat{X}_1^k(t), \hat{X}_2^k(t))'$  and  $U^k(t) = (U_1^k(t), U_2^k(t))'$  for  $t \geq 0$ . We consider work conserving and preemptive scheduling policies resulting in constant controls at the limit, i.e.  $\{U^k(t)\}_{t \geq 0} \xrightarrow[k \rightarrow \infty]{J_1} \{V(t)\}_{t \geq 0}$ , where  $V(t) = v$  for  $t \geq 0$  with  $v \in \mathbb{R}^2$  being a probability vector. Then, as in [7, Theorem 4.1], it can shown that  $\{\hat{X}^k(t)\}_{t \geq 0} \xrightarrow[k \rightarrow \infty]{M_1} \{X(t)\}_{t \geq 0}$ , where the limit process  $\{X(t)\}_{t \geq 0}$  is a solution to the following two-dimensional degenerate  $\alpha$ -stable driven SDE:

$$\begin{aligned} dX_1(t) &= \left( l_1 - m_1(X_1(t) - \langle e, X(t) \rangle^+ v_1) - \gamma_1 \langle e, X(t) \rangle^+ v_1 \right) dt, \\ dX_2(t) &= \left( l_2 - m_2(X_2(t) - \langle e, X(t) \rangle^+ v_2) - \gamma_2 \langle e, X(t) \rangle^+ v_2 \right) dt + d\hat{A}_2(t), \end{aligned}$$

which is (3.1) with  $l = (l_1, l_2)'$ ,  $M = \text{diag}(m_1, m_2)$ ,  $\Gamma = \text{diag}(\gamma_1, \gamma_2)$ ,  $\sigma(x) = (0, 0)'$ , and  $L(t) = (0, \hat{A}_2(t))'$  for  $t \geq 0$ . Observe that the process  $\{X(t)\}_{t \geq 0}$  does not fall into any of the four categories in [7, Theorem 3.1]. In fact, one can consider multiple classes of jobs with all heavy-tailed arrival processes that have different scaling parameters  $\alpha_i$ 's for  $i = 1, \dots, \bar{k}$ , in their corresponding FCLTs. The centered queueing process should be scaled as  $k^{-1/\alpha}$ , where  $\alpha := \min_{i=1, \dots, \bar{k}} \{\alpha_i\}$ , and the limit process has the components  $\{X_i(t)\}_{t \geq 0}$  driven by independent  $\alpha$ -stable processes if the arrival process of class  $i$  has the parameter  $\alpha_i$  equal to the minimum  $\alpha$ , and the other components are degenerate without stochastic driving terms.

We remark here that without assuming irreducibility and aperiodicity, establishing subgeometric ergodicity in the case  $\Gamma v = 0$  is difficult. Consider the following example. Let  $n = 1$ ,  $\sigma(x) \equiv 0$ ,  $L(t) \equiv 0$  for  $t \geq 0$ , and

$$\bar{b}(x) = \begin{cases} -1, & x \geq 0, \\ -1 - x, & x \leq 0. \end{cases}$$

Clearly,  $\bar{b}(x)$  satisfies all the assumptions in [7], and

$$X^x(t) = x + \int_0^t \bar{b}(X^x(s)) ds, \quad t \geq 0, \quad x \in \mathbb{R}.$$

A straightforward calculation shows that

$$X^x(t) = \begin{cases} \begin{cases} x - t, & 0 \leq t \leq x \\ e^{x-t} - 1, & t \geq x \end{cases}, & x \geq 0, \\ -1 + e^{-t} + x e^{-t}, & x \leq 0. \end{cases}$$

Let

$$d(x, y) := \frac{|x - y|}{1 + |x - y|}, \quad x, y \in \mathbb{R}.$$

Then it is easy to see that the conditions (1)–(3) in [12, Theorem 2.4] hold. However, condition (4) does not hold. Namely, for arbitrary  $t_0 > 0$  let  $x, y > t_0$ . Then,  $d(X^x(t), X^y(t)) = d(x, y)$  for all  $t_0 \leq t \leq x \wedge y$ .

Let us mention that ergodic properties of piecewise Ornstein-Uhlenbeck processes with jumps in the total variation norm have been considered in [7, 23, 70].

### 3.4. Piecewise Ornstein-Uhlenbeck processes with jumps under stationary Markov controls

In [Subsection 3.3](#) we consider a model with a constant control, i.e. with the vector  $v \in \Delta := \{u \in \mathbb{R}^n : u \geq 0, \langle e, u \rangle = 1\}$  being constant and fixed. If the scheduling policy (control) is a function of the state of the system, then  $v(x)$  in the limiting SDE (3.1) is, in general, a Borel measurable map from  $\mathbb{R}^n$  to  $\Delta$ . We call such a  $v(x)$  a stationary Markov control and denote the set of such controls by  $\mathfrak{U}_{\text{SM}}$ . If  $L_t \equiv 0$  for  $t \geq 0$ , or it is a compound Poisson process, it follows from the results in [37] that, under any  $v \in \mathfrak{U}_{\text{SM}}$ , (3.1) admits a unique conservative strong solution which is a strong Markov process with càdlàg sample paths. In the general case, we consider the subclass of stationary Markov controls for which

$$\bar{b}_v(x) = l - M(x - \langle e, x \rangle^+ v(x)) - \langle e, x \rangle^+ \Gamma v(x),$$

is locally Lipschitz continuous. We let  $\tilde{\mathfrak{U}}_{\text{sm}}$  denote the class of such controls. Clearly, for any  $v \in \tilde{\mathfrak{U}}_{\text{sm}}$ , the drift  $\bar{b}_v(x)$  has at most linear growth. Other parameters are as in [Subsection 3.3](#). Again, the SDE of the form (3.1), with  $\bar{b}(x)$  replaced by  $\bar{b}_v(x)$ , admits a unique conservative strong solution  $\{X(t)\}_{t \geq 0}$  which is a strong Markov process with càdlàg sample paths. Also, it is an Itô process satisfying (MP) with  $b(x) = b_L + \bar{b}_v(x)$ ,  $a(x) = \sigma(x)\sigma(x)'$ , and  $\nu(x, dy) = \nu_L(dy)$ .

Recently, in [6] the authors have studied ergodic properties with respect to the total variation norm of this model with  $\{L(t)\}_{t \geq 0}$  being either (or a combination of) a rotationally invariant  $\alpha$ -stable Lévy process, an anisotropic Lévy process consisting of independent one-dimensional symmetric  $\alpha$ -stable components, or a compound Poisson process. Observe that in this situation we cannot follow the procedure from the constant control case. Namely, the matrices  $Q \in \mathcal{M}_+$  used in constructing the appropriate Lyapunov functions  $\mathcal{V}(x)$  depend on  $v$ .

**Proposition 3.6.** *Grant the assumptions of [7, Theorem 3.1], and suppose that  $M = \text{diag}(m_1, \dots, m_n)$ , with  $m_i > 0$  for  $i = 1, \dots, n$ .*

- (i) *Assume that the diagonal components of  $\Gamma$  are strictly positive,  $a(x)$  satisfies (3.5), and  $\{L(t)\}_{t \geq 0}$  is either a rotationally invariant  $\alpha$ -stable Lévy process, an anisotropic Lévy process consisting of independent one-dimensional symmetric  $\alpha$ -stable components (in both cases we assume that  $\alpha \in (1, 2)$ ), or a compound Poisson process satisfying  $1 \in \Theta_\nu$ . We allow  $\{L(t)\}_{t \geq 0}$  to have a drift. Then, for any  $v \in \tilde{\mathfrak{U}}_{\text{sm}}$  and  $\theta \in [1, \theta_\nu] \cap \Theta_\nu$ , the assertions of [Theorem 1.1](#) (iii) hold true with  $\eta = \theta$ , and  $\mathcal{V}(x) = (\tilde{\mathcal{V}}(x))^\theta + 1$ , where  $\tilde{\mathcal{V}} \in C^2(\mathbb{R}^n)$  (given explicitly in [6, Definition 1]) is bounded from below away from zero, is Lipschitz continuous, and satisfies*

$$0 < \liminf_{|x| \rightarrow \infty} \frac{\tilde{\mathcal{V}}(x)}{|x|} \leq \limsup_{|x| \rightarrow \infty} \frac{\tilde{\mathcal{V}}(x)}{|x|} < \infty.$$

- (ii) Assume  $\langle e, M^{-1}\tilde{l} \rangle < 0$ , where  $\tilde{l}$  is given in (3.2),  $a(x)$  satisfies (3.3), and  $\{L(t)\}_{t \geq 0}$  is a pure-jump Lévy process (possibly with drift) satisfying  $2 \in \Theta_\nu$ . Then, for any  $v \in \tilde{\mathfrak{U}}_{\text{sm}}$  and  $\theta \in [2, \theta_\nu] \cap \Theta_\nu$ , the assertions of Theorem 1.1 (i) and (ii) hold true with  $\phi(t) = t^{(\theta-1)/\theta}$ ,  $\eta = \theta - 1$ , and  $\mathcal{V}(x)$  as in (i).
- (iii) In addition to the assumptions in (ii) assume that  $\theta_\nu = \theta \in (2, \infty)$ , where  $\theta_\nu$  is given in (3.4). Then, due to (ii), for any  $v \in \tilde{\mathfrak{U}}_{\text{sm}}$ ,  $\{X(t)\}_{t \geq 0}$  admits a unique invariant  $\pi_v \in \mathcal{P}_{\theta_\nu-1-\iota}(\mathbb{R}^n)$  for  $\iota \in (0, \theta_\nu - 1)$ . Next, fix  $\rho \in (0, (\theta_\nu - 2) \wedge 1)$  and  $\varepsilon \in (\rho, 1)$ . Then, for any  $v \in \tilde{\mathfrak{U}}_{\text{sm}}$  such that  $\Gamma v(x) = 0$  a.e.,  $p \in [1, \theta_\nu - \rho - 1]$  and  $\iota \in (0, 1 - \varepsilon)$ , there exist  $\bar{c} > 0$  and a diverging increasing sequence  $\{t_n\}_{n \in \mathbb{N}} \subset [0, \infty)$ , depending on these parameters, such that (1.9) holds for the corresponding  $\pi_v(dx)$  with  $\theta = \theta_\nu - \rho$ ,  $\vartheta = \theta - 1$ , and  $\mathcal{V}(x)$  as above.

*Proof.* (i) Observe first that in the case when  $\{L(t)\}_{t \geq 0}$  is a rotationally invariant  $\alpha$ -stable Lévy process or an anisotropic Lévy process consisting of independent one-dimensional symmetric  $\alpha$ -stable components,  $\Theta_\nu = [0, \alpha)$ . In [6, Theorem 3 and the discussion after Theorem 5] it has been shown that for any  $v \in \tilde{\mathfrak{U}}_{\text{sm}}$  and  $\theta \in [1, \theta_\nu] \cap \Theta_\nu$  there exist  $\bar{c} = \bar{c}(\theta, v) > 0$  and  $\tilde{c} = \tilde{c}(\theta, v) > 0$ , such that

$$\mathcal{L}(\bar{\mathcal{V}}^\theta)(x) \leq \bar{c} - \tilde{c} (\bar{\mathcal{V}}(x))^\theta \quad \forall x \in \mathbb{R}^n.$$

It is easy to see that the above relation implies that there exist  $r > 0$ ,  $\hat{c} > 0$ , and  $\check{c} > 0$ , such that

$$\mathcal{L}\mathcal{V}(x) \leq \hat{c} \mathbb{1}_{\bar{\mathcal{B}}_r}(x) - \check{c} \mathcal{V}(x) \quad \forall x \in \mathbb{R}^n.$$

The assertion then follows from Theorem 1.1 (iii), together with [2, Proposition 4.3], [24, Theorem 3.4], and [79, Theorems 5.1 and 7.1].

- (ii) In Theorem 5 and the discussion following the proof of this theorem in [6] it has been shown that for any  $v \in \tilde{\mathfrak{U}}_{\text{sm}}$  and  $\theta \in (1, \theta_\nu] \cap \Theta_\nu$  there exist  $r = r(\theta, v) > 0$ ,  $\bar{c} = \bar{c}(\theta, v) > 0$ , and  $\tilde{c} = \tilde{c}(\theta, v) > 0$ , such that

$$\mathcal{L}(\bar{\mathcal{V}}^\theta)(x) \leq \bar{c} \mathbb{1}_{\bar{\mathcal{B}}_r} - \tilde{c} (\bar{\mathcal{V}}(x))^{\theta-1} \quad \forall x \in \mathbb{R}^n.$$

It is easy to see that the above relation implies that there exist  $\hat{r} > 0$ ,  $\hat{c} > 0$ , and  $\check{c} > 0$ , such that

$$\mathcal{L}\mathcal{V}(x) \leq \hat{c} \mathbb{1}_{\bar{\mathcal{B}}_{\hat{r}}}(x) - \check{c} (\mathcal{V}(x))^{(\theta-1)/\theta} \quad \forall x \in \mathbb{R}^n,$$

with  $\mathcal{V}(x)$  given as above. The assertion now follows from Theorem 1.1 (i) and (ii), together with the results from [2, 24, 79] cited in part (i).

- (iii) Clearly,  $\vartheta + \varepsilon > \theta_\nu - 1$ . Thus, according to [7, Lemma 5.7 (b)],

$$\int_{\mathbb{R}^n} (\langle e, M^{-1}x \rangle^+)^{\vartheta+\varepsilon} \pi_v(dx) = \infty.$$

The assertion now follows from (the proof of) (ii) (together with the results from [2, 24, 79] cited in part (i)), and Theorem 1.2 by setting  $L(x) = \bar{\mathcal{V}}(x)$  and  $\phi(t) = t^{(\theta-1)/\theta}$ .  $\square$

As discussed in Subsection 3.3, the hypothesis that  $\theta_\nu = \theta$  is true if  $\{L(t)\}_{t \geq 0}$  is a compound Poisson process (possibly with drift) with Lévy measure  $\nu_L(dy)$  supported on a half-line of the form  $\{tw : t \in [0, \infty)\}$  with  $\langle e, M^{-1}w \rangle > 0$ .

Ergodic properties in the  $f$ -norm of piecewise Ornstein-Uhlenbeck processes with jumps under stationary Markov controls have been considered in [5, 6].

### 3.5. State-space models

Let  $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$  be continuous, and such that  $|F(x)| \leq c|x|$  for some  $c > 0$  and all  $x \in \mathbb{R}^n$ . Further, let  $X(0)$  be an  $\mathbb{R}^n$ -valued random variable, and let  $\{W(k)\}_{k \in \mathbb{N}}$  be a sequence of i.i.d.  $\mathbb{R}^n$ -valued random variables independent of  $X(0)$ . Assume that the common distribution of  $\{W(k)\}_{k \in \mathbb{N}}$  has a nontrivial absolutely continuous component which is bounded away from zero in a neighborhood of the origin. Then the Markov process defined by

$$X(k+1) = F(X(k)) + W(k+1), \quad k \geq 0,$$

is irreducible, aperiodic, and all compact sets are petite (see [78, Proposition 5.2]). Further, assume that there exist constants  $l \in \mathbb{N}$ ,  $l \geq 2$ ,  $\varepsilon \in (0, 1)$ , and  $\bar{c}, r > 0$ , such that

$$\mathbb{E}[|W_1|^l] < \infty, \quad \text{and} \quad |F(x)| \leq c|x| - \bar{c}|x|^{1-\varepsilon} \quad \forall x \in \mathcal{B}_{\bar{r}}^c.$$

**Proposition 3.7.** *Under the above assumptions, the assertions of Theorem 1.1 (i) and (ii) hold with  $\mathcal{V}(x) = |x|^l$ ,  $\phi(t) = t^{(l-1)/l}$ , and  $\eta = l - 1$ .*

*Proof.* In [78, Proposition 5.2] it has been proved that the Foster-Lyapunov condition in (1.1) holds with  $\mathcal{V}(x)$  and  $\phi(t)$  as above, and  $C = \mathcal{B}_{\bar{r}}$  for some  $\bar{r} > 0$ . The result now follows from Theorem 1.1 (i) and (ii).  $\square$

Ergodic properties of state-space models in the  $f$ -norm have been studied in [32, 78].

### 3.6. Backward recurrence time chain

Let  $\{p_i\}_{i \geq 0} \subset (0, \infty)$  be such that  $p_0 = 1$ ,  $p_i < 1$  for  $i \in \mathbb{N}$ , and  $\prod_{j=0}^i p_j \rightarrow 0$ , as  $i \rightarrow \infty$ . Let  $\{X(k)\}_{k \geq 0}$  be a Markov process on  $\{0, 1, \dots\}$  defined by the transition kernel  $p(i, i+1) = 1 - p(i, 0) := p_i$  for  $i \geq 0$ . The process  $\{X(k)\}_{k \geq 0}$  is irreducible and aperiodic, and it admits a unique invariant  $\pi \in \mathcal{P}(\{0, 1, \dots\})$  if, and only, if

$$c := \sum_{i=1}^{\infty} \prod_{j=1}^i p_j < \infty.$$

In this case,  $\pi(0) = \pi(1) = (2+c)^{-1}$ , and  $\pi(i) = (2+c)^{-1} \prod_{j=0}^{i-1} p_j$  for  $i \geq 2$ .

**Proposition 3.8.** (i) *If there exist  $i_0 \in \mathbb{N}$  and  $\alpha > 1$ , such that  $p_i = \frac{1+\alpha}{i}$  for  $i \geq i_0$ , then the assertions of Theorem 1.1 (i) and (ii) hold with*

$$\mathcal{V}(i) = i^{\beta(1+\alpha)} + 1, \quad \phi(t) = t^{1 - \frac{1}{\beta(1+\alpha)}}, \quad \text{and} \quad \eta = \beta(1+\alpha) - 1 \quad \text{for} \quad \beta \in [2/(1+\alpha), 1).$$

(ii) *Under the assumptions in (i),  $\pi \in \mathcal{P}_{\alpha-\iota}(\{0, 1, \dots\})$  for  $\iota \in (0, \alpha)$ . Next, fix  $\rho \in (0, (\alpha - 1) \wedge 1)$  and  $\varepsilon \in [\rho, 1)$ . Then, for every  $p \in [1, \alpha - \rho]$  and  $\iota \in (0, 1 - \varepsilon)$  there exist a positive constant  $c$  and a diverging increasing sequence  $\{t_n\}_{n \in \mathbb{N}} \subset [0, \infty)$ , depending on these parameters, such that (1.9) holds with  $\mathcal{V}(i)$  as above,  $\theta = 1 + \alpha - \rho$ , and  $\vartheta = \alpha - \rho$ .*

*Proof.* (i) In [25, Section 3] it has been shown that the Foster-Lyapunov condition in (1.1) holds with a Lyapunov function  $\tilde{\mathcal{V}}(i)$  which asymptotically behaves like  $\mathcal{V}(i)$ ,  $\phi(t)$  as above, and  $C$  being a finite set for any  $\alpha > 0$  and  $\beta \in (0, 1)$ . Taking into account (1.2), the assertion follows.

(ii) From the assumptions on the sequence  $\{p_i\}_{i \geq 0}$  we see that  $\lim_{i \rightarrow \infty} i^{1+\alpha} \pi(i) > 0$ . Now, since  $\vartheta + \varepsilon - 1 - \alpha \geq -1$ , we have  $\sum_{i=0}^{\infty} i^{\vartheta+\varepsilon} \pi(i) = \infty$ . The assertion now follows from Theorem 1.2 by taking  $L(i) = i$ .  $\square$

## 4. Concluding Remarks

We remark on some other approaches in the study of exponential or subexponential ergodicity of Markov processes. By analyzing polynomial moments of hitting times of compact sets directly, polynomial ergodicity results are established in [80, Theorem 6] for a class of irreducible (with respect to the Lebesgue measure) and aperiodic diffusion processes. In a follow-up work, by using analogous techniques, the same author established polynomial ergodicity of a class of diffusion processes without directly assuming irreducibility and aperiodicity of the process, but employing instead a so-called (local) Dobrushin condition (also known as Markov-Dobrushin condition) [81, Theorem 6]. This approach is based on a Foster-Lyapunov condition of the form (1.1), and instead of assuming irreducibility and aperiodicity of  $\{X(t)\}_{t \geq 0}$ , it is assumed that (i)  $\mathcal{V}(x)$  has precompact sub-level sets, and (ii) for every  $\delta > 0$  there exists  $t_\delta \in \mathbb{T} \setminus \{0\}$  such that

$$\sup_{(x,y) \in \{(u,v) : \mathcal{V}(u) + \mathcal{V}(v) \leq \delta\}} \|p(t_\delta, x, dz) - p(t_\delta, y, dz)\|_{\text{TV}} < 1,$$

(see [53, Chapter 3]). Observe that this condition actually means that for each  $(x, y)$  satisfying  $\mathcal{V}(x) + \mathcal{V}(y) \leq \delta$  the probability measures  $p(t_\delta, x, dz)$  and  $p(t_\delta, y, dz)$  are not mutually singular. Intuitively, the Dobrushin condition encodes irreducibility and aperiodicity of  $\{X(t)\}_{t \geq 0}$ , and petiteness of sub-level sets of  $\mathcal{V}(x)$ . Based on these assumptions, and using an appropriate Markov coupling of  $\{M(t)\}_{t \geq 0}$ , it follows that the  $\Phi^{-1}$ -modulated moment of the corresponding coupling time is finite and controlled by  $\mathcal{V}(x) + \mathcal{V}(y)$ . This then implies (sub)geometric ergodicity of  $\{X(t)\}_{t \geq 0}$  in the total variation norm (see [38, Theorem 4.1] or [53, Chapter 3]).

We remark that irreducibility and aperiodicity (together with (1.1)) imply that the Dobrushin condition holds on the Cartesian product of any petite set with itself. Namely, according to [65, Proposition 6.1], for any petite set  $C$  there exists  $t_C \in \mathbb{T} \setminus \{0\}$  such that for the measure  $\chi(dt)$  (in the definition of petiteness) the Dirac measure in  $t_C$  can be taken (together with some non-trivial measure  $\nu_\chi(dx)$ ). Thus,  $p(t_C, x, B) \geq \nu_\chi(B)$  for any  $x \in C$  and  $B \in \mathfrak{B}(\mathbb{X})$ , which implies that

$$\sup_{(x,y) \in C \times C} \|p(t_C, x, dz) - p(t_C, y, dz)\|_{\text{TV}} < 1. \tag{4.1}$$

If, in addition,  $\{X(t)\}_{t \geq 0}$  is  $C_b$ -Feller (i.e.  $x \mapsto \int_{\mathbb{X}} f(y) p(t, x, dy)$  is continuous and bounded for any  $t \in \mathbb{T}$  and any continuous and bounded function  $f(x)$ ), and the support of the corresponding irreducibility measure has nonempty interior, then every compact set is petite (see [79, Theorems 5.1 and 7.1]) and thus (4.1) holds for any bounded set  $C$ . This shows that, at least in this particular situation, the approach based on the Dobrushin condition is more general than the approach based on irreducibility and aperiodicity. Situations where it has a clear advantage are discussed in [1, 54]. In [54], the author considers a Markov process obtained as a solution to a Lévy-driven SDE with highly irregular coefficients and noise term; while in [1], 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 irreducibility and aperiodicity of the processes, whereas the authors obtain (4.1) for any compact set  $C$  (see [54, Theorem 1.3] and [1, Lemma 3]). For additional results on ergodic properties of Markov processes based on the Dobrushin condition we refer the readers to [38, 53].

### Acknowledgements

We thank the anonymous referee for the helpful comments that have led to significant improvements of the results in the article. This research was supported in part by the Army Research Office through grant W911NF-17-1-001, and in part by the National Science Foundation through grants DMS-1715210, CMMI-1635410, and DMS-1715875. Financial support through the Alexander von Humboldt Foundation (No. HRV 1151902 HFST-E) and the Croatian Science Foundation under the project 8958 (for N. Sandrić) is gratefully acknowledged.

### References

- [1] ABOURASHCHI, N. and VERETENNIKOV, A. Y. (2010). On stochastic averaging and mixing. *Theory Stoch. Process.* **16** 111–129. [MR2779833](#)
- [2] ALBEVERIO, S., BRZEŃNIAK, 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](#)
- [3] ANDRIEU, C., FORT, G. and VIHOLA, M. (2015). Quantitative convergence rates for subgeometric Markov chains. *J. Appl. Probab.* **52** 391–404. [MR3372082](#)
- [4] ARAPOSTATHIS, A., BORKAR, V. S. and GHOSH, M. K. (2012). *Ergodic control of diffusion processes. Encyclopedia of Mathematics and its Applications* **143**. Cambridge University Press, Cambridge. [MR2884272](#)
- [5] ARAPOSTATHIS, A., HMEDI, H. and PANG, G. (2020). On uniform exponential ergodicity of Markovian multiclass many-server queues in the Halfin-Whitt regime. *Math. Oper. Res.* (to appear).
- [6] ARAPOSTATHIS, A., HMEDI, H., PANG, G. and SANDRIĆ, N. (2019). Uniform polynomial rates of convergence for a class of Lévy-driven controlled SDEs arising in multiclass many-server queues. In

- Modeling, stochastic control, optimization, and applications. IMA Vol. Math. Appl.* **164** 1–20. Springer, Cham. [MR3970163](#)
- [7] 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](#)
  - [8] ARISAWA, M. (2009). Homogenization of a class of integro-differential equations with Lévy operators. *Comm. Partial Differential Equations* **34** 617–624. [MR2560294](#)
  - [9] BASS, R. F. and CRANSTON, M. (1986). The Malliavin calculus for pure jump processes and applications to local time. *Ann. Probab.* **14** 490–532. [MR832021](#)
  - [10] BLUMENTHAL, R. M. and GETTOOR, R. K. (1968). *Markov processes and potential theory. Pure and Applied Mathematics, Vol. 29.* Academic Press, New York-London. [MR0264757](#)
  - [11] 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](#)
  - [12] BUTKOVSKY, O. (2014). Subgeometric rates of convergence of Markov processes in the Wasserstein metric. *Ann. Appl. Probab.* **24** 526–552. [MR3178490](#)
  - [13] ÇINLAR, E. and JACOD, J. (1981). Representation of semimartingale Markov processes in terms of Wiener processes and Poisson random measures. In *Seminar on Stochastic Processes, 1981 (Evanston, Ill., 1981). Progr. Probab. Statist.* **1** 159–242. Birkhäuser, Boston, Mass. [MR647786](#)
  - [14] CHEN, M.-F. (2005). *Eigenvalues, inequalities, and ergodic theory. Probability and its Applications.* Springer-Verlag London, Ltd., London. [MR2105651](#)
  - [15] CHEN, X., CHEN, Z. Q. and WANG, J. (2020). Heat kernel for non-local operators with variable order. *Stochastic Process. Appl.* **130** 3574–3647. [MR4092415](#)
  - [16] CHEN, Z.-Q., HU, E., XIE, L. and ZHANG, X. (2017). Heat kernels for non-symmetric diffusion operators with jumps. *J. Differential Equations* **263** 6576–6634. [MR3693184](#)
  - [17] CHEN, Z.-Q. and ZHANG, X. (2016). Heat kernels and analyticity of non-symmetric jump diffusion semigroups. *Probab. Theory Related Fields* **165** 267–312. [MR3500272](#)
  - [18] CHEN, Z.-Q. and ZHANG, X. (2018). Heat kernels for time-dependent non-symmetric stable-like operators. *J. Math. Anal. Appl.* **465** 1–21. [MR3806688](#)
  - [19] CONNOR, S. B. and FORT, G. (2009). State-dependent Foster-Lyapunov criteria for subgeometric convergence of Markov chains. *Stochastic Process. Appl.* **119** 4176–4193. [MR2565563](#)
  - [20] DENG, C. S. (2020). Subgeometric Rates of Convergence for Discrete-Time Markov Chains Under Discrete-Time Subordination. *J. Theoret. Probab.* **33** 522–532. [MR4064311](#)
  - [21] 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](#)
  - [22] 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](#)
  - [23] DIEKER, A. B. and GAO, X. (2013). Positive recurrence of piecewise Ornstein-Uhlenbeck processes and common quadratic Lyapunov functions. *Ann. Appl. Probab.* **23** 1291–1317. [MR3098433](#)
  - [24] 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](#)
  - [25] 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](#)
  - [26] DOWN, D., MEYN, S. P. and TWEEDIE, R. L. (1995). Exponential and uniform ergodicity of Markov processes. *Ann. Probab.* **23** 1671–1691. [MR1379163](#)
  - [27] DUAN, G.-R. and PATTON, R. J. (1998). A note on Hurwitz stability of matrices. *Automatica J. IFAC* **34** 509–511. [MR1626135](#)
  - [28] DURMUS, A., FORT, G. and MOULINES, E. (2016). Subgeometric rates of convergence in Wasserstein distance for Markov chains. *Ann. Inst. Henri Poincaré Probab. Stat.* **52** 1799–1822. [MR3573296](#)
  - [29] EBERLE, A. (2011). Reflection coupling and Wasserstein contractivity without convexity. *C. R. Math. Acad. Sci. Paris* **349** 1101–1104. [MR2843007](#)
  - [30] EBERLE, A. (2016). Reflection couplings and contraction rates for diffusions. *Probab. Theory Related Fields* **166** 851–886. [MR3568041](#)
  - [31] ETHIER, S. N. and KURTZ, T. G. (1986). *Markov processes. Characterization and convergence. Wiley Series in Probability and Mathematical Statistics.* John Wiley & Sons, Inc., New York. [MR838085](#)

- [32] FORT, G. and MOULINES, E. (2003). Polynomial ergodicity of Markov transition kernels. *Stochastic Process. Appl.* **103** 57–99. [MR1947960](#)
- [33] FORT, G. and ROBERTS, G. O. (2005). Subgeometric ergodicity of strong Markov processes. *Ann. Appl. Probab.* **15** 1565–1589. [MR2134115](#)
- [34] FRIESEN, M., JIN, P., KREMER, J. and RÜDIGER, B. (2019). Exponential ergodicity for stochastic equations of nonnegative processes with jumps. *arXiv e-prints* **1902.02833**.
- [35] GOLOMOZIŇ, V. V. (2009). A subgeometric estimate for the stability of time-homogeneous Markov chains. *Teor. Ľmovĭr. Mat. Stat.* **81** 31–45. [MR2667308](#)
- [36] GRZYWNY, T. and SZCZYPKOWSKI, K. (2019). Heat kernels of non-symmetric Lévy-type operators. *J. Differential Equations* **267** 6004–6064. [MR3996792](#)
- [37] GYÖNGY, I. and KRYLOV, N. (1996). Existence of strong solutions for Itô’s stochastic equations via approximations. *Probab. Theory Related Fields* **105** 143–158. [MR1392450](#)
- [38] HAIRER, M. (2016). Convergence of Markov processes. *Lecture notes, University of Warwick*.
- [39] ISHIKAWA, Y. (2001). Density estimate in small time for jump processes with singular Lévy measures. *Tohoku Math. J. (2)* **53** 183–202. [MR1829978](#)
- [40] JACOD, J. and SHIRYAEV, A. N. (2003). *Limit theorems for stochastic processes*, second ed. *Grundlehren der Mathematischen Wissenschaften* **288**. Springer-Verlag, Berlin. [MR1943877](#)
- [41] KEVEI, P. (2018). Ergodic properties of generalized Ornstein-Uhlenbeck processes. *Stochastic Process. Appl.* **128** 156–181. [MR3729534](#)
- [42] KIM, P. and LEE, J. (2019). Heat kernels of non-symmetric jump processes with exponentially decaying jumping kernel. *Stochastic Process. Appl.* **129** 2130–2173. [MR3958426](#)
- [43] KIM, P., SONG, R. and VONDRAČEK, Z. (2018). Heat kernels of non-symmetric jump processes: beyond the stable case. *Potential Anal.* **49** 37–90. [MR3817130](#)
- [44] KNOPOVA, V. and KULIK, A. (2018). Parametrix construction of the transition probability density of the solution to an SDE driven by  $\alpha$ -stable noise. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** 100–140. [MR3765882](#)
- [45] 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](#)
- [46] 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](#)
- [47] KOLOKOLTSOV, V. (2000). Symmetric stable laws and stable-like jump-diffusions. *Proc. London Math. Soc. (3)* **80** 725–768. [MR1744782](#)
- [48] KOLOKOLTSOV, V. N. (2011). *Markov processes, semigroups and generators*. *De Gruyter Studies in Mathematics* **38**. Walter de Gruyter & Co., Berlin. [MR2780345](#)
- [49] KOMOROWSKI, T. and WALCZUK, A. (2012). Central limit theorem for Markov processes with spectral gap in the Wasserstein metric. *Stochastic Process. Appl.* **122** 2155–2184. [MR2921976](#)
- [50] KÜHN, F. (2018). On martingale problems and Feller processes. *Electron. J. Probab.* **23** Paper No. 13, 18. [MR3771750](#)
- [51] KÜHN, F. (2018). Random time changes of Feller processes. *ArXiv e-prints* **1705.02830**.
- [52] KÜHN, F. (2020). Existence of (Markovian) solutions to martingale problems associated with Lévy-type operators. *Electron. J. Probab.* **25** 1–26.
- [53] KULIK, A. (2018). *Ergodic behavior of Markov processes, with applications to limit theorems*. *De Gruyter Studies in Mathematics* **67**. De Gruyter, Berlin. [MR3791835](#)
- [54] KULIK, A. M. (2009). Exponential ergodicity of the solutions to SDE’s with a jump noise. *Stochastic Process. Appl.* **119** 602–632. [MR2494006](#)
- [55] KWON, Y. and LEE, C. (1999). Strong Feller property and irreducibility of diffusions with jumps. *Stochastics Stochastics Rep.* **67** 147–157. [MR1717795](#)
- [56] LAZIĆ, P. and SANDRIĆ, N. (2021). On sub-geometric ergodicity of diffusion processes. *Bernoulli* **27** 348–380. [MR4177373](#)
- [57] LIANG, M., MAJKA, M. B. and WANG, J. (2021). Exponential ergodicity for SDEs and McKean-Vlasov processes with Lévy noise. *Ann. Inst. Henri Poincaré Probab. Stat.* **57** 1665–1701. [MR4291453](#)
- [58] LIU, Y., ZHANG, H. and ZHAO, Y. (2010). Subgeometric ergodicity for continuous-time Markov chains. *J. Math. Anal. Appl.* **368** 178–189. [MR2609268](#)
- [59] LUO, D. and WANG, J. (2016). Exponential convergence in  $L^p$ -Wasserstein distance for diffusion pro-

- cesses without uniformly dissipative drift. *Math. Nachr.* **289** 1909–1926. [MR3563910](#)
- [60] MAJKA, M. B. (2017). Coupling and exponential ergodicity for stochastic differential equations driven by Lévy processes. *Stochastic Process. Appl.* **127** 4083–4125. [MR3718107](#)
- [61] MASUDA, H. (2004). On multidimensional Ornstein-Uhlenbeck processes driven by a general Lévy process. *Bernoulli* **10** 97–120. [MR2044595](#)
- [62] MASUDA, H. (2007). Ergodicity and exponential  $\beta$ -mixing bounds for multidimensional diffusions with jumps. *Stochastic Process. Appl.* **117** 35–56. [MR2287102](#)
- [63] MASUDA, H. (2009). Erratum to: “Ergodicity and exponential  $\beta$ -mixing bound for multidimensional diffusions with jumps” [*Stochastic Process. Appl.* 117 (2007) 35–56] [[MR2287102](#)]. *Stochastic Process. Appl.* **119** 676–678. [MR2494009](#)
- [64] MEYN, S. and TWEEDIE, R. L. (2009). *Markov chains and stochastic stability*, Second ed. Cambridge University Press, Cambridge. [MR2509253](#)
- [65] 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](#)
- [66] 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](#)
- [67] 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](#)
- [68] PICARD, J. (1996). On the existence of smooth densities for jump processes. *Probab. Theory Related Fields* **105** 481–511. [MR1402654](#)
- [69] PICARD, J. (2010). Erratum to: On the existence of smooth densities for jump processes [[MR1402654](#)]. *Probab. Theory Related Fields* **147** 711–713. [MR2639720](#)
- [70] REED, J. and ZWART, B. (2011). A piecewise linear stochastic differential equation driven by a Lévy process. *J. Appl. Probab.* **48A** 109–119. [MR2865620](#)
- [71] SANDRIĆ, N. (2016). Long-time behavior for a class of Feller processes. *Trans. Amer. Math. Soc.* **368** 1871–1910. [MR3449227](#)
- [72] SANDRIĆ, N., VALENTIĆ, I. and WANG, J. (2021). Periodic homogenization of a Lévy-type process with small jumps. *J. Evol. Equ.* **21** 771–803. [MR4238225](#)
- [73] SATO, K.-I. and YAMAZATO, M. (1984). Operator-self-decomposable distributions as limit distributions of processes of Ornstein-Uhlenbeck type. *Stochastic Process. Appl.* **17** 73–100. [MR738769](#)
- [74] SCHILLING, R. L., SONG, R. and VONDRAČEK, Z. (2012). *Bernstein functions. Theory and applications*, second ed. *De Gruyter Studies in Mathematics* **37**. Walter de Gruyter & Co., Berlin. [MR2978140](#)
- [75] SCHILLING, R. L. and WANG, J. (2012). On the coupling property and the Liouville theorem for Ornstein-Uhlenbeck processes. *J. Evol. Equ.* **12** 119–140. [MR2891204](#)
- [76] 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](#)
- [77] STROOCK, D. W. (1975). Diffusion processes associated with Lévy generators. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **32** 209–244. [MR433614](#)
- [78] TUOMINEN, P. and TWEEDIE, R. L. (1994). Subgeometric rates of convergence of  $f$ -ergodic Markov chains. *Adv. in Appl. Probab.* **26** 775–798. [MR1285459](#)
- [79] TWEEDIE, R. L. (1994). Topological conditions enabling use of Harris methods in discrete and continuous time. *Acta Appl. Math.* **34** 175–188. [MR1273853](#)
- [80] VERETENNIKOV, A. Y. (1997). On polynomial mixing bounds for stochastic differential equations. *Stochastic Process. Appl.* **70** 115–127. [MR1472961](#)
- [81] VERETENNIKOV, A. Y. (1999). On polynomial mixing and the rate of convergence for stochastic differential and difference equations. *Teor. Veroyatnost. i Primenen.* **44** 312–327. [MR1751475](#)
- [82] VILLANI, C. (2009). *Optimal transport. Grundlehren der Mathematischen Wissenschaften* **338**. Springer-Verlag, Berlin. [MR2459454](#)
- [83] WANG, F.-Y. (2011). Coupling for Ornstein-Uhlenbeck processes with jumps. *Bernoulli* **17** 1136–1158. [MR2854767](#)
- [84] WANG, J. (2012). On the exponential ergodicity of Lévy-driven Ornstein-Uhlenbeck processes. *J. Appl. Probab.* **49** 990–1004. [MR3058984](#)
- [85] WANG, J. (2016).  $LP$ -Wasserstein distance for stochastic differential equations driven by Lévy processes. *Bernoulli* **22** 1598–1616. [MR3474827](#)

- [86] XI, F. and ZHU, C. (2019). Jump type stochastic differential equations with non-Lipschitz coefficients: non-confluence, Feller and strong Feller properties, and exponential ergodicity. *J. Differential Equations* **266** 4668–4711. [MR3912730](#)
- [87] ZHANG, X. (2014). Densities for SDEs driven by degenerate  $\alpha$ -stable processes. *Ann. Probab.* **42** 1885–1910. [MR3262494](#)