Bergische Universität Wuppertal

Fachbereich Mathematik und Naturwissenschaften

Institute of Mathematical Modelling, Analysis and Computational Mathematics (IMACM)

Preprint BUW-IMACM 19/04

Martin Friesen, Peng Jin, Jonas Kremer and Barbara Rüdiger

Exponential ergodicity for stochastic equations of nonnegative processes with jumps

February 11, 2019

http://www.math.uni-wuppertal.de

EXPONENTIAL ERGODICITY FOR STOCHASTIC EQUATIONS OF NONNEGATIVE PROCESSES WITH JUMPS

MARTIN FRIESEN, PENG JIN*, JONAS KREMER, AND BARBARA RÜDIGER

Abstract. In this work, we study ergodicity of continuous time Markov processes on state space $\mathbb{R}_{\geq 0} := [0,\infty)$ obtained as unique strong solutions to stochastic equations with jumps. Our first main result establishes exponential ergodicity in the Wasserstein distance, provided the stochastic equation satisfies a comparison principle and the drift is dissipative. In particular, it is applicable to continuous-state branching processes with immigration (shorted as CBI processes), possibly with nonlinear branching mechanisms or in Lévy random environments. Our second main result establishes exponential ergodicity in total variation distance for subcritical CBI processes under a first moment condition on the jump measure for branching and a log-moment condition on the jump measure for immigration.

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 will 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., [30, 40]), possibly in Lévy random environments (see [18, 38]), *continuous-state nonlinear branching processes* (see [29]), and *TCP processes* (see, e.g., [3, 8]). All these processes just mentioned belong to the class of Markov processes with state space $\mathbb{R}_{\geq 0}$ whose Markov generator is, for twice continuously differentiable functions f with compact support, i.e., $f \in C_c^2(\mathbb{R}_{\geq 0})$, of the form

$$\begin{aligned} Lf(x) &= b(x)f'(x) + \frac{1}{2}\int_E \sigma(x,u)^2 \varkappa(\mathrm{d}u) f''(x) + \int_{U_1} \left(f(x+g_1(x,u)) - f(x)\right)\mu_1(\mathrm{d}u) \\ &\quad + \int_{U_0} \big(f(x+g_0(x,u)) - f(x) - g_0(x,u)f'(x)\big)\,\mu_0(\mathrm{d}u), \qquad x \geq 0. \end{aligned} \tag{1.1}$$

Here E, U_0, U_1 are complete, separable metric spaces, $\varkappa$, μ_0, μ_1 are σ-finite measures and b, σ, g_0, g_1 should satisfy certain restrictions such that the corresponding Markov process exists. A pathwise construction for this type of Markov processes in terms of strong solutions to stochastic equations were developed in the works of Fu and Li [17], Dawson and Li [10], and Li and Pu [33]. Additional related results for stochastic equations on $\mathbb{R}_{\geq 0}$ can be found in Li and Mytnik [32] as well as Fournier [15].

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

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

Date: February 11, 2019.

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.

*Peng Jin is supported by the STU Scientific Research Foundation for Talents (No. NTF18023).

Existence of invariant distributions is often shown by various compactness arguments, see, e.g., Section 9 in Chapter 4 of [13] for some sufficient conditions. Unlike existence, uniqueness of the invariant distribution may be a more demanding mathematical problem. Once existence and uniqueness of an invariant distribution π is shown, it is then natural to study the convergence of $P_t(x, \mathrm{d}y)$ to π. In order to study such convergence, let us define, for ϱ, $\widetilde{\varrho} \in \mathcal{P}(\mathbb{R}_{\geq 0})$, the Wasserstein distance

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

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. We will collect some basic properties of W_d in Section 2, while a detailed treatment of Wasserstein distances is provided in the monograph of Villani [45].

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 [36, 37]. 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., [7, 12, 41, 46].

In this work we provide a simple approach to exponential ergodicity in the Wasserstein distance W_d with $d(x,y) = |x-y|$ for Markov processes on $\mathbb{R}_{\geq 0}$ which can be constructed as strong solutions of stochastic equations satisfying the comparison principle. In particular, based on the comparison principle we estimate the trajectories of the Markov process and deduce from that the existence, uniqueness of invariant distributions, and exponential ergodicity in the Wasserstein distance. The corresponding main result is formulated in Theorem 3.2. Similar ideas have been previously applied in [16] to affine processes which include the specific case of continuous-state branching processes with immigration. To illustrate the usage of Theorem 3.2, we apply it to three classes of Markov processes on $\mathbb{R}_{\geq 0}$ which we next explain in more details. We will also discuss limitations and possible improvements of Theorem 3.2 in these particular cases.

1.1. **Continuous-state nonlinear branching processes with immigration.** *Continuous-state nonlinear branching processes (CNB process)* were recently introduced and studied by Li *et al.* (2017) [29]. 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)*. The corresponding Markov generator for the class of CNBI processes is, for $f \in C_c^2(\mathbb{R}_{\geq 0})$, of the form

$$\begin{aligned} Lf(x) &= \gamma_0(x) f'(x) + \frac{\gamma_1(x)}{2} f''(x) \\ &\quad + \gamma_2(x) \int_{\mathbb{R}_{\geq 0}} (f(x+z) - f(x) - z f'(x)) m(\mathrm{d}z) + \int_{\mathbb{R}_{\geq 0}} (f(x+z) - f(x))\, \nu(\mathrm{d}z), \end{aligned} \tag{1.2}$$

where γ_0, γ_1, γ_2 are Borel-functions on $\mathbb{R}_{\geq 0}$, $\gamma_0(0) \geq 0$ satisfying $\gamma_1, \gamma_2 \geq 0$, and m, ν are Borel measures on $(0, \infty)$ satisfying

$$\int_{\mathbb{R}_{\geq 0}} (z \wedge z^2) m(\mathrm{d}z) + \int_{\mathbb{R}_{\geq 0}} (1 \wedge z) \nu(\mathrm{d}z) < \infty. \tag{1.3}$$

If γ_0, γ_1, γ_2 are locally Lipschitz continuous on $(0, \infty)$ and $\nu \equiv 0$, then a pathwise construction of the corresponding Markov process (called a CNB process) is established in [29]. More precisely, the authors identified the CNB process with a unique strong solution to a certain stochastic equation with the additional restriction that 0 and ∞ are traps. The last requirement is obligatory to treat particular cases as, e.g., $\gamma_2(x) = x^p$ with $p > 0$. The authors then studied extinction, explosion, and coming down from infinity behaviors of the CNB process. In Section 4 we provide some simple sufficient conditions on the parameters γ_0, γ_1, γ_2 such that the CNBI process is exponentially ergodic in the Wasserstein distance W_1 (see Theorem 4.2). To the best of our knowledge, it is the first ergodicity result for general CNBI processes.

1.2. **Continuous-state branching processes with immigration.** *Continuous-state branching processes with immigration (CBI processes)* are particular cases of CNBI processes. They have been first introduced by Feller (1950) [14] and Jiřina (1958) [23] and then developed by Kawazu and Watanabe (1971) [24]. For a detailed treatment of CBI processes encompassing a concise introduction we refer to the monographs of Li [30] and Pardoux [40]. Following [24], CBI processes are Feller processes whose Markov generator is, for $f \in C_c^2(\mathbb{R}_{\geq 0})$, given by

$$\begin{aligned} Lf(x) &= (\beta - bx) f'(x) + \frac{\sigma^2}{2} x f''(x) \\ &\quad + x \int_0^\infty \big(f(x+z) - f(x) - z f'(x)\big) m(\mathrm{d}z) + \int_0^\infty (f(x+z) - f(x)) \nu(\mathrm{d}z). \end{aligned} \tag{1.4}$$

Here $(\beta, b, \sigma, m, \nu)$ are *admissible parameters* in the sense that $\beta \geq 0$, $b \in \mathbb{R}$, $\sigma \geq 0$, m and ν satisfy (1.3). In Section 5 we also briefly recall other characterizations of CBI processes in terms of their Laplace transforms and strong solutions to stochastic equations with jumps motivated by [9, 17].

Previously, a number of authors investigated the long-time behavior of CBI processes. Pinsky [42] announced the existence of a limit distribution for subcritical ($b > 0$) CBI processes under the condition

$$\int_{\{z>1\}} \log(z) \nu(\mathrm{d}z) < \infty. \tag{1.5}$$

It was shown subsequently in [30, Theorem 3.20 and Corollary 3.21] and [25, Theorem 3.16] that for subcritical CBI processes condition (1.5) is equivalent to the weak convergence of the associated transition probabilities towards a limiting distribution. The limit distribution was also shown to be the unique invariant distribution for the CBI process. Properties of this distribution were investigated in [26]. A multidimensional version of Pinsky's result was recently studied in [20], while in [16] exponential ergodicity in different Wasserstein distances was derived under reasonable integrability conditions on ν.

In the setting of CBI processes, our Theorem 3.2 is applicable but the obtained result is not particularly strong, compared with [16, Theorem 1.6]. On the other hand, the exponential ergodicity in the total variation distance is yet under current investigation. Based on the approach of Meyn and Tweedie, particular examples have been studied in [19, 21, 22, 35] using the condition that

$$\exists\, \varepsilon > 0 \text{ such that } \int_{\{z>1\}} z^{\varepsilon} \nu(\mathrm{d}z) < \infty \tag{1.6}$$

in order to derive a Lyapunov drift criteria inherent in the approach of Meyn and Tweedie. An alternative approach based on the construction of a successful coupling was recently established by Li and Ma [31]. Following [31, Theorem 2.5], a subcritical CBI process with admissible parameters $(b, \beta, \sigma, m, \nu \equiv 0)$ is exponentially ergodic in the total variation distance, provided Grey's condition on the immigration mechanism is satisfied (see condition (5.a) below).

In Section 5 we extend the aforementioned results and establish exponential ergodicity in total variation distance for general subcritical CBI processes (including the case $\nu \not\equiv 0$) under the weaker integrability condition (1.5) (compared with (1.6)). The corresponding main result is formulated in Theorem 5.7 and it seems to the first result establishing a convergence rate in total variation distance merely under (1.5).

Exponential ergodicity often plays an essential role in statistical estimation of the parameters of the underlying process, as demonstrated in several articles, see [1, 5, 4, 31] and the references therein. Eventually, to illustrate some applications of Theorem 5.7, we add at the end of section 5 a strong law of large numbers as well as a functional central limit theorem for CBI processes (see Corollaries 5.9 and 5.10 below).

1.3. **Continuous-state branching processes with immigration in the Lévy random environment.** A *continuous-state branching process with immigration in a Lévy random environment (CBIRE process)* is a Markov process on $\mathbb{R}_{\geq 0}$ with generator $L = L_0 + L_1$ acting on $C_c^2(\mathbb{R}_{\geq 0})$, where L_0 is given by (1.4) and

$$
\begin{aligned}
L_1 f(x) = b_E x f'(x) + \frac{\sigma_E^2}{2} x^2 f''(x) &+ \int_{[-1,1]^c} \left(f(xe^z) - f(x) \right) \mu_E(\mathrm{d}z) \\
&+ \int_{[-1,1]} \left(f(xe^z) - f(x) - x(e^z - 1) f'(x) \right) \mu_E(\mathrm{d}z).
\end{aligned} \tag{1.7}
$$

Here, $b_E \in \mathbb{R}$, $\sigma_E \geq 0$ and μ_E is a Lévy measure on $\mathbb{R}$. We refer to He *et al.* [18] for the general theory and a comprehensive introduction of CBIRE processes, see also the works of Palau and Pardo [38, 39]. In [18] the authors gave a necessary and sufficient condition for the existence of invariant distributions of CBIRE processes. It is our aim in Section 6 to prove ergodicity in both the Wasserstein and total variation distance for these processes (see Theorems 6.1 and 6.2 below). Those results are obtained with the help of our previously obtained ergodic results in Sections 3 and 5.

1.4. **Structure of the work.** In Section 2 we recall some properties of Wasserstein distances. Exponential ergodicity in the Wasserstein distance for Markov processes with generator (1.1) is established in Section 3. The particular example of CNBI processes is then discussed in Section 4. Next, in Section 5 we study the exponential ergodicity in total variation distance for CBI processes. Finally, in Section 6 an application of our previous results to CBIRE processes is given.

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 $\varrho, \widetilde{\varrho} \in \mathcal{P}(\mathbb{R}_{\geq 0})$, 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 separable metric space and define

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

The *Wasserstein distance* on $\mathcal{P}_d(\mathbb{R}_{\geq 0})$ is defined by

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

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 [45, p.95]), i.e., there exists $H \in \mathcal{H}(\varrho, \widetilde{\varrho})$ such that

$$W_d(\varrho, \widetilde{\varrho}) = \int_{\mathbb{R}_{\geq 0}\times\mathbb{R}_{\geq 0}} d(x,y) H(\mathrm{d}x, \mathrm{d}y). \tag{2.2}$$

Since $(\mathbb{R}_{\geq 0}, d)$ is supposed to be a complete separable metric space, according to [45, 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*

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

is the total variation distance.

(b) *The Wasserstein-1-distance corresponds to* $d_1(x,y) = |x-y|$, *where*

$$\mathcal{P}_{d_1}(\mathbb{R}_{\geq 0}) := \mathcal{P}_1(\mathbb{R}_{\geq 0}) := \left\{ \varrho \in \mathcal{P}(\mathbb{R}_{\geq 0}) \,:\, \int_{\mathbb{R}_{\geq 0}} x\varrho(\mathrm{d}x) < \infty \right\}.$$

In this case we adopt the shorthand $W_1 := W_{d_1}$.

(c) *If* $d_{\log}(x,y) = \log(1+|x-y|)$, *then*

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

and $W_{\log} := W_{d_{\log}}$ *is suited for CBI processes studied in Section 5.*

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

Lemma 2.2. *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*

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

For a proof we refer the reader to [45, 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 2.3. *Let d be a metric such that $(\mathbb{R}_{\geq 0}, d)$ is a complete separable metric space. Suppose that*

$$d(x+y, \widetilde{x}+y) \leq d(x, \widetilde{x}), \qquad x, \widetilde{x}, y \geq 0.$$

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. We define $\|h\|_{\mathrm{Lip}} = \sup_{x\neq y} \frac{|h(x)-h(y)|}{d(x,y)}$. Using the Kantorovich-Duality we obtain

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

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

where we used that $h_g(x) = \int_{\mathbb{R}_{\ge 0}} h(x+y) g\,(\mathrm{d}y)$ satisfies $\|h_g\|_{\mathrm{Lip}} \le 1$. □

Although we formulated Lemma 2.2 and 2.3 on the state space $\mathbb{R}_{\ge 0}$, it is clear that they naturally extend to more abstract state spaces. In particular, Lemma 2.3 can be easily obtained for arbitrary convex cones.

3. Stochastic equations of nonnegative processes with jumps

Let E, U_0, and U_1 be complete separable metric spaces. Following Dawson and Li [10], we say that the parameters (b, σ, g_0, g_1) are *admissible* if:

- $b(0) \ge 0$ and $b(x) = b_1(x) - b_2(x)$ is defined on $\mathbb{R}_{\ge 0}$, where $x \mapsto b_1(x)$ is a continuous function, and $x \mapsto b_2(x)$ is a continuous and nondecreasing function;
- $(x, u) \mapsto \sigma(x, u)$ is a Borel function on $\mathbb{R}_{\ge 0} \times E$ satisfying $\sigma(0, u) = 0$ for $u \in E$;
- $(x, u) \mapsto g_0(x, u)$ is a Borel function on $\mathbb{R}_{\ge 0} \times U_0$ satisfying $g_0(0, u) = 0$ and $g_0(x, u) + x \ge 0$ for $x > 0$ and $u \in U_0$;
- $(x, u) \mapsto g_1(x, u)$ is a Borel function on $\mathbb{R}_{\ge 0} \times U_1$ satisfying $g_1(x, u) + x \ge 0$ for $x \in \mathbb{R}_{\ge 0}$ and $u \in U_1$.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge 0}, \mathbb{P})$ be a filtered probability space satisfying the usual hypotheses, i.e. $(\Omega, \mathcal{F}, \mathbb{P})$ is complete, the filtration $(\mathcal{F}_t)_{t\ge 0}$ is right-continuous, and $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets in $\mathcal{F}$. Let $\varkappa(\mathrm{d}z)$, $\mu_0(\mathrm{d}u)$, and $\mu_1(\mathrm{d}u)$ be σ-finite measures on E, U_0, and U_1, respectively. Let $W(\mathrm{d}t, \mathrm{d}u)$ be a $(\mathcal{F}_t)_{t\ge 0}$-Gaussian white noise on $\mathbb{R}_{\ge 0} \times E$ 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 $(\mathcal{F}_t)_{t\ge 0}$-Poisson random measures on $\mathbb{R}_{\ge 0} \times U_0$ and $\mathbb{R}_{\ge 0} \times U_1$ with intensities $\mu_0(\mathrm{d}u)$ and $\mu_1(\mathrm{d}u)$, 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 $W(\mathrm{d}s, \mathrm{d}u)$, $N_0(\mathrm{d}s, \mathrm{d}u)$, and $N_1(\mathrm{d}s, \mathrm{d}u)$ are mutually independent.

We consider a stochastic process $\{X_t : t \ge 0\}$ with state space $\mathbb{R}_{\ge 0}$ determined by the stochastic equation

$$
\begin{aligned}
X_t &= X_0 + \int_0^t b\,(X_s)\,\mathrm{d}s + \int_0^t \int_E \sigma\,(X_s, u)\,W\,(\mathrm{d}s, \mathrm{d}u) \\
&\quad + \int_0^t \int_{U_0} g_0\,(X_{s-}, u)\,\widetilde{N}_0\,(\mathrm{d}s, \mathrm{d}u) + \int_0^t \int_{U_1} g_1\,(X_{s-}, u)\,N_1\,(\mathrm{d}s, \mathrm{d}u), \quad t \ge 0,
\end{aligned} \tag{3.1}
$$

where $X_0 \ge 0$ is $\mathcal{F}_0$-measurable. A *strong solution* to (3.1) is, by definition, a nonnegative càdlàg and $(\mathcal{F}_t)_{t\ge 0}$-adapted[1] process $\{X_t : t \ge 0\}$ satisfying the equation (3.1) almost surely for every $t \ge 0$. Existence and uniqueness of strong solutions to (3.1) were studied by Dawson and Li [10], respectively, where the following conditions have been introduced:

(3.a) there is a constant $K \ge 0$ so that

$$
|b(x)| + \int_{U_1} |g_1(x, u)|\,\mu_1\,(\mathrm{d}u) \le K(1 + x), \quad x \ge 0;
$$

[1] Adapted to the augmented natural filtration generated by $W(\mathrm{d}s, \mathrm{d}u)$, $N_0(\mathrm{d}s, \mathrm{d}u)$, and $N_1(\mathrm{d}s, \mathrm{d}u)$, see, e.g., Situ [44, p.76].

(3.b) for each $u \in U_1$ the function $x \mapsto g_1(x,u) + x$ is nondecreasing and for each $m \geq 1$ there is a nondecreasing concave function $z \mapsto r_m(z)$ on $\mathbb{R}_{\geq 0}$ such that $\int_{0+} r_m(z)^{-1}\mathrm{d}z = \infty$ and

$$|b_1(x) - b_1(y)| + \int_{U_1} |g_1(x,u) - g_1(y,u)| \, \mu_1(\mathrm{d}u) \leq r_m(|x-y|), \quad 0 \leq x,\, y \leq m;$$

(3.c) for each $u \in U_0$ the function $g_0(x,u)$ is nondecreasing, and for each $m \geq 1$ there is a nonnegative and nondecreasing function $z \mapsto \rho_m(z)$ on $\mathbb{R}_{\geq 0}$ so that $\int_{0+} \rho_m(z)^{-2}\mathrm{d}z = \infty$ and

$$\int_E |\sigma(x,u) - \sigma(y,u)|^2 \, \varkappa(\mathrm{d}u) + \int_{U_0} \left(|g_0(x,u) - g_0(y,u)| \wedge |g_0(x,u) - g_0(y,u)|^2 \right) \mu_0(\mathrm{d}u)$$
$$\leq \rho_m(|x-y|)^2, \qquad 0 \leq x,\, y \leq m.$$

The next result summarizes Theorems 2.3 and 2.5 of Dawson and Li [10].

Proposition 3.1. *Suppose that (b, σ, g_0, g_1) are admissible parameters satisfying conditions (3.a)-(3.c).*

(a) *Then, for each $\mathcal{F}_0$-measurable random variable with $X_0 \geq 0$ almost surely, there exists a unique strong solution $\{X_t : t \geq 0\}$ to* (3.1).

(b) *Let X_0 and Y_0 be two $\mathcal{F}_0$-measurable nonnegative random variables. Denote by $\{X_t : t \geq 0\}$ and $\{Y_t : t \geq 0\}$ the corresponding strong solutions to* (3.1). *If $\mathbb{P}(X_0 \leq Y_0) = 1$, then $\mathbb{P}(X_t \leq Y_t$ for all $t \geq 0) = 1$.*

In the following, we write $\{X_t^x : t \geq 0\}$ for the unique strong solution of (3.1) to indicate that the process X_t starts with initial variable $X_0 = x \geq 0$ almost surely. Denote by $\mathcal{B}_b(\mathbb{R}_{\geq 0})$ the Banach space of all real-valued, bounded and Borel-measurable functions on $\mathbb{R}_{\geq 0}$ endowed with the norm $\|f\|_\infty := \sup_{x \in \mathbb{R}_{\geq 0}} |f(x)|$. Arguinig as in [28] (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 and has transition probabilities $P_t(x, \mathrm{d}y)$, i.e., it holds

$$P_t f(x) := \mathbb{E}\left[f\left(X_t^x\right)\right] = \int_{\mathbb{R}_{\geq 0}} f(y) P_t(x, \mathrm{d}y), \quad f \in \mathcal{B}_b(\mathbb{R}_{\geq 0}).$$

Applying Itô's formula one finds that $\{X_t^x : t \geq 0\}$ solves the local martingale problem with generator L given by (1.1) and domain $C_c^2(\mathbb{R}_{\geq 0})$. The adjoint 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 $0 \leq s \leq t$. Let us formulate the following conditions in addition to (2.a)-(2.c):

(3.d) it holds that $\{M_t : t \geq 0\}$ defined by

$$M_t = \int_0^t \int_E \sigma\left(X_s^x, u\right) W(\mathrm{d}s, \mathrm{d}u) + \int_0^t \int_{U_0} g_0\left(X_{s-}^x, u\right) \widetilde{N}_0(\mathrm{d}s, \mathrm{d}u), \quad t \geq 0,$$

is a martingale with respect to the filtration $(\mathcal{F}_t)_{t \geq 0}$;

(3.e) there exists a constant $A > 0$ such that, for $\widetilde{b}(x) := b(x) - \int_{U_1} g_1(x,u)\mu_1(\mathrm{d}u)$, we have

$$\widetilde{b}(y) - \widetilde{b}(x) \leq -A(y-x), \quad 0 \leq x \leq y.$$

Under the given conditions (3.a)-(3.e) we are able to show that the corresponding Markov process is exponentially ergodic in the Wasserstein distance W_1.

Theorem 3.2. *Let (b, σ, g_0, g_1) be admissible parameters satisfying conditions (3.a)-(3.e). Then, for all ϱ, $\widetilde{\varrho} \in \mathcal{P}_1(\mathbb{R}_{\geq 0})$, we have*

$$W_1\left(P_t^* \varrho, P_t^* \widetilde{\varrho}\right) \leq \mathrm{e}^{-At} W_1\left(\varrho, \widetilde{\varrho}\right), \quad t \geq 0. \tag{3.2}$$

In particular, there exists a unique invariant distribution $\pi \in \mathcal{P}(\mathbb{R}_{\geq 0})$. 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. The proof is divided into several steps.

Step 1: Let $\{X_t^x : t \geq 0\}$ and $\{X_t^y : t \geq 0\}$ be strong solutions of (3.1) with $0 \leq x \leq y$. Using Proposition 3.1 (b) together with (2.d) and (2.e), we obtain

$$\begin{aligned} \mathbb{E}\left[|X_t^x - X_t^y|\right] &= \mathbb{E}\left[X_t^y\right] - \mathbb{E}\left[X_t^x\right] \\ &= y - x + \int_0^t \mathbb{E}\left[\widetilde{b}\left(X_s^y\right) - \widetilde{b}\left(X_s^x\right)\right] \mathrm{d}s \\ &\leq |x - y| - A \int_0^t \mathbb{E}\left[|X_s^x - X_s^y|\right] \mathrm{d}s, \quad t \geq 0. \end{aligned}$$

Applying Gronwall's lemma yields

$$\mathbb{E}\left[|X_t^x - X_t^y|\right] \leq |x - y| \mathrm{e}^{-At}, \quad t \geq 0. \tag{3.3}$$

Step 2: Let us prove (3.2). We denote by δ_x and δ_y the Dirac measure concentrated in x and y, respectively. Assume first $0 \leq x \leq y$. 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 from (3.3),

$$W_1\left(P_t^* \delta_x, P_t^* \delta_y\right) \leq \mathbb{E}\left[|X_t^x - X_t^y|\right] \leq |x - y| \mathrm{e}^{-At}, \quad t \geq 0. \tag{3.4}$$

Let now H be any coupling of $(\varrho, \widetilde{\varrho})$ satisfying (2.2). Using the convexity of W_1 and (3.4), we get

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

Step 3: We prove existence of π. 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 (3.2). 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 $k \to \infty$.

We proceed to show invariance of π, i.e. $P_h^* \pi = \pi$ for all $h > 0$. Fix $h > 0$. Using the semigroup property and (3.2) 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 $W_1(P_h^* \pi, \pi) = 0$.

Step 4, uniqueness of π: Let $\widehat{\pi} \in \mathcal{P}(\mathbb{R}_{\geq 0})$ be any invariant distribution. Let $W_1^{\leq 1}$ be the Wasserstein distance given by (2.1) with $d(x, y) = 1 \wedge |x - y|$. 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

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

where the last inequality follows from (3.3). By dominated convergence we see that the right-hand side vanishes as $t \to \infty$. Consequently, $W_1^{\leq 1}(\pi, \widehat{\pi}) = 0$ which implies that $\pi = \widehat{\pi}$. The proof is complete. □

Combining the characterization of convergence with respect to W_1 and [43, Theorem 1.2 and Corollary 1.3], we deduce the following.

Corollary 3.3. *Suppose that the same conditions as in Theorem 3.2 are satisfied. Let* $\{X_t : t \geq 0\}$ *be the corresponding unique solution to* (3.1). *Then the following assertions hold:*

(a) $\lim_{t\to\infty} \mathbb{E}[X_t^x] = \int_{\mathbb{R}_{\geq 0}} x\pi(\mathrm{d}x)$.

(b) *For any* $p \in [1, \infty)$ *and* $f \in L^p(\mathbb{R}_{\geq 0}, \pi)$, *we have*

$$\frac{1}{T}\int_0^T f(X_t)\mathrm{d}t \to \int_0^\infty f(x)\pi(\mathrm{d}x), \quad T \to \infty,$$

in $L^p(\Omega, \mathcal{F}, \mathbb{P})$.

4. Continuous-state nonlinear branching processes with immigration

In this section we constitute an application of Theorem 3.2 to the Markov process with generator (1.2). For this purpose, we first provide a construction of this process as a strong solution to a certain stochastic equation.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a stochastic basis with the usual conditions rich enough to support a $(\mathcal{F}_t)_{t\geq 0}$-Gaussian white noise $W(\mathrm{d}t, \mathrm{d}u)$ with intensity measure $\mathrm{d}t\mathrm{d}u$ on $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$ and $(\mathcal{F}_t)_{t\geq 0}$-Poisson random measures $N_0(\mathrm{d}t, \mathrm{d}z, \mathrm{d}u)$ and $N_1(\mathrm{d}t, \mathrm{d}z)$ with intensity measures $\mathrm{d}tm(\mathrm{d}z)\mathrm{d}u$ on $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$ and $\mathrm{d}t\nu(\mathrm{d}z)$ on $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$, where m and ν satisfy (1.3). Further, suppose 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. Denote by $\widetilde{N}_0(\mathrm{d}t, \mathrm{d}z, \mathrm{d}u) = N_0(\mathrm{d}t, \mathrm{d}z, \mathrm{d}u) - \mathrm{d}t\mathrm{d}m(\mathrm{d}u)$ the corresponding compensated Poisson random measure. Below we provide reasonable conditions on γ_0, γ_1, and γ_2 such that

$$\begin{aligned} X_t = X_0 &+ \int_0^t \gamma_0(X_s)\,\mathrm{d}s + \int_0^t\int_0^\infty \mathbb{1}_{\{u\leq\gamma_1(X_s)\}} W(\mathrm{d}s, \mathrm{d}u) \\ &+ \int_0^t\int_0^\infty\int_0^\infty \mathbb{1}_{\{u\leq\gamma_2(X_{s-})\}} z\widetilde{N}_0(\mathrm{d}s, \mathrm{d}z, \mathrm{d}u) + \int_0^t\int_0^\infty zN_1(\mathrm{d}s, \mathrm{d}z), \end{aligned} \tag{4.1}$$

has a pathwise unique strong solution so that the results of Section 3 are applicable.

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

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

(ii) *there exists a constant* $K \geq 0$ *such that* $|\gamma_0(x)| \leq K(1 + x)$ *for all* $x \geq 0$;

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

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

Then, for any $\mathcal{F}_0$-measurable and nonnegative initial value X_0, there exists a unique strong solution $\{X_t : t \geq 0\}$ to (4.1)*.*

Proof. We are going to apply Proposition 3.1 with the following choices:

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

Now, it is easy to see that conditions (3.a) and (3.b) are satisfied. We turn to check condition (3.c). Define $l_0(x,y,u) := \mathbb{1}_{\{u \leq \gamma_2(x)\}} - \mathbb{1}_{\{u \leq \gamma_2(y)\}}$. For each $m \geq 1$, we estimate

$$\begin{aligned}\int_0^\infty \left|\mathbb{1}_{\{u \leq \gamma_1(x)\}} - \mathbb{1}_{\{u \leq \gamma_1(y)\}}\right|^2 \mathrm{d}u + \int_0^\infty \int_0^\infty \left(|zl_0(x,y,u)| \wedge |zl_0(x,y,u)|^2\right) m\,(\mathrm{d}z)\,\mathrm{d}u \\ \leq |\gamma_1(x) - \gamma_1(y)| + |\gamma_2(y) - \gamma_2(x)| \int_0^\infty \left(z \wedge z^2\right) m\,(\mathrm{d}z) \\ \leq c'_m |x - y|,\end{aligned}$$

for all $0 \leq x$, $y \leq m$ and some constant $c'_m > 0$, yielding that condition (3.c) holds for $\rho_m(z) = c_m \sqrt{z}$. □

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

Theorem 4.2. *Suppose that conditions (i) – (iii) of Theorem 4.1 are satisfied and assume the following:*

(a) *there exists a constant $A > 0$ such that*

$$\gamma_0(y) - \gamma_0(x) \leq -A(y - x), \quad 0 \leq x \leq y;$$

(b) *there exists $\lambda \in [1,2]$ and $K > 0$ such that $\gamma_1(x) + \gamma_2(x) \leq K(1+x)^\lambda$, $x \geq 0$, and*

$$\int_{\{z>1\}} z^2 m(\mathrm{d}z) + \int_{\{z>1\}} z^\lambda \nu(\mathrm{d}z) < \infty.$$

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

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

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

$$W_1\left(P_t^* \varrho, \pi\right) \leq \mathrm{e}^{-At} W_1\left(\varrho, \pi\right), \quad t \geq 0. \tag{4.2}$$

Proof. In light of Theorem 3.2 it only suffices to show that condition (3.d) is satisfied, i.e.

$$M_t := \int_0^t \int_0^\infty \mathbb{1}_{\{u \leq \gamma_1(X_s^x)\}} W(\mathrm{d}s, \mathrm{d}u) + \int_0^t \int_0^\infty \int_0^\infty z \mathbb{1}_{\{u \leq \gamma_2(X_{s-}^x)\}} \widetilde{N}\left(\mathrm{d}s, \mathrm{d}z, \mathrm{d}u\right), \quad t \geq 0$$

is a martingale with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$. In order to prove that $\{M_t : t \geq 0\}$ is a martingale, it is enough to show that $\int_0^t \mathbb{E}\left[\gamma_1(X_s^x) + \gamma_2(X_s^x)\right] ds < \infty$ which in turn is true if

$$\sup_{s\in[0,t]} \mathbb{E}[(X_s^x)^\lambda] < \infty, \quad t,\, x \geq 0. \tag{4.3}$$

Define $V_\lambda(x) = (1+x)^\lambda$, $x \geq 0$. By Itô's formula, we obtain

$$V_\lambda(X_t^x) = V(x) + \int_0^t (LV_\lambda)(X_s^x) ds + M_t\left(V_\lambda\right), \tag{4.4}$$

where, by abuse of notation, we continue to write LV_λ which is given in (1.2) and $\{M_t(V_\lambda) : t \geq 0\}$ is a local martingale given by

$$M_t(V_\lambda) = \int_0^t \int_0^\infty V_\lambda'(X_s^x) \mathbb{1}_{\{u \leq \gamma_1(X_s^x)\}} W(\mathrm{d}s, \mathrm{d}u) + \int_0^t \int_0^\infty \int_0^\infty \left(V_\lambda\left(X_s^x + z \mathbb{1}_{\{u \leq \gamma_2(X_{s-}^x)\}} \right) - V_\lambda(X_s^x) \right) \widetilde{N}(\mathrm{d}s, \mathrm{d}z, \mathrm{d}u).$$

Define, for $n \in \mathbb{N}$, a stopping time $\tau_n = \inf\{t \in \mathbb{R}_{\geq 0} : X_t > n\}$. It is easy to see that $\{M_{t\wedge\tau_n}(V_\lambda) : t \geq 0\}$ is a martingale with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$, for any $n \in \mathbb{N}$. Hence, taking the expectation in (4.4) and using Lemma A.1 from the appendix yields

$$\mathbb{E}\left[V_\lambda(X_{t\wedge\tau_n}^x)\right] \leq V_\lambda(x) + C \int_0^t \mathbb{E}\left[V_\lambda(X_{s\wedge\tau_n}^x)\right] \mathrm{d}s, \quad t \geq 0,$$

where $C > 0$ is a constant. By means of Gronwall's lemma, we estimate $\mathbb{E}\left[V(X_{t\wedge\tau_n}^x)\right] \leq V_\lambda(x)\exp(Ct)$. Noting that $\{X_t^x : t \geq 0\}$ has càdlàg paths and C is independent of n, we can take the limit $n \to \infty$ and apply the Lemma of Fatou to get $\mathbb{E}[V_\lambda(X_t^x)] \leq V_\lambda(x)\exp(Ct)$. This proves (4.3). □

We end this section by constituting 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* $\alpha \in [1, 2]$*;*
- $\gamma_2(x) = x^\delta$ *with* $\delta \in [1, 2]$.

The reader may wonder why we have apparently excluded the cases when α and δ either belong to $(0, 1)$ or $(2, \infty)$. If α, $\delta \in (0, 1)$, then we may loose uniqueness, while for α, $\delta \in (2, \infty)$ the corresponding process may have an explosion. In both cases we may try, following [29], to study solutions having a trap at 0 and/or at ∞. However, since in this case δ_0 and δ_∞ would be invariant distributions, (4.2) cannot hold in general.

5. Continuous-state branching processes with immigration

Recall that $(\beta, b, \sigma, m, \nu)$ are admissible parameters, if $\beta \geq 0$, $b \in \mathbb{R}$, and $\sigma \geq 0$ are constants, and $m(\mathrm{d}z)$ and $\nu(\mathrm{d}z)$ are Lévy-measures on $\mathbb{R}_{\geq 0}$ satisfying (1.3). For $\lambda \geq 0$, define the *branching mechanism*

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

and *immigration mechanism*

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

A Markov process with state space $\mathbb{R}_{\geq 0}$ is called a CBI process with admissible parameters $(\beta, b, \sigma, m, \nu)$ if its transition semigroup $\{P_t : t \geq 0\}$ has the representation

$$\int_0^\infty \mathrm{e}^{-\lambda z} P_t(x, \mathrm{d}z) = \exp\left(-x v_t(\lambda) - \int_0^t \psi(v_s(\lambda))\, \mathrm{d}s \right), \quad x, \lambda \geq 0, \tag{5.1}$$

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

$$\frac{\partial}{\partial t} v_t(\lambda) = -\phi(v_t(\lambda)), \quad v_0(\lambda) = \lambda. \tag{5.2}$$

It can be shown that the corresponding Markov process is a Feller process, $C_c^\infty(\mathbb{R}_{\geq 0})$ is a core for its generator, and the action of the generator is given by (1.4). Moreover, it is well-known that each CBI process can be obtained as the unique strong solution to a certain stochastic equation with jumps, see, e.g., [9, 17]. In what follows, we briefly describe this equation.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a stochastic basis satisfying the usual conditions rich enough to support the following random objects: a $(\mathcal{F}_t)_{t\geq 0}$-Gaussian white noise $W(\mathrm{d}s, \mathrm{d}u)$ with intensity $\mathrm{d}t\mathrm{d}u$, $(\mathcal{F}_t)_{t\geq 0}$-Poisson random measures $N_0(\mathrm{d}t, \mathrm{d}z, \mathrm{d}u)$ and $N_1(\mathrm{d}t, \mathrm{d}z)$ on $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$ and $\mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0}$ with intensities $\mathrm{d}tm(\mathrm{d}z)\mathrm{d}u$ and $\mathrm{d}t\nu(\mathrm{d}z)$, respectively. We denote by $\widetilde{N}_0(\mathrm{d}t, \mathrm{d}z, \mathrm{d}u)$ the compensated measure of $N_0(\mathrm{d}t, \mathrm{d}z, \mathrm{d}u)$. Suppose that $W(\mathrm{d}s, \mathrm{d}u)$, $N_0(\mathrm{d}s, \mathrm{d}z, \mathrm{d}u)$, and $N_1(\mathrm{d}s, \mathrm{d}z)$ are mutually independent. Then, for each $\mathcal{F}_0$-measurable $X_0 \geq 0$, there exists a unique strong solution to

$$\begin{aligned} X_t = X_0 &+ \int_0^t (\beta - bX_s)\,\mathrm{d}s + \sigma \int_0^t \int_0^\infty \mathbb{1}_{\{u\leq X_s\}} W(\mathrm{d}s, \mathrm{d}u) \\ &+ \int_0^t \int_0^\infty \int_0^\infty z \mathbb{1}_{\{u \leq X_{s-}\}} \widetilde{N}_0(\mathrm{d}s, \mathrm{d}z, \mathrm{d}u) + \int_0^t \int_0^\infty z N_1(\mathrm{d}s, \mathrm{d}z). \end{aligned} \tag{5.3}$$

In order to deduce this result one may, e.g., apply Proposition 3.1. Itô's formula shows that $\{X_t : t \geq 0\}$ is a Markov process whose generator is given by (1.4). Conversely, the law of any CBI process with admissible parameters $(\beta, b, \sigma, m, \nu)$ can be obtained from (5.3).

The following is a particular case of [2, Lemma 3.4].

Remark 5.1. *Let $(\beta, b, \sigma, m, \nu)$ be admissible parameters with ν satisfying*

$$\int_{\{z>1\}} z\nu(\mathrm{d}z) < \infty. \tag{5.4}$$

Then X_t^x has finite first moment given by

$$\mathbb{E}[X_t^x] = \mathrm{e}^{-bt}x + \left(\beta + \int_{\mathbb{R}_{\geq 0}} z\nu(\mathrm{d}z)\right) \int_0^t \mathrm{e}^{-bs}\mathrm{d}s, \quad t \geq 0.$$

5.1. **Exponential ergodicity in Wasserstein distance.** The aim of this subsection is to derive exponential ergodicity in the Wasserstein distance $W_{\log}$ for CBI processes. Let us start with a general characterization of the existence and uniqueness of the invariant distribution. The next theorem is proved in [30].

Theorem 5.2 (Li [30], Theorem 3.20)**.** *Let $(\beta, b, \sigma, m, \nu)$ be admissible parameters satisfying $b \geq 0$ and $\phi(\lambda) \neq 0$ for $\lambda > 0$. Then the following are equivalent:*

(a) *There exists $x \geq 0$ and $\pi \in \mathcal{P}(\mathbb{R}_{\geq 0})$ such that $P_t(x, \cdot) \to \pi$ weakly as $t \to \infty$.*
(b) *The branching and immigration mechanisms satisfy*

$$\int_0^\lambda \frac{\psi(u)}{\phi(u)}\mathrm{d}u < \infty \quad \text{for some } \lambda > 0. \tag{5.5}$$

Moreover, if either (a) or (b) is satisfied, then $P_t(x, \cdot)$ converges weakly to π as $t \to \infty$ for all $x \geq 0$, π is the unique invariant distribution and its Laplace transform is given by

$$\int_0^\infty \mathrm{e}^{-\lambda x}\pi(\mathrm{d}x) = \exp\left(-\int_0^\lambda \frac{\psi(u)}{\phi(u)}\mathrm{d}u\right), \quad \lambda \geq 0. \tag{5.6}$$

The following is due to [30, Corollary 3.21].

Remark 5.3. *If $b > 0$, then* (5.5) *is equivalent to* (1.5)*.*

As a consequence of Theorem 3.2, we find the following ergodicity result for CBI processes in the Wasserstein distance W_1.

Remark 5.4. *Let $(\beta, b, \sigma, m, \nu)$ be admissible parameters with $b > 0$ and ν satisfying (5.4). Then all assumptions of Theorem 3.2 are satisfied implying that the unique invariant distribution π given by Theorem 5.2 has finite first moment and satisfies, for all $\varrho \in \mathcal{P}_1(\mathbb{R}_{\geq 0})$,*

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

An analogue result for affine processes was recently established in [16]. Note that the class of affine processes include the (multidimensional) CBI processes as a special case. Moreover, it was shown in [16] that under the analogue of (1.5) the statement of Remark 5.4 holds true for affine processes even in the Wasserstein distance $W_{\log}$. Motivated by the fact that our main result (see Theorem 5.7 below) requires the existence of log-moments for the invariant distribution π, we recall the result of [16, Theorem 1.6 (a)] for our purpose.

Theorem 5.5. *Let $(\beta, b, \sigma, m, \nu)$ be admissible parameters and assume that $b > 0$ and (1.5) are satisfied. Let $\{P_t : t \geq 0\}$ be the corresponding transition semigroup. Then there exists a constant $C > 0$ such that, for all ϱ, $\widetilde{\varrho} \in \mathcal{P}_{\log}(\mathbb{R}_{\geq 0})$, we have*

$$W_{\log}\left(P_t^*\varrho, P_t^*\widetilde{\varrho}\right) \leq C \min\left\{\mathrm{e}^{-bt}, W_{\log}\left(\varrho, \widetilde{\varrho}\right)\right\} + C\mathrm{e}^{-bt} W_{\log}\left(\varrho, \widetilde{\varrho}\right), \quad t \geq 0. \tag{5.7}$$

In particular, the unique invariant distribution π belongs to $\mathcal{P}_{\log}(\mathbb{R}_{\geq 0})$ and satisfies

$$W_{\log}\left(P_t^*\varrho, \pi\right) \leq C \min\left\{\mathrm{e}^{-bt}, W_{\log}\left(\varrho, \pi\right)\right\} + C\mathrm{e}^{-bt} W_{\log}\left(\varrho, \pi\right), \quad t \geq 0.$$

In contrast to [16, Theorem 1.6 (a)], the proof can be significantly simplified in our context, i.e., when dealing with one-dimensional CBI processes. For convenience of the reader and in order to keep this work self-contained, we provide a sketch of the proof.

Proof of Theorem 5.5. We divide the proof in three steps.

Step 1: We show that $P_t^*(\mathcal{P}_{\log}(\mathbb{R}_{\geq 0})) \subset \mathcal{P}_{\log}(\mathbb{R}_{\geq 0})$, $t \geq 0$. Let $V(x) := \log(1+x)$, $x \geq 0$. Then it suffices to find a constant $C > 0$ such that

$$\mathbb{E}\left[V\left(X_t\right)\right] \leq Ct + \mathbb{E}[V(X_0)], \quad t \geq 0. \tag{5.8}$$

We apply Itô's formula to $V(x)$, and get

$$V\left(X_t\right) = V\left(X_0\right) + \int_0^t LV\left(X_s\right) \mathrm{d}s + M_t(V), \quad t \geq 0, \tag{5.9}$$

where LV is informally defined by (1.4) and

$$\begin{aligned} M_t(V) := \sigma & \int_0^t \int_0^\infty \mathbb{1}_{\{u \leq X_s\}} W\left(\mathrm{d}s, \mathrm{d}u\right) \\ & + \int_0^t \int_0^\infty \int_0^\infty \left(V\left(X_{s-} + z\mathbb{1}_{\{u \leq X_{s-}\}}\right) - V\left(X_{s-}\right)\right) \widetilde{N}_0\left(\mathrm{d}s, \mathrm{d}z, \mathrm{d}u\right) \\ & + \int_0^t \int_0^\infty \left(V\left(X_{s-} + z\right) - V\left(X_{s-}\right)\right) \widetilde{N}_1\left(\mathrm{d}s, \mathrm{d}z\right), \quad t \geq 0, \end{aligned}$$

where $\widetilde{N}_1(\mathrm{d}s, \mathrm{d}z)$ denotes the compensated $(\mathcal{F}_t)_{t\geq 0}$-Poisson random measure of $N_1(\mathrm{d}s, \mathrm{d}z)$. For $n \in \mathbb{N}$, define a stopping time by $\tau_n = \inf\{t \in \mathbb{R}_{\geq 0} : X_t > n\}$. Then $\{M_{t\wedge\tau_n} : t \geq 0\}$ is a martingale with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$ for any $n \in \mathbb{N}$. Hence, taking expectations in (5.9) and using that $LV \leq C$, for all $x \geq 0$ and a constant $C > 0$, see Lemma A.2 in the appendix, gives

$$\mathbb{E}\left[V\left(X_{t\wedge\tau_n}\right)\right] = \mathbb{E}[V(X_0)] + \mathbb{E}\left[\int_0^t LV\left(X_{s\wedge\tau_n}\right) \mathrm{d}s\right] \leq \mathbb{E}[V(X_0)] + Ct.$$

Noting that $\{X_t : t \geq 0\}$ has càdlàg paths and C is independent of n, we can take the limit $n \to \infty$ and apply Fatou's lemma to conclude with (5.8).

Step 2: From now on, we can proceed very close to the proof of Theorem 3.2, albeit with some slightly different estimates. The details are as follows: let $\{Q_t : t \geq 0\}$ be the transition semigroup with admissible parameters $b > 0$, $\beta = 0$, m, and $\nu \equiv 0$. Hence, for $\lambda \geq 0$, we have

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

and consequently, $P_t(x,\cdot) = Q_t(x,\cdot) * P_t(0,\cdot)$. Noting that Q_t satisfies the conditions of Theorem 3.2, we obtain from Step 1 and 2 of its proof

$$W_1(Q_t^*\delta_x, Q_t^*\delta_y) \leq \mathrm{e}^{-bt}|x-y|, \quad t \geq 0,$$

where $\{Q_t^* : t \geq 0\}$ denotes the dual semigroup of Q_t. With this, and applying Lemma 2.3 (a), we get

$$\begin{aligned} W_{\log}(P_t^*\delta_x, P_t^*\delta_y) &\leq W_{\log}(Q_t^*\delta_x, Q_t^*\delta_y) \\ &\leq \log(1 + W_1(Q_t^*\delta_x, Q_t^*\delta_y)) \leq \log\left(1 + \mathrm{e}^{-bt}|x-y|\right). \end{aligned} \tag{5.10}$$

Moreover, for a, $d \geq 0$, we use the elementary inequality

$$\log(1 + a\cdot d) \leq C\min\{a, \log(1+d)\} + Ca\log(1+d), \tag{5.11}$$

where $C > 0$ is a generic constant, to obtain from (5.10)

$$W_{\log}(P_t^*\delta_x, P_t^*\delta_y) \leq C\min\{e^{-bt}, \log(1+|x-y|)\} + Ce^{-bt}\log(1+|x-y|).$$

By using the convexity of $W_{\log}$ we get, for any $H \in \mathcal{H}(\varrho, \widetilde{\varrho})$,

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

Finally, taking H as the optimal coupling of $(\varrho, \widetilde{\varrho})$, we deduce that

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

Step 3: Let $\varrho \in \mathcal{P}_{\log}(\mathbb{R}_{\geq 0})$. Arguing similar to the proof of Theorem 3.2, we see that $(P_t^*\varrho)_{k\in\mathbb{N}} \subset \mathcal{P}_{\log}(\mathbb{R}_{\geq 0})$ is a Cauchy sequence and, thus, has a limit $\widehat{\pi} \in \mathcal{P}_{\log}(\mathbb{R}_{\geq 0})$. Clearly, $\widehat{\pi}$ is an invariant distribution and hence, by uniqueness of the invariant distribution, $\pi = \widehat{\pi} \in \mathcal{P}_{\log}(\mathbb{R}_{\geq 0})$. Eventually, noting that

$$W_{\log}(P_t^*\varrho, \pi) = W_{\log}(P_t^*\varrho, P_t^*\pi) \leq C\min\left\{\mathrm{e}^{-bt}, W_{\log}(\varrho, \pi)\right\} + C\mathrm{e}^{-bt}W_{\log}(\varrho, \pi),$$

we conclude with (5.7). This completes the proof. □

5.2. **Exponential ergodicity in total variation distance.** A general exponential ergodicity result in the total variation distance was recently obtained by Li and Ma [31], where Grey's condition was used:

(5.a) there exists $\theta > 0$ such that $\phi(\lambda) > 0$ for $\lambda > \theta$ and

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

The following was shown in [31].

Theorem 5.6 ([31]). *Let $(\beta, b, \sigma, m, \nu)$ be admissible parameters. Suppose that $b > 0$, $\nu \equiv 0$, and (5.a) is satisfied. Let $\{P_t^0 : t \geq 0\}$ be the corresponding transition semigroup given by*

$$\int_{\mathbb{R}_{\geq 0}} \mathrm{e}^{-\lambda z} P_t^0 (x, \mathrm{d}z) = \exp\left(-x v_t(\lambda) - \beta \int_0^t v_s(\lambda) \mathrm{d}s\right), \quad \lambda \geq 0, \tag{5.12}$$

where v_t is determined by (5.2). *Denote by π^0 the corresponding unique invariant distribution given by Theorem 5.2. Then there exists a constant $C > 0$ such that, for all t, x, $y \geq 0$,*

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

and

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

are satisfied. In particular, $\{P_t^0 : t \geq 0\}$ has the strong Feller property and is exponentially ergodic in the total variation distance.

Our main theorem extends the result of Li and Ma [31] to CBI processes with non-vanishing jump measure ν for immigration.

Theorem 5.7. *Let $(\beta, b, \sigma, m, \nu)$ be admissible parameters with $b > 0$ and suppose that* (1.5) *and (5.a) are satisfied. Let $\{P_t : t \geq 0\}$ be the transition semigroup given by* (5.1) *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 given by (5.12) and let $\{P_t^1 : t \geq 0\}$ be given by

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

where in both cases v_t is obtained from (5.2). By definition of the immigration mechanism we have, for all $\lambda \geq 0$,

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

yielding that $P_t(x, \cdot) = P_t^0(x, \cdot) * P_t^1(0, \cdot)$ for all t, $x \geq 0$. Combining the latter with Lemma 2.3 (b), 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 Theorem 5.6.

(b) From Theorem 5.5 we know that $\pi \in \mathcal{P}_{\log}(\mathbb{R}_{\geq 0})$. Let $\varrho \in \mathcal{P}_{\log}(\mathbb{R}_{\geq 0})$ and $H \in \mathcal{H}(\varrho, \pi)$ such that the infimum is attained (see (2.2)). Here and below we let $C > 0$ be a generic constant which may vary from line to line. Using the invariance of π combined with the convexity of the Wasserstein distance (see Lemma 2.2), 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 follows from statement (a) and $1 \wedge a \leq \log(2)^{-1}\log(1+a)$ for all $a \geq 0$. Finally, using (5.11) and the same estimates as in Step 2 of the proof of Theorem 5.5, we readily deduce that

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

□

The following remark shows that the obtained convergence has indeed exponential rate.

Remark 5.8. *Under the assumptions of Theorem 5.7, we obtain for all $x, t \geq 0$*

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

In particular, $\{P_t : t \geq 0\}$ is exponentially ergodic in the total variation distance.

5.3. **Functional central limit theorem.** A direct consequence of our ergodic result is the following strong law of large numbers in accordance with the discussion after [6, Proposition 2.5].

Corollary 5.9. *Under the conditions of Theorem 5.7, for all Borel functions $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ with $\int_0^\infty |f(x)|\pi(\mathrm{d}x) < \infty$, it holds*

$$\frac{1}{t}\int_0^t f(X_s^x)\mathrm{d}s \to \int_0^\infty f(x)\pi(\mathrm{d}x) \qquad a.s. \;\; as \; t \to \infty.$$

The latter convergence may be very useful for parameter estimation of one-dimensional CBI processes. A further consequence of our ergodicity result is the functional central limit theorem which is stated below.

Recall that the Feller semigroup $\{P_t : t \geq 0\}$ given by (5.1) has infinitesimal generator $(L, \mathrm{dom}(L))$ of the form (1.4) acting on $C_c^2(\mathbb{R}_{\geq 0})$. By virtue of [11, Theorem 2.7], $C_c^\infty(\mathbb{R}_{\geq 0})$ is a core of L and $C_c^2(\mathbb{R}_{\geq 0}) \subset \mathrm{dom}(L)$. Since $\|f\|_{L^2(\mathbb{R}_{\geq 0}, \pi)} \leq \|f\|_\infty$ for $f \in C_0(\mathbb{R}_{\geq 0})$, $C_0(\mathbb{R}_{\geq 0}) \subset L^2(\mathbb{R}_{\geq 0}, \pi)$ is dense, and P_t satisfies $\|P_t f\|_{L^2(\mathbb{R}_{\geq 0}, \pi)} \leq \|f\|_{L^2(\mathbb{R}_{\geq 0}, \pi)}$ for $f \in C_0(\mathbb{R}_{\geq 0})$, there exists a unique extension $\{\widehat{P}_t : t \geq 0\}$ on $L^2(\mathbb{R}_{\geq 0}, \pi)$. This extension is again a strongly continuous semigroup. Let $(\widehat{L}, \mathrm{dom}(\widehat{L}))$ be its infinitesimal generator. Then $\mathrm{dom}(L) \subset \mathrm{dom}(\widehat{L})$ and $Lf = \widehat{L}f$ for all $f \in \mathrm{dom}(L)$. Define $\mathrm{range}(\widehat{L}) = \{\widehat{L}f : f \in \mathrm{dom}(\widehat{L})\}$. The next result is the announced functional central limit theorem for CBI processes.

Corollary 5.10. *Under the conditions of Theorem 5.7, for all $f \in \mathrm{range}(\widehat{L})$, it holds*

$$n^{-1/2}\int_0^{nt} f(X_s^x)\,\mathrm{d}s \to W. \qquad weakly \; as \; n \to \infty,$$

where W is a one-dimensional Wiener process with zero drift and variance parameter γ given by

$$\gamma^2 = -2\int_{\mathbb{R}_{\geq 0}} \widehat{L}f(y)\cdot f(y)\pi(\mathrm{d}y)\,. \tag{5.13}$$

In general it is unlikely to find an explicit formula for (5.13) in terms of its admissible parameters. However, for the particular case $Lf_\lambda \in \mathrm{range}(\widehat{L})$, where $f_\lambda(y)=\exp(-\lambda y)$, we obtain the following.

Example 5.11. *For each $\lambda>0$ and $x\geq 0$ we see that $n^{-1/2}\int_0^{nt}(\widehat{L}f_\lambda)(X_s^x)\mathrm{d}s$ converges weakly as $n\to\infty$ to a Wiener process with zero drift and variance parameter γ given by*

$$\gamma^2 = -2\int_{\mathbb{R}_{\geq 0}} \widehat{L}f_\lambda(y)\cdot f_\lambda(y)\pi(\mathrm{d}y)\,.$$

Furthermore, an easy calculation shows that $\widehat{L}f_\lambda(y) = Lf_\lambda(y) = f_\lambda(y)(-\psi(\lambda)+y\phi(\lambda))$ and thereby

$$\gamma^2 = 2\psi(\lambda)\int_{\mathbb{R}_{\geq 0}} \mathrm{e}^{-2\lambda y}\pi(\mathrm{d}y) - 2\phi(\lambda)\int_{\mathbb{R}_{\geq 0}} y\mathrm{e}^{-2\lambda y}\pi(\mathrm{d}y)\,.$$

Recall that the Laplace transform of π is given in (5.6)*. By a change of variables $u=v_s(\lambda)$, we see that*

$$\int_0^\infty \mathrm{e}^{-\lambda y}\pi(\mathrm{d}y) = \exp\left(-\int_0^\lambda \frac{\psi(u)}{\phi(u)}\mathrm{d}u\right), \quad u\in\mathbb{R}_{\geq 0},$$

see also [27, Formula (3.27)]*. Therefore, we have*

$$\begin{aligned}
\gamma^2 &= 2\psi(\lambda)\exp\left(-\int_0^{2\lambda}\frac{\psi(u)}{\phi(u)}\mathrm{d}u\right) + \phi(\lambda)\frac{\mathrm{d}}{\mathrm{d}\lambda}\exp\left(-\int_0^{2\lambda}\frac{\psi(u)}{\phi(u)}\mathrm{d}u\right)\\
&= \left(2\psi(\lambda)+\frac{\phi(\lambda)\psi(2\lambda)}{\phi(2\lambda)}\right)\exp\left(-\int_0^{2\lambda}\frac{\psi(u)}{\phi(u)}\mathrm{d}u\right).
\end{aligned}$$

6. Continuous-state branching processes with immigration in Lévy random environments

Let (β,b,σ,m,ν) be admissible parameters, $b_E\in\mathbb{R}$, $\sigma_E\geq 0$ and μ_E a Lévy measure on $\mathbb{R}$. We start with a brief description of CBIRE processes. Let $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq 0},\mathbb{P})$ be a stochastic basis satisfying the usual conditions rich enough to support

- a $(\mathcal{F}_t)_{t\geq 0}$-Gaussian white noise $W(\mathrm{d}t,\mathrm{d}u)$ with intensity $\mathrm{d}t\mathrm{d}u$;
- a $(\mathcal{F}_t)_{t\geq 0}$-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 $\mathrm{d}tm(\mathrm{d}z)$;
- a $(\mathcal{F}_t)_{t\geq 0}$-Poisson random measure $N_1(\mathrm{d}t,\mathrm{d}z)$ on $\mathbb{R}_{\geq 0}\times\mathbb{R}_{\geq 0}$ with intensity $\mathrm{d}t\nu(\mathrm{d}z)$;
- a $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion $\{B_t : t\geq 0\}$;
- a $(\mathcal{F}_t)_{t\geq 0}$-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 two $(\mathcal{F}_t)_{t\geq 0}$-Lévy processes $\{\xi_t : t\geq 0\}$ and $\{Z_t : t\geq 0\}$ by

$$\begin{aligned}
\xi_t &= a_Et+\sigma_EB_t+\int_0^t\int_{[-1,1]} z\widetilde{M}(\mathrm{d}s,\mathrm{d}z)+\int_0^t\int_{[-1,1]^c} zM(\mathrm{d}s,\mathrm{d}z),\\
Z_t &= b_Et+\sigma_EB_t+\int_0^t\int_{[-1,1]} (e^z-1)\widetilde{M}(\mathrm{d}s,\mathrm{d}z)+\int_0^t\int_{[-1,1]^c} (e^z-1)M(\mathrm{d}s,\mathrm{d}z),
\end{aligned}$$

where $[-1,1]^c = \mathbb{R}\backslash[-1,1]$, $\widetilde{M}(\mathrm{d}s,\mathrm{d}z) := M(\mathrm{d}s,\mathrm{d}z) - \mathrm{d}t\mu_E(\mathrm{d}z)$, and the drift coefficients b_E and a_E are related by

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

Note that $\{Z_t : t \geq 0\}$ has no jump less than -1. According to [18, Theorem 5.1], we have that

$$\begin{aligned} X_t = X_0 &+ \int_0^t (\beta - bX_s)\,\mathrm{d}s + \sigma \int_0^t \int_0^\infty \mathbb{1}_{\{u\leq X_s\}} W(\mathrm{d}s,\mathrm{d}u) \\ &+ \int_0^t \int_0^\infty \int_0^\infty z\mathbb{1}_{\{u\leq X_{s-}\}} \widetilde{N}_0(\mathrm{d}s,\mathrm{d}z,\mathrm{d}u) + \int_0^t \int_0^\infty zN_1(\mathrm{d}s,\mathrm{d}z) + \int_0^t X_{s-}\mathrm{d}Z_s \end{aligned} \tag{6.1}$$

has for each $\mathcal{F}_0$-measurable random variable $X_0 \geq 0$ a pathwise unique nonnegative strong solution $\{X_t : t \geq 0\}$. It is not difficult to see that the Markov generator of $\{X_t : t \geq 0\}$ acting on $C_c^\infty(\mathbb{R}_{\geq 0})$ is given by $L_0 + L_1$, where L_0 is defined by (1.4) and L_1 by (1.7), respectively. In view of [18, Theorem 5.4], the Markov process $\{X_t : t \geq 0\}$ has Feller transition semigroup $\{P_t : t \geq 0\}$ and its transition probabilities $P_t(x,\mathrm{d}y)$ satisfy

$$\int_{\mathbb{R}_{\geq 0}} \mathrm{e}^{-\lambda y} P_t(x,\mathrm{d}y) = \mathbb{E}\left[\exp\left(-xv_{0,t}^\xi(\lambda) - \int_0^t \psi\left(v_{s,t}^\xi(\lambda)\right)\mathrm{d}s\right)\right], \tag{6.2}$$

where ϕ and ψ are the corresponding branching and immigration mechanisms for the CBI process with admissible parameters (β,b,σ,m,ν) and $r \mapsto v_{r,t}^\xi(\lambda)$ is the pathwise unique nonnegative solution to

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

Existence of limiting distribution was recently characterized in [18]. In the following we present sufficient conditions for both the ergodicity in the Wasserstein and total variation distance based on an application of Theorem 3.2. We start with the simpler case of ergodicity in Wasserstein distance W_1.

Theorem 6.1. *Let (β,b,σ,m,ν) be admissible parameters, $b_E \in \mathbb{R}$, $\sigma_E \geq 0$ and μ_E a Lévy measure on $\mathbb{R}$. Let $\{P_t : t \geq 0\}$ be the transition semigroup with transition probabilities defined by (6.2) and denote by $\{P_t^* : t \geq 0\}$ the dual semigroup. Suppose that*

$$\int_{(1,\infty)} z\nu(\mathrm{d}z) + \int_{(1,\infty)} \mathrm{e}^z\mu_E(\mathrm{d}z) < \infty \quad \text{and} \quad b > \mathbb{E}[Z_1]. \tag{6.3}$$

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

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

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 $\varrho \in \mathcal{P}_1(\mathbb{R}_{\geq 0})$,

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

Proof. Let us first verify that (6.1) is a particular case of (3.1). 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 $\varkappa(\mathrm{d}y,\mathrm{d}u) = \delta_1(\mathrm{d}y)\mathrm{d}u + \delta_2(\mathrm{d}y)\delta_0(\mathrm{d}u)$ on E. Then $\mathcal{W}(\mathrm{d}s,\mathrm{d}y,\mathrm{d}u) := \delta_1(\mathrm{d}y)W(\mathrm{d}s,\mathrm{d}u) + \mathrm{d}B_s\delta_2(\mathrm{d}y)\delta_0(\mathrm{d}u)$ defines an $(\mathcal{F}_t)_{t\geq 0}$-Gaussian white noise on $\mathbb{R}_{\geq 0}\times E$ with intensity measure $\mathrm{d}s\varkappa(\mathrm{d}y,\mathrm{d}u)$. Let $\sigma(x,y,u) := \sigma\mathbb{1}_{\{y=1\}}\mathbb{1}_{\{u\leq x\}} + \sigma_E\mathbb{1}_{\{y=2\}}\mathbb{1}_{\{0\}}(u)x$ for $(x,y,u) \in \mathbb{R}_{\geq 0} \times E$. We see that

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

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

$$\begin{aligned}\int_0^t \int_{U_0} g_0(X_s, y, z, u)\widetilde{\mathcal{N}}_0(\mathrm{d}s, \mathrm{d}y, \mathrm{d}z, \mathrm{d}u) &= \int_0^t \int_0^\infty \int_0^\infty z\mathbb{1}_{\{u\leq X_s\}}\widetilde{N}_0(\mathrm{d}s, \mathrm{d}z, \mathrm{d}u) \\ &\quad + \int_0^t \int_{-1}^1 (\mathrm{e}^z - 1)\, X_s\widetilde{M}(\mathrm{d}s, \mathrm{d}z)\end{aligned}$$

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

$$\int_0^t \int_{U_1} g_1(X_s, y, z)\mathcal{N}_1(\mathrm{d}s, \mathrm{d}y, \mathrm{d}z) = \int_0^t \int_0^\infty zN_1(\mathrm{d}s, \mathrm{d}z) + \int_0^t \int_{[-1,1]^c} (\mathrm{e}^z - 1)\, X_sM(\mathrm{d}s, \mathrm{d}z).$$

This shows that (6.1) is indeed a particular case of (3.1). It is not difficult to see that conditions (3.a)-(3.e) are satisfied which completes the proof. □

Theorem 6.2. *Let $(\beta, b, \sigma, m, \nu)$ be admissible parameters, $b_E \in \mathbb{R}$, $\sigma_E \geq 0$ and μ_E a Lévy measure on $\mathbb{R}$. Let $\{P_t : t \geq 0\}$ be the transition semigroup with transition probabilities defined by (6.2) and denote by $\{P_t^* : t \geq 0\}$ the dual semigroup. Suppose that (6.3) and Grey's condition (5.a) is satisfied. Let π be the unique invariant distribution. Then, for any $\varrho \in \mathcal{P}_1(\mathbb{R}_{\geq 0})$,*

$$\|P_t^*\varrho - \pi\|_{TV} \leq 2\mathbb{E}\left[\overline{v}_{0,t}^{\xi}\right] W_1(\varrho, \pi), \quad t \geq 0, \tag{6.4}$$

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

$$\lim_{t\to\infty} \|P_t^*\varrho - \pi\|_{TV} = 0. \tag{6.5}$$

Proof. As a consequence of Grey's condition, [18, 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 literally as in the proof of [18, Theorem 4.5], we observe that, for $0 \leq x \leq y$,

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

Taking the supremum over all $f \in \mathcal{B}_b(\mathbb{R}_{\geq 0})$ 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 $\lim_{t\to\infty} \overline{v}_{0,t}^{\xi} = 0$ in view of [18, Corollary 4.4] and, thus, (6.5) follows from dominated convergence. Finally, we conclude with the estimate (6.4) by estimating

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

where we chose 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 [38] 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 [34].

Appendix

Lemma A.1. *Let γ_0, γ_1, γ_2 be Borel functions on $\mathbb{R}_{\geq 0}$ and m, ν Borel measures on $\mathbb{R}_{\geq 0}$ satisfying* (1.3)*. Suppose that conditions (i) – (iii) of Theorem 4.1 and condition (b) of Theorem 4.2 are satisfied. For $\lambda \in [1,2]$, define $V_\lambda(x) = (1+x)^\lambda$, $x \geq 0$. Then there exists a constant $C > 0$ such that*

$$LV_\lambda(x) \leq CV_\lambda(x), \quad x \geq 0,$$

where the operator L is given in (1.2)*.*

Proof. Defining the operators

$$\begin{aligned} DV_\lambda(x) &:= \gamma_0(x)V_\lambda'(x) + \frac{\gamma_1(x)}{2}V_\lambda''(x), \\ J_mV_\lambda(x) &:= \gamma_2(x)\int_{\mathbb{R}_{\geq 0}} \left(V_\lambda(x+z) - V_\lambda(x) - zV_\lambda'(x)\right) m\,(\mathrm{d}z)\,, \\ J_\nu V_\lambda(x) &:= \int_{\mathbb{R}_{\geq 0}} \left(V_\lambda(x+z) - V_\lambda(x)\right) \nu\,(\mathrm{d}z)\,, \end{aligned}$$

we see that $LV_\lambda = DV_\lambda + J_mV_\lambda + J_\nu V_\lambda$. Moreover, it holds that $V_\lambda'(x) = \lambda(1+x)^{\lambda-1}$ and $V_\lambda''(x) = \lambda(\lambda-1)(1+x)^{\lambda-2}$. In the following $C > 0$ denotes some generic constant which may vary from line to line. The drift can be easily estimated by

$$DV_\lambda(x) \leq C(1+x)(1+x)^{\lambda-1} + C(1+x)^\lambda(1+x)^{\lambda-2} \leq CV_\lambda(x).$$

For the state-dependent jumps, by using the mean-value theorem twice, we get

$$V_\lambda(x+z) - V_\lambda(x) - zV_\lambda'(x) = z^2\int_0^1 (1-s)V_\lambda''(x+sz)\mathrm{d}s \leq Cz^2,$$

and hence $J_mV_\lambda(x) \leq CV_\lambda(x)$. Turning to the state-independent jumps, we use the mean-value theorem to show that

$$V_\lambda(x+z) - V_\lambda(x) \leq C\left(\mathbb{1}_{(0,1]}(z)z + \mathbb{1}_{[1,\infty)}(z)z^\lambda\right)V_\lambda(x),$$

which implies $J_\nu V_\lambda(x) \leq CV_\lambda(x)$. Collecting the estimates for DV_λ, J_mV_λ, and $J_\nu V_\lambda$ proves the asserted. □

Lemma A.2. *Let $(\beta, b, \sigma, m, \nu)$ be admissible parameters and suppose that* (1.5) *holds. Let $V(x) := \log(1+x)$, $x \geq 0$. Then there exists a constant $C > 0$ such that*

$$LV(x) \leq C, \quad x \geq 0,$$

where the operator L is given in (1.4)*.*

Proof. Let us introduce the operators

$$\begin{aligned} DV(x) &:= (\beta - bx)\,V'(x) + \sigma xV''(x); \\ J_mV(x) &:= x\int_0^\infty \left(V(x+z) - V(x) - zV'(x)\right) m\,(\mathrm{d}z)\,; \\ J_\nu V(x) &:= \int_0^\infty \left(V(x+z) - V(x)\right) \nu\,(\mathrm{d}z)\,. \end{aligned}$$

So $LV = DV + J_m V + J_\nu V$. We now estimate DV, $J_m V$, and $J_\nu V$ separately. Concerning DV, it is easy to see that $DV(x) \le \frac{\beta}{1+x} - \frac{bx}{1+x} \le \beta + |b|$. Turning to $J_\nu V$, we first note that

$$V(x+z) - V(x) = \log\left(1 + \frac{z}{1+x}\right) \le \frac{z}{1+x} \le z \tag{A.1}$$

on the one hand and

$$V(x+z) - V(x) = \log\left(1 + \frac{z}{1+x}\right) \le \log(1+z)$$

on the other hand. Having established the latter inequalities, we obtain

$$J_\nu V(x) \le \int_0^\infty \left(z\mathbb{1}_{\{0<z\le 1\}} + \log(1+z)\mathbb{1}_{\{z>1\}}\right) \nu(\mathrm{d}z) < \infty.$$

For $J_m V$, we decompose it further as $J_m V = J_{m,*}V + J_m^* V$, where

$$J_{m,*}V(x) := x\int_{\{0<z\le 1\}} \left(V(x+z) - V(x) - zV'(x)\right) m(\mathrm{d}z),$$

$$J_m^* V(x) := x\int_{\{z>1\}} \left(V(x+z) - V(x) - zV'(x)\right) m(\mathrm{d}z).$$

For J_m^* we use (A.1) to obtain

$$J_m^* V(x) \le \frac{x}{1+x}\int_{\{z>1\}} zm(\mathrm{d}z) \le \int_{\{z>1\}} zm(\mathrm{d}z) < \infty.$$

For $J_{m,*}$ we use the mean value theorem and

$$V'(x+rz) - V'(x) = \frac{1}{x+rz} - \frac{1}{x} = -\frac{rz}{x(x+rz)} \le 0,$$

where $r \in [0,1]$, to obtain

$$J_{m,*}V(x) = x\int_{\{0<z\le 1\}} \left(\int_0^1 \left(V'(x+rz) - V'(x)\right) z\mathrm{d}r\right) m(\mathrm{d}z) \le 0.$$

Combining the estimates for DV, $J_{m,*}V$, $J_m^* V$, and $J_\nu V$ yields the asserted estimate. □

References

1. Mátyás Barczy, Mohamed Ben Alaya, Ahmed Kebaier, and Gyula Pap, *Asymptotic properties of maximum likelihood estimator for the growth rate for a jump-type CIR process based on continuous time observations*, Stochastic Process. Appl. **128** (2018), no. 4, 1135–1164. MR 3769658
2. Mátyás Barczy, Zenghu Li, and Gyula Pap, *Stochastic differential equation with jumps for multi-type continuous state and continuous time branching processes with immigration*, ALEA Lat. Am. J. Probab. Math. Stat. **12** (2015), no. 1, 129–169. MR 3340375
3. Jean-Baptiste Bardet, Alejandra Christen, Arnaud Guillin, Florent Malrieu, and Pierre-André Zitt, *Total variation estimates for the TCP process*, Electron. J. Probab. **18** (2013), no. 10, 21. MR 3035738
4. Mohamed Ben Alaya and Ahmed Kebaier, *Parameter estimation for the square-root diffusions: ergodic and nonergodic cases*, Stoch. Models **28** (2012), no. 4, 609–634. MR 2995525
5. ______, *Asymptotic behavior of the maximum likelihood estimator for ergodic and nonergodic square-root diffusions*, Stoch. Anal. Appl. **31** (2013), no. 4, 552–573. MR 3175784
6. R. N. Bhattacharya, *On the functional central limit theorem and the law of the iterated logarithm for Markov processes*, Z. Wahrsch. Verw. Gebiete **60** (1982), no. 2, 185–201. MR 663900
7. Oleg Butkovsky, *Subgeometric rates of convergence of markov processes in the Wasserstein metric*, Ann. Appl. Probab. **24** (2014), no. 2, 526–552. MR 3178490
8. Djalil Chafaï, Florent Malrieu, and Katy Paroux, *On the long time behavior of the TCP window size process*, Stochastic Process. Appl. **120** (2010), no. 8, 1518–1534. MR 2653264
9. D. A. Dawson and Zenghu Li, *Skew convolution semigroups and affine Markov processes*, Ann. Probab. **34** (2006), no. 3, 1103–1142. MR 2243880

10. Donald A. Dawson and Zenghu Li, *Stochastic equations, flows and measure-valued processes*, Ann. Probab. **40** (2012), no. 2, 813–857. MR 2952093
11. D. Duffie, D. Filipović, and W. Schachermayer, *Affine processes and applications in finance*, Ann. Appl. Probab. **13** (2003), no. 3, 984–1053.
12. Andreas Eberle, *Reflection couplings and contraction rates for diffusions*, Probab. Theory Related Fields **166** (2016), no. 3-4, 851–886. MR 3568041
13. Stewart N. Ethier and Thomas G. Kurtz, *Markov processes: Characterization and convergence*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1986. MR 838085 (88a:60130)
14. William Feller, *Diffusion processes in genetics*, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950, University of California Press, Berkeley and Los Angeles, 1951, pp. 227–246. MR 0046022
15. Nicolas Fournier, *On pathwise uniqueness for stochastic differential equations driven by stable Lévy processes*, Ann. Inst. Henri Poincaré Probab. Stat. **49** (2013), no. 1, 138–159. MR 3060151
16. Martin Friesen, Peng Jin, and Barbara Rüdiger, *Stochastic equation and exponential ergodicity in Wasserstein distances for affine processes*, (2018).
17. Zongfei Fu and Zenghu Li, *Stochastic equations of non-negative processes with jumps*, Stochastic Process. Appl. **120** (2010), no. 3, 306–330. MR 2584896 (2011d:60178)
18. Hui He, Zenghu Li, and Wei Xu, *Continuous-state branching processes in Lévy random environments*, J. Theoret. Probab. **31** (2018), no. 4, 1952–1974. MR 3866603
19. Peng Jin, Jonas Kremer, and Barbara Rüdiger, *Moments and ergodicity of the jump-diffusion CIR process*, arXiv e-prints (2017), arXiv:1709.00969.
20. ———, *Existence of limiting distribution for affine processes*, arXiv e-prints (2018), arXiv:1812.05402.
21. Peng Jin, Barbara Rüdiger, and Chiraz Trabelsi, *Exponential ergodicity of the jump-diffusion CIR process*, Stochastics of environmental and financial economics—Centre of Advanced Study, Oslo, Norway, 2014–2015, Springer Proc. Math. Stat., vol. 138, Springer, Cham, 2016, pp. 285–300. MR 3451177
22. ———, *Positive Harris recurrence and exponential ergodicity of the basic affine jump-diffusion*, Stoch. Anal. Appl. **34** (2016), no. 1, 75–95. MR 3437080
23. Miloslav Jiřina, *Stochastic branching processes with continuous state space*, Czechoslovak Math. J. **8 (83)** (1958), 292–313. MR 0101554
24. K. Kawazu and S. Watanabe, *Branching processes with immigration and related limit theorems*, Theory of Probability & Its Applications **16** (1971), no. 1, 3654.
25. Martin Keller-Ressel, *Moment explosions and long-term behavior of affine stochastic volatility models*, Math. Finance **21** (2011), no. 1, 73–98. MR 2779872 (2012e:91126)
26. Martin Keller-Ressel and Aleksandar Mijatović, *On the limit distributions of continuous-state branching processes with immigration*, Stochastic Process. Appl. **122** (2012), no. 6, 2329–2345. MR 2922631
27. Martin Keller-Ressel and Thomas Steiner, *Yield curve shapes and the asymptotic short rate distribution in affine one-factor models*, Finance Stoch. **12** (2008), no. 2, 149–172. MR 2390186 (2009c:60222)
28. Pei-Sen Li, *A continuous-state polynomial branching process*, Stochastic Processes and their Applications (2018).
29. Pei-Sen Li, Xu Yang, and Xiaowen Zhou, *A general continuous-state nonlinear branching process*, arXiv e-prints (2017), arXiv:1708.01560.
30. Zenghu Li, *Measure-valued branching Markov processes*, Probability and its Applications (New York), Springer, Heidelberg, 2011. MR 2760602
31. Zenghu Li and Chunhua Ma, *Asymptotic properties of estimators in a stable Cox-Ingersoll-Ross model*, Stochastic Process. Appl. **125** (2015), no. 8, 3196–3233. MR 3343292
32. Zenghu Li and Leonid Mytnik, *Strong solutions for stochastic differential equations with jumps*, Ann. Inst. Henri Poincaré Probab. Stat. **47** (2011), no. 4, 1055–1067. MR 2884224
33. Zenghu Li and Fei Pu, *Strong solutions of jump-type stochastic equations*, Electron. Commun. Probab. **17** (2012), no. 33, 13. MR 2965746
34. Zenghu Li and Wei Xu, *Asymptotic results for exponential functionals of Lévy processes*, Stochastic Process. Appl. **128** (2018), no. 1, 108–131. MR 3729532
35. Hiroki Masuda, *Ergodicity and exponential β-mixing bounds for multidimensional diffusions with jumps*, Stochastic Process. Appl. **117** (2007), no. 1, 35–56. MR 2287102
36. Sean Meyn and Richard L. Tweedie, *Markov chains and stochastic stability*, second ed., Cambridge University Press, Cambridge, 2009, With a prologue by Peter W. Glynn. MR 2509253
37. Sean P. Meyn and R. L. Tweedie, *Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes*, Adv. in Appl. Probab. **25** (1993), no. 3, 518–548. MR 1234295 (94g:60137)

38. S. Palau and J. C. Pardo, *Continuous state branching processes in random environment: the Brownian case*, Stochastic Process. Appl. **127** (2017), no. 3, 957–994. MR 3605717
39. ______, *Branching processes in a Lévy random environment*, Acta Appl. Math. **153** (2018), 55–79. MR 3745730
40. Étienne Pardoux, *Probabilistic models of population evolution*, Mathematical Biosciences Institute Lecture Series. Stochastics in Biological Systems, vol. 1, Springer, [Cham]; MBI Mathematical Biosciences Institute, Ohio State University, Columbus, OH, 2016, Scaling limits, genealogies and interactions. MR 3496029
41. Xuhui Peng and Rangrang Zhang, *Exponential ergodicity for SDEs under the total variation*, J. Evol. Equ. **18** (2018), no. 3, 1051–1067. MR 3859440
42. Mark A. Pinsky, *Limit theorems for continuous state branching processes with immigration*, Bull. Amer. Math. Soc. **78** (1972), 242–244. MR 0295450
43. Nikola Sandrić, *A note on the Birkhoff ergodic theorem*, Results Math. **72** (2017), no. 1-2, 715–730. MR 3684455
44. Rong Situ, *Theory of stochastic differential equations with jumps and applications*, Mathematical and Analytical Techniques with Applications to Engineering, Springer, New York, 2005, Mathematical and analytical techniques with applications to engineering. MR 2160585 (2006e:60004)
45. Cédric Villani, *Optimal transport*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 338, Springer-Verlag, Berlin, 2009, Old and new. MR 2459454
46. Jian Wang, *L^p-Wasserstein distance for stochastic differential equations driven by Lévy processes*, Bernoulli **22** (2016), no. 3, 1598–1616. MR 3474827

(Martin Friesen) Fakultät für Mathematik und Naturwissenschaften, Bergische Universität Wuppertal, 42119 Wuppertal, Germany
E-mail address, Martin Friesen: `friesen@math.uni-wuppertal.de`

(Peng Jin) Department of Mathematics, Shantou University, Shantou, Guangdong 515063, China
E-mail address: `pjin@stu.edu.cn`

(Jonas Kremer) Fakultät für Mathematik und Naturwissenschaften, Bergische Universität Wuppertal, 42119 Wuppertal, Germany
E-mail address: `kremer@math.uni-wuppertal.de`

(Barbara Rüdiger) Fakultät für Mathematik und Naturwissenschaften, Bergische Universität Wuppertal, 42119 Wuppertal, Germany
E-mail address: `ruediger@uni-wuppertal.de`