ALEA, Lat. Am. J. Probab. Math. Stat. 18, 1–4 (2021)
DOI: 10.30757/ALEA.v18-

Exponential ergodicity for stochastic equations of nonnegative processes with jumps

Martin Friesen, Peng Jin, Jonas Kremer and Barbara Rüdiger

(Martin Friesen) School of Mathematical Sciences, Dublin City University, Dublin, Ireland
E-mail address: `martin.friesen@dcu.ie`

(Peng Jin) Guangdong Provincial Key Laboratory of IRADS, BNU-HKBU United International College, Zhuhai 519087, China; Department of Mathematics, Shantou University, Shantou, Guangdong 515063, China
E-mail address: `pengjin@uic.edu.cn`

(Jonas Kremer) School of Mathematical and Natural Sciences, University of Wuppertal, 42119 Wuppertal, Germany
E-mail address: `jonkremer@web.de`

(Barbara Rüdiger) School of Mathematical and Natural Sciences, University of Wuppertal, 42119 Wuppertal, Germany
E-mail address: `ruediger@uni-wuppertal.de`

Abstract. We study the long-time behavior for continuous-time Markov processes on the state space $\mathbb{R}_{\geq 0} := [0, \infty)$, which arise as unique strong solutions to stochastic equations with jumps. We establish, under a global dissipativity condition combined with a comparison principle, exponential ergodicity in various Wasserstein distances on $\mathbb{R}_{\geq 0}$. Our main emphasis lies on the derivation of these estimates under minimal moment conditions to be imposed on the associated Lévy measures of the noises. We apply our method to continuous-state branching processes with immigration (shorted as CBI processes), to nonlinear CBI processes, and finally to CBI processes in Lévy random environments.

1. Introduction

The study of long-time behavior for continuous-time Markov processes is a classical and still popular topic in probability theory. In this paper we investigate this problem for jump-diffusions on the state space $\mathbb{R}_{\geq 0} := [0, \infty)$, which include interesting classes of processes such as *continuous-state branching processes with immigration* (see, e.g., Li (2011); Pardoux (2016)), possibly in Lévy random environments (see He et al. (2018); Palau and Pardo (2017)), *continuous-state nonlinear branching processes* (see Li et al. (2019)), and *TCP processes* (see, e.g., Bardet et al. (2013); Chafaï et al.

2010 *Mathematics Subject Classification.* Primary 60J80; Secondary 60J25, 60G10, 60F17, 60H20.

Key words and phrases. Branching process, random environment, invariant distribution, ergodicity, coupling, Wasserstein distance.

The research of Peng Jin is supported by the Guangdong Basic and Applied Basic Research Foundation (No. 2020A1515010436), the Guangdong Provincial Key Laboratory of IRADS, BNU-HKBU United International College (2022B1212010006), the Guangdong Higher Education Upgrading Plan (2021-2025) (UIC R0400024-21), the UIC Start-up Research Fund (No. R72021102) and NSFC (Nos. 11861029, 12071499) .

(2010)). More precisely, we investigate the long-time behaviour for the Markov process determined by the stochastic equation

$$\begin{aligned} X_t = X_0 &+ \int_0^t b\left(X_s\right) \mathrm{d}s + \int_0^t \int_E \sigma\left(X_s, u\right) W\left(\mathrm{d}s, \mathrm{d}u\right) \\ &+ \int_0^t \int_{U_0} g_0\left(X_{s-}, u\right) \widetilde{N}_0\left(\mathrm{d}s, \mathrm{d}u\right) + \int_0^t \int_{U_1} g_1\left(X_{s-}, u\right) N_1\left(\mathrm{d}s, \mathrm{d}u\right), \quad t \geq 0, \end{aligned} \tag{1.1}$$

where $W(\mathrm{d}s, \mathrm{d}u)$ is a Gaussian white noise on $\mathbb{R}_{\geq 0} \times E$, $N_0(\mathrm{d}s, \mathrm{d}u)$ is a Poisson random measure on $\mathbb{R}_{\geq 0} \times U_0$, $N_1(\mathrm{d}s, \mathrm{d}u)$ is a Poisson random measure on $\mathbb{R}_{\geq 0} \times U_1$, and $\widetilde{N}_0$ denotes the compensated Poisson random measure associated with N_0. The precise conditions we impose on the coefficients b, σ, g_0, g_1 and on the noise terms are given in Section 3. A *strong solution* to (1.1) is, by definition, a nonnegative càdlàg process $\{X_t : t \geq 0\}$ that is adapted to the augmented natural filtration generated by $W(\mathrm{d}s, \mathrm{d}u)$, $N_0(\mathrm{d}s, \mathrm{d}u)$ and $N_1(\mathrm{d}s, \mathrm{d}u)$ and that satisfies the equation (1.1) almost surely for every $t \geq 0$. The existence and uniqueness of strong solutions were studied, e.g., in Fu and Li Fu and Li (2010), Dawson and Li Dawson and Li (2012), and Li and Pu Li and Pu (2012). Additional related results for stochastic equations on $\mathbb{R}_{\geq 0}$ can be found in Li and Mytnik Li and Mytnik (2011) as well as in Fournier Fournier (2013).

Let $\{P_t(x, \mathrm{d}y) : t, x \geq 0\}$ be the transition probabilities of the Markov process with state space $\mathbb{R}_{\geq 0}$ obtained from (1.1) and let $\mathcal{P}(\mathbb{R}_{\geq 0})$ be the space of all Borel probability measures over $\mathbb{R}_{\geq 0}$. We call $\pi \in \mathcal{P}(\mathbb{R}_{\geq 0})$ an *invariant measure* for $\{P_t(x, \mathrm{d}y) : t, x \geq 0\}$, if

$$\int_{\mathbb{R}_{\geq 0}} P_t(x, \mathrm{d}y)\pi\left(\mathrm{d}x\right) = \pi(\mathrm{d}y), \quad t \geq 0.$$

Existence of invariant measures can be shown by a compactness argument, see, e.g., Section 9 in Chapter 4 of Ethier and Kurtz (1986) for some sufficient condition and often relies on a Lyapunov drift condition. Unlike existence, uniqueness of the invariant measure is a more demanding mathematical problem and can be deduced, e.g., from the convergence of $P_t(x, \mathrm{d}y)$ to π. In order to study such convergence, let us define, for $\varrho, \widetilde{\varrho} \in \mathcal{P}(\mathbb{R}_{\geq 0})$, the Wasserstein distance

$$W_d(\varrho, \widetilde{\varrho}) = \inf \left\{ \int_{\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}} d(x, y) H(\mathrm{d}x, \mathrm{d}y) \ : \ H \text{ is a coupling of } (\varrho, \widetilde{\varrho}) \right\}, \tag{1.2}$$

where d is a suitably chosen metric on $\mathbb{R}_{\geq 0}$. Natural examples for d, among others, are $d(x, y) = \mathbb{1}_{\{x \neq y\}}$ corresponding to the total variation distance and $d(x, y) = |x - y|$ in accordance with the Kantorovich-Rubinstein distance. In Section 2 we collect some basic properties of W_d while a detailed treatment of Wasserstein distances is provided in the monograph of Villani Villani (2009).

We call a Markov process *exponentially ergodic in* W_d, if there exists a constant $A > 0$ and a function $K : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ satisfying

$$W_d(P_t(x, \cdot), \pi) \leq K(x)\mathrm{e}^{-At}, \quad t, x \geq 0.$$

A widely used approach for the study of exponential ergodicity in the total variation distance (i.e. $d(x, y) = \mathbb{1}_{\{x \neq y\}}$) is due to Meyn and Tweedie Meyn and Tweedie (2009, 1993). The essential obstacle when applying their approach lies within the "irreducibility of a skeleton chain". To prove this, it is sufficient and customary to verify that $P_t(x, \mathrm{d}y)$ has a jointly continuous density which is strictly positive. While such an approach is suitable for diffusion processes, the situation is more delicate and requires a custom-tailored analysis when dealing with Markov processes with jumps. Albeit being challenging, the approach of Meyn and Tweedie has already been successfully applied to diverse Markov processes with jumps. Another approach to prove ergodicity of Markov processes is based on the construction of successful couplings. Such construction is usually closely related with the mathematical model at hand and often a difficult task, see, e.g., Butkovsky (2014); Eberle (2016); Peng and Zhang (2018); Wang (2016).

In this work we provide a simple approach to exponential ergodicity in the Wasserstein distance W_d with $d(x,y) = |x-y|^\lambda$, $\lambda \in (0,1]$ or $d(x,y) = \log(1+|x-y|)$ for Markov processes on $\mathbb{R}_{\geq 0}$ which can be constructed as strong solutions of stochastic equations satisfying the comparison principle. Indeed, using the comparison principle we construct a monotone coupling for (X_t^x, X_t^y) with $X_0^x = x$ and $X_0^y = y$ obtained from (1.1). This coupling allows us to estimate the Wasserstein distance and subsequently prove the existence, uniqueness of invariant measures, and exponential ergodicity in the Wasserstein distance. The corresponding results are formulated in Section 3 and extend the method developed in Friesen et al. (2020b), where affine processes on the canonical state space have been studied, to a general class of Markov processes on $\mathbb{R}_{\geq 0}$. For simplicity, we consider only the one-dimensional case. However, using the arguments based on the comparison principle similar to Friesen et al. (2020a), one may extend most of the results of this paper also to multi-dimensional settings, i.e., to Markov processes on state space $\mathbb{R}_+^m = \{x \in \mathbb{R}^m \ : \ x_k \geq 0, \ \forall k = 1,\dots,m\}$.

To illustrate the usage of our results obtained in Section 3, we apply them in Section 4 to continuous-state branching processes and nonlinear continuous-state branching processes, as well as in Section 5 to continuous-state branching processes in Lévy environments. All these processes have in common that, the coefficients in front of the noise terms (continuous and pure-jump noise) are typically degenerate at the boundary 0 and hence the methods from Douc et al. (2009); Eberle et al. (2019); Kulik (2009); Veretennikov (1987) do not directly apply here.

2. Some basic properties of Wasserstein distances

By $\mathcal{P}(\mathbb{R}_{\geq 0})$ we denote the space of all Borel probability measures over $\mathbb{R}_{\geq 0}$. Given ϱ, $\widetilde{\varrho} \in \mathcal{P}(\mathbb{R}_{\geq 0})$, a coupling H of $(\varrho, \widetilde{\varrho})$ is a Borel probability measure on $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$ which has marginals ϱ and $\widetilde{\varrho}$, respectively. We write $\mathcal{H}(\varrho, \widetilde{\varrho})$ for the collection of all such couplings. Let d be a metric on $\mathbb{R}_{\geq 0}$ such that $(\mathbb{R}_{\geq 0}, d)$ is a complete seperable metric space and define

$$\mathcal{P}_d\left(\mathbb{R}_{\geq 0}\right) = \left\{ \varrho \in \mathcal{P}\left(\mathbb{R}_{\geq 0}\right) \ : \ \int_{\mathbb{R}_{\geq 0}} d(x,0)\varrho\left(\mathrm{d}x\right) < \infty \right\}.$$

The *Wasserstein distance* on $\mathcal{P}_d(\mathbb{R}_{\geq 0})$ is defined by (1.2) Note that, since ϱ and $\widetilde{\varrho}$ belong to $\mathcal{P}_d(\mathbb{R}_{\geq 0})$, the expression $W_d(\varrho, \widetilde{\varrho})$ is finite. Moreover, it can be shown that this infimum is attained (see Villani (2009, p.95)), i.e., there exists $H \in \mathcal{H}(\varrho, \widetilde{\varrho})$ such that

$$W_d\left(\varrho, \widetilde{\varrho}\right) = \int_{\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}} d(x,y) H\left(\mathrm{d}x, \mathrm{d}y\right). \tag{2.1}$$

Since $(\mathbb{R}_{\geq 0}, d)$ is assumed to be a complete separable metric space, according to Villani (2009, Theorem 6.16), $(\mathcal{P}_d(\mathbb{R}_{\geq 0}), W_d)$ is also a complete separable metric space. In the remainder of the article, we will use the following particular examples.

Example 2.1.

(a) If $d_{TV}(x,y) = \mathbb{1}_{\{x \neq y\}}$, then $\mathcal{P}_{d_{TV}}(\mathbb{R}_{\geq 0}) = \mathcal{P}(\mathbb{R}_{\geq 0})$ and

$$W_{d_{TV}}(\varrho, \widetilde{\varrho}) = \frac{1}{2}\|\varrho - \widetilde{\varrho}\|_{TV} := \frac{1}{2}\sup\left\{|\varrho(A) - \widetilde{\varrho}(A)| \ : \ A \subset \mathbb{R} \text{ Borel set}\right\}$$

is the total variation distance. Note that in this case $(\mathbb{R}_{\geq}, d_{TV})$ is not seperable. However, $W_{d_{TV}}$ still defines a metric on $\mathcal{P}_{d_{TV}}(\mathbb{R}_{\geq})$ with respect to which this space is complete.

(b) For $\lambda \in (0,1]$, the Wasserstein-λ-distance corresponds to $d_\lambda(x,y) = |x-y|^\lambda$, where

$$\mathcal{P}_{d_\lambda}(\mathbb{R}_{\geq 0}) := \left\{ \varrho \in \mathcal{P}(\mathbb{R}_{\geq 0}) \ : \ \int_{\mathbb{R}_{\geq 0}} x^\lambda \varrho(\mathrm{d}x) < \infty \right\}.$$

For simplicity, we write $(\mathcal{P}_\lambda(\mathbb{R}_{\geq 0}), W_\lambda)$ instead of $(\mathcal{P}_{d_\lambda}(\mathbb{R}_{\geq 0}), W_{d_\lambda})$.

(c) The limiting case $\lambda = 0$ should naturally correspond to the case where $d_{\log}(x,y) = \log(1 + |x-y|)$. In such a case we have

$$\mathcal{P}_{d_{\log}}(\mathbb{R}_{\geq 0}) := \left\{ \varrho \in \mathcal{P}(\mathbb{R}_{\geq 0}) \, : \, \int_{\{x>1\}} \log(x)\varrho(\mathrm{d}x) < \infty \right\}.$$

Similarly as in (b), we write $(\mathcal{P}_{\log}(\mathbb{R}_{\geq 0}), W_{\log})$ instead of $(\mathcal{P}_{d_{\log}}(\mathbb{R}_{\geq 0}), W_{d_{\log}})$.

(d) Finally, we will use the following auxilliary metric defined by $d_1^{\leq 1}(x,y) = 1 \wedge |x-y|$. The corresponding Wasserstein distance is then well-defined on $\mathcal{P}_{d_1^{\leq 1}}(\mathbb{R}_{\geq 0}) = \mathcal{P}(\mathbb{R}_{\geq 0})$.

We will need the following simple relation between the distances W_λ and $W_{\log}$.

Lemma 2.2. *Let $\lambda \in (0,1]$. For all $\varrho, \widetilde{\varrho} \in \mathcal{P}(\mathbb{R}_{\geq 0})$ it holds that*

$$W_{\log}(\varrho, \widetilde{\varrho}) \leq \log(1 + W_1(\varrho, \widetilde{\varrho})) \quad \text{and} \quad W_\lambda(\varrho, \widetilde{\varrho}) \leq (W_1(\varrho, \widetilde{\varrho}))^\lambda.$$

Proof: Let H be the optimal coupling for $W_1(\varrho, \widetilde{\varrho})$. Then using that $[0,\infty) \ni x \longmapsto \log(1+x)$ is concave, we get

$$\begin{aligned} W_{\log}(\varrho, \widetilde{\varrho}) &\leq \int_{\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}} \log(1 + |x-y|) H(\mathrm{d}x, \mathrm{d}y) \\ &\leq \log\left(1 + \int_{\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}} |x-y| H(\mathrm{d}x, \mathrm{d}y)\right) \\ &= \log(1 + W_1(\varrho, \widetilde{\varrho})). \end{aligned}$$

This proves the first inequality. The second inequality can be shown in the same way, since $[0,\infty) \ni x \longmapsto x^\lambda$ is also concave. □

One simple but important property of Wasserstein distances is their convexity as formulated below.

Lemma 2.3. *Let d be a metric such that $(\mathbb{R}_{\geq 0}, d)$ is a complete separable metric space. Let ϱ, $\widetilde{\varrho} \in \mathcal{P}_d(\mathbb{R}_{\geq 0})$ and suppose that $P_t(x, \mathrm{d}y)$ is a Markov kernel on $\mathbb{R}_{\geq 0}$. Then, for any $H \in \mathcal{H}(\varrho, \widetilde{\varrho})$, we have*

$$W_d\left(\int_{\mathbb{R}_{\geq 0}} P(x, \cdot)\varrho(\mathrm{d}x), \int_{\mathbb{R}_{\geq 0}} P(x, \cdot)\widetilde{\varrho}(\mathrm{d}x) \right) \leq \int_{\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}} W_d(P(x,\cdot), P(y,\cdot)) H(\mathrm{d}x, \mathrm{d}y).$$

Note that this assertion also holds for the total variation distance. For a proof we refer the reader to Villani (2009, Theorem 4.8). The convolution between measures ϱ and $\widetilde{\varrho}$ on $\mathbb{R}_{\geq 0}$ is denoted by $\varrho * \widetilde{\varrho}$. We close the presentation with a useful convolution estimate for Wasserstein distances.

Lemma 2.4. *Let d be a metric such that $(\mathbb{R}_{\geq 0}, d)$ is a complete separable metric space. Suppose that $d(x+y, \widetilde{x}+y) \leq d(x, \widetilde{x})$ for $x, \widetilde{x}, y \geq 0$. Let ϱ, $\widetilde{\varrho}$, $g \in \mathcal{P}_d(\mathbb{R}_{\geq 0})$. Then $W_d(\varrho * g, \widetilde{\varrho} * g) \leq W_d(\varrho, \widetilde{\varrho})$.*

Proof: For a function h on $\mathbb{R}_{\geq 0}$, we define $\|h\|_{\mathrm{Lip}} = \sup_{x \neq y} \frac{|h(x)-h(y)|}{d(x,y)}$ and a new function h_g as $h_g(x) = \int_{\mathbb{R}_{\geq 0}} h(x+y) g(\mathrm{d}y)$, $x \geq 0$. Note that if $\|h\|_{\mathrm{Lip}} \leq 1$, then

$$\begin{aligned} \left| h_g(x) - h_g(x') \right| &\leq \int_{\mathbb{R}_{\geq 0}} \left| h(x+y) - h(x'+y) \right| g(\mathrm{d}y) \\ &\leq \|h\|_{\mathrm{Lip}} \int_{\mathbb{R}_{\geq 0}} d(x+y, x'+y) g(\mathrm{d}y) \\ &\leq \|h\|_{\mathrm{Lip}} \int_{\mathbb{R}_{\geq 0}} d(x, x') g(\mathrm{d}y) \leq d(x, x'), \end{aligned}$$

showing that $\|h_g\|_{\mathrm{Lip}} \leq 1$. Combining this observation with the Kantorovich-Duality, we obtain

$$\begin{aligned} W_d(\varrho * g, \widetilde{\varrho} * g) &= \sup_{\|h\|_{\mathrm{Lip}}\leq 1} \left| \int_{\mathbb{R}_{\geq 0}} h(x)(\varrho * g)(\mathrm{d}x) - \int_{\mathbb{R}_{\geq 0}} h(x)(\widetilde{\varrho} * g)(\mathrm{d}x) \right| \\ &= \sup_{\|h\|_{\mathrm{Lip}}\leq 1} \left| \int_{\mathbb{R}_{\geq 0}} h_g(x)\varrho(\mathrm{d}x) - \int_{\mathbb{R}_{\geq 0}} h_g(x)\widetilde{\varrho}(\mathrm{d}x) \right| \\ &\leq \sup_{\|h\|_{\mathrm{Lip}}\leq 1} \left| \int_{\mathbb{R}_{\geq 0}} h(x)\varrho(\mathrm{d}x) - \int_{\mathbb{R}_{\geq 0}} h(x)\widetilde{\varrho}(\mathrm{d}x) \right| = W_d(\varrho, \widetilde{\varrho}). \end{aligned}$$

□

Although we formulated Lemma 2.3 and 2.4 on the state space $\mathbb{R}_{\geq 0}$, it is clear that they naturally extend to more abstract state spaces.

3. Stochastic equations of nonnegative processes with jumps

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space with the usual hypotheses, let E, U_0, and U_1 be complete separable metric spaces, $\varkappa(\mathrm{d}u)$, $\mu_0(\mathrm{d}u)$ and $\mu_1(\mathrm{d}u)$ be σ-finite measures on E, U_0 and U_1, respectively. Then let $W(\mathrm{d}t, \mathrm{d}u)$ be an $(\mathcal{F}_t)_{t\geq 0}$-Gaussian white noise on $\mathbb{R}_{\geq 0} \times E$ with intensity measure $\mathrm{d}t\varkappa(\mathrm{d}z)$, and let $N_0(\mathrm{d}t, \mathrm{d}u)$ and $N_1(\mathrm{d}t, \mathrm{d}u)$ be $(\mathcal{F}_t)_{t\geq 0}$-Poisson random measures on $\mathbb{R}_{\geq 0} \times U_0$ and $\mathbb{R}_{\geq 0} \times U_1$ with intensities $\mathrm{d}t\mu_0(\mathrm{d}u)$ and $\mathrm{d}t\mu_1(\mathrm{d}u)$, respectively. Denote by $\widetilde{N}_0(\mathrm{d}t, \mathrm{d}u) := N_0(\mathrm{d}t, \mathrm{d}u) - \mathrm{d}t\mu_0(\mathrm{d}u)$ the compensated Poisson random measure of $N_0(\mathrm{d}t, \mathrm{d}u)$. Suppose that the random objects $W(\mathrm{d}t, \mathrm{d}u)$, $N_0(\mathrm{d}t, \mathrm{d}u)$, and $N_1(\mathrm{d}t, \mathrm{d}u)$ are mutually independent. We call parameters (b, σ, g_0, g_1) *admissible*, if they satisfy the following set of conditions:

- $x \mapsto b(x)$ is continuous on $\mathbb{R}_{\geq 0}$ and satisfies $b(0) \geq 0$;
- $(x, u) \mapsto \sigma(x, u)$ is a Borel function on $\mathbb{R}_{\geq 0} \times E$ satisfying $\sigma(0, u) = 0$ for $u \in E$;
- $(x, u) \mapsto g_0(x, u)$ is a Borel function on $\mathbb{R}_{\geq 0} \times U_0$ satisfying $g_0(0, u) = 0$ and $g_0(x, u) + x \geq 0$ for $x > 0$ and $u \in U_0$.
- $(x, u) \mapsto g_1(x, u)$ is a Borel function on $\mathbb{R}_{\geq 0} \times U_1$ satisfying $g_1(x, u) + x \geq 0$ for $x \in \mathbb{R}_{\geq 0}$ and $u \in U_1$.

While the above conditions are natural to ensure solutions of (1.1) to be non-negative, existence and uniqueness of solutions do require a set of additional technical conditions introduced below.

(3.a) There is a constant $K_0 \geq 0$ and a Borel function $(x, u) \longmapsto \overline{g}_1(x, u)$ on $\mathbb{R}_{\geq 0} \times U_1$ such that $\sup_{y\in[0,x]} |g_1(y, u)| \leq \overline{g}_1(x, u)$ and for $x \geq 0$

$$b(x) + \int_{U_1} 1 \wedge |\overline{g}_1(x, u)| \,\mu_1 (\mathrm{d}u) \leq K_0(2 + x);$$

(3.b) There exist continuous functions b_1, b_2 such that $b(x) = b_1(x) - b_2(x)$ and $x \longmapsto b_2(x)$ is nondecreasing. Finally, for each $j \geq 1$ there is a nondecreasing, continuous, concave function $z \mapsto r_j(z)$ on $\mathbb{R}_{\geq 0}$ such that $\int_{0+} r_j(z)^{-1}\mathrm{d}z = \infty$ and for all $0 \leq x, y \leq j$

$$|b_1(x) - b_1(y)| + \int_{U_1} |g_1(x, u) - g_1(y, u)| \,\mu_1 (\mathrm{d}u) \leq r_j (|x - y|)\,;$$

(3.c) For each $j \geq 1$ there is a constant $A_j > 0$ such that for $0 \leq x, y \leq j$

$$\int_E |\sigma(x, u) - \sigma(y, u)|^2 \varkappa(\mathrm{d}u) \leq A_j|x - y|,$$

the function $x \longmapsto \int_E \sigma(x,u)\varkappa(\mathrm{d}u)$ is continuous, and there exists a constant $K_1 > 0$ such that for $x \geq 0$

$$\int_E \sigma(x,u)^2 \varkappa(\mathrm{d}u) \leq K_1(2+x)^2, \qquad x \geq 0;$$

(3.d) There is a non-decreasing function $x \longmapsto L(x)$ on $\mathbb{R}_{\geq 0}$ and a Borel function $(x,u) \longmapsto \overline{g}_0(x,u)$ on $\mathbb{R}_{\geq 0} \times U_0$ so that $\sup_{y\in[0,x]} |g_0(y,u)| \leq \overline{g}_0(x,u)$ and

$$\int_{U_0} \left(\overline{g}_0(x,u) \wedge \overline{g}_0(x,u)^2\right) \mu_0\,(\mathrm{d}u) \leq L(x)$$

for every $x \geq 0$, and there exists a constant $K_2 \geq 0$ such that

$$\int_{U_0} \mathbb{1}_{\{|g_0(x,u)|\leq 1\}} |g_0(x,u)|^2 \mu_0(\mathrm{d}u) \leq K_2(2+x)^2$$

holds for each $x \geq 0$. Moreover, for each $j \geq 1$ there exists a constant $B_j \geq 0$ such that

$$\int_{U_0} \left(|g_0(x,u) - g_0(y,u)| \wedge |g_0(x,u) - g_0(y,u)|^2\right) \mu_0\,(\mathrm{d}u) \leq B_j |x-y|$$

holds for all $0 \leq x,\, y \leq j$.

(3.e) The functions $x \longmapsto x + g_0(x,u)$ and $x \longmapsto x + g_1(x,u)$ are nondecreasing for each $u \in U_0$ or $u \in U_1$, respectively.

Definition 3.1. Equation (1.1) is said to have the comparison property, if for any non-negative $\mathcal{F}_0$-measurable random variables X_0, Y_0 and corresponding solutions $\{X_t : t \geq 0\}$ and $\{Y_t : t \geq 0\}$ to (1.1), it holds that: if $\mathbb{P}(X_0 \leq Y_0) = 1$, then $\mathbb{P}(X_t \leq Y_t \text{ for all } t \geq 0) = 1$.

Before we formulate our main statement, we introduce the following conditions on the big jumps of the process.

$(m)_\lambda$: There exists $\lambda > 0$ and $K_\lambda > 0$ such that for all $x \geq 0$,

$$\int_{U_1} \mathbb{1}_{\{|g_1(x,u)|>1\}} |g_1(x,u)|^\lambda \mu_1(\mathrm{d}u) \leq K_\lambda (2+x)^{\min\{1,\lambda\}}, \tag{3.1}$$

$$\int_{U_0} \mathbb{1}_{\{|g_0(x,u)|>1\}} |g_0(x,u)|^{\max\{1,\lambda\}} \mu_0(\mathrm{d}u) \leq K_\lambda (2+x)^{\rho(\lambda)},$$

where $\rho(\lambda) = \lambda \mathbb{1}_{(0,1]}(\lambda) + \mathbb{1}_{(1,2)}(\lambda) + 2\mathbb{1}_{[2,\infty)}$.

$(m)_{\log}$: There exists a constant $K_{\log} > 0$ such that for all $x \geq 0$,

$$\int_{U_1} \mathbb{1}_{\{|g_1(x,u)|>1\}} \log(1 + |g_1(x,u)|) \mu_1(\mathrm{d}u) \leq K_{\log}(2+x), \tag{3.2}$$

$$\int_{U_0} \mathbb{1}_{\{|g_0(x,u)|>1\}} |g_0(x,u)| \mu_0(\mathrm{d}u) \leq K_{\log}(2+x).$$

Remark 3.2. If condition $(m)_\lambda$ holds for $\lambda \in (0,2]$, then also condition $(m)_{\log}$ holds true.

The next theorem provides the existence, uniqueness and comparison property for solutions of (1.1).

Theorem 3.3. *Let (b, σ, g_0, g_1) be admissible parameters, suppose that conditions (3.a) – (3.e) are satisfied. Suppose that one of the following two conditions holds:*
(i) condition $(m)_\lambda$ holds for some $\lambda > 0$ and $\mathbb{E}\left[X_0^\lambda\right] < \infty$;
(ii) condition $(m)_{\log}$ is satisfied and $\mathbb{E}\left[\log(2 + X_0)\right] < \infty$.
Then there exists a unique strong solution $\{X_t : t \geq 0\}$ of (1.1). This solution satisfies, under condition (i) the estimate

$$\mathbb{E}\left[X_t^\lambda\right] \leq \mathbb{E}\left[(2+X_0)^\lambda\right] \exp\left(c(\lambda)t\right), \qquad t \geq 0,$$

where $c(\lambda) > 0$ is some constant, while under condition (ii) it satisfies

$$\mathbb{E}\left[\log(2 + X_t)\right] \leq \mathbb{E}\left[\log(2 + X_0)\right] + c_{\log} t, \qquad t \geq 0,$$

where $c_{\log} > 0$ is some constant. Moreover, (1.1) has the comparison property.

Proof: The existence and uniqueness of a $[0, \infty]$-valued solution of (1.1) follows from Palau and Pardo (2018, Proposition 1). Applying then Lemma A.3 from the appendix proves that this unique solution is actually a.s. finite and satisfies the desired moment bounds. Finally, the comparison property can be shown by following the proof of Dawson and Li (2012, Theorem 2.3) with the only difference that our conditions (3.a), (3.d) are weaker than those imposed in Dawson and Li (2012, Theorem 2.3). However, this does not affect the key argument. This completes the proof of this statement. □

Let us remark that the constants appearing in the moment estimates can be computed explicitly. Below we provide the semimartingale characteristics for the process.

Lemma 3.4. *Let (b, σ, g_0, g_1) be admissible parameters and suppose that conditions (3.a) – (3.e) are satisfied, and that condition $(m)_\lambda$ holds for some $\lambda > 0$ or condition $(m)_{\log}$ is satisfied. Let $\{X_t \ : \ t \geq 0\}$ be the unique strong solution to (1.1), and let h be a continuous bounded function on $\mathbb{R}$ such that $h(x) = x$ holds for $|x| \leq 1$. Then X is the unique semimartingale with characteristics (B, C, K) with respect to h given by*

$$\begin{aligned}
B_t &= \int_0^t \widetilde{b}(X_s)\mathrm{d}s, \qquad \widetilde{b}(x) = b(x) - \int_{U_0} \widetilde{h}(g_0(x,u))\mu_0(\mathrm{d}u) + \int_{U_1} h(g_1(x,u))\mu_1(\mathrm{d}u), \qquad (3.3)\\
C_t &= \int_0^t c(X_s)\mathrm{d}s, \qquad c(x) = \int_E \sigma(x,u)^2 \varkappa(\mathrm{d}u),\\
K(x, A) &= \int_{U_0} \mathbb{1}_{A\setminus\{0\}}(g_0(x,u))\mu_0(\mathrm{d}u) + \int_{U_1} \mathbb{1}_{A\setminus\{0\}}(g_1(x,u))\mu_1(\mathrm{d}u),
\end{aligned}$$

where $\widetilde{h}(x) = x - h(x)$.

Proof: If $f \in C_b^2(\mathbb{R})$, then an application of the Itô formula shows that $f(X_t) = f(X_0) + \mathcal{A}_t + \mathcal{M}_t$, where $(\mathcal{A}_t)_{t\geq 0}$ is of locally finite variation and $(\mathcal{M}_t)_{t\geq 0}$ is a local martingale, respectively given by

$$\begin{aligned}
\mathcal{M}_t &= \int_0^t \int_E \sigma(X_s,u) f'(X_s) W(\mathrm{d}s, \mathrm{d}u)\\
&\quad + \int_0^t \int_{U_0} \left(f(X_{s-} + g_0\left(X_{s-}, u\right)) - f(X_{s-})\right) \widetilde{N}_0\left(\mathrm{d}s, \mathrm{d}u\right)\\
&\quad + \int_0^t \int_{U_1} \left(f(X_{s-} + g_1\left(X_{s-}, u\right)) - f(X_{s-})\right) \widetilde{N}_1\left(\mathrm{d}s, \mathrm{d}u\right),\\
\mathcal{A}_t &= \int_0^t b(X_s) f'(X_s)\mathrm{d}s + \frac{1}{2}\int_0^t \int_E \sigma(X_s,u)^2 f''(X_s)\varkappa(\mathrm{d}u)\mathrm{d}s\\
&\quad + \int_0^t \int_{U_0} \big(f(X_s + g_0(X_s,u)) - f(X_s) - g_0(X_s,u) f'(X_s)\big)\, \mu_0(\mathrm{d}u)\mathrm{d}s\\
&\quad + \int_0^t \int_{U_1} \left(f(X_s + g_1(X_s,u)) - f(X_s)\right) \mu_1(\mathrm{d}u)\mathrm{d}s.
\end{aligned}$$

A short computation shows that

$$\begin{aligned}
\mathcal{A}_t &= \int_0^t \widetilde{b}(X_s)f'(X_s)\mathrm{d}s + \frac{1}{2}\int_0^t\int_E \sigma(X_s,u)^2 f''(X_s)\varkappa(\mathrm{d}u)\mathrm{d}s \\
&\quad + \int_0^t\int_{U_0}\big(f(X_s+g_0(X_s,u)) - f(X_s) - h(g_0(X_s,u))f'(X_s)\big)\,\mu_0(\mathrm{d}u)\mathrm{d}s \\
&\quad + \int_0^t\int_{U_1}\big(f(X_s+g_1(X_s,u)) - f(X_s) - h(g_1(X_s,u))f'(X_s)\big)\,\mu_1(\mathrm{d}u)\mathrm{d}s \\
&= \int_0^t \widetilde{b}(X_s)f'(X_s)\mathrm{d}s + \frac{1}{2}\int_0^t\int_E \sigma(X_s,u)^2 f''(X_s)\varkappa(\mathrm{d}u)\mathrm{d}s \\
&\quad + \int_0^t\int_{U_0\times U_1\times\{0,1\}}\Big(f(X_s+G(X_s,u_0,u_1,l)) - f(X_s) \\
&\qquad\qquad - h(G(X_s,u_0,u_1,l))f'(X_s)\Big)F(\mathrm{d}u_0,\mathrm{d}u_1,\mathrm{d}l)\mathrm{d}s \\
&= \int_0^t \widetilde{b}(X_s)f'(X_s)\mathrm{d}s + \frac{1}{2}\int_0^t\int_E \sigma(X_s,u)^2 f''(X_s)\varkappa(\mathrm{d}u)\mathrm{d}s \\
&\quad + \int_0^t\int_{U_0\times U_1\times\{0,1\}}\Big(f(X_s+z) - f(X_s) - h(z)f'(X_s)\Big)K(X_s,\mathrm{d}z)\mathrm{d}s,
\end{aligned}$$

where

$$\begin{aligned}
G(x,u_0,u_1,l) &= \mathbb{1}_{\{l=0\}}g_0(x,u) + \mathbb{1}_{\{l=1\}}g_1(x,u), \\
F(\mathrm{d}u_0,\mathrm{d}u_1,\mathrm{d}l) &= \mu_0(\mathrm{d}u_0)\delta_0(\mathrm{d}u_1)\delta_0(\mathrm{d}l) + \delta_0(\mathrm{d}u_0)\mu_1(\mathrm{d}u_1)\delta_1(\mathrm{d}l),
\end{aligned}$$

and

$$\begin{aligned}
K(x,A) &= \int_{U_0}\mathbb{1}_{A\setminus\{0\}}(g_0(x,u))\mu_0(\mathrm{d}u) + \int_{U_1}\mathbb{1}_{A\setminus\{0\}}(g_1(x,u))\mu_1(\mathrm{d}u) \\
&= \int_{U_0\times U_1\times\{0,1\}}\mathbb{1}_{A\setminus\{0\}}(G(x,u_0,u_1,l))F(\mathrm{d}u_0,\mathrm{d}u_1,\mathrm{d}l).
\end{aligned}$$

The particular form of the characteristics is now a consequence of Kurtz (2011, Chapter II, Theorem 2.42), while the uniqueness of semimartingale characteristics is a consequence of Kurtz (2011, Chapter III, Theorem 2.26). This proves the assertion. □

In the following, we write $\{X_t^x : t \geq 0\}$ for the unique strong solution of (1.1) to indicate that the process X_t starts with deterministic initial variable $X_0 = x \geq 0$. Arguing as in Li (2019) (see the proof of Theorem 1.1 therein) one can show that $\{X_t^x : t \geq 0\}$ is a strong $(\mathcal{F}_t)_{t\geq 0}$-Markov process with transition kernel $P_t(x,\mathrm{d}y)$. The adjoint action of the transition semigroup on $\mathcal{P}(\mathbb{R}_{\geq 0})$ is defined by

$$P_t^*\varrho\,(\mathrm{d}y) = \int_{\mathbb{R}_{\geq 0}} P_t\,(x,\mathrm{d}y)\,\varrho\,(\mathrm{d}x)\,, \quad t\geq 0,\ \varrho\in\mathcal{P}\,(\mathbb{R}_{\geq 0})\,.$$

By the Markov property we have that $P_{t+s}^* = P_t^*P_s^*$ for all $0\leq s\leq t$. Finally, in order to study the long-time behaviour we impose the following uniform dissipativity condition of the drift:

(3.f) It holds that $\int_{U_1}|g_1(x,u)|\mu_1(\mathrm{d}u) < \infty$ for all $x\geq 0$ and there exists a constant $A>0$ such that

$$b(y) - b(x) + \int_{U_1}(g_1(y,u) - g_1(x,u))\mu_1(\mathrm{d}u) \leq -A(y-x)$$

holds for all $x,y\geq 0$ with $x\leq y$.

Under the given conditions we are able to show that the corresponding Markov process is exponentially ergodic in the Wasserstein distance W_1.

Theorem 3.5. *Let (b,σ,g_0,g_1) be admissible parameters and suppose that conditions (3.a)–(3.f) hold, and that $(m)_\lambda$ holds for $\lambda=1$. Then, for all ϱ, $\widetilde{\varrho}\in\mathcal{P}_1(\mathbb{R}_{\geq 0})$, we have*

$$W_1\left(P_t^*\varrho, P_t^*\widetilde{\varrho}\right) \leq \mathrm{e}^{-At} W_1\left(\varrho,\widetilde{\varrho}\right), \quad t\geq 0. \tag{3.4}$$

In particular, there exists a unique invariant measure $\pi\in\mathcal{P}(\mathbb{R}_{\geq 0})$. Moreover, we have $\pi\in\mathcal{P}_1(\mathbb{R}_{\geq 0})$ and, for all $\varrho\in\mathcal{P}_1(\mathbb{R}_{\geq 0})$,

$$W_1\left(P_t^*\varrho,\pi\right)\leq \mathrm{e}^{-At}W_1\left(\varrho,\pi\right), \quad t\geq 0.$$

Proof: First observe that $\{X_t^x : t\geq 0\}$ has finite first moment due to Lemma A.3.(a). Define for $n\in\mathbb{N}$ the stopping time $\tau_n(x)=\inf\{t\geq 0 \;:\; X_t^x\geq n\}$. In the proof of Lemma A.3.(a), we have shown that $\tau_n(x)\to\infty$ a.s. as $n\to\infty$ and that $(M_t^{n,x})_{t\geq 0}$ is a martingale for each $n\in\mathbb{N}$, where

$$M_t^{n,x}=\int_0^{t\wedge\tau_n(x)}\int_E \sigma\left(X_s^x,u\right)W\left(\mathrm{d}s,\mathrm{d}u\right)+\int_0^{t\wedge\tau_n(x)}\int_{U_0} g_0\left(X_{s-}^x,u\right)\widetilde{N}_0\left(\mathrm{d}s,\mathrm{d}u\right).$$

Let $0\leq x\leq y$ and define $\tau_n=\tau_n(x)\wedge\tau_n(y)$. Using the comparison principle, optional stopping, and then condition (3.f) we find that

$$\begin{aligned}
\mathbb{E}\left[|X_{t\wedge\tau_n}^x - X_{t\wedge\tau_n}^y|\right] &= \mathbb{E}[X_{t\wedge\tau_n}^y]-\mathbb{E}[X_{t\wedge\tau_n}^x]\\
&= y-x+\mathbb{E}\left[\int_0^{t\wedge\tau_n}(\widetilde{b}(X_s^y)-\widetilde{b}(X_s^x))\mathrm{d}s\right]\\
&\leq |x-y| - A\mathbb{E}\left[\int_0^{t\wedge\tau_n}(X_s^y-X_s^x)\mathrm{d}s\right]\\
&= |x-y| - A\mathbb{E}\left[\int_0^{t\wedge\tau_n}|X_s^x-X_s^y|\mathrm{d}s\right],
\end{aligned}$$

where $\widetilde{b}(x)=b(x)+\int_{U_1}g_1(x,u)\mu_1(\mathrm{d}u)$. Let $D:=\{t\geq 0 \;:\; \mathbb{P}[X_t^x=X_{t-}^x, X_t^y=X_{t-}^y]=1\}$. Taking the limit $n\to\infty$ and invoking $\tau_n\to\infty$ a.s., yields, due to cádlág sample paths of X, that for $t\in D$,

$$\begin{aligned}
\mathbb{E}[|X_t^x-X_t^y|] &= \mathbb{E}[|X_{t-}^x-X_{t-}^y|]\\
&\leq \liminf_{n\to\infty}\mathbb{E}[|X_{t\wedge\tau_n}^x - X_{t\wedge\tau_n}^y|]\\
&\leq |x-y| - A\mathbb{E}\left[\int_0^t |X_s^x-X_s^y|\mathrm{d}s\right].
\end{aligned}$$

For $t\notin D$, we find $D\ni t_n\downarrow t$ and then let $n\to\infty$ to obtain the above estimate also for $t\notin D$. Applying Gronwall's lemma yields $\mathbb{E}\left[|X_t^x-X_t^y|\right]\leq |x-y|\mathrm{e}^{-At}$ for $t\geq 0$. Let us now derive (3.4). We denote by δ_x and δ_y the Dirac measure concentrated in x and y, respectively. Assume $0\leq x\leq y$. Since the joint distribution of (X_t^x,X_t^y) belongs to $\mathcal{H}(P_t^*\delta_x,P_t^*\delta_y)$, we obtain

$$W_1\left(P_t^*\delta_x,P_t^*\delta_y\right)\leq \mathbb{E}\left[|X_t^x-X_t^y|\right]\leq |x-y|\mathrm{e}^{-At}, \quad t\geq 0. \tag{3.5}$$

Let H be any coupling of $(\varrho,\widetilde{\varrho})$ satisfying (2.1). Using the convexity of W_1 and (3.5), we get

$$\begin{aligned}
W_1\left(P_t^*\varrho,P_t^*\widetilde{\varrho}\right) &\leq \int_{\mathbb{R}_{\geq 0}\times\mathbb{R}_{\geq 0}} W_1\left(P_t^*\delta_x,P_t^*\delta_y\right)H\left(\mathrm{d}x,\mathrm{d}y\right)\\
&\leq \mathrm{e}^{-At}\int_{\mathbb{R}_{\geq 0}\times\mathbb{R}_{\geq 0}}|x-y|\,H\left(\mathrm{d}x,\mathrm{d}y\right)=e^{-At}W_1(\varrho,\widetilde{\varrho}).
\end{aligned}$$

Next we prove the existence of an invariant measure $\pi \in \mathcal{P}_1(\mathbb{R}_{\geq 0})$. We fix any $\varrho \in \mathcal{P}_1(\mathbb{R}_{\geq 0})$. Then, for $k, l \in \mathbb{N}$ with $k < l$,

$$W_1\left(P_k^*\varrho, P_l^*\varrho\right) \leq \sum_{s=k}^{l-1} W_1\left(P_{s+1}^*\varrho, P_s^*\varrho\right) \leq \sum_{s=k}^{l-1} \mathrm{e}^{-As} W_1\left(P_1^*\varrho, \varrho\right),$$

where we have used the semigroup property of P_{s+1}^* together with (3.4). Since the right-hand side tends to zero as $k, l \to \infty$, $(P_k^*\varrho)_{k\in\mathbb{N}} \subset \mathcal{P}_1(\mathbb{R}_{\geq 0})$ is a Cauchy sequence. As a consequence, there exists $\pi \in \mathcal{P}_1(\mathbb{R}_{\geq 0})$ such that $W_1(P_k^*\varrho, \pi)$ converges to zero as $k \to \infty$. To prove the invariance of π, i.e. $P_h^*\pi = \pi$, for all $h > 0$, let us fix $h > 0$. Using the semigroup property and (3.4) it follows

$$\begin{aligned} W_1\left(P_h^*\pi, \pi\right) &\leq W_1\left(P_h^*\pi, P_h^*P_k^*\varrho\right) + W_1\left(P_k^*P_h^*\varrho, P_k^*\varrho\right) + W_1\left(P_k^*\varrho, \pi\right) \\ &\leq \mathrm{e}^{-Ah} W_1\left(\pi, P_k^*\varrho\right) + \mathrm{e}^{-Ak} W_1\left(P_h^*\varrho, \varrho\right) + W_1\left(P_k^*\varrho, \pi\right), \end{aligned}$$

and the right-hand side tends to zero as $k \to \infty$. Hence, we see that $W_1(P_h^*\pi, \pi) = 0$.

Finally, we prove the uniqueness of invariant measures. Let $W_1^{\leq 1}$ be the Wasserstein distance given by (1.2) with $d(x, y) = 1 \wedge |x - y|$ as introduced in Example 2.1.(d). Using the invariance of π, $\widehat{\pi}$, and the convexity of $W_1^{\leq 1}$, for any $H \in \mathcal{H}(\pi, \widehat{\pi})$, we derive

$$\begin{aligned} W_1^{\leq 1}\left(\pi, \widehat{\pi}\right) &\leq \int_{\mathbb{R}_{\geq 0}\times\mathbb{R}_{\geq 0}} W_1^{\leq 1}\left(P_t^*\delta_x, P_t^*\delta_y\right) H\left(\mathrm{d}x, \mathrm{d}y\right) \\ &\leq \int_{\mathbb{R}_{\geq 0}\times\mathbb{R}_{\geq 0}} \mathbb{E}\left[|X_t^x - X_t^y| \wedge 1\right] H\left(\mathrm{d}x, \mathrm{d}y\right) \\ &\leq \int_{\mathbb{R}_{\geq 0}\times\mathbb{R}_{\geq 0}} \left(1 \wedge \mathbb{E}\left[|X_t^x - X_t^y|\right]\right) H\left(\mathrm{d}x, \mathrm{d}y\right) \\ &\leq \int_{\mathbb{R}_{\geq 0}\times\mathbb{R}_{\geq 0}} \left(1 \wedge \left(|x-y|\mathrm{e}^{-At}\right)\right) H\left(\mathrm{d}x, \mathrm{d}y\right). \end{aligned}$$

By dominated convergence we see that the right-hand side vanishes as $t \to \infty$. Consequently, $W_1^{\leq 1}(\pi, \widehat{\pi}) = 0$ and this implies that $\pi = \widehat{\pi}$. □

In our next result we prove that a similar ergodicity statement also holds true, if we replace $(m)_1$ either by $(m)_\lambda$ with $\lambda \in (0,1)$ or by $(m)_{\log}$. However, as in such a case the process does not have finite first moment, we cannot directly use the above arguments. Our idea is to approximate the process by truncating the big jumps. In this way, the approximating process has finite first moments and hence previous results can be applied. Using the properties of the Wasserstein distance and passing to the limit where the truncation vanishes, we are able to derive the desired result formulated at the end of this section. Below we first prove that by truncating the big jumps, we obtain an approximation of the process.

Proposition 3.6. *Let (b, σ, g_0, g_1) be admissible parameters, that conditions (3.a) – (3.e), $(m)_{\log}$ holds, and that*

$$x \longmapsto g_0(x, u_0), \quad \textit{and} \quad x \longmapsto g_1(x, u_1) \tag{3.6}$$

are continuous for μ_0-a.a. $u_0 \in U_0$ and μ_1-a.a. $u_1 \in U_1$. Finally, assume that there exists a sequence of subsets $(U_1^k)_{k\geq 1}$ of U_1 such that $|g_1(x, u)| \leq k$ holds for each $x \geq 0$, $k \geq 1$, $u \in U_1^k$, and $U_1^k \nearrow U_1$. Define $g_1^k(x, u) = \mathbb{1}_{U_1^k}(u) g_1(x, u)$ for $x \geq 0$ and $u \in U_1$. Then (b, σ, g_0, g_1^k) are admissible parameters which satisfy conditions (3.a) – (3.e), and $(m)_1$. Let $\{X_t^k : t \geq 0\}$ be the associated sequence of strong solutions to (1.1) with g_1 replaced by g_1^k but the same initial condition. Then $X^k \Rightarrow X$ holds in law as $k \to \infty$, where X is the unique strong solution to (1.1) with admissible parameters (b, σ, g_0, g_1).

Proof: It is not difficult to see that (b,σ,g_0,g_1^k) are admissible and satisfy conditions (3.a) – (3.e), and $(m)_1$. Let

$$h(x) = \begin{cases} x, & x \in [-1,1], \\ 1, & x \geq 1, \\ -1, & x \leq -1, \end{cases} \qquad \widetilde{h}(x) = x - h(x) = \begin{cases} 0, & x \in [-1,1], \\ x-1, & x \geq 1, \\ x+1, & x \leq -1. \end{cases}$$

Applying Lemma 3.4, we observe that X and X^k are semimartingales with characteristics (B,C,K) with respect to the truncation function h given by (3.3), and (B^k,C^k,K^k) given by

$$\begin{aligned} B_t^k &= \int_0^t \widetilde{b}^k(X_s^k)\mathrm{d}s, \qquad \widetilde{b}^k(x) = b(x) - \int_{U_0} \widetilde{h}(g_0(x,u))\mu_0(\mathrm{d}u) + \int_{U_1} h(g_1^k(x,u))\mu_1(\mathrm{d}u), \\ C_t^k &= \int_0^t c(X_s^k)\mathrm{d}s, \qquad c(x) = \int_E \sigma(x,u)^2 \varkappa(\mathrm{d}u), \\ K^k(x,A) &= \int_{U_0} \mathbb{1}_{A\backslash\{0\}}(g_0(x,u))\mu_0(\mathrm{d}u) + \int_{U_1} \mathbb{1}_{A\backslash\{0\}}(g_1^k(x,u))\mu_1(\mathrm{d}u). \end{aligned}$$

Finally define

$$\widetilde{c}(x) = c(x) + \int_{\mathbb{R}} h(y)^2 K(x,\mathrm{d}y), \qquad \widetilde{c}^k(x) = c(x) + \int_{\mathbb{R}} h(y)^2 K^k(x,\mathrm{d}y).$$

The assertion now follows from Kurtz (2011, Chapter IX, Theorem 4.8) and the characterization of weak solutions to (1.1) in terms of semi-martingale characteristics (see Kurtz (2011, Chapter III, Theorem 2.26)), provided we can show the following set of conditions:

(a) $\lim_{a\nearrow\infty}\sup_{x\in[0,R]} K(x,[-a,a]^c) = 0$ for each $R>0$;

(b) The functions $x \longmapsto \widetilde{b}(x), \widetilde{c}(x), \int_{\mathbb{R}} f(y)K(x,\mathrm{d}y)$ are continuous for each continuous and bounded function f satisfying $|f(y)| \leq C_f(1\wedge y^2)$ for some constant $C_f>0$;

(c) The convergence $\widetilde{b}^k \to \widetilde{b}$, $\widetilde{c}^k \to \widetilde{c}$, and $\int_{\mathbb{R}} f(y)K^k(\cdot,\mathrm{d}y) \to \int_{\mathbb{R}} f(y)K(\cdot,\mathrm{d}y)$ hold locally uniformly for each continuous and bounded function f satisfying $|f(y)| \leq C_f(1\wedge y^2)$, where $C_f>0$ is some constant.

Condition (a) follows from $(m)_{\log}$ and the estimates

$$\begin{aligned} K(x,[-a,a]^c) &= \mu_0(\{|g_0(x,\cdot)|>a\}) + \mu_1(\{|g_1(x,\cdot)|>a\}) \\ &\leq \frac{1}{a}\int_{U_0} \mathbb{1}_{\{|g_0(x,u)|>1\}}|g_0(x,u)|\mu_0(\mathrm{d}u) \\ &\quad + \frac{1}{\log(1+a)}\int_{U_1} \mathbb{1}_{\{|g_1(x,u)|>1\}}\log(1+|g_1(x,u)|)\mu_1(\mathrm{d}u) \\ &\leq \frac{K_{\log}(2+x)}{a} + \frac{K_{\log}(2+x)}{\log(1+a)}, \end{aligned}$$

where $a>1$. For condition (b), we let $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}_{\geq 0}$ such that $x_n \to x$. Then

$$\begin{aligned} |\widetilde{b}(x_n) - \widetilde{b}(x)| &\leq |b(x_n)-b(x)| + \int_{U_0} |\widetilde{h}(g_0(x_n,u)) - \widetilde{h}(g_0(x,u))|\mu_0(\mathrm{d}u) \\ &\quad + \int_{U_1} |h(g_1(x_n,u)) - h(g_1(x,u))|\mu_1(\mathrm{d}u) \\ &\leq J_1 + J_2 + J_3. \end{aligned}$$

The first term J_1 converges to zero due to the continuity of b. The second term tends to zero due to (3.6), the continuity of $\widetilde{h}$, and the dominated convergence theorem. The latter one is applicable in view of $|\widetilde{h}(g_0(x_n,u))| \leq \mathbb{1}_{\{|g_0(x_n,u)|>1\}}(|g_0(x_n,u)|+1) \leq \mathbb{1}_{\{\overline{g}_0(j,u)>1\}}2\overline{g}_0(j,u)$ and condition (3.d), i.e., $\int_{U_0} \mathbb{1}_{\{\overline{g}_0(j,u)>1\}} \overline{g}_0(j,u)\mu_0(\mathrm{d}u) \leq L(j) < \infty$, where $j \in \mathbb{N}$ is such that $x_n, x \leq j$ for each $n \in \mathbb{N}$. For the last term J_3 we use condition (3.b) to find that

$$J_3 \leq \int_{U_1} |g_1(x_n,u) - g_1(x,u)|\mu_1(\mathrm{d}u) \leq r_j(|x_n - x|),$$

which tends to zero due to the continuity of r_j. Next, for $\widetilde{c}$ we obtain

$$\begin{aligned} |\widetilde{c}(x_n) - \widetilde{c}(x)| &\leq \int_E |\sigma(x_n,u)^2 - \sigma(x,u)^2|\varkappa(\mathrm{d}u) \\ &\quad + \int_{U_0} \left|h(g_0(x_n,u))^2 - h(g_0(x,u))^2\right| \mu_0(\mathrm{d}u) \\ &\quad + \int_{U_1} \left|h(g_1(x_n,u))^2 - h(g_1(x,u))^2\right| \mu_1(\mathrm{d}u) \\ &= I_1 + I_2 + I_3. \end{aligned}$$

For the first term we use the Hölder inequality and then condition (3.c) to get

$$\begin{aligned} I_1 &\leq \int_E (|\sigma(x_n,u)| + |\sigma(x,u)|)|\sigma(x_n,u) - \sigma(x,u)|\varkappa(\mathrm{d}u) \\ &\leq \left(\int_E (|\sigma(x_n,u)| + |\sigma(x,u)|)^2 \varkappa(\mathrm{d}u)\right)^{1/2} \left(\int_E |\sigma(x_n,u) - \sigma(x,u)|^2 \varkappa(\mathrm{d}u)\right)^{1/2} \\ &\leq \left(2K_1(2+x_n)^2 + 2K_1(2+x)^2\right)^{1/2} \sqrt{A_j|x_n - x|} \to 0. \end{aligned}$$

Similarly, using conditions (3.d), $|h(x)| \leq 1 \wedge |x|$, and $|h(x) - h(y)| \leq 2 \wedge |x-y| \leq 2(1 \wedge |x-y|)$ we find that

$$\begin{aligned} I_2 &\leq \left(\int_{U_0} \left(2h(g_0(x_n,u))^2 + 2h(g_0(x,u))^2\right) \mu_0(\mathrm{d}u)\right)^{1/2} \\ &\qquad \cdot \left(\int_{U_0} |h(g_0(x_n,u)) - h(g_0(x,u))|^2 \mu_0(\mathrm{d}u)\right)^{1/2} \\ &\leq 2\sqrt{2}\left(\int_{U_0} \left(1 \wedge |g_0(x_n,u)|^2 + 1 \wedge |g_0(x,u)|^2\right) \mu_0(\mathrm{d}u)\right)^{1/2} \\ &\qquad \cdot \left(\int_{U_0} 1 \wedge |g_0(x_n,u) - g_0(x,u)|^2 \mu_0(\mathrm{d}u)\right)^{1/2} \\ &\leq 2\sqrt{2}\left(\int_{U_0} \left(|g_0(x_n,u)| \wedge |g_0(x_n,u)|^2 + |g_0(x,u)| \wedge |g_0(x,u)|^2\right) \mu_0(\mathrm{d}u)\right)^{1/2} \\ &\qquad \cdot \left(\int_{U_0} |g_0(x_n,u) - g_0(x,u)| \wedge |g_0(x_n,u) - g_0(x,u)|^2 \mu_0(\mathrm{d}u)\right)^{1/2} \\ &\leq 2\sqrt{2}\left(L(x_n) + L(x)\right)^{1/2} \sqrt{B_j|x_n - x|}, \end{aligned}$$

which tends to zero as $n \to \infty$. Similarly, for I_3 we find by conditions (3.a) and (3.e) that

$$
\begin{aligned}
I_3 &\leq \left(\int_{U_1} \left(2h(g_1(x_n,u))^2 + 2h(g_1(x,u))^2 \right) \mu_1(\mathrm{d}u) \right)^{1/2} \\
&\qquad \cdot \left(\int_{U_1} |h(g_1(x_n,u)) - h(g_1(x,u))|^2 \, \mu_1(\mathrm{d}u) \right)^{1/2} \\
&\leq 2\sqrt{2} \left(\int_{U_1} \left(1 \wedge |g_1(x_n,u)|^2 + 1 \wedge |g_1(x,u)|^2 \right) \mu_1(\mathrm{d}u) \right)^{1/2} \\
&\qquad \cdot \left(\int_{U_1} 1 \wedge |g_1(x_n,u) - g_1(x,u)|^2 \, \mu_1(\mathrm{d}u) \right)^{1/2} \\
&\leq 2\sqrt{2} \sqrt{K_0((2+x_n) + (2+x)) r_j(|x_n - x|)},
\end{aligned}
$$

which tends to zero as $n \to \infty$. Finally, we find that

$$
\begin{aligned}
\left| \int_{\mathbb{R}} f(y) K(x_n, \mathrm{d}y) - \int_{\mathbb{R}} f(y) K(x, \mathrm{d}y) \right| &\leq \left| \int_{U_0} (f(g_0(x_n,u)) - f(g_0(x,u))) \mu_0(\mathrm{d}u) \right| \\
&\quad + \left| \int_{U_1} (f(g_1(x_n,u)) - f(g_1(x,u))) \mu_1(\mathrm{d}u) \right|.
\end{aligned}
$$

For the integral against μ_0 we use (3.6) combined with the dominated convergence theorem, which is applicable due to $|f(g_0(x_n,u))| \leq C_f(1 \wedge |g_0(x_n,u)|^2) \leq C_f(1 \wedge |\overline{g}_0(j,u)|^2)$ and condition (3.d), i.e.,

$$
\int_{U_0} (1 \wedge |\overline{g}_0(j,u)|^2) \mu_0(\mathrm{d}u) \leq \int_{U_0} (|\overline{g}_0(j,u)| \wedge |\overline{g}_0(j,u)|^2) \mu_0(\mathrm{d}u) \leq L(j) < \infty.
$$

For the second term we use again the dominated convergence theorem, which is applicable due to (3.6) and condition (3.a), since $|f(g_1(x_n,u))| \leq C_f(1 \wedge |g_1(x_n,u)|^2) \leq C_f(1 \wedge |g_1(x_n,u)|) \leq C_f(1 \wedge |\overline{g}_1(j,u)|)$ and $\int_{U_1} (1 \wedge |\overline{g}_1(j,u)|) \mu_1(\mathrm{d}u) \leq K_0(2+j) < \infty$. Thus we have shown that $\int_{\mathbb{R}} f(y) K(x_n, \mathrm{d}y) \longmapsto \int_{\mathbb{R}} f(y) K(x, \mathrm{d}y)$, which completes the proof of property (b). To prove property (c), let us first observe that

$$
\begin{aligned}
|\widetilde{b}^k(x) - \widetilde{b}(x)| &\leq \left| \int_{U_1} \left(h(g_1^k(x,u)) - h(g_1(x,u)) \right) \mu_1(\mathrm{d}u) \right| \\
&= \left| \int_{U_1 \setminus U_1^k} h(g_1(x,u)) \mu_1(\mathrm{d}u) \right| \\
&\leq \int_{U_1 \setminus U_1^k} (1 \wedge |g_1(x,u)|) \mu_1(\mathrm{d}u) \\
&\leq \int_{U_1 \setminus U_1^k} (1 \wedge |\overline{g}_1(j,u)|) \mu_1(\mathrm{d}u),
\end{aligned}
$$

where $j \in \mathbb{N}$ is such that $x \in [0, j]$. The convergence to zero is now a direct consequence of $U_1^k \nearrow U_1$, condition (3.a), and the dominated convergence theorem. For the second term we observe that

$$\begin{aligned}|\widetilde{c}^k(x) - \widetilde{c}(x)| &\le \int_{U_1 \setminus U_1^k} h(g_1(x,u))^2 \mu_1(\mathrm{d}u) \\ &\le \int_{U_1 \setminus U_1^k} (1 \wedge |g_1(x,u)|^2)\mu_1(\mathrm{d}u) \\ &\le \int_{U_1 \setminus U_1^k} (1 \wedge |\overline{g}_1(j,u)|)\mu_1(\mathrm{d}u).\end{aligned}$$

This shows that $\widetilde{c}^k \longrightarrow \widetilde{c}$ locally uniformly in x. Finally, let f be a continuous and bounded function satisfying $|f(y)| \le C_f(1 \wedge |y|^2)$. Then

$$\begin{aligned}\left| \int_{\mathbb{R}} f(y) K^k(x, \mathrm{d}y) - \int_{\mathbb{R}} f(y) K(x, \mathrm{d}y) \right| &= \left| \int_{U_1} \left(f(g_1^k(x,u)) - f(g_1(x,u)) \right) \mu_1(\mathrm{d}u) \right| \\ &= \left| \int_{U_1 \setminus U_1^k} f(g_1(x,u))\mu_1(\mathrm{d}u) \right| \\ &\le C_f \int_{U_1 \setminus U_1^k} (1 \wedge |\overline{g}_1(j,u)|)\mu_1(\mathrm{d}u).\end{aligned}$$

This completes the proof of property (c) and hence completes the proof of this Proposition. □

Remark 3.7. By adapting the methods from Dawson and Li (2012), it should be possible to derive a stronger statement where X and the approximation X^k are both realized on the same probability space and the convergence holds at least in the sense that $\mathbb{E}[1 \wedge |X_t^k - X_t|]$ converges to zero. Since we do not need such a stronger statement, we leave these details for the interested reader.

Below we introduce a localized version of the dissipativity condition (3.f).

(3.g) There exists a sequence of subsets $(U_1^k)_{k \ge 1}$ of U_1 such that $|g_1(x,u)| \le k$ holds for each $x \ge 0$, $k \ge 1$ and $u \in U_1^k$, and $U_1^k \nearrow U_1$. Moreover, there exists a constant $A > 0$ such that

$$b(y) - b(x) + \int_{U_1^k} (g_1(y,u) - g_1(x,u))\mu_1(\mathrm{d}u) \le -A(y-x)$$

holds for all $0 \le x \le y$ and $k \ge 1$.

Now we are prepared for our second main result of this section.

Theorem 3.8. *Let (b, σ, g_0, g_1) be admissible parameters. Suppose that conditions (3.a) – (3.e), (3.g), and (3.6) hold. Then the following assertions hold:*

(a) *If $(m)_\lambda$ holds for some $\lambda \in (0,1)$, then*

$$W_{\lambda'}(P_t^* \varrho, P_t^* \widetilde{\varrho}) \le \mathrm{e}^{-\lambda' A t} W_{\lambda'}(\varrho, \widetilde{\varrho}), \qquad t \ge 0$$

holds for each $\lambda' \in (0, \lambda)$. In particular, there exists a unique invariant measure $\pi \in \mathcal{P}(\mathbb{R}_{\ge 0})$, this measure belongs to $\mathcal{P}_{\lambda'}(\mathbb{R}_{\ge 0})$, and for all $\varrho \in \mathcal{P}_{\lambda'}(\mathbb{R}_{\ge 0})$ one has

$$W_{\lambda'}(P_t^* \varrho, \pi) \le \mathrm{e}^{-\lambda' A t} W_{\lambda'}(\varrho, \pi), \quad t \ge 0.$$

(b) *If $(m)_{\log}$ holds, then for all $\varrho, \widetilde{\varrho} \in \mathcal{P}(\mathbb{R}_{\ge 0})$ one has $\lim_{t \to \infty} W_1^{\le 1}(P_t^* \varrho, P_t^* \widetilde{\varrho}) = 0$, and there exists a constant $c > 0$ such that*

$$W_1^{\le 1}(P_t^* \varrho, P_t^* \widetilde{\varrho}) \le c \left(\mathrm{e}^{-At} W_{\log}(\varrho, \widetilde{\varrho}) + \min\{\mathrm{e}^{-At}, W_{\log}(\varrho, \widetilde{\varrho})\} \right).$$

In particular, there exists a unique invariant measure π, this measure belongs to $\mathcal{P}_{\log}(\mathbb{R}_{\geq 0})$, and for all $\varrho \in \mathcal{P}(\mathbb{R}_{\geq 0})$ one has $\lim_{t\to\infty} W_1^{\leq 1}(P_t^*\varrho, \pi) = 0$ and

$$W_1^{\leq 1}(P_t^*\varrho, \pi) \leq c\left(\mathrm{e}^{-At} W_{\log}(\varrho, \pi) + \min\{\mathrm{e}^{-At}, W_{\log}(\varrho, \pi)\}\right).$$

Proof: Let $(P_t^*)_{t\geq 0}$ be the transition semigroup associated with $\{X_t^x : t \geq 0\}$, and let $(P_t^{*,k})_{t\geq 0}$ be the transition semigroup associated with the approximation $\{X_t^{x,k} : t \geq 0\}$ constructed in Proposition 3.6. Then $\{X_t^{x,k} : t \geq 0\}$ satisfies all conditions of Theorem 3.5.

(a) Lemma 2.2 and Theorem 3.5 applied to the approximating process $\{X_t^{x,k} : t \geq 0\}$ gives $W_{\lambda'}(P_t^{*,k}\delta_x, P_t^{*,k}\delta_y) \leq \left(W_1(P_t^{*,k}\delta_x, P_t^{*,k}\delta_y)\right)^{\lambda'} \leq \mathrm{e}^{-A\lambda' t}|x-y|^{\lambda'}$. This yields

$$\begin{aligned} W_{\lambda'}(P_t^*\delta_x, P_t^*\delta_y) &\leq W_{\lambda'}(P_t^*\delta_x, P_t^{*,k}\delta_x) + W_{\lambda'}(P_t^{*,k}\delta_x, P_t^{*,k}\delta_y) + W_{\lambda'}(P_t^{*,k}\delta_y, P_t^*\delta_y) \\ &\leq W_{\lambda'}(P_t^*\delta_x, P_t^{*,k}\delta_x) + \mathrm{e}^{-A\lambda' t}|x-y|^{\lambda'} + W_{\lambda'}(P_t^{*,k}\delta_y, P_t^*\delta_y). \end{aligned}$$

Let $D := \{t \geq 0 \ : \ \mathbb{P}[X_t^x = X_{t-}^x] = 1\}$. Since the process X has cádlág paths, the set $\mathbb{R}_{\geq 0}\backslash D$ is at most countable (see Ethier and Kurtz (1986, Chapter 3, Lemma 7.7)). Moreover, since $X^k \Rightarrow X$, the characterization of convergence in the Skorohod space implies that $X_t^{x,k} \Rightarrow X_t^x$ holds for all $t \in D$ (see Ethier and Kurtz (1986, Chapter 3, Theorem 7.8)). Since $\sup_{k\geq 1}\mathbb{E}[(X_t^k)^\lambda] < \infty$, we conclude that $X_t^{x,k} \Rightarrow X_t^x$ also holds in $W_{\lambda'}$ for $t \in D$. Hence taking the limit $k \to \infty$ gives $W_{\lambda'}(P_t^*\delta_x, P_t^*\delta_y) \leq \mathrm{e}^{-\lambda' At}|x-y|$ for $t \in D$. Now let $t \in \mathbb{R}_{\geq 0}\backslash D$, then using Ethier and Kurtz (1986, Chapter 3, Theorem 7.8.(a)) we find a sequence $(t_n)_{n\in\mathbb{N}} \subset \mathbb{R}_{\geq 0}\backslash D$ such that $t_n \to t$, $t_n > t$. Then

$$\begin{aligned} W_{\lambda'}(P_t^*\delta_x, P_t^*\delta_y) &\leq W_{\lambda'}(P_t^*\delta_x, P_{t_n}^*\delta_x) + W_{\lambda'}(P_{t_n}^*\delta_x, P_{t_n}^*\delta_y) + W_{\lambda'}(P_{t_n}^*\delta_y, P_t^*\delta_y) \\ &\leq W_{\lambda'}(P_t^*\delta_x, P_{t_n}^*\delta_x) + \mathrm{e}^{-\lambda' At}|x-y| + W_{\lambda'}(P_{t_n}^*\delta_y, P_t^*\delta_y). \end{aligned}$$

It is not difficult to show that the first and last term tend to zero as $n \to \infty$. This proves the assertion.

(b) In this case we find that

$$W_1^{\leq 1}(P_t^{*,k}\delta_x, P_t^{*,k}\delta_y) \leq 1 \wedge \mathbb{E}[|X_t^{x,k} - X_t^{y,k}|] \leq \min\left\{1, \mathrm{e}^{-At}|x-y|\right\}$$

and hence

$$W_1^{\leq 1}(P_t^*\delta_x, P_t^*\delta_y) \leq W_1^{\leq 1}(P_t^*\delta_x, P_t^{*,k}\delta_x) + \min\{1, \mathrm{e}^{-At}|x-y|\} + W_1^{\leq 1}(P_t^{*,k}\delta_y, P_t^*\delta_y).$$

Since $X_t^{x,k} \Rightarrow X_t^x$, we find that $W_1^{\leq 1}(P_t^*\delta_x, P_t^{*,k}\delta_x) \longrightarrow 0$ as $k \to \infty$ for $t \in D$. Hence taking the limit $k \to \infty$ gives $W_1^{\leq 1}(P_t^*\delta_x, P_t^*\delta_y) \leq \min\{1, \mathrm{e}^{-At}|x-y|\}$ first for $t \in D$, and then using the same arguments as in part (a) also for all $t \geq 0$. Using now that $1 \wedge a \leq \log(2)^{-1}\log(1+a)$ for all $a \geq 0$, combined with the inequality

$$\log(1 + a\cdot d) \leq C\min\{a, \log(1+d)\} + Ca\log(1+d), \tag{3.7}$$

where $C > 0$ is a constant and a, $d \geq 0$ (see Friesen et al. (2020b, Lemma A.5)), we find for some new constant $C > 0$

$$W_1^{\leq 1}(P_t^*\delta_x, P_t^*\delta_y) \leq C\left(e^{-At}\log(1+|x-y|) + \min\left\{e^{-At}, \log(1+|x-y|)\right\}\right).$$

This readily implies

$$W_1^{\leq 1}(P_t^*\varrho, P_t^*\widetilde{\varrho}) \leq c\left(\mathrm{e}^{-At}W_{\log}(\varrho, \widetilde{\varrho}) + \min\{\mathrm{e}^{-At}, W_{\log}(\varrho, \widetilde{\varrho})\}\right).$$

The second part of the assertion can be now deduced by a standard Cauchy argument similarly to the proof of Friesen et al. (2020b, Lemma 7.2). □

At this point it is worthwhile to mention that although condition ($3.f$) (or ($3.g$)) seems to be restrictive, there are abundant classes of processes fulfilling it. In the subsequent sections we discuss the particular classes of continuous-state branching processes with immigration and their extensions to nonlinear branching mechanisms or Lévy random environments.

However, condition ($3.g$) requires a global estimate valid for all $x, y \geq 0$ which essentially rules out models with strongly non-linear drifts such as $b(x) = x - x^3$. Assuming that the coefficients in front of the noise terms are non-degenerate, it was shown in Douc et al. (2009); Eberle et al. (2019); Kulik (2009); Veretennikov (1987) that already a local version of condition ($3.g$) is sufficient for the exponential ergodicity. Thus one may hope to replace the global drift condition ($3.g$) by a local drift condition, provided that the noise terms are non-degenerate in a certain sense. However, our applications of interest studied in the subsequent sections do not satisfy such non-degeneracy conditions, but still satisfy the global drift condition ($3.g$). Therefore, we focus in this work on this specific setting.

4. Branching processes

4.1. *Nonlinear branching processes.* Let γ_0, γ_1, γ_2 be continuous functions on $\mathbb{R}_{\geq 0}$ satisfying $\gamma_0(0) \geq 0$, $\gamma_1, \gamma_2 \geq 0$, $\gamma_1(0) = \gamma_2(0) = 0$, and let m, ν be Borel measures on $\mathbb{R}_{\geq 0}$ satisfying $m(\{0\}) = \nu(\{0\}) = 0$ and the moment condition

$$\int_{\mathbb{R}_{\geq 0}} z \wedge z^2 m(\mathrm{d}z) + \int_{\mathbb{R}_{\geq 0}} 1 \wedge z\nu(\mathrm{d}z) < \infty. \tag{4.1}$$

A CNBI process is obtained as the unique $\mathbb{R}_{\geq 0}$-valued strong solution of

$$\begin{aligned} X_t = X_0 &+ \int_0^t \gamma_0(X_s)\,\mathrm{d}s + \int_0^t \int_0^\infty \mathbb{1}_{\{u \leq \gamma_1(X_s)\}} W(\mathrm{d}s, \mathrm{d}u) \\ &+ \int_0^t \int_0^\infty \int_0^\infty \mathbb{1}_{\{u \leq \gamma_2(X_{s-})\}} z \widetilde{N}_0(\mathrm{d}s, \mathrm{d}z, \mathrm{d}u) + \int_0^t \int_0^\infty z N_1(\mathrm{d}s, \mathrm{d}z), \end{aligned} \tag{4.2}$$

where $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ is a stochastic basis with the usual conditions, $W(\mathrm{d}t, \mathrm{d}u)$ is an $(\mathcal{F}_t)_{t\geq 0}$-Gaussian white noise with intensity measure $\mathrm{d}t\mathrm{d}u$ on $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$, $N_0(\mathrm{d}t, \mathrm{d}z, \mathrm{d}u)$ and $N_1(\mathrm{d}t, \mathrm{d}z)$ are $(\mathcal{F}_t)_{t\geq 0}$-Poisson random measures with intensity measures $\mathrm{d}t m(\mathrm{d}z)\mathrm{d}u$ on $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$ and $\mathrm{d}t\nu(\mathrm{d}z)$ on $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$. Further, it is assumed that $W(\mathrm{d}t, \mathrm{d}u)$, $N_0(\mathrm{d}t, \mathrm{d}z, \mathrm{d}u)$, and $N_1(\mathrm{d}t, \mathrm{d}z)$ are mutually independent and $\widetilde{N}_0(\mathrm{d}t, \mathrm{d}z, \mathrm{d}u) = N_0(\mathrm{d}t, \mathrm{d}z, \mathrm{d}u) - \mathrm{d}t m(\mathrm{d}z)\mathrm{d}u$ denotes the corresponding compensated Poisson random measure.

Continuous-state nonlinear branching processes (CNB process) were recently introduced and studied by Li *et al.* Li et al. (2019) and correspond to (4.2) where $\nu = 0$, i.e., to absence of immigration. In this paper we add to the CNB process a general immigration mechanism and therefore call it a continuous-state nonlinear branching process with immigration (CNBI process). Below we apply our main results from Section 3 to the case of CNBI and CBI processes which leads to new insights on the existence, uniqueness of invariant measures as well as convergence of transition probabilities towards the unique invariant measure.

Theorem 4.1. *Suppose that the functions γ_i, $i = 0, 1, 2$, satisfy the following:*

(i) $\gamma_0(0) \geq 0$, γ_1, $\gamma_2 \geq 0$ *are continuous, and γ_2 is nondecreasing;*

(ii) *there exists a constant $K \geq 0$ such that for all $x \geq 0$,*

$$\gamma_0(x) \leq K(1 + x), \qquad \gamma_1(x) + \gamma_2(x) \leq K(1 + x)^2;$$

(iii) *for each $j \geq 1$ there exists a constant $c_j > 0$ such that, for all $0 \leq x, y \leq j$,*

$$|\gamma_0(x) - \gamma_0(y)| + |\gamma_1(x) - \gamma_1(y)| + |\gamma_2(x) - \gamma_2(y)| \leq c_j|x - y|.$$

(iv) *It holds that $\int_{\{|z|>1\}} \log(1 + z)\nu(\mathrm{d}z) < \infty$.*

Then there exists a unique non-negative strong solution of (4.2)*, and this solution has the comparison property.*

Proof: The assertion is a direct consequence of Theorem 3.3 applied to the following choices:

- $E = \mathbb{R}_{\geq 0}$, $U_0 = \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$, $U_1 = \mathbb{R}_{\geq 0}$;
- $b(x) = b_1(x) = \gamma_0(x)$, $\sigma(x,u) = \mathbb{1}_{\{u \leq \gamma_1(x)\}}$, $g_0(x,z,u) = z\mathbb{1}_{\{u \leq \gamma_2(x)\}}$, $g_1(x,z) = z$;
- $\varkappa(\mathrm{d}u) = \mathrm{d}u$, $\mu_0(\mathrm{d}z, \mathrm{d}u) = m(\mathrm{d}z)\mathrm{d}u$, $\mu_1(\mathrm{d}z) = \nu(\mathrm{d}z)$.

Indeed, it is routine to check the conditions (3.a)-(3.e), and $(m)_{\log}$ is clearly true because of assumption (iv). This proves the assertion. □

As a consequence of Theorem 4.1 the unique solution to (4.2) is a strong $(\mathcal{F}_t)_{t\geq 0}$-Markov process which is the desired CNBI process. Let $\{P_t : t \geq 0\}$ be its transition semigroup and $\{P_t^* : t \geq 0\}$ be the dual semigroup. Ergodicity of the CNBI process in different Wasserstein distances is obtained below.

Theorem 4.2. *Suppose that conditions (i) – (iii) of Theorem 4.1 are satisfied and assume that there exists a constant $A > 0$ such that*

$$\gamma_0(y) - \gamma_0(x) \leq -A(y-x), \quad 0 \leq x \leq y; \tag{4.3}$$

Then the following assertions hold:

(a) *If $\int_{\{z>1\}} \log(1+z)\nu(\mathrm{d}z) < \infty$ and $\gamma_2(x) \leq K(1+x)$ hold for all $x \geq 0$, then for all $\varrho, \widetilde{\varrho} \in \mathcal{P}(\mathbb{R}_{\geq 0})$ one has $\lim_{t\to\infty} W_1^{\leq 1}(P_t^*\varrho, P_t^*\widetilde{\varrho}) = 0$, and there exists a constant $c > 0$ such that*
$$W_1^{\leq 1}(P_t^*\varrho, P_t^*\widetilde{\varrho}) \leq c\left(\mathrm{e}^{-At} W_{\log}(\varrho, \widetilde{\varrho}) + \min\{\mathrm{e}^{-At}, W_{\log}(\varrho, \widetilde{\varrho})\}\right).$$
In particular, there exists a unique invariant measure π, this measure belongs to $\mathcal{P}_{\log}(\mathbb{R}_{\geq 0})$, and for all $\varrho \in \mathcal{P}(\mathbb{R}_{\geq 0})$ one has $\lim_{t\to\infty} W_1^{\leq 1}(P_t^*\varrho, \pi) = 0$ and
$$W_1^{\leq 1}(P_t^*\varrho, \pi) \leq c\left(\mathrm{e}^{-At} W_{\log}(\varrho, \pi) + \min\{\mathrm{e}^{-At}, W_{\log}(\varrho, \pi)\}\right).$$

(b) *If there exists $\lambda \in (0,1)$ such that $\int_{\{z>1\}} z^\lambda \nu(\mathrm{d}z) < \infty$ and $\gamma_2(x) \leq K(1+x)^\lambda$ for all $x \geq 0$, then*
$$W_{\lambda'}(P_t^*\varrho, P_t^*\widetilde{\varrho}) \leq \mathrm{e}^{-\lambda' At} W_{\lambda'}(\varrho, \widetilde{\varrho}), \qquad t \geq 0$$
holds for each $\lambda' \in (0,\lambda)$. In particular, there exists a unique invariant measure $\pi \in \mathcal{P}(\mathbb{R}_{\geq 0})$, this measure belongs to $\mathcal{P}_{\lambda'}(\mathbb{R}_{\geq 0})$, and for all $\varrho \in \mathcal{P}_{\lambda'}(\mathbb{R}_{\geq 0})$ one has
$$W_{\lambda'}\left(P_t^*\varrho, \pi\right) \leq \mathrm{e}^{-\lambda' At} W_{\lambda'}\left(\varrho, \pi\right), \quad t \geq 0.$$

(c) *If $\int_{\{z>1\}} z\nu(\mathrm{d}z) < \infty$ and $\gamma_2(x) \leq K(1+x)$ for all $x \geq 0$, then*
$$W_1(P_t^*\varrho, P_t^*\widetilde{\varrho}) \leq \mathrm{e}^{-At} W_1(\varrho, \widetilde{\varrho}), \qquad t \geq 0.$$
In particular, there exists a unique invariant measure $\pi \in \mathcal{P}(\mathbb{R}_{\geq 0})$, this measure belongs to $\mathcal{P}_1(\mathbb{R}_{\geq 0})$, and for all $\varrho \in \mathcal{P}_1(\mathbb{R}_{\geq 0})$ one has
$$W_1\left(P_t^*\varrho, \pi\right) \leq \mathrm{e}^{-At} W_1\left(\varrho, \pi\right), \quad t \geq 0.$$

Proof: The assertion follows from either Theorem 3.5 or Theorem 3.8. Indeed, choosing $U_1^k = [0,k]$, for (a) condition $(m)_{\log}$ is satisfied, (b) implies $(m)_\lambda$, (c) implies $(m)_1$. □

We conclude this section with the following example for γ_0, γ_1, γ_2.

Example 4.3. Theorem 4.2 is applicable for the particular choice:

- $\gamma_0(x) = \beta - bx$ where $\beta \geq 0$ and $b > 0$ are constants;
- $\gamma_1(x) = x^\alpha$ with $\alpha \in [1,2]$;
- $\gamma_2(x) = x$ or $\gamma_2(x) = x \wedge x^\delta$ with $\delta \in [0,1]$.

Note that the particular choice $\gamma_0(x) = x - x^3$ would not satisfy the condition (4.3) and hence our result from Section 3 is not applicable in this case. The treatment of all these cases requires different methods and is left for future research.

4.2. *Continuous-state branching processes with immigration.* Next we study the special case where $\gamma_0, \gamma_1, \gamma_2$ are affine linear, i.e., the case of CBI processes in more detail. Namely, we suppose that

$$\gamma_0(x) = \beta + bx, \quad \gamma_1(x) = \sigma^2 x, \quad \text{and} \quad \gamma_2(x) = x.$$

In this case the corresponding process obtained from (4.2) is a continuous-state branching processes with immigration (CBI processes) as first introduced by Feller (1950) Feller (1951) and Jiřina (1958) Jiřina (1958) and then developed by Kawazu and Watanabe (1971) Kawazu and Watanabe (1971). For a detailed treatment of CBI processes encompassing a concise introduction we refer to the monographs of Li Li (2011).

Following the notion on CBI processes, we call $(\beta, b, \sigma, m, \nu)$ *admissible parameters* if $\beta \geq 0$, $b \in \mathbb{R}$, $\sigma \geq 0$, m and ν are Borel measures on $\mathbb{R}_{\geq 0}$ satisfying $m(\{0\}) = \nu(\{0\}) = 0$, and (4.1). CBI processes have the crucial property that their Laplace transform can be explicitly expressed in terms of a solution to a ordinary differential equations (the so-called Riccatti equation). Namely, for $\lambda \geq 0$, define the *branching mechanism*

$$\phi(\lambda) = b\lambda + \frac{1}{2}\sigma^2\lambda^2 + \int_0^\infty \left(\mathrm{e}^{-\lambda z} - 1 + \lambda z\right) m\left(\mathrm{d}z\right),$$

and *immigration mechanism*

$$\psi(\lambda) = \beta\lambda + \int_0^\infty \left(1 - \mathrm{e}^{-\lambda z}\right) \nu\left(\mathrm{d}z\right).$$

Then the transition kernel of the CBI process $\{P_t : t \geq 0\}$ has the Laplace transform

$$\int_0^\infty \mathrm{e}^{-\lambda z} P_t\left(x, \mathrm{d}z\right) = \exp\left(-x v_t(\lambda) - \int_0^t \psi\left(v_s(\lambda)\right) \mathrm{d}s\right), \quad x,\, \lambda \geq 0, \tag{4.4}$$

where $t \mapsto v_t(\lambda)$ is the unique nonnegative solution of the ODE

$$\frac{\partial}{\partial t} v_t(\lambda) = -\phi\left(v_t\left(\lambda\right)\right), \quad v_0\left(\lambda\right) = \lambda. \tag{4.5}$$

Pinsky Pinsky (1972) announced the existence of a limit distribution for subcritical ($b > 0$) CBI processes under the condition

$$\int_{\{z>1\}} \log(z)\nu\left(\mathrm{d}z\right) < \infty. \tag{4.6}$$

It was shown subsequently in Li (2011, Theorem 3.20 and Corollary 3.21) and Keller-Ressel (2011, Theorem 3.16) that for subcritical CBI processes condition (4.6) is equivalent to the weak convergence of process towards a limiting distribution, which is the unique invariant measure. Properties of this distribution were investigated in Keller-Ressel and Mijatović (2012) while a multidimensional version of Pinsky's result was recently studied in Jin et al. (2020). For the more general class of affine processes on the canonical state space, which includes CBI processes, convergence in the Wasserstein distances W_λ with $\lambda \in (0, 1]$ and $W_{\log}$ was recently shown in Friesen et al. (2020b, Theorem 1.5). Thus, our results from Theorem 4.2 are close to optimal. Below we focus on convergence with respect to the stronger total variation distance.

Based on the approach of Meyn and Tweedie, the exponential ergodicity in the total variation distance was for particular examples studied in Jin et al. (2019, 2016a,b); Masuda (2007). Using the construction of a successful couling, the exponential ergodicity in the total variation distance for a subcritical CBI process (that is $b > 0$) with admissible parameters $(b, \beta, \sigma, m, \nu \equiv 0)$ was recently shown in Li and Ma (2015), provided Grey's condition is satisfied, i.e., (4.a) there exists $\theta > 0$ such that $\phi(\lambda) > 0$ for $\lambda > \theta$ and

$$\int_\theta^\infty \phi(\lambda)^{-1}\mathrm{d}\lambda < \infty.$$

Our next theorem extends the result of Li and Ma (2015, Theorem 2.4 and Theorem 2.5) to CBI processes with non-vanishing jump measure ν for immigration.

Theorem 4.4. *Let $(\beta, b, \sigma, m, \nu)$ be admissible parameters with $b > 0$ and suppose that (4.6) and (4.a) are satisfied. Let $\{P_t : t \geq 0\}$ be the transition kernel given by (4.4) and let π be the unique invariant distribution. Then the following holds:*

(a) *There exists a constant $C > 0$ such that*

$$\|P_t(x,\cdot) - P_t(y,\cdot)\|_{TV} \leq C \min\left\{1, \mathrm{e}^{-bt}|x-y|\right\}.$$

In particular, $\{P_t : t \geq 0\}$ has the strong Feller property.

(b) *There exists a constant $C > 0$ such that, for all $\varrho \in \mathcal{P}_{\log}(\mathbb{R}_{\geq 0})$ and $t \geq 0$, we have*

$$\|P_t^*\varrho - \pi\|_{TV} \leq C \min\{\mathrm{e}^{-bt}, W_{\log}(\varrho,\pi)\} + C\mathrm{e}^{-bt}W_{\log}(\varrho,\pi).$$

Proof: (a) Let $\{P_t^0 : t \geq 0\}$ be the transition kernel with admissible parameters $(\beta, b, \sigma, m, \nu = 0)$, and let $\{P_t^1 : t \geq 0\}$ be the transition kernel with admissible parameters $(\beta = 0, b, \sigma, m, \nu)$. In particular,

$$\int_{\mathbb{R}_{\geq 0}} \mathrm{e}^{-\lambda z} P_t^1(0, \mathrm{d}z) = \exp\left(-\int_0^t \int_0^\infty \left(1 - \mathrm{e}^{-v_s(\lambda)z}\right)\nu(\mathrm{d}z)\,\mathrm{d}s\right), \quad \lambda \geq 0,$$

where v_t is obtained from (4.5). For all $\lambda \geq 0$,

$$\int_{\mathbb{R}_{\geq 0}} \mathrm{e}^{-\lambda z} P_t(x, \mathrm{d}z) = \int_{\mathbb{R}_{\geq 0}} \mathrm{e}^{-\lambda z} P_t^0(x, \mathrm{d}z) \int_{\mathbb{R}_{\geq 0}} \mathrm{e}^{-\lambda z} P_t^1(0, \mathrm{d}z),$$

yielding that $P_t(x,\cdot) = P_t^0(x,\cdot) * P_t^1(0,\cdot)$ holds for all t, $x \geq 0$. Combining the latter with Example 2.1 (a) and Lemma 2.4, we deduce

$$\|P_t(x,\cdot) - P_t(x,\cdot)\|_{TV} \leq \left\|P_t^0(x,\cdot) - P_t^0(x,\cdot)\right\|_{TV} \leq C\min\left\{1, \mathrm{e}^{-bt}|x-y|\right\},$$

where the last inequality follows from Li and Ma (2015, Theorem 2.4) and the proof of Li and Ma (2015, Theorem 2.5).

(b) From Friesen et al. (2020b, Theorem 1.5) we know that the invariant measure satisfies $\pi \in \mathcal{P}_{\log}(\mathbb{R}_{\geq 0})$. Hence for each $\varrho \in \mathcal{P}_{\log}(\mathbb{R}_{\geq 0})$ one finds that $W_{\log}(\varrho,\pi)$ is finite. Let $H \in \mathcal{H}(\varrho,\pi)$ be the optimal coupling of (ϱ,π) such that

$$W_{\log}(\varrho,\pi) = \int_{\mathbb{R}_{\geq 0}\times\mathbb{R}_{\geq 0}} \log(1 + |x-y|)\,H(\mathrm{d}x,\mathrm{d}y). \tag{4.7}$$

Here and below we let $C > 0$ be a generic constant. Using the invariance of π combined with the convexity of the Wasserstein distance (see Lemma 2.3) shows that

$$\begin{aligned}\|P_t^*\varrho - \pi\|_{TV} &\leq \int_{\mathbb{R}_{\geq 0}\times\mathbb{R}_{\geq 0}} \|P_t(x,\cdot) - P_t(y,\cdot)\|_{TV}\,H(\mathrm{d}x,\mathrm{d}y) \\ &\leq C\int_{\mathbb{R}_{\geq 0}\times\mathbb{R}_{\geq 0}} \log\left(1 + \mathrm{e}^{-bt}|x-y|\right)H(\mathrm{d}x,\mathrm{d}y),\end{aligned}$$

where the last inequality following from statement (a) and $1 \wedge a \leq \log(2)^{-1} \log(1+a)$ for $a \geq 0$. Using again (3.7) gives

$$\begin{aligned}\|P_t^* \varrho - \pi\|_{TV} &\leq C \min\left\{ \mathrm{e}^{-bt}, \int_{\mathbb{R}_{\geq 0}\times\mathbb{R}_{\geq 0}} \log\left(1+|x-y|\right) H\left(\mathrm{d}x, \mathrm{d}y\right) \right\} \\ &\quad + C\mathrm{e}^{-bt} \int_{\mathbb{R}_{\geq 0}\times\mathbb{R}_{\geq 0}} \log\left(1+|x-y|\right) H\left(\mathrm{d}x, \mathrm{d}y\right) \\ &= C \min\left\{ \mathrm{e}^{-bt}, W_{\log}(\varrho, \pi) \right\} + C\mathrm{e}^{-bt} W_{\log}(\varrho, \pi),\end{aligned}$$

where the last equality follows from (4.7). □

Assertion (a) can also be derived from the method in Li and Ma (2015, Theorem 2.5) without the use of a convolution argument. However, assertion (b) can not be obtained by the method of Li and Ma (2015), since it relies on the fact that $\int_{\mathbb{R}_{\geq 0}} \log(1+z)\pi(\mathrm{d}z) < \infty$, which was first shown in Friesen et al. (2020b, Theorem 1.5).

4.3. *Q-process of a CBI process conditioned to be never extinct.* In this section we consider a CBI process with *branching mechanism*

$$\phi(\lambda) = b\lambda + \frac{1}{2}\sigma^2\lambda^2 + \int_0^\infty \left(\mathrm{e}^{-\lambda z} - 1 + \lambda z \right) m\left(\mathrm{d}z\right),$$

and *immigration mechanism* $\psi(\lambda) = 0$. Here $b \in \mathbb{R}$, $\sigma \geq 0$, and m is a Borel measure on $\mathbb{R}_{\geq 0}$ satisfying $m(\{0\}) = 0$ and (4.1), i.e., $\int_0^\infty z \wedge z^2 m(\mathrm{d}z) < \infty$. Thus, the admissible parameters are given by $(\beta = 0, b, \sigma, m, \nu = 0)$. In the literature such a process is called a CB process. Note that it has no immigration component. Applying our previous results to this particular case shows that this process has a unique invariant distribution whenever $b > 0$. Moreover, exponential ergodicity holds in the Wasserstein-1 distance. Assuming additionally Grey's condition, one also obtains convergence in the total variation distance. However, the absence of immigration implies that 0 is a trap for the process and hence its invariant distribution is given by δ_0, i.e., the population gets extinct.

As a next step one may study the long-time behavior of the process when conditioned on non-extinction. Let $(X_t^x)_{t\geq 0}$ be the corresponding CB process starting at $x \geq 0$ and define its extinction time by

$$T_x = \inf\{s \geq 0 \ : \ X_s^x = 0\}.$$

Note that $X_t^0 = 0$ for all $t \geq 0$, so that we may assume that $x > 0$. Following Lambert (2007, Theorem 4.1) it was shown that under $\phi'(0) = b > 0$ the conditional laws satisfy

$$\mathbb{P}(\cdot \mid T_x > s) = \mathbb{P}^{\nearrow}(\Theta), \qquad \Theta \in \mathcal{F}_t, \qquad t \geq 0.$$

It was also shown that $\mathbb{P}^{\nearrow}$ uniquely determines a CBI process $(Z_t^x)_{t\geq 0}$ with branching mechanism ϕ as above, and immigration mechanism

$$\psi(\lambda) = \phi'(\lambda) - \phi'(0).$$

Since $\phi'(0) = b$, we find that

$$\psi(\lambda) = \sigma^2\lambda + \int_0^\infty \left(1 - e^{-\lambda z}\right) z m(dz)$$

and hence its admissible parameters are given by $(\beta, b, \sigma, m, \nu)$ with $\beta = \sigma^2$ and $\nu(\mathrm{d}z) = zm(\mathrm{d}z)$. The process $(Z_t^x)_{t\geq 0}$ is called Q-process.

It follows from Lambert (2007, Theorem 4.2) that this process converges a.s. to $+\infty$ when $b = 0$ or $b > 0$ but $\int_1^\infty z\log(z)m(\mathrm{d}z) = +\infty$. Finally, if $b > 0$ and $\int_1^\infty z\log(z)m(\mathrm{d}z) < \infty$, then the Q-process $(Z_t^x)_{t\geq 0}$ converges in distribution as $t \to \infty$ towards a limit distribution π. It can be shown that π is a size-biased Yaglom distribution associated with the CB process $(X_t^x)_{t\geq 0}$. Note that the condition $\int_1^\infty z\log(z)m(\mathrm{d}z) < \infty$ is simply $\int_1^\infty \log(z)\nu(\mathrm{d}z) < \infty$, so that the convergence in law is actually a particular case of the general convergence results discussed in previous section. Below we strengthen Lambert (2007, Theorem 4.2 (ii).b) towards convergence in Wasserstein and total variation distance.

Theorem 4.5. *Let $b > 0$, $\sigma \geq 0$, and m be a Borel measure on $\mathbb{R}_{\geq 0}$ satisfying $m(\{0\}) = 0$ and*

$$\int_0^1 z^2 m(\mathrm{d}z) < \infty, \qquad \int_1^\infty z\log(z)m(\mathrm{d}z) < \infty.$$

Let $(Z_t^x)_{t\geq 0}$ be the coresponding Q-process and let π be its limit distribution. Finally, let p_t denote its transition probabilities. Then there exists a constant $C > 0$ such that

$$W_{log}(p_t(x,\cdot),\pi) \leq Ce^{-bt}\left(1 + \log(1+x) + \int_0^\infty \log(1+z)\pi(dz)\right), \qquad \forall x, t > 0,$$

where the right-hand side is finite. If m satisfies the stronger integrability condition

$$\int_1^\infty z^{1+\lambda} m(\mathrm{d}z) < \infty$$

for some $\lambda \in (0,1)$, then we even have

$$W_\lambda(p_t(x,\cdot),\pi) \leq e^{-b\lambda t}\left(x^\lambda + \int_0^\infty z^\lambda \pi(dz)\right) < \infty. \tag{4.8}$$

Finally, if Grey's condition (4.a) is satisfied, then

$$\|p_t(x,\cdot) - \pi\|_{TV} \leq Ce^{-bt}\left(1 + \log(1+x) + \int_0^\infty \log(1+z)\pi(dz)\right).$$

Proof: First note that $(Z_t^x)_{t\geq 0}$ is a CBI process with admissible parameters $(\beta, b, \sigma, m, \nu)$ with $\beta = \sigma^2$ and $\nu(\mathrm{d}z) = zm(\mathrm{d}z)$. Assume first that $\int_1^\infty z^2 m(\mathrm{d}z) < \infty$ which implies that the process has finite first moment. Hence it satisfies all assumptions of Theorem 3.5 which yields estimate (4.8) for $\lambda = 1$. Moreover, the proof thereof shows that

$$W_1(p_t(x,\cdot), p_t(y,\cdot)) \leq e^{-bt}|x-y|, \qquad x, y \geq 0,\ t \geq 0. \tag{4.9}$$

Assuming now that m satisfies the weaker condition $\int_1^\infty z^{1+\lambda} m(\mathrm{d}z) < \infty$ for some $\lambda \in (0,1)$, we may use a similar convolution trick to the proof of Theorem 4.4, i.e., write $p_t(x,\cdot) = p_t^0(x,\cdot) * p_t(0,\cdot)$, where p_t^0 denote the transition probabilities of a CBI process with admissible parameters $(\beta = 0, b, \sigma, m, \nu = 0)$. Then using (4.9) for p_t^0 yields

$$\begin{aligned} W_\lambda(p_t(x,\cdot), p_t(y,\cdot)) &\leq W_\lambda(p_t^0(x,\cdot), p_t^0(y,\cdot)) \\ &\leq (W_1(p_t^0(x,\cdot), p_t^0(y,\cdot))^\lambda \\ &\leq e^{-b\lambda t}|x-y|^\lambda. \end{aligned}$$

From this and the inequality $W_\lambda(\delta_x, \pi) \leq \int_{\mathbb{R}_{\geq 0}} |x-z|^\lambda \pi(\mathrm{d}z) \leq x^\lambda + \int_{\mathbb{R}_{\geq 0}} z\pi(\mathrm{d}z)$ we readily deduce (4.8). Let us now assume that m satisfies $\int_1^\infty z\log(1+z)m(\mathrm{d}z) < \infty$. Using again the convexity of the Wasserstein distance and the same decomposition $p_t(x,\cdot) = p_t^0(x,\cdot) * p_t(0,\cdot)$, we find that

$$W_{\log}(p_t(x,\cdot), p_t(y,\cdot)) \leq \log(1 + W_1(p_t^0(x,\cdot), p_t^0(y,\cdot))) \leq \log(1 + \mathrm{e}^{-bt}|x-y|).$$

The assertion now follows by similar arguments to the proof of Theorem 4.4 with the Wasserstein distance instead of the TV distance, i.e.,

$$\begin{aligned} W_{log}(p_t(x,\cdot),\pi) &\leq C\min\{\mathrm{e}^{-bt}, W_{\log}(\delta_x,\pi)\} + C\mathrm{e}^{-bt}W_{\log}(\delta_x,\pi) \\ &\leq C\mathrm{e}^{-bt}\left(1+W_{\log}(\delta_x,\pi)\right) \\ &\leq C\mathrm{e}^{-bt}\left(1+\log(1+x)+\int_{\mathbb{R}_{\geq 0}}\log(1+z)\pi(\mathrm{d}z)\right), \end{aligned}$$

where the last inequality follows from $W_{\log}(\delta_x,\pi) \leq \int_{\mathbb{R}_{\geq 0}}\log(1+|x-z|)\pi(\mathrm{d}z) \leq \log(1+x) + \int_{\mathbb{R}_{\geq 0}}\log(1+z)\pi(\mathrm{d}z)$. Finally, convergence in the TV distance is a particular case of Theorem 4.4 combined with $W_{\log}(\delta_x,\pi) \leq \log(1+x)+\int_{\mathbb{R}_{\geq 0}}\log(1+z)\pi(\mathrm{d}z)$. □

4.4. *Continuous-state branching processes with immigration in Lévy random environments.* Following He *et al.* He et al. (2018) and Palau and Pardo Palau and Pardo (2017, 2018), below we introduce a continuous-state branching process with immigration in a Lévy random environment (CBIRE process). Namely, let $\beta, b_E \in \mathbb{R}$, $b \geq 0$, $\sigma, \sigma_E \geq 0$, and m, ν, μ_E be Lévy measures on $\mathbb{R}$ such that

$$m((-\infty,0]) = 0, \quad \nu((-\infty,0]) = 0, \quad \int_{\mathbb{R}_{\geq 0}} 1\wedge z\nu(\mathrm{d}z) < \infty.$$

Let $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq 0},\mathbb{P})$ be a stochastic basis satisfying the usual conditions rich enough to support

- an $(\mathcal{F}_t)_{t\geq 0}$-Gaussian white noise $W(\mathrm{d}t,\mathrm{d}u)$ with intensity $\mathrm{d}t\mathrm{d}u$;
- an $(\mathcal{F}_t)_{t\geq 0}$-Poisson random measure $N_0(\mathrm{d}t,\mathrm{d}z,\mathrm{d}u)$ on $\mathbb{R}_{\geq 0}\times\mathbb{R}_{\geq 0}$ with intensity $\mathrm{d}tm(\mathrm{d}z)\mathrm{d}u$;
- an $(\mathcal{F}_t)_{t\geq 0}$-Poisson random measure $N_1(\mathrm{d}t,\mathrm{d}z)$ on $\mathbb{R}_{\geq 0}\times\mathbb{R}_{\geq 0}$ with intensity $\mathrm{d}t\nu(\mathrm{d}z)$;
- an $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion $\{B_t : t\geq 0\}$;
- an $(\mathcal{F}_t)_{t\geq 0}$-Poisson random measure $M(\mathrm{d}t,\mathrm{d}z)$ on $\mathbb{R}_{\geq 0}\times\mathbb{R}$ with intensity $\mathrm{d}t\mu_E(\mathrm{d}z)$.

Suppose that these random objects are mutually independent. Define an $(\mathcal{F}_t)_{t\geq 0}$-Lévy process $\{Z_t : t\geq 0\}$ by

$$Z_t = b_E t + \sigma_E B_t + \int_0^t\int_{[-1,1]}(\mathrm{e}^z-1)\widetilde{M}(\mathrm{d}s,\mathrm{d}z) + \int_0^t\int_{[-1,1]^c}(\mathrm{e}^z-1)M(\mathrm{d}s,\mathrm{d}z), \tag{4.10}$$

where $[-1,1]^c = \mathbb{R}\backslash[-1,1]$, $\widetilde{M}(\mathrm{d}s,\mathrm{d}z) := M(\mathrm{d}s,\mathrm{d}z) - \mathrm{d}s\mu_E(\mathrm{d}z)$. Note that $\{Z_t : t\geq 0\}$ has no jump less than -1. Then it was shown in He et al. (2018, Theorem 5.1) and Palau and Pardo (2018, Theorem 1) that

$$\begin{aligned} X_t = X_0 &+ \int_0^t(\beta - bX_s)\,\mathrm{d}s + \sigma\int_0^t\int_0^\infty \mathbb{1}_{\{u\leq X_s\}}W(\mathrm{d}s,\mathrm{d}u) \\ &+ \int_0^t\int_0^\infty\int_0^\infty z\mathbb{1}_{\{u\leq X_{s-}\}}\widetilde{N}_0(\mathrm{d}s,\mathrm{d}z,\mathrm{d}u) + \int_0^t\int_0^\infty zN_1(\mathrm{d}s,\mathrm{d}z) + \int_0^t X_{s-}\mathrm{d}Z_s \end{aligned} \tag{4.11}$$

has for each $\mathcal{F}_0$-measurable random variable $X_0\geq 0$ a pathwise unique nonnegative strong solution $\{X_t : t\geq 0\}$. We call the corresponding Markov process CBIRE process.

In view of He et al. (2018, Theorem 5.4), the Markov process $\{X_t : t\geq 0\}$ has Feller transition semigroup $\{P_t : t\geq 0\}$ and its transition kernels $P_t(x,\mathrm{d}y)$ satisfy

$$\int_{\mathbb{R}_{\geq 0}}\mathrm{e}^{-\lambda y}P_t(x,\mathrm{d}y) = \mathbb{E}\left[\exp\left(-xv_{0,t}^{\xi}(\lambda) - \int_0^t\psi\left(v_{s,t}^{\xi}(\lambda)\right)\mathrm{d}s\right)\right]. \tag{4.12}$$

Here ϕ and ψ are the corresponding branching and immigration mechanisms for the CBI process with admissible parameters $(\beta, b, \sigma, m, \nu)$ and $r \mapsto v^{\xi}_{r,t}(\lambda)$ is the pathwise unique nonnegative solution to

$$v^{\xi}_{r,t}(\lambda) = \mathrm{e}^{\xi(t)-\xi(r)}\lambda - \int_r^t \mathrm{e}^{\xi(s)-\xi(r)}\phi\left(v^{\xi}_{s,t}(\lambda)\right)\mathrm{d}s, \quad 0 \le r \le t,$$

where $\{\xi_t : t \geq 0\}$ is an $(\mathcal{F}_t)_{t\geq 0}$-Lévy process defined by

$$\xi_t = a_E t + \sigma_E B_t + \int_0^t \int_{[-1,1]} z\widetilde{M}(\mathrm{d}s, \mathrm{d}z) + \int_0^t \int_{[-1,1]^c} zM(\mathrm{d}s, \mathrm{d}z),$$

with the drift coefficients b_E and a_E being related by

$$b_E = a_E + \frac{\sigma_E^2}{2} + \int_{[-1,1]} (\mathrm{e}^z - 1 - z)\mu_E(\mathrm{d}z).$$

The next remark provides an alternative representation for (4.11) on which our results are based.

Remark 4.6. In view of (4.10), equation (4.11) is equivalent to

$$\begin{aligned}X_t = X_0 &+ \int_0^t (\beta + (b_E - b)X_s)\,\mathrm{d}s + \int_0^t \int_0^\infty zN_1(\mathrm{d}s, \mathrm{d}z) + \int_0^t \int_{[-1,1]^c} X_{s-}(e^z - 1)M(\mathrm{d}s, \mathrm{d}z)\\ &+ \sigma \int_0^t \int_0^\infty \mathbb{1}_{\{u \le X_s\}} W(\mathrm{d}s, \mathrm{d}u) + \sigma_E \int_0^t X_s \mathrm{d}B_s\\ &+ \int_0^t \int_0^\infty \int_0^\infty z\mathbb{1}_{\{u\le X_{s-}\}}\widetilde{N}_0(\mathrm{d}s, \mathrm{d}z, \mathrm{d}u) + \int_0^t \int_{[-1,1]} X_{s-}(\mathrm{e}^z - 1)\widetilde{M}(\mathrm{d}s, \mathrm{d}z).\end{aligned}$$

So far, the comparison principle has not been obtained for the CBIRE process. Below we show that (4.11) is a special case of (1.1) which satisfies the conditions of Theorem 3.3 and hence also satisfies the comparison principle.

Proposition 4.7. *Assume that*

$$\int_{(1,\infty)} z\nu(\mathrm{d}z) + \int_{(1,\infty)} zm(\mathrm{d}z) + \int_{(1,\infty)} \mathrm{e}^z\mu_E(\mathrm{d}z) < \infty. \tag{4.13}$$

Then (4.11) *has the comparison property.*

Proof: Let us first show that (4.11) is a particular case of (1.1). The drift coefficient is simply given by $b(x) = b_1(x) - b_2(x)$ with $b_1(x) = \beta - (b - b_E)x$ and $b_2(x) = 0$ for $x \in \mathbb{R}_{\geq 0}$. For the diffusion component set $E = \{1, 2\} \times \mathbb{R}_{\geq 0}$ and $\varkappa(\mathrm{d}k, \mathrm{d}u) = \delta_1(\mathrm{d}k)\mathrm{d}u + \delta_2(\mathrm{d}k)\delta_0(\mathrm{d}u)$ on E. Then $\mathcal{W}(\mathrm{d}s, \mathrm{d}k, \mathrm{d}u) := \delta_1(\mathrm{d}k)W(\mathrm{d}s, \mathrm{d}u) + \mathrm{d}B_s\delta_2(\mathrm{d}k)\delta_0(\mathrm{d}u)$ defines an $(\mathcal{F}_t)_{t\geq 0}$-Gaussian white noise on $\mathbb{R}_{\geq 0} \times E$ with intensity measure $\mathrm{d}s\varkappa(\mathrm{d}k, \mathrm{d}u)$. Let $\sigma(x, k, u) := \sigma\mathbb{1}_{\{k=1\}}\mathbb{1}_{\{u\le x\}} + \sigma_E\mathbb{1}_{\{k=2\}}\mathbb{1}_{\{0\}}(u)x$ for $(x, k, u) \in \mathbb{R}_{\geq 0} \times E$. We see that

$$\int_0^t \int_E \sigma(X_s, k, u)\,\mathcal{W}(\mathrm{d}s, \mathrm{d}k, \mathrm{d}u) = \sigma \int_0^t \int_{\mathbb{R}_{\geq 0}} \mathbb{1}_{\{u\le X_s\}}W(\mathrm{d}s, \mathrm{d}u) + \sigma_E \int_0^t X_s\mathrm{d}B_s.$$

Turning to the jump components, define $U_0 = \{1, 2\} \times \mathbb{R} \times \mathbb{R}_{\geq 0}$ and further $\mu_0(\mathrm{d}k, \mathrm{d}z, \mathrm{d}u) = \mathbb{1}_{\mathbb{R}_{\geq 0}}(z)\delta_1(\mathrm{d}k)m(\mathrm{d}z)\mathrm{d}u + \delta_2(\mathrm{d}k)\mu_E(\mathrm{d}z)\delta_0(\mathrm{d}u)$ on U_0. Then we have that $\mathcal{N}_0(\mathrm{d}s, \mathrm{d}k, \mathrm{d}z, \mathrm{d}u) := \mathbb{1}_{\mathbb{R}_{\geq 0}}(z)\delta_1(\mathrm{d}k)N_0(\mathrm{d}s, \mathrm{d}z, \mathrm{d}u) + \delta_2(\mathrm{d}k)M(\mathrm{d}s, \mathrm{d}z)\delta_0(\mathrm{d}u)$ defines an $(\mathcal{F}_t)_{t\geq 0}$-Poisson random measure on $\mathbb{R}_{\geq 0} \times U_0$ with intensity $\mathrm{d}s\mu_0(\mathrm{d}k, \mathrm{d}z, \mathrm{d}u)$. Letting $g_0(x, k, z, u) = \mathbb{1}_{\{k=1\}}\mathbb{1}_{\{u\le x\}}\mathbb{1}_{\mathbb{R}_{\geq 0}}(z)z +$

$\mathbb{1}_{\{k=2\}}\mathbb{1}_{[-1,1]}(z)(\exp(z)-1)x$ for $(x,k,z,u)\in\mathbb{R}_{\geq 0}\times U_0$ yields

$$\int_0^t\int_{U_0} g_0(X_{s-},k,z,u)\widetilde{\mathcal{N}}_0(\mathrm{d}s,\mathrm{d}k,\mathrm{d}z,\mathrm{d}u) = \int_0^t\int_0^\infty\int_0^\infty z\mathbb{1}_{\{u\leq X_{s-}\}}\widetilde{N}_0(\mathrm{d}s,\mathrm{d}z,\mathrm{d}u) + \int_0^t\int_{-1}^1 \left(\mathrm{e}^z-1\right)X_{s-}\widetilde{M}(\mathrm{d}s,\mathrm{d}z)$$

Finally, let $U_1=\{1,2\}\times\mathbb{R}$ and define $\mu_1(\mathrm{d}k,\mathrm{d}z)=\mathbb{1}_{\mathbb{R}_{\geq 0}}(z)\delta_1(\mathrm{d}k)\nu(\mathrm{d}z)+\delta_2(\mathrm{d}k)\mu_E(\mathrm{d}z)$ on U_1. Then $\mathcal{N}_1(\mathrm{d}s,\mathrm{d}k,\mathrm{d}z)=\delta_1(\mathrm{d}k)N_1(\mathrm{d}s,\mathrm{d}z)+\delta_2(\mathrm{d}k)M(\mathrm{d}s,\mathrm{d}z)$ defines an $(\mathcal{F}_t)_{t\geq 0}$-Poisson random measure on $\mathbb{R}_{\geq 0}\times U_1$ with intensity $\mathrm{d}s\mu_1(\mathrm{d}k,\mathrm{d}z)$. Letting $g_1(x,k,z)=\mathbb{1}_{\{k=1\}}\mathbb{1}_{\mathbb{R}_{\geq 0}}(z)z+\mathbb{1}_{\{k=2\}}\mathbb{1}_{[-1,1]^c}(z)(\exp(z)-1)x$ for $(x,k,z)\in\mathbb{R}_{\geq 0}\times U_1$ yields

$$\int_0^t\int_{U_1} g_1(X_{s-},k,z)\mathcal{N}_1(\mathrm{d}s,\mathrm{d}k,\mathrm{d}z) = \int_0^t\int_0^\infty zN_1(\mathrm{d}s,\mathrm{d}z)+\int_0^t\int_{[-1,1]^c}\left(\mathrm{e}^z-1\right)X_{s-}M(\mathrm{d}s,\mathrm{d}z).$$

This shows that (4.11) is indeed a particular case of (1.1).

Next we verify conditions (3.a) – (3.e) and show that $(m)_1$ holds. Writing $(k,z)=u$, condition (3.a) follows from

$$\sup_{y\in[0,x]}|g_1(y,u)|\leq \mathbb{1}_{\{k=1\}}\mathbb{1}_{\mathbb{R}_{\geq 0}}(z)z+\mathbb{1}_{\{k=2\}}\mathbb{1}_{[-1,1]^c}(z)|\exp(z)-1|x =: \overline{g}_1(x,u)$$

and

$$\begin{aligned} &b(x)+\int_{U_1}|\overline{g}_1(x,u)|\mu_1(\mathrm{d}u)\\ &=\beta-(b-b_E)x+\int_{\mathbb{R}_{\geq 0}} z\nu(\mathrm{d}z)+\int_{[-1,1]^c}|\exp(z)-1|x\mu_E(\mathrm{d}z)\\ &\leq\left(\beta+\int_{\mathbb{R}_{\geq 0}} z\nu(\mathrm{d}z)\right)+x\left((b+|b_E|)+\int_{[-1,1]^c}|\exp(z)-1|\mu_E(\mathrm{d}z)\right), \end{aligned}$$

where both integrals are finite due to (4.13). Condition (3.b) follows directly from

$$\begin{aligned} &|b_1(x)-b_1(y)|+\int_{U_1}|g_1(x,u)-g_1(y,u)|\mu_1(\mathrm{d}u)\\ &\leq |b-b_E||x-y|+|x-y|\int_{[-1,1]^c}|\exp(z)-1|\mu_E(\mathrm{d}z) =: r_j(|x-y|). \end{aligned}$$

Condition (3.c) follows from

$$\int_E\sigma(x,u)^2\varkappa(\mathrm{d}u)=\sigma^2\int_{\mathbb{R}_{\geq 0}}\mathbb{1}_{\{u\leq x\}}\mathrm{d}u+\sigma_E^2x^2\leq(\sigma^2+\sigma_E^2)(2+x)^2$$

and for $j\geq 1$ with $0\leq x,y\leq j$

$$\begin{aligned}\int_E|\sigma(x,u)-\sigma(y,u)|^2\varkappa(\mathrm{d}u) &\leq \sigma^2\int_{\mathbb{R}_{\geq 0}}\left|\mathbb{1}_{\{u\leq x\}}-\mathbb{1}_{\{u\leq y\}}\right|\mathrm{d}u+\sigma_E^2|x-y|^2\\ &\leq\left(\sigma^2+2j\sigma_E^2\right)|x-y|.\end{aligned}$$

For condition (3.d) we observe that for $0\leq y\leq x$,

$$|g_0(y,k,z,u)|\leq\mathbb{1}_{\{k=1\}}\mathbb{1}_{\{u\leq x\}}\mathbb{1}_{\mathbb{R}_{\geq 0}}(z)z+\mathbb{1}_{\{k=2\}}\mathbb{1}_{[-1,1]}(z)|\exp(z)-1|x =: \overline{g}_0(x,k,z,u).$$

One has

$$
\begin{aligned}
\int_{U_0} \left(\overline{g}_0(x,u) \wedge \overline{g}_0(x,u)^2\right) \mu_0(\mathrm{d}u) &= \int_{\mathbb{R}^2_{\geq 0}} \mathbb{1}_{\{u \leq x\}} \left(z \wedge z^2\right) m(\mathrm{d}z)\mathrm{d}u \\
&\quad + \int_{[-1,1]} \min\{|\exp(z)-1|x, |\exp(z)-1|^2 x^2\} \mu_E(\mathrm{d}z) \\
&\leq (2+x)^2 \left(\int_{\mathbb{R}_{\geq 0}} z \wedge z^2 m(\mathrm{d}z) + \int_{[-1,1]} |\exp(z)-1|^2 \mu_E(\mathrm{d}z) \right).
\end{aligned}
$$

Moreover, it holds that

$$
\begin{aligned}
\int_{U_0} \mathbb{1}_{\{|g_0(x,u)|\leq 1\}} |g_0(x,u)|^2 \mu_0(\mathrm{d}u) &\leq x \int_{(0,1]} z^2 m(\mathrm{d}z) + x^2 \int_{[-1,1]} |\exp(z)-1|^2 \mu_E(\mathrm{d}z) \\
&\leq \left(\int_{(0,1]} z^2 m(\mathrm{d}z) + \int_{[-1,1]} |\exp(z)-1|^2 \mu_E(\mathrm{d}z) \right) (2+x)^2
\end{aligned}
$$

and for each $j \geq 1$ with $0 \leq x, y \leq j$

$$
\begin{aligned}
&\int_{U_0} \left(|g_0(x,u) - g_0(y,u)| \wedge |g_0(x,u) - g_0(y,u)|^2\right) \mu_0(\mathrm{d}u) \\
&\leq |x-y| \int_{\mathbb{R}_{\geq 0}} \left(z \wedge z^2\right) m(\mathrm{d}z) + \int_{[-1,1]} \left[(|x-y||\exp(z)-1|) \wedge (|x-y|^2 |\exp(z)-1|^2) \right] \mu_E(\mathrm{d}z) \\
&\leq |x-y| \int_{\mathbb{R}_{\geq 0}} \left(z \wedge z^2\right) m(\mathrm{d}z) + (1+2j)|x-y| \int_{[-1,1]} |\exp(z)-1|^2 \mu_E(\mathrm{d}z).
\end{aligned}
$$

This proves (3.d). Condition (3.e) follows from

$$
\begin{aligned}
x + g_0(x,k,z,u) &= \mathbb{1}_{\{k=1\}} \mathbb{1}_{\{u \leq x\}} \mathbb{1}_{\mathbb{R}_{\geq 0}}(z) z + \mathbb{1}_{\{k=2\}} \mathbb{1}_{[-1,1]}(z)(\exp(z)-1)x + x \\
&= \mathbb{1}_{\{k=1\}} \mathbb{1}_{\{u \leq x\}} \mathbb{1}_{\mathbb{R}_{\geq 0}}(z) z + \mathbb{1}_{\{k=2\}} \mathbb{1}_{[-1,1]}(z) \exp(z) x + \left(1 - \mathbb{1}_{\{k=2\}} \mathbb{1}_{[-1,1]}(z)\right) x
\end{aligned}
$$

and likewise

$$
x + g_1(x,k,z) = \mathbb{1}_{\{k=1\}} \mathbb{1}_{\mathbb{R}_{\geq 0}}(z) z + \mathbb{1}_{\{k=2\}} \mathbb{1}_{[-1,1]^c}(z) \exp(z) x + \left(1 - \mathbb{1}_{\{k=2\}} \mathbb{1}_{[-1,1]^c}(z)\right) x.
$$

Finally, the moment condition $(m)_1$ follows from

$$
\begin{aligned}
&\int_{U_1} \mathbb{1}_{\{|g_1(x,u)|>1\}} |g_1(x,u)| \mu_1(\mathrm{d}u) \\
&\leq \int_{\mathbb{R}_{\geq 0}} z \nu(\mathrm{d}z) + x \int_{[-1,1]^c} |\exp(z)-1| \mu_E(\mathrm{d}z)
\end{aligned}
$$

and

$$
\begin{aligned}
&\int_{U_0} \mathbb{1}_{\{|g_0(x,u)|>1\}} |g_0(x,u)| \mu_0(\mathrm{d}u) \\
&\leq x \int_{\{z>1\}} z m(\mathrm{d}z) + \int_{[-1,1]} \mathbb{1}_{\{x|\exp(z)-1|>1\}} x |\exp(z)-1| \mu_E(\mathrm{d}z) \\
&\leq x \int_{\{z>1\}} z m(\mathrm{d}z) + x \int_{[-1,1]} |\exp(z)-1| \mu_E(\mathrm{d}z).
\end{aligned}
$$

$\square$

Existence of limiting distributions (and hence invariant measures) was recently characterized in He et al. (2018). In the following we present sufficient conditions for both the ergodicity in the

Wasserstein and total variation distance based on an application of Theorem 3.5. Ergodicity in the Wasserstein distance W_1 is stated below.

Theorem 4.8. *Let $\{P_t : t \geq 0\}$ be the transition semigroup with transition kernels defined by (4.12) and denote by $\{P_t^* : t \geq 0\}$ the dual semigroup. Suppose that (4.13) holds and that*

$$b > \mathbb{E}[Z_1]. \tag{4.14}$$

Then, for all ϱ, $\widetilde{\varrho} \in \mathcal{P}_1(\mathbb{R}_{\geq 0})$, we have

$$W_1\left(P_t^*\varrho, P_t^*\widetilde{\varrho}\right) \leq \mathrm{e}^{-(b-\mathbb{E}[Z_1])t} W_1\left(\varrho, \widetilde{\varrho}\right), \quad t \geq 0.$$

In particular, there exists a unique invariant distribution $\pi \in \mathcal{P}(\mathbb{R}_{\geq 0})$. Moreover, π belongs to $\mathcal{P}_1(\mathbb{R}_{\geq 0})$ and, for all $\varrho \in \mathcal{P}_1(\mathbb{R}_{\geq 0})$,

$$W_1\left(P_t^*\varrho, \pi\right) \leq \mathrm{e}^{-(b-\mathbb{E}[Z_1])t} W_1\left(\varrho, \pi\right), \quad t \geq 0.$$

Proof: It remains to verify condition (3.f). First observe that

$$\begin{aligned}\widetilde{b}(x) &:= b(x) + \int_{U_1} g_1(x,u)\mu_1(\mathrm{d}u) \\ &= \beta + \int_{\mathbb{R}_{\geq 0}} z\nu(\mathrm{d}z) + \left(\int_{[-1,1]^c} (\exp(z)-1)\mu_E(\mathrm{d}z) - (b-b_E)\right) x\end{aligned}$$

and hence for $0 \leq x \leq y$

$$\widetilde{b}(y) - \widetilde{b}(x) = -\left(b - b_E - \int_{[-1,1]^c} (\exp(z)-1)\mu_E(\mathrm{d}z)\right)(y-x).$$

Using $\mathbb{E}[Z_1] = b_E + \int_{[-1,1]^c}(\exp(z)-1)\mu_E(\mathrm{d}z)$ we conclude condition (3.f) from (4.14). □

Theorem 4.9. *Let $\{P_t : t \geq 0\}$ be the transition semigroup with transition kernels defined by (4.12) and denote by $\{P_t^* : t \geq 0\}$ the dual semigroup. Suppose that (4.14) and Grey's condition (4.a) is satisfied. Let π be the unique invariant distribution. Then, for any $\varrho \in \mathcal{P}_1(\mathbb{R}_{\geq 0})$,*

$$\|P_t^*\varrho - \pi\|_{TV} \leq 2\mathbb{E}\left[\overline{v}_{0,t}^{\xi}\right] W_1(\varrho, \pi), \quad t \geq 0, \tag{4.15}$$

where $\overline{v}_{0,t}^{\xi} := \lim_{\lambda\to\infty} v_{0,t}^{\xi}(\lambda) \in [0,\infty)$. If, in addition, $\liminf_{t\to\infty} \xi(t) = -\infty$ almost surely, then

$$\lim_{t\to\infty} \|P_t^*\varrho - \pi\|_{TV} = 0. \tag{4.16}$$

Proof: As a consequence of Grey's condition, He et al. (2018, Theorem 4.1) applies, yielding that $\overline{v}_{0,t}^{\xi} \in [0,\infty)$ almost surely for all $t > 0$. Let $f \in \mathcal{B}_b(\mathbb{R}_{\geq 0})$ be arbitrary. Arguing as in the proof of He et al. (2018, Theorem 4.5), we observe that, for $0 \leq x \leq y$,

$$|P_t f(x) - P_t f(y)| \leq 2\|f\|_\infty \mathbb{E}\left[1 - \mathrm{e}^{-(y-x)\overline{v}_{0,t}^{\xi}}\right] \leq 2\|f\|_\infty \mathbb{E}\left[\min\left\{1, |x-y|\overline{v}_{0,t}^{\xi}\right\}\right].$$

Taking the supremum over all $f \in \mathcal{B}_b(\mathbb{R}_{\geq 0})$ with $\|f\|_\infty \leq 1$ shows that $\|P_t(x,\cdot) - P_t(y,\cdot)\|_{TV} \leq 2\mathbb{E}[\min\{1, |x-y|\overline{v}_{0,t}^{\xi}\}]$. Let now $\varrho \in \mathcal{P}_1(\mathbb{R}_{\geq 0})$ and let H be any coupling of (ϱ, π). By convexity of the Wasserstein distance, we obtain

$$\begin{aligned}\|P_t^*\varrho - \pi\|_{TV} &\leq \int_{\mathbb{R}_{\geq 0}\times\mathbb{R}_{\geq 0}} \|P_t(x,\cdot) - P_t(y,\cdot)\|_{TV}\, H(\mathrm{d}x, \mathrm{d}y) \\ &\leq 2\int_{\mathbb{R}_{\geq 0}\times\mathbb{R}_{\geq 0}} \mathbb{E}\left[\min\left\{1, |x-y|\overline{v}_{0,t}^{\xi}\right\}\right] H(\mathrm{d}x, \mathrm{d}y).\end{aligned}$$

If $\liminf_{t\to\infty} \xi(t) = -\infty$, then $\lim_{t\to\infty} \overline{v}^{\xi}_{0,t} = 0$ in view of He et al. (2018, Corollary 4.4) and, thus, (4.16) follows from dominated convergence. Finally, we conclude the estimate (4.15) by estimating

$$\int_{\mathbb{R}_{\geq 0}\times\mathbb{R}_{\geq 0}} \mathbb{E}\left[\min\left\{1, |x-y|\overline{v}^{\xi}_{0,t}\right\}\right] H(\mathrm{d}x,\mathrm{d}y) \leq \mathbb{E}\left[\overline{v}^{\xi}_{0,t}\right] \int_{\mathbb{R}_{\geq 0}\times\mathbb{R}_{\geq 0}} |x-y| H(\mathrm{d}x,\mathrm{d}y),$$

where we choose H as the optimal coupling of (ϱ, π) with respect to W_1. □

The decay rate for $\overline{v}^{\xi}_{0,t}$ as $t \to \infty$ was studied by Palau and Pardo Palau and Pardo (2017) for a continuous-state branching process in Brownian random environment with stable branching. For the same class of processes but in a general Lévy environment this problem was studied by Li and Xu Li and Xu (2018).

Acknowledgements

The author would like to thank an anonymous referee for pointing out several mistakes in a preliminary version.

Appendix

Lemma A.1. *Let $\lambda > 0$ and set $V_\lambda(x) = (2+x)^\lambda$. Then the following assertions hold.*

(a) *There exists a constant $C_\lambda > 0$ such that for $z \in \mathbb{R}$ and $x \geq 0$ with $x+z \geq 0$ it holds that*

$$\begin{aligned}|V_\lambda(x+z) - V_\lambda(x)| &\leq C_\lambda(2+x)^{\lambda-1}\mathbb{1}_{\{|z|\leq 1\}}|z| \\ &\quad + C_\lambda\left(\mathbb{1}_{(0,1]}(\lambda) + \mathbb{1}_{(1,\infty)}(\lambda)(2+x)^{\lambda-1}\right)\mathbb{1}_{\{|z|>1\}}|z|^\lambda.\end{aligned}$$

(b) *There exists another constant $\widetilde{C}_\lambda > 0$ such that for $z \in \mathbb{R}$ and $x \geq 0$ with $x+z \geq 0$ it holds that*

$$\begin{aligned}&\left|V_\lambda(x+z) - V_\lambda(x) - zV_\lambda'(x)\right| \\ &\leq \widetilde{C}_\lambda(2+x)^{\lambda-2}\mathbb{1}_{\{|z|\leq 1\}}|z|^2 \\ &\quad + \widetilde{C}_\lambda\left(\mathbb{1}_{(0,1]}(\lambda) + \mathbb{1}_{(1,2)}(\lambda)(2+x)^{\lambda-1} + \mathbb{1}_{[2,\infty)}(\lambda)(2+x)^{\lambda-2}\right)\mathbb{1}_{\{|z|>1\}}|z|^{\max\{1,\lambda\}}.\end{aligned}$$

Proof: (a) Using the mean-value theorem we find some ξ lying between 0 and z such that $V_\lambda(x+z) - V_\lambda(x) = \lambda z(2+x+\xi)^{\lambda-1}$. Consider first the case where $\lambda > 1$. For $|z| \leq 1$ we use $|(2+x+\xi)^{\lambda-1}| \leq (2+x+|z|)^{\lambda-1} \leq (3+x)^{\lambda-1} \leq 2^{\lambda-1}(2+x)^{\lambda-1}$ to find that

$$|V_\lambda(x+z) - V_\lambda(x)| \leq \lambda 2^{\lambda-1}(2+x)^{\lambda-1}|z|.$$

For $|z| > 1$ we use the inequality $2+a+b \leq (2+a)(2+b)$ for $a, b \geq 0$ to find that

$$\begin{aligned}|V_\lambda(x+z) - V_\lambda(x)| &= \lambda|z|(2+x+\xi)^{\lambda-1} \\ &\leq \lambda|z|(2+x+|z|)^{\lambda-1} \\ &\leq \lambda|z|(2+x)^{\lambda-1}(2+|z|)^{\lambda-1} \\ &\leq 3^{\lambda-1}\lambda(2+x)^{\lambda-1}|z|^\lambda.\end{aligned}$$

This proves the assertion for $\lambda > 1$. Now assume that $\lambda \in (0,1]$. If $|z| \leq 1$, then $2+\xi \geq 2-|z| \geq 1$ and hence $2+x+\xi \geq 1+x$. This yields

$$\begin{aligned}|V_\lambda(x+z) - V_\lambda(x)| &= \lambda|z|(2+x+\xi)^{\lambda-1}\\ &\leq \lambda(1+x)^{\lambda-1}|z|\\ &= 2^{1-\lambda}\lambda(2+2x)^{\lambda-1}|z|\\ &\leq 2^{1-\lambda}\lambda(2+x)^{\lambda-1}|z|.\end{aligned}$$

Finally, if $|z| > 1$, then we find that

$$\begin{aligned}|V_\lambda(x+z) - V_\lambda(x)| &= \left|(2+x+z)^\lambda - (2+x)^\lambda\right|\\ &\leq |(2+x+z)-(2+x)|^\lambda\\ &= |z|^\lambda.\end{aligned}$$

Combining all estimates proves the assertion.

(b) To prove the second inequality, we use the mean-value theorem twice to find that

$$\begin{aligned}V_\lambda(x+z) - V_\lambda(x) - zV_\lambda'(x) &= z^2 \int_0^1 \int_0^1 rV_\lambda''(x+rsz)\mathrm{d}r\mathrm{d}s\\ &= z^2\lambda(\lambda-1)\int_0^1\int_0^1 r(2+x+rsz)^{\lambda-2}\mathrm{d}r\mathrm{d}s,\end{aligned}$$

where the integrand is non-negative and well-defined since $2+x+rsz \geq 2$ holds for all $z \in \mathbb{R}$ and $x \geq 0$ satisfying $x+z \geq 0$. Suppose first that $|z| \leq 1$. If $\lambda \in (0,2]$, then we use that $2+x+rsz \geq 2+x-|z| \geq 1+x$ to find that

$$\begin{aligned}\left|V_\lambda(x+z) - V_\lambda(x) - zV_\lambda'(x)\right| &\leq \lambda|\lambda-1||z|^2(1+x)^{\lambda-2}\\ &\leq 2^{2-\lambda}\lambda|\lambda-1||z|^2(2+x)^{\lambda-2}.\end{aligned}$$

If $\lambda > 2$, then we use that $2+x+rsz \leq 3+x$ to find that

$$\begin{aligned}\left|V_\lambda(x+z) - V_\lambda(x) - zV_\lambda'(x)\right| &\leq \lambda(\lambda-1)(3+x)^{\lambda-2}|z|^2\\ &\leq \lambda(\lambda-1)2^{\lambda-2}(2+x)^{\lambda-2}|z|^2.\end{aligned}$$

The last two estimates prove the assertion for the case $|z| \leq 1$. Now suppose that $|z| > 1$. For $\lambda \geq 2$ we find that

$$(2+x+rsz)^{\lambda-2} \leq (2+x+|z|)^{\lambda-2} \leq (2+x)^{\lambda-2}(2+|z|)^{\lambda-2},$$

which gives

$$\begin{aligned}\left|V_\lambda(x+z) - V_\lambda(x) - zV_\lambda'(x)\right| &\leq \lambda(\lambda-1)(2+x)^{\lambda-2}|z|^2(2+|z|)^{\lambda-2}\\ &\leq \lambda(\lambda-1)3^{\lambda-2}(2+x)^{\lambda-2}|z|^\lambda.\end{aligned}$$

Finally, if $\lambda \in (1,2)$, then we use part (a) to find that

$$\begin{aligned}\left|V_\lambda(x+z) - V_\lambda(x) - zV_\lambda'(x)\right| &\leq |V_\lambda(x+z) - V_\lambda(x)| + |z|V_\lambda'(x)\\ &\leq C_\lambda(2+x)^{\lambda-1}|z|^\lambda + \lambda|z|(2+x)^{\lambda-1}\\ &\leq (C_\lambda+\lambda)(2+x)^{\lambda-1}|z|^\lambda,\end{aligned}$$

while for $\lambda \in (0,1]$ we obtain

$$\begin{aligned}\left|V_\lambda(x+z) - V_\lambda(x) - zV_\lambda'(x)\right| &\le |V_\lambda(x+z) - V_\lambda(x)| + |z|V_\lambda'(x)\\ &\le C_\lambda |z|^\lambda + \lambda |z|(2+x)^{\lambda-1}\\ &\le (C_\lambda + \lambda)\,|z|^{\max\{1,\lambda\}}.\end{aligned}$$

This completes the proof of part (b).

□

Lemma A.2. *For $x \ge 0$ set $V(x) = \log(2+x)$. Then there exists a constant $C_{\log} > 0$ such that for $z \in \mathbb{R}$ and $x \ge 0$ with $x+z \ge 0$ it holds that*

$$|V(x+z) - V(x)| \le \mathbb{1}_{\{|z|\le 1\}} \frac{|z|}{1+x} + C_{\log} \mathbb{1}_{\{|z|>1\}} \frac{1+\log(1+|z|)}{2+x}$$

and

$$\left|V(x+z) - V(x) - zV'(x)\right| \le \mathbb{1}_{\{|z|\le 1\}} \frac{|z|^2}{(1+x)^2} + C_{\log} \mathbb{1}_{\{|z|>1\}} \frac{|z|}{2+x}.$$

Proof: For the first inequality we first consider the case $|z| \le 1$. Using the mean-value theorem we find some ξ lying between 0 and z such that

$$|V(x+z) - V(x)| = \left|\frac{z}{2+x+\xi}\right| = \frac{|z|}{2+x+\xi} \le \frac{|z|}{1+x},$$

where we have used that $2+x+\xi \ge 1+x$. For $|z| > 1$ we use the monotonicity of the logarithm and then for a, $d \ge 0$ the inequality $\log(1 + a\cdot d) \le C\min\{a, \log(1+d)\} + Ca\log(1+d)$, where $C > 0$ is a constant (see Friesen et al. (2020b, Lemma A.5)), to find that

$$\begin{aligned}|V(x+z) - V(x)| &= \log\left(1 + \frac{|z|}{2+x}\right)\\ &\le C\min\left\{\frac{1}{2+x}, \log(1+|z|)\right\} + \frac{C}{2+x}\log(1+|z|)\\ &\le \frac{C}{2+x} + C\frac{\log(1+|z|)}{2+x}.\end{aligned}$$

Combining both estimates proves the first inequality. For the second inequality we observe that

$$\begin{aligned}V(x+z) - V(x) - zV'(x) &= z^2 \int_0^1\int_0^1 rV_\lambda''(x+rsz)\mathrm{d}r\mathrm{d}s\\ &= -z^2\int_0^1\int_0^1 \frac{r}{(2+x+rsz)^2}\mathrm{d}r\mathrm{d}s.\end{aligned}$$

Since $2+x+rsz \ge 1+x$ holds for $|z| \le 1$, we conclude that

$$\left|V(x+z) - V(x) - zV'(x)\right| \le \frac{|z|^2}{(1+x)^2}.$$

Finally, if $|z| > 1$ then we use the previously shown inequality to find that

$$\begin{aligned}\left|V(x+z) - V(x) - zV'(x)\right| &\le |V(x+z) - V(x)| + |z||V'(x)|\\ &\le C\frac{1}{2+x} + \frac{C}{2+x}\log(1+|z|) + \frac{|z|}{2+x}\\ &\le \frac{C + (C+1)|z|}{2+x} \le \frac{(2C+1)|z|}{2+x}.\end{aligned}$$

This proves the second inequality. □

Lemma A.3. *Let (b, σ, g_0, g_1) be admissible parameters and that conditions (3.a) – (3.e) are satisfied. Let $\{X_t : t \geq 0\}$ be the unique strong $[0,\infty]$-valued solution of (1.1). Then the following assertions hold:*

(a) *If condition $(m)_\lambda$ holds for some $\lambda > 0$, then there exists a constant $c(\lambda) > 0$ such that*

$$\mathbb{E}\left[X_t^\lambda\right] \leq \mathbb{E}\left[(2+X_0)^\lambda\right] \exp\left(c(\lambda)t\right), \qquad t \geq 0.$$

In particular, if $\mathbb{E}\left[(2+X_0)^\lambda\right] < \infty$, then the unique solution of (1.1) is a.s. finite.

(b) *If condition $(m)_{\log}$ holds, then there exists a constant $c_{\log} > 0$ such that*

$$\mathbb{E}\left[\log(2+X_t)\right] \leq \mathbb{E}\left[\log(2+X_0)\right] + c_{\log} t, \qquad t \geq 0.$$

In particular, if $\mathbb{E}\left[\log(2+X_0)\right] < \infty$, then the unique solution of (1.1) is a.s. finite.

Proof: (a) Let $V_\lambda(x) = (2+x)^\lambda$ and set $\tau_n = \inf\{t \geq 0 : X_t \geq n\}$. Then $X_{t-} \leq n$ for $0 \leq t \leq \tau_n$. Applying the Itô formula to $V_\lambda(X_t)$ and stopping by τ_n yield

$$V_\lambda(X_{t\wedge\tau_n}) = V_\lambda(X_0) + A_{t\wedge\tau_n} + M_{t\wedge\tau_n}, \tag{A.17}$$

where $(M_{t\wedge\tau_n})_{t\geq 0}$ is a local martingale and $(A_{t\wedge\tau_n})_{t\geq 0}$ is of finite variation given by

$$\begin{aligned}
M_{t\wedge\tau_n} &= \int_0^{t\wedge\tau_n}\int_E \sigma(X_{s-},u)V_\lambda'(X_{s-})W(\mathrm{d}s,\mathrm{d}u) \\
&\quad + \int_0^{t\wedge\tau_n}\int_{U_0} \left(V_\lambda(X_{s-}+g_0(X_{s-},u)) - V_\lambda(X_{s-})\right)\widetilde{N}_0(\mathrm{d}s,\mathrm{d}u), \\
A_{t\wedge\tau_n} &= \int_0^{t\wedge\tau_n} b(X_{s-})V_\lambda'(X_{s-})\mathrm{d}s + \frac{1}{2}\int_0^{t\wedge\tau_n}\int_E \sigma(X_{s-},u)^2 V_\lambda''(X_{s-})\varkappa(\mathrm{d}u)\mathrm{d}s \\
&\quad + \int_0^{t\wedge\tau_n}\int_{U_1} \left(V_\lambda(X_{s-}+g_1(X_{s-},u)) - V_\lambda(X_{s-})\right)N_1(\mathrm{d}s,\mathrm{d}u) \\
&\quad + \int_0^{t\wedge\tau_n}\int_{U_0} \big(V_\lambda(X_{s-}+g_0(X_{s-},u)) - V_\lambda(X_{s-}) - V_\lambda'(X_{s-})g_0(X_{s-},u)\big)\,\mu_0(\mathrm{d}u)\mathrm{d}s.
\end{aligned}$$

To estimate the local martingale part, we use the Burkholder-Davis-Gundy inequality, Lemma A.1, conditions (3.c) and (3.d), and $(m)_\lambda$, to find that

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]}|M_{t\wedge\tau_n}|\right]\\
&\leq \mathbb{E}\left[\left(\int_0^{T\wedge\tau_n} V_\lambda'(X_{s-})^2\left(\int_E \sigma(X_{s-},u)^2\varkappa(\mathrm{d}u)\right)\mathrm{d}s\right)^{1/2}\right]\\
&\quad+\mathbb{E}\left[\left(\int_0^{T\wedge\tau_n}\int_{U_0}\mathbb{1}_{\{|g_0(X_{s-},u)|\leq 1\}}\left(V_\lambda(X_{s-}+g_0(X_{s-},u))-V_\lambda(X_{s-})\right)^2 N_0(\mathrm{d}u,\mathrm{d}s)\right)^{1/2}\right]\\
&\quad+2\mathbb{E}\left[\int_0^{T\wedge\tau_n}\int_{U_0}\mathbb{1}_{\{|g_0(X_{s-},u)|>1\}}\left|V_\lambda(X_{s-}+g_0(X_{s-},u))-V_\lambda(X_{s-})\right|\mu_0(\mathrm{d}u)\mathrm{d}s\right]\\
&\leq \lambda\mathbb{E}\left[\left(\int_0^{T\wedge\tau_n}(2+X_{s-})^{2(\lambda-1)}\left(\int_E\sigma(X_{s-},u)^2\varkappa(\mathrm{d}u)\right)\mathrm{d}s\right)^{1/2}\right]\\
&\quad+C_\lambda\left(\mathbb{E}\left[\left(\int_0^{T\wedge\tau_n}\int_{U_0}(2+X_{s-})^{2(\lambda-1)}\mathbb{1}_{\{|g_0(X_{s-},u)|\leq 1\}}|g_0(X_{s-},u)|^2\mu_0(\mathrm{d}u)\mathrm{d}s\right)\right]\right)^{1/2}\\
&\quad+2C_\lambda\mathbb{E}\left[\int_0^{T\wedge\tau_n}\left(\mathbb{1}_{(0,1]}(\lambda)+\mathbb{1}_{(1,\infty)}(\lambda)(2+X_{s-})^{\lambda-1}\right)\left(\int_{U_0}\mathbb{1}_{\{|g_0(X_{s-},u)|>1\}}|g_0(X_{s-},u)|^\lambda\mu_0(\mathrm{d}u)\right)\mathrm{d}s\right]\\
&\leq(\lambda\sqrt{K_1}+C_\lambda\sqrt{K_2})\mathbb{E}\left[\left(\int_0^{T\wedge\tau_n}(2+X_{s-})^{2\lambda}\mathrm{d}s\right)^{1/2}\right]\\
&\quad+2C_\lambda K_\lambda\mathbb{E}\left[\int_0^{T\wedge\tau_n}\left(\mathbb{1}_{(0,2)}(\lambda)(2+X_{s-})^\lambda+\mathbb{1}_{[2,\infty)}(\lambda)(2+X_{s-})^{\lambda+1}\right)\mathrm{d}s\right]\\
&\leq(\lambda\sqrt{K_1}+C_\lambda\sqrt{K_2})\sqrt{T}(2+n)^\lambda+2C_\lambda K_\lambda T\left(\mathbb{1}_{(0,2)}(\lambda)(2+n)^\lambda+\mathbb{1}_{[2,\infty)}(\lambda)(2+n)^{\lambda+1}\right).
\end{aligned}
$$

This shows that $(M_{t\wedge\tau_n})_{t\geq 0}$ is a martingale for each fixed $n\in\mathbb{N}$. For the process $(A_{t\wedge\tau_n})_{t\geq0}$ we use Lemma A.1 combined with conditions (3.a), (3.c), (3.d), and $(m)_\lambda$, to find that

$$
\begin{aligned}
A_{t\wedge\tau_n}&\leq\lambda\int_0^{t\wedge\tau_n}b(X_{s-})(2+X_{s-})^{\lambda-1}\mathrm{d}s+\frac{\lambda(\lambda-1)}{2}\int_0^{t\wedge\tau_n}(2+X_{s-})^{\lambda-2}\left(\int_E\sigma(X_{s-},u)^2\varkappa(\mathrm{d}u)\right)\mathrm{d}s\\
&\quad+C_\lambda\int_0^{t\wedge\tau_n}(2+X_{s-})^{\lambda-1}\left(\int_{U_1}\mathbb{1}_{\{|g_1(X_{s-},u)|\leq1\}}|g_1(X_{s-},u)|\mu_1(\mathrm{d}u)\right)\mathrm{d}s\\
&\quad+C_\lambda\int_0^{t\wedge\tau_n}\left(\mathbb{1}_{(0,1]}(\lambda)+\mathbb{1}_{(1,\infty)}(\lambda)(2+X_{s-})^{\lambda-1}\right)\left(\int_{U_1}\mathbb{1}_{\{|g_1(X_{s-},u)|>1\}}|g_1(X_{s-},u)|^\lambda\mu_1(\mathrm{d}u)\right)\mathrm{d}s\\
&\quad+\widetilde{C}_\lambda\int_0^{t\wedge\tau_n}(2+X_{s-})^{\lambda-2}\left(\int_{U_0}\mathbb{1}_{\{|g_0(X_{s-},u)|\leq1\}}|g_0(X_{s-},u)|^2\mu_0(\mathrm{d}u)\right)\mathrm{d}s\\
&\quad+\widetilde{C}_\lambda\int_0^{t\wedge\tau_n}\left(\mathbb{1}_{(0,1]}(\lambda)+\mathbb{1}_{(1,2)}(\lambda)(2+X_{s-})^{\lambda-1}+1_{[2,\infty)}(\lambda)(2+X_{s-})^{\lambda-2}\right)\cdot\\
&\qquad\qquad\cdot\left(\int_{U_0}\mathbb{1}_{\{|g_0(X_{s-},u)|>1\}}|g_0(X_{s-},u)|^{\max\{1,\lambda\}}\mu_0(\mathrm{d}u)\right)\mathrm{d}s\\
&\leq\left(K_0\lambda+\frac{K_1\lambda(\lambda-1)}{2}+K_0C_\lambda+K_\lambda C_\lambda+K_2\widetilde{C}_\lambda+K_\lambda\widetilde{C}_\lambda\right)\int_0^{t\wedge\tau_n}(2+X_{s-})^\lambda\mathrm{d}s
\end{aligned}
$$

Hence taking expectations in (A.17) yields

$$\begin{aligned}\mathbb{E}[V_\lambda(X_{t\wedge\tau_n})] &= \mathbb{E}[V_\lambda(X_0)] + \mathbb{E}[A_{t\wedge\tau_n}]\\ &\le \mathbb{E}[V_\lambda(X_0)] + c(\lambda)\int_0^t \mathbb{E}[V_\lambda(X_{s\wedge\tau_n})]\mathrm{d}s,\end{aligned}$$

which gives $\mathbb{E}[V_\lambda(X_{t\wedge\tau_n})] \le \mathbb{E}[V_\lambda(X_0)]e^{c(\lambda)t}$. By the right-continuity of X we have $X_{\tau_n} \ge n$ and hence

$$\begin{aligned}\mathbb{P}[\tau_n \le t] &= \mathbb{P}\left[V_\lambda(X_{t\wedge\tau_n}) \ge V_\lambda(n),\ t \ge \tau_n\right]\\ &\le \mathbb{P}\left[V_\lambda(X_{t\wedge\tau_n}) \ge V_\lambda(n)\right]\\ &\le \frac{1}{V_\lambda(n)}\mathbb{E}\left[V_\lambda(X_{t\wedge\tau_n})\right]\\ &\le \frac{1}{V_\lambda(n)}\mathbb{E}\left[V_\lambda(X_0)\right]e^{c(\lambda)t}.\end{aligned}$$

Hence $\tau_n \nearrow \infty$ a.s. as $n \to \infty$, which implies that X is a.s. finite. The Lemma of Fatou finally yields $\mathbb{E}[V_\lambda(X_t)] \le \mathbb{E}[V_\lambda(X_0)]e^{c(\lambda)t}$.

(b) Let τ_n be as in part (a). Applying here the Itô formula to $\log(2+X_{t\wedge\tau_n})$ yields $\log(2+X_{t\wedge\tau_n}) = \log(2+X_0) + A^{\log}_{t\wedge\tau_n} + M^{\log}_{t\wedge\tau_n}$, where $(M^{\log}_{t\wedge\tau_n})_{t\ge0}$ and $(A^{\log}_{t\wedge\tau_n})_{t\ge0}$ are given by

$$\begin{aligned}M^{\log}_{t\wedge\tau_n} &= \int_0^{t\wedge\tau_n}\int_E \sigma(X_{s-},u)\frac{1}{2+X_{s-}}W(\mathrm{d}s,\mathrm{d}u)\\ &\quad+ \int_0^{t\wedge\tau_n}\int_{U_0}\left(\log(2+X_{s-}+g_0(X_{s-},u)) - \log(2+X_{s-})\right)\widetilde{N}_0(\mathrm{d}s,\mathrm{d}u),\\ A^{\log}_{t\wedge\tau_n} &= \int_0^{t\wedge\tau_n} b(X_{s-})\frac{1}{2+X_{s-}}\mathrm{d}s - \frac{1}{2}\int_0^{t\wedge\tau_n}\int_E \sigma(X_{s-},u)^2\frac{1}{(2+X_{s-})^2}\varkappa(\mathrm{d}u)\mathrm{d}s\\ &\quad+ \int_0^{t\wedge\tau_n}\int_{U_1}\left(\log(2+X_{s-}+g_1(X_{s-},u)) - \log(2+X_{s-})\right)N_1(\mathrm{d}u,\mathrm{d}s)\\ &\quad+ \int_0^{t\wedge\tau_n}\int_{U_0}\left(\log(2+X_{s-}+g_0(X_{s-},u)) - \log(2+X_{s-}) - \frac{g_0(X_{s-},u)}{2+X_{s-}}\right)\mu_0(\mathrm{d}u)\mathrm{d}s.\end{aligned}$$

Similarly as in part (a), using the Burkholder-Davis-Gundy inequality, Lemma A.2, conditions (3.c), (3.d), and $(m)_{\log}$ we find that

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,T]}|M^{\log}_{t\wedge\tau_n}|\right] &\le \mathbb{E}\left[\left(\int_0^{T\wedge\tau_n}\frac{1}{(2+X_{s-})^2}\left(\int_E\sigma(X_{s-},u)^2\varkappa(\mathrm{d}u)\right)\mathrm{d}s\right)^{1/2}\right]\\ &\quad+\left(\mathbb{E}\left[\int_0^{T\wedge\tau_n}\frac{1}{(1+X_{s-})^2}\left(\int_{U_0}\mathbb{1}_{\{|g_0(X_{s-},u)|\le1\}}|g_0(X_{s-},u)|^2\mu_0(\mathrm{d}u)\right)\mathrm{d}s\right]\right)^{1/2}\\ &\quad+2C_{\log}\mathbb{E}\left[\int_0^{T\wedge\tau_n}\frac{\mu_0\left(\{|g_0(X_{s-},\cdot)|>1\}\right)}{2+X_{s-}}\mathrm{d}s\right]\\ &\quad+2C_{\log}\mathbb{E}\left[\int_0^{T\wedge\tau_n}\frac{1}{2+X_{s-}}\left(\int_{U_0}\mathbb{1}_{\{|g_0(X_s,u)|>1\}}|g_0(X_{s-},u)|\mu_0(\mathrm{d}u)\right)\mathrm{d}s\right]\\ &\le(\sqrt{K_1T}+2\sqrt{K_2T})+4C_{\log}K_{\log}T,\end{aligned}$$

where we have used that $\log(1+|g_0(X_{s-},u)|) \le |g_0(X_{s-},u)|$ and

$$\mu_0\left(\{|g_0(x,\cdot)|>1\}\right) \le \int_{U_0}\mathbb{1}_{\{|g_0(x,u)|>1\}}|g_0(x,u)|\mu_0(\mathrm{d}u) \le K_{\log}(2+x).$$

This shows that $(M^{\log}_{t\wedge\tau_n})_{t\geq 0}$ is a martingale. Secondly, using also condition (3.a), we observe that

$$\begin{aligned}
|A^{\log}_{t\wedge\tau_n}| \leq K_0 \int_0^{t\wedge\tau_n} \frac{2+X_{s-}}{2+X_{s-}}\mathrm{d}s + \frac{1}{2}\int_0^{t\wedge\tau_n} \frac{1}{(2+X_{s-})^2}\left(\int_E \sigma(X_{s-},u)^2\varkappa(\mathrm{d}u)\right)\mathrm{d}s \\
+ C_{\log}\int_0^{t\wedge\tau_n} \frac{1}{1+X_{s-}}\left(\int_{U_1} \mathbb{1}_{\{|g_1(X_{s-},u)|\leq 1\}}|g_1(X_{s-},u)|\mu_1(\mathrm{d}u)\right)\mathrm{d}s \\
+ C_{\log}\int_0^{t\wedge\tau_n} \frac{\mu_1\left(\{|g_1(X_{s-},\cdot)|>1\}\right)}{2+X_{s-}}\mathrm{d}s \\
+ C_{\log}\int_0^{t\wedge\tau_n} \frac{1}{2+X_{s-}}\left(\int_{U_1} \mathbb{1}_{\{|g_1(X_{s-},u)|>1\}}\log(1+|g_1(X_{s-},u))|)\mu_1(\mathrm{d}u)\right)\mathrm{d}s \\
+ C_{\log}\int_0^{t\wedge\tau_n} \frac{1}{(1+X_{s-})^2}\left(\int_{U_0} \mathbb{1}_{\{|g_0(X_{s-},u)|\leq 1\}}|g_0(X_{s-},u))|^2\mu_0(\mathrm{d}u)\right)\mathrm{d}s \\
+ C_{\log}\int_0^{t\wedge\tau_n} \frac{\mu_0\left(\{|g_0(X_{s-},\cdot)|>1\}\right)}{2+X_{s-}}\mathrm{d}s \\
+ C_{\log}\int_0^{t\wedge\tau_n} \frac{1}{2+X_{s-}}\left(\int_{U_0} \mathbb{1}_{\{|g_0(X_{s-},u)|>1\}}|g_0(X_{s-},u))|\mu_0(\mathrm{d}u)\right)\mathrm{d}s \\
\leq ((1+2C_{\log})K_0 + K_1/2 + 3C_{\log}K_{\log} + C_{\log}K_2)t,
\end{aligned}$$

where we have used $\mu_1\left(\{|g_1(x,\cdot)|>1\}\right) \leq K_{\log}(2+x)$. Thus, taking expectations of $\log(2+X_{t\wedge\tau_n}) = \log(2+X_0) + M^{\log}_{t\wedge\tau_n} + A^{\log}_{t\wedge\tau_n}$ and using the above estimate proves the assertion. $\square$

References

Bardet, J.-B., Christen, A., Guillin, A., Malrieu, F., and Zitt, P.-A. Total variation estimates for the TCP process. *Electron. J. Probab.*, **18**, no. 10, 21 (2013). doi:10.1214/EJP.v18-1720. URL https://doi.org/10.1214/EJP.v18-1720.

Butkovsky, O. Subgeometric rates of convergence of Markov processes in the Wasserstein metric. *Ann. Appl. Probab.*, **24** (2), 526–552 (2014). doi:10.1214/13-AAP922. URL https://doi.org/10.1214/13-AAP922.

Chafaï, D., Malrieu, F., and Paroux, K. On the long time behavior of the TCP window size process. *Stochastic Process. Appl.*, **120** (8), 1518–1534 (2010). doi:10.1016/j.spa.2010.03.019. URL https://doi.org/10.1016/j.spa.2010.03.019.

Dawson, D. A. and Li, Z. Stochastic equations, flows and measure-valued processes. *Ann. Probab.*, **40** (2), 813–857 (2012). doi:10.1214/10-AOP629. URL https://doi.org/10.1214/10-AOP629.

Douc, R., Fort, G., and Guillin, A. Subgeometric rates of convergence of f-ergodic strong Markov processes. *Stochastic Process. Appl.*, **119** (3), 897–923 (2009). doi:10.1016/j.spa.2008.03.007. URL https://doi.org/10.1016/j.spa.2008.03.007.

Eberle, A. Reflection couplings and contraction rates for diffusions. *Probab. Theory Related Fields*, **166** (3-4), 851–886 (2016). doi:10.1007/s00440-015-0673-1. URL https://doi.org/10.1007/s00440-015-0673-1.

Eberle, A., Guillin, A., and Zimmer, R. Quantitative Harris-type theorems for diffusions and McKean-Vlasov processes. *Trans. Amer. Math. Soc.*, **371** (10), 7135–7173 (2019). doi:10.1090/tran/7576. URL https://doi.org/10.1090/tran/7576.

Ethier, S. N. and Kurtz, T. G. *Markov processes: Characterization and convergence.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York (1986). ISBN 0-471-08186-8. doi:10.1002/9780470316658. URL http://dx.doi.org/10.1002/9780470316658.

Feller, W. Diffusion processes in genetics. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950*, pp. 227–246. University of California Press, Berkeley and Los Angeles (1951).

Fournier, N. On pathwise uniqueness for stochastic differential equations driven by stable Lévy processes. *Ann. Inst. Henri Poincaré Probab. Stat.*, **49** (1), 138–159 (2013). doi:10.1214/11-AIHP420. URL https://doi.org/10.1214/11-AIHP420.

Friesen, M., Jin, P., and Rüdiger, B. On the boundary behavior of multi-type continuous-state branching processes with immigration. *Electron. Commun. Probab.*, **25**, Paper No. 84, 14 (2020a). doi:10.3390/mca25010013. URL https://doi.org/10.3390/mca25010013.

Friesen, M., Jin, P., and Rüdiger, B. Stochastic equation and exponential ergodicity in Wasserstein distances for affine processes. *Ann. Appl. Probab.*, **30** (5), 2165–2195 (2020b). doi:10.1214/19-AAP1554. URL https://doi.org/10.1214/19-AAP1554.

Fu, Z. and Li, Z. Stochastic equations of non-negative processes with jumps. *Stochastic Process. Appl.*, **120** (3), 306–330 (2010). doi:10.1016/j.spa.2009.11.005. URL http://dx.doi.org/10.1016/j.spa.2009.11.005.

He, H., Li, Z., and Xu, W. Continuous-state branching processes in Lévy random environments. *J. Theoret. Probab.*, **31** (4), 1952–1974 (2018). doi:10.1007/s10959-017-0765-1. URL https://doi.org/10.1007/s10959-017-0765-1.

Jin, P., Kremer, J., and Rüdiger, B. Moments and ergodicity of the jump-diffusion CIR process. *Stochastics*, **91** (7), 974–997 (2019). doi:10.1080/17442508.2019.1576686. URL https://doi.org/10.1080/17442508.2019.1576686.

Jin, P., Kremer, J., and Rüdiger, B. Existence of limiting distribution for affine processes. *J. Math. Anal. Appl.*, **486** (2), 123912, 31 (2020). doi:10.1016/j.jmaa.2020.123912. URL https://doi.org/10.1016/j.jmaa.2020.123912.

Jin, P., Rüdiger, B., and Trabelsi, C. Exponential ergodicity of the jump-diffusion CIR process. In *Stochastics of environmental and financial economics—Centre of Advanced Study, Oslo, Norway, 2014–2015*, volume 138 of *Springer Proc. Math. Stat.*, pp. 285–300. Springer, Cham (2016a). doi:10.1007/978-3-319-23425-0_11. URL https://doi.org/10.1007/978-3-319-23425-0_11.

Jin, P., Rüdiger, B., and Trabelsi, C. Positive Harris recurrence and exponential ergodicity of the basic affine jump-diffusion. *Stoch. Anal. Appl.*, **34** (1), 75–95 (2016b). doi:10.1080/07362994.2015.1105752. URL https://doi.org/10.1080/07362994.2015.1105752.

Jiřina, M. Stochastic branching processes with continuous state space. *Czechoslovak Math. J.*, **8 (83)**, 292–313 (1958).

Kawazu, K. and Watanabe, S. Branching Processes with Immigration and Related Limit Theorems. *Theory of Probability & Its Applications*, **16** (1), 36–54 (1971). doi:10.1137/1116003. URL https://doi.org/10.1137/1116003.

Keller-Ressel, M. Moment explosions and long-term behavior of affine stochastic volatility models. *Math. Finance*, **21** (1), 73–98 (2011). doi:10.1111/j.1467-9965.2010.00423.x. URL http://dx.doi.org/10.1111/j.1467-9965.2010.00423.x.

Keller-Ressel, M. and Mijatović, A. On the limit distributions of continuous-state branching processes with immigration. *Stochastic Process. Appl.*, **122** (6), 2329–2345 (2012). doi:10.1016/j.spa.2012.03.012. URL https://doi.org/10.1016/j.spa.2012.03.012.

Kulik, A. M. Exponential ergodicity of the solutions to SDE's with a jump noise. *Stochastic Process. Appl.*, **119** (2), 602–632 (2009). doi:10.1016/j.spa.2008.02.006. URL https://doi.org/10.1016/j.spa.2008.02.006.

Kurtz, T. G. *Equivalence of stochastic equations and martingale problems.* Springer Heidelberg (2011). doi:10.1007/978-3-642-15358-76.

Lambert, A. Quasi-stationary distributions and the continuous-state branching process conditioned to be never extinct. *Electron. J. Probab.*, **12**, no. 14, 420–446 (2007). doi:10.1214/EJP.v12-402. URL https://doi.org/10.1214/EJP.v12-402.

Li, P.-S. A continuous-state polynomial branching process. *Stochastic Process. Appl.*, **129** (8), 2941–2967 (2019). doi:10.1016/j.spa.2018.08.013. URL https://doi.org/10.1016/j.spa.2018.08.013.

Li, P.-S., Yang, X., and Zhou, X. A general continuous-state nonlinear branching process. *Ann. Appl. Probab.*, **29** (4), 2523–2555 (2019). doi:10.1214/18-AAP1459. URL https://doi.org/10.1214/18-AAP1459.

Li, Z. *Measure-valued branching Markov processes.* Probability and its Applications (New York). Springer, Heidelberg (2011). ISBN 978-3-642-15003-6. doi:10.1007/978-3-642-15004-3. URL https://doi.org/10.1007/978-3-642-15004-3.

Li, Z. and Ma, C. Asymptotic properties of estimators in a stable Cox-Ingersoll-Ross model. *Stochastic Process. Appl.*, **125** (8), 3196–3233 (2015). doi:10.1016/j.spa.2015.03.002. URL https://doi.org/10.1016/j.spa.2015.03.002.

Li, Z. and Mytnik, L. Strong solutions for stochastic differential equations with jumps. *Ann. Inst. Henri Poincaré Probab. Stat.*, **47** (4), 1055–1067 (2011). doi:10.1214/10-AIHP389. URL https://doi.org/10.1214/10-AIHP389.

Li, Z. and Pu, F. Strong solutions of jump-type stochastic equations. *Electron. Commun. Probab.*, **17**, no. 33, 13 (2012). doi:10.1214/ECP.v17-1915. URL https://doi.org/10.1214/ECP.v17-1915.

Li, Z. and Xu, W. Asymptotic results for exponential functionals of Lévy processes. *Stochastic Process. Appl.*, **128** (1), 108–131 (2018). doi:10.1016/j.spa.2017.04.005. URL https://doi.org/10.1016/j.spa.2017.04.005.

Masuda, H. Ergodicity and exponential β-mixing bounds for multidimensional diffusions with jumps. *Stochastic Process. Appl.*, **117** (1), 35–56 (2007). doi:10.1016/j.spa.2006.04.010. URL https://doi.org/10.1016/j.spa.2006.04.010.

Meyn, S. and Tweedie, R. L. *Markov chains and stochastic stability.* Cambridge University Press, Cambridge, second edition (2009). ISBN 978-0-521-73182-9. doi:10.1017/CBO9780511626630. URL https://doi.org/10.1017/CBO9780511626630. With a prologue by Peter W. Glynn.

Meyn, S. P. and Tweedie, R. L. Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.*, **25** (3), 518–548 (1993). doi:10.2307/1427522. URL http://dx.doi.org/10.2307/1427522.

Palau, S. and Pardo, J. C. Continuous state branching processes in random environment: the Brownian case. *Stochastic Process. Appl.*, **127** (3), 957–994 (2017). doi:10.1016/j.spa.2016.07.006. URL https://doi.org/10.1016/j.spa.2016.07.006.

Palau, S. and Pardo, J. C. Branching processes in a Lévy random environment. *Acta Appl. Math.*, **153**, 55–79 (2018). doi:10.1007/s10440-017-0120-7. URL https://doi.org/10.1007/s10440-017-0120-7.

Pardoux, E. *Probabilistic models of population evolution*, volume 1 of *Mathematical Biosciences Institute Lecture Series. Stochastics in Biological Systems.* Springer, [Cham]; MBI Mathematical Biosciences Institute, Ohio State University, Columbus, OH (2016). ISBN 978-3-319-30326-0; 978-3-319-30328-4. doi:10.1007/978-3-319-30328-4. URL https://doi.org/10.1007/978-3-319-30328-4. Scaling limits, genealogies and interactions.

Peng, X. and Zhang, R. Exponential ergodicity for SDEs under the total variation. *J. Evol. Equ.*, **18** (3), 1051–1067 (2018). doi:10.1007/s00028-018-0429-3. URL https://doi.org/10.1007/s00028-018-0429-3.

Pinsky, M. A. Limit theorems for continuous state branching processes with immigration. *Bull. Amer. Math. Soc.*, **78**, 242–244 (1972). doi:10.1090/S0002-9904-1972-12938-0. URL https://doi.org/10.1090/S0002-9904-1972-12938-0.

Veretennikov, A. Y. Estimates of the mixing rate for stochastic equations. *Teor. Veroyatnost. i Primenen.*, **32** (2), 299–308 (1987).

Villani, C. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin (2009). ISBN 978-3-540-71049-3. doi:10.1007/978-3-540-71050-9. URL https://doi.org/10.1007/978-3-540-71050-9. Old and new.

Wang, J. L^p-Wasserstein distance for stochastic differential equations driven by Lévy processes. *Bernoulli*, **22** (3), 1598–1616 (2016). doi:10.3150/15-BEJ705. URL https://doi.org/10.3150/15-BEJ705.