J. Appl. Probab. **60,** 812–834 (2023)
*doi:*10.1017/jpr.2022.100

STRONG FELLER AND ERGODIC PROPERTIES OF THE (1+1)-AFFINE PROCESS

SHUKAI CHEN,* *Fujian Normal University*
ZENGHU LI,** *Beijing Normal University*

Abstract

We prove some estimates for the variations of transition probabilities of the (1+1)-affine process. From these estimates we deduce the strong Feller and the ergodic properties of the total variation distance of the process. The key strategy is the pathwise construction and analysis of several Markov couplings using strong solutions of stochastic equations.

Keywords: Affine process; strong Feller property; ergodicity; total variation distance; coupling

2020 Mathematics Subject Classification: Primary 60J80
Secondary 60J85; 60H15

1. Introduction

Let $m \geq 0$ and $n \geq 0$ be integers. A time-homogeneous $(m+n)$-dimensional Markov processes $\{X_t : t \geq 0\} = \{(Y_t, Z_t) : t \geq 0\}$ taking values in $D := \mathbb{R}_+^m \times \mathbb{R}^n$ is called an affine Markov process if its characteristic function satisfies

$$\mathbb{E}\big(e^{i\langle X_t, u\rangle} | X_0 = x\big) = \exp\{\langle x, \psi(t, iu)\rangle + \phi(t, iu)\}, \quad x \in D, u \in \mathbb{R}^{m+n}, \tag{1.1}$$

where ϕ and ψ satisfy certain generalized Riccati differential equations. The affine property means roughly that the logarithm of the characteristic function is affine with respect to the initial state. In this case, it is known that the m-dimensional process $\{Y_t : t \geq 0\}$ is a *continuous-state branching process with immigration* (CBI process). The n-dimensional process $\{Z_t : t \geq 0\}$ can be regarded as an Ornstein–Uhlenbeck-type process (OU-type process) depending on $\{Y_t : t \geq 0\}$. Then the above formulation includes as special cases both the CBI process and the OU-type process. A one-dimensional CBI process first appeared in the scaling limit theorem for discrete Galton–Watson branching processes with immigration established in Kawazu and Watanabe [12]. Compared with the discrete model, the CBI process is easier to deal with because its time and state spaces are both smooth, and the distributions that appear are infinitely divisible. For general treatments and backgrounds on branching processes in continuous state spaces, the reader may refer to Kyprianou [14] and Li [15, 17]. The affine processes involve rich common mathematical structures and have found interesting connections and applications in several areas.

Received 16 November 2020; revision received 22 August 2022.
* Postal address: School of Mathematics and Statistics, Fujian Normal University, Fuzhou 350007, People's Republic of China. Email address: skchen@mail.bnu.edu.cn
** Postal address: Laboratory of Mathematics and Complex Systems, School of Mathematical Sciences, Beijing Normal University, Beijing 100875, People's Republic of China. Email address: lizh@bnu.edu.cn

The general theory of finite-dimensional affine Markov processes, including several equivalent characterizations and common financial applications, was given by Duffie *et al.* [6] under a regularity assumption. The regularity problem asks whether this property holds automatically for stochastically continuous affine processes. This property was established in Dawson and Li [4] under the first moment condition. The problem was finally settled in Keller-Ressel *et al.* [13], where it was proved that any stochastically continuous affine Markov process is regular. The connection of the regularity problem with Hilbert's fifth problem is explained in Keller-Ressel *et al.* [13].

The strong Feller and ergodic properties of the CBI and OU-type processes have been studied by a number of authors. In particular, a sufficient and necessary integrability condition for the ergodicity of a one-dimensional subcritical or critical CBI process was announced in Pinsky [23]; see Li [15] for a proof of the result. The strong Feller property and exponential ergodicity in the total variation distance of one-dimensional CBI processes were established in Li and Ma [16] by using a coupling process constructed using strong solutions of a stochastic equation; see also Li [17]. The analytic properties of a finite-dimensional stable jump-type CBI process were studied by Friesen and Jin [7], who proved that the transition kernel of the process satisfies an a priori bound in a weighted anisotropic Besov norm. From this regularity they deduced the strong Feller property and proved in the subcritical case the exponential ergodicity in the total variation distance; see also Jin *et al.* [10]. The strong Feller and ergodic properties of Dawson–Watanabe superprocesses with or without immigration were proved in the recent work of Li [18] using coupling methods, which generalize the work of Li and Ma [16].

It was proved in Sato and Yamazato [26] that a finite-dimensional OU-type process is ergodic if and only if the eigenvalues of its coefficient matrix have strictly negative real parts. The coupling property and strong Feller property of finite-dimensional OU-type processes were studied in Priola and Zabczyk [24] and Wang [28]. The ergodicity and exponential ergodicity of such processes in the total variation distance were proved in Schilling and Wang [27] and Wang [29].

Barczy *et al.* [3] studied the existence and uniqueness of a stationary distribution for a special subcritical two-factor affine process, where the first factor was an α-stable CBI process and the second one was driven by a Brownian motion. The exponential ergodicity of the process in the total variation distance was established in Jin *et al.* [9]. For a subcritical two-factor affine process driven by Lévy stable processes, the exponential ergodicity in the L^1-Wasserstein distance was established in Bao and Wang [2] by a coupling approach. For general finite-dimensional affine Markov processes, Jin *et al.* [11] proved a sufficient condition for the ergodicity in weak convergence, which covers partially the results of Pinsky [23] and Sato and Yamazato [26]. The necessity of the condition of Jin *et al.* [11] was still an open problem. The exponential ergodicities in two suitably chosen Wasserstein distances for the process were established in Friesen *et al.* [8] by coupling methods. Some results on the ergodicity and exponential ergodicity in the total variation distance for affine processes driven by Brownian motions and compound Poisson processes were given by Zhang and Glynn [30]. For general affine processes on cones, the exponential ergodicity was studied by Mayerhofer *et al.* [19] under certain irreducibility, aperiodicity, and finite second moment assumptions.

In this paper, we prove some estimates for the variations of transition probabilities of the (1+1)-affine process. From those estimates we deduce the strong Feller and the ergodic properties in the total variation distance of the process. The key strategy is to construct several Markov couplings using strong solutions of stochastic equations, which naturally extend those of the CBI process and the OU-type process introduced by Li and Ma [16], Schilling and Wang

https://doi.org/10.1017/jpr.2022.100 Published online by Cambridge University Press

[27], and Wang [29]. The stochastic equations established by Dawson and Li [4, 5] provide an efficient method for the pathwise construction and analysis of the couplings, which are of interest in themselves; see, e.g., Friesen *et al.* [8, p. 2170] and Jin *et al.* [9, p. 1145]. For simplicity, here we only discuss the (1+1)-dimensional process. The method can be modified to treat general finite-dimensional affine processes by some extra work, which will be addressed separately.

The paper is organized as follows. In Section 2, we give the definition and some basic properties of the (1+1)-affine process. The key estimates for the variations of the transition probabilities are established in Section 3, where the strong Feller property and the exponential ergodicity are also deduced. In Section 4, an ergodicity result is proved under a weaker condition.

2. The affine process

Let us introduce more precisely the (1+1)-affine process. Here we adopt the framework of Duffie *et al.* [6]; see also Dawson and Li [4]. Let $D=\mathbb{R}_+\times\mathbb{R}$ be endowed with its Borel σ-algebra.

Definition 2.1. A set of parameters $(a, (\alpha_{ij}), (b_1, b_2), (\beta_{ij}), \mu, \nu)$ is called admissible if the following hold:

(i) $a\in\mathbb{R}_+$ is a constant;

(ii) $\alpha=(\alpha_{ij})$ is a symmetric nonnegative definite (2×2) matrix;

(iii) $b=(b_1, b_2)\in D$ is a vector;

(iv) $\beta=(\beta_{ij})$ is a (2×2) matrix with $\beta_{12}=0$;

(v) $\mu(\mathrm{d}v)=\mu(\mathrm{d}v_1, \mathrm{d}v_2)$ is a σ-finite measure on D, supported on $D\setminus\{0\}$, such that

$$\int_D \big(v_1\wedge v_1^2+|v_2|\wedge|v_2|^2\big)\mu(\mathrm{d}v)<\infty;$$

(vi) $\nu(\mathrm{d}v)=\nu(\mathrm{d}v_1, \mathrm{d}v_2)$ is a σ-finite measure on D, supported on $D\setminus\{0\}$, such that

$$\int_D \big(v_1+|v_2|\wedge|v_2|^2\big)\nu(\mathrm{d}v)<\infty.$$

Let $U=\mathbb{C}_-\times i\mathbb{R}$, where $\mathbb{C}_-=\{a+ib: a\in\mathbb{R}_-, b\in\mathbb{R}\}$ and $i\mathbb{R}=\{ia: a\in\mathbb{R}\}$. Given a set of admissible parameters $(a, (\alpha_{ij}), (b_1, b_2), (\beta_{ij}), \mu, \nu)$, we define the functions F and R on U by the Lévy–Khintchine type representations

$$F(u)=\langle b, u\rangle+au_2^2+\int_D \big(\mathrm{e}^{\langle u,v\rangle}-1-v_2u_2\big)\nu(\mathrm{d}v) \tag{2.1}$$

and

$$R(u)=\langle \beta_{\cdot 1}, u\rangle+\langle u, \alpha u\rangle+\int_D \big(\mathrm{e}^{\langle u,v\rangle}-1-\langle u, v\rangle\big)\mu(\mathrm{d}v). \tag{2.2}$$

The *(1+1)-affine process* $\{X_t: t\geq 0\}$ is a Markov process with state space D and Feller transition semigroup $(P_t)_{t\geq 0}$ defined by

$$\int_D \mathrm{e}^{\langle u,v\rangle}P_t(x, \mathrm{d}v)=\exp\big\{\langle x, \psi(t, u)\rangle+\phi(t, u)\big\}, \quad u\in U, \tag{2.3}$$

https://doi.org/10.1017/jpr.2022.100 Published online by Cambridge University Press

where (ϕ, ψ) is the unique solution of the system of equations

$$\begin{cases} \partial_t \phi(t, u) = F(\psi(t, u)), & \phi(0, u) = 0, \\ \partial_t \psi_1(t, u) = R(\psi(t, u)), & \psi_1(0, u) = u_1, \\ \psi_2(t, u) = \mathrm{e}^{\beta_{22} t} u_2. \end{cases} \tag{2.4}$$

The uniqueness of the solution implies that

$$\psi(s + t, u) = \psi(s, \psi(t, u)), \quad s, t \geq 0, \; u \in U. \tag{2.5}$$

Clearly, we can rewrite (2.3) as

$$\int_D \mathrm{e}^{\langle u, v\rangle} P_t(x, \mathrm{d}v) = \exp\left\{\langle x, \psi(t, u)\rangle + \int_0^t F(\psi(t, u))\mathrm{d}s\right\}, \quad u \in U. \tag{2.6}$$

The Feller property implies that $\{X_t : t \geq 0\}$ has a càdlàg realization.

Suppose that $\{X_t : t \geq 0\} = \{(Y_t, Z_t) : t \geq 0\}$ is a (1+1)-affine process with transition semigroup $(P_t)_{t\geq 0}$ defined by (2.3) and (2.4). Then $\{Y_t : t \geq 0\}$ is a Markov process on $\mathbb{R}_+$ with Feller transition semigroup $(P_t^{(1)})_{t\geq 0}$ defined by

$$\int_{\mathbb{R}_+} \mathrm{e}^{-\lambda v} P_t^{(1)}(x, \mathrm{d}v) = \exp\left\{x\psi_1(t, -\lambda, 0) + \int_0^t F(\psi_1(t, -\lambda, 0), 0)\mathrm{d}s\right\}, \tag{2.7}$$

where $\lambda \geq 0$. It is known that $\{Y_t : t \geq 0\}$ is a *continuous-state branching process with immigration* (CBI process) with *branching mechanism* $\lambda \mapsto -R(-\lambda, 0)$ and *immigration mechanism* $\lambda \mapsto -F(-\lambda, 0)$. It is known that

$$\int_{\mathbb{R}_+} v P_t^{(1)}(x, \mathrm{d}v) = \mathrm{e}^{\beta_{11} t} y + \left[b_1 + m_1(\nu)\right] \int_0^t \mathrm{e}^{\beta_{11} s}\mathrm{d}s, \tag{2.8}$$

where

$$m_1(\nu) = \int_D v_1 \nu(\mathrm{d}v);$$

see, e.g., the formula (79) in Li [17].

In particular, when $F(-\lambda, 0) \equiv 0$ for all $\lambda \geq 0$, the CBI process $\{Y_t : t \geq 0\}$ reduces to a *continuous-state branching process* (CB process) with *branching mechanism* $\lambda \mapsto R(-\lambda, 0)$. Then a CB process has Feller transition semigroup $(Q_t)_{t\geq 0}$ defined by

$$\int_{\mathbb{R}_+} \mathrm{e}^{-\lambda v} Q_t(x, \mathrm{d}v) = \exp\left\{x\psi_1(t, -\lambda, 0)\right\}, \quad \lambda \geq 0. \tag{2.9}$$

As a special case of (2.8) we have

$$\int_{\mathbb{R}_+} v Q_t(y, \mathrm{d}v) = y\mathrm{e}^{\beta_{11} t}, \quad t \geq 0, \; y \geq 0. \tag{2.10}$$

Then Jensen's inequality implies

$$-\psi_1(t, -\lambda, 0) \leq \mathrm{e}^{\beta_{11} t}\lambda, \quad t \geq 0, \; \lambda \geq 0. \tag{2.11}$$

https://doi.org/10.1017/jpr.2022.100 Published online by Cambridge University Press

From (2.9) it is easy to see that zero is a trap for the CB process. For a càdlàg realization of the CB process $\{Y_t : t \geq 0\}$, we define its *extinction time*

$$\tau_0 := \inf\{t \geq 0 : Y_t = 0\}.$$

The reader may refer to Kyprianou [14] and Li [15, 17] for compact treatments of CB and CBI processes.

Condition 2.2. (Grey's condition.) *There exists a constant* $\lambda_0 > 0$ *such that*

$$-R(-\lambda, 0) > 0 \ \textit{for} \ \lambda \geq \lambda_0 \ \textit{and} \ -\int_{\lambda_0}^{\infty} R(-\lambda, 0)^{-1} \mathrm{d}\lambda < \infty.$$

Proposition 2.3. *Suppose that* $F(-\lambda, 0) \equiv 0$ *and Condition 2.2 holds. Let* $\{Y_t : t \geq 0\}$ *be a càdlàg realization of the CB process with* $Y_0 = y$. *Then we have*

$$\mathbb{P}(\tau_0 > t) = \mathbb{P}(Y_t > 0) = 1 - \mathrm{e}^{-y\bar{v}_t}, \qquad t > 0, \tag{2.12}$$

where $t \mapsto \bar{v}_t := -\lim_{\lambda\to\infty} \psi_1(t, -\lambda, 0)$ *is the unique positive solution of*

$$\partial_t \bar{v}_t = -R(-\bar{v}_t, 0), \qquad \bar{v}_{0+} = \infty.$$

The above proposition follows from Theorem 3.4 and Corollary 3.14 in Li [17]. By (2.5) and (2.11), for any $t \geq \delta > 0$ we have

$$\begin{aligned} \bar{v}_t &= -\lim_{\lambda\to\infty} \psi_1(t-\delta, -\psi_1(\delta, -\lambda, 0), 0) \\ &= -\psi_1(t-\delta, -\bar{v}_\delta, 0) \leq \mathrm{e}^{\beta_{11}(t-\delta)} \bar{v}_\delta. \end{aligned} \tag{2.13}$$

Then for $\beta_{11} < 0$ the probability in (2.12) decays exponentially fast as $t \to \infty$.

A càdlàg realization of the general (1+1)-affine process can be constructed as the unique strong solution to a system of stochastic integral equations. Let $\sigma_0 = \sqrt{a}$, and let (σ_{ij}) be a (2×2) matrix satisfying $(\sigma_{ij}) = (\alpha_{ij})(\alpha_{ij})^\tau$. Suppose that $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P})$ is a filtered probability space satisfying the usual hypotheses. Let $W_0(t)$ be a standard $(\mathscr{F}_t)$-Brownian motion. Let $W_i(\mathrm{d}s, \mathrm{d}u)$, $i = 1, 2$, be $(\mathscr{F}_t)$-Gaussian white noises on $(0, \infty)^2$ with intensity $\mathrm{d}s\mathrm{d}u$. Let $M(\mathrm{d}s, \mathrm{d}u, \mathrm{d}v)$ be an $(\mathscr{F}_t)$-Poisson random measure on $(0, \infty)^2 \times D$ with intensity $\mathrm{d}s\mathrm{d}u\mu(\mathrm{d}v)$, and let $N(\mathrm{d}s, \mathrm{d}v)$ be an $(\mathscr{F}_t)$-Poisson random measure on $(0, \infty) \times D$ with intensity $\mathrm{d}s\nu(\mathrm{d}v)$. The corresponding compensated measures are denoted by $\tilde{M}(\mathrm{d}s, \mathrm{d}u, \mathrm{d}v)$ and $\tilde{N}(\mathrm{d}s, \mathrm{d}v)$. We assume these random elements are independent of each other. Given an $\mathscr{F}_0$-measurable random variable $(Y_0, Z_0) \in D$, we consider the following system of stochastic integral equations:

$$\begin{aligned} Y_t = Y_0 &+ \int_0^t (b_1 + \beta_{11} Y_s)\mathrm{d}s + \sqrt{2}\sigma_{11} \int_0^t \int_0^{Y_s} W_1(\mathrm{d}s, \mathrm{d}u) \\ &+ \sqrt{2}\sigma_{12} \int_0^t \int_0^{Y_s} W_2(\mathrm{d}s, \mathrm{d}u) + \int_0^t \int_0^{Y_{s-}} \int_D v_1 \tilde{M}(\mathrm{d}s, \mathrm{d}u, \mathrm{d}v) \\ &+ \int_0^t \int_D v_1 N(\mathrm{d}s, \mathrm{d}v) \end{aligned} \tag{2.14}$$

https://doi.org/10.1017/jpr.2022.100 Published online by Cambridge University Press

and

$$
\begin{aligned}
Z_t = Z_0 &+ \int_0^t (b_2 + \beta_{21} Y_s + \beta_{22} Z_s)\mathrm{d}s + \sqrt{2}\sigma_0 W_0(t) \\
&+ \sqrt{2}\sigma_{21} \int_0^t \int_0^{Y_s} W_1(\mathrm{d}s, \mathrm{d}u) + \sqrt{2}\sigma_{22} \int_0^t \int_0^{Y_s} W_2(\mathrm{d}s, \mathrm{d}u) \\
&+ \int_0^t \int_0^{Y_{s-}} \int_D v_2 \tilde{M}(\mathrm{d}s, \mathrm{d}u, \mathrm{d}v) + \int_0^t \int_D v_2 \tilde{N}(\mathrm{d}s, \mathrm{d}v).
\end{aligned} \tag{2.15}
$$

Here and in the sequel, we understand that, for any $a \le b \in \mathbb{R}$,

$$
\int_a^b = \int_{(a,b]} \quad \text{and} \quad \int_a^\infty = \int_{(a,\infty)}.
$$

The existence and pathwise uniqueness of the solution to (2.14) follows from Theorem 8.5 in Li [17]. A weakly equivalent stochastic equation was first introduced by Dawson and Li [4]; see also Dawson and Li [5]. The existence and pathwise uniqueness of the solution to (2.15) are straightforward. In fact, by (2.14)–(2.15), one can see using integration by parts that

$$
\begin{aligned}
\mathrm{e}^{-\beta_{22}t} Z_t = Z_0 &+ \int_0^t \mathrm{e}^{-\beta_{22}s}(b_2 + \beta_{21} Y_s)\mathrm{d}s + \sqrt{2}\sigma_0 \int_0^t \mathrm{e}^{-\beta_{22}s}\mathrm{d}W_0(s) \\
&+ \sqrt{2}\sigma_{21} \int_0^t \int_0^{Y_s} \mathrm{e}^{-\beta_{22}s} W_1(\mathrm{d}s, \mathrm{d}u) + \sqrt{2}\sigma_{22} \int_0^t \int_0^{Y_s} \mathrm{e}^{-\beta_{22}s} W_2(\mathrm{d}s, \mathrm{d}u) \\
&+ \int_0^t \int_0^{Y_{s-}} \int_D \mathrm{e}^{-\beta_{22}s} v_2 \tilde{M}(\mathrm{d}s, \mathrm{d}u, \mathrm{d}v) + \int_0^t \int_D \mathrm{e}^{-\beta_{22}s} v_2 \tilde{N}(\mathrm{d}s, \mathrm{d}v).
\end{aligned} \tag{2.16}
$$

By Theorem 6.2 of Dawson and Li [4], the process $\{(Y_t, Z_t) : t \ge 0\}$ defined by (2.14)–(2.15) is a (1+1)-affine process with transition semigroup $(P_t)_{t\ge 0}$ given by (2.3) and (2.4).

Let us remark that $\{Z_t : t \ge 0\}$ reduces to a one-dimensional OU-type process if $\sigma_{21} = \sigma_{22} = \beta_{21} = 0$ and μ is supported on $(0, \infty) \times \{0\}$. For instance, in this case, from (2.16) we have

$$
Z_t = Z_0 + \int_0^t (b_2 + \beta_{22} Z_s)\mathrm{d}s + \sqrt{2}\sigma_0 W_0(t) + \int_0^t \int_D v_2 \tilde{N}(\mathrm{d}s, \mathrm{d}v). \tag{2.17}
$$

A number of moment estimates for general finite-dimensional affine processes were given in Friesen *et al.* [8]. Since more accurate estimates are needed in this work, we here present the following result.

Proposition 2.4. *Let $\{(Y_t, Z_t) : t \ge 0\}$ be a (1+1)-affine process with $Y_0 = y \in \mathbb{R}_+$ and $Z_0 = z \in \mathbb{R}$. Let $D_1 = \mathbb{R}_+ \times [-1, 1]$ and $D_1^c = D \setminus D_1$. Then we have*

$$
\begin{aligned}
\mathbb{E}(|Z_t|) \le \mathrm{e}^{\beta_{22}t}|z| &+ \Big(|b_2| + 2\int_{D_1^c} |v_2|\nu(\mathrm{d}v)\Big) \int_0^t \mathrm{e}^{\beta_{22}(t-s)}\mathrm{d}s \\
&+ \Big(|\beta_{21}| + 2\int_{D_1^c} |v_2|\mu(\mathrm{d}v)\Big) \int_0^t \mathrm{e}^{\beta_{22}(t-s)}\mathbb{E}(Y_s)\mathrm{d}s
\end{aligned}
$$

https://doi.org/10.1017/jpr.2022.100 Published online by Cambridge University Press

$$
\begin{aligned}
&+\Big[\sqrt{2}\sigma_0+\Big(\int_{D_1} v_2^2\nu(\mathrm{d}v)\Big)^{1/2}\Big]\Big(\int_0^t \mathrm{e}^{2\beta_{22}(t-s)}\mathrm{d}s\Big)^{1/2}\\
&+\sqrt{2}(\sigma_{21}+\sigma_{22})\Big(\int_0^t \mathrm{e}^{2\beta_{22}(t-s)}\mathbb{E}(Y_s)\mathrm{d}s\Big)^{1/2}\\
&+\Big(\int_{D_1} v_2^2\mu(\mathrm{d}v)\Big)^{1/2}\Big(\int_0^t \mathrm{e}^{2\beta_{22}(t-s)}\mathbb{E}(Y_s)\mathrm{d}s\Big)^{1/2}.
\end{aligned}
\tag{2.18}
$$

Proof. We may assume that $\{(Y_t, Z_t) : t \ge 0\}$ is defined by (2.14)–(2.15). In view of (2.16) we have

$$
\begin{aligned}
\mathbb{E}\big(\mathrm{e}^{-\beta_{22}t}|Z_t|\big) \le{}& |z|+|\beta_{21}|\int_0^t \mathrm{e}^{-\beta_{22}s}\mathbb{E}(Y_s)\mathrm{d}s+\sqrt{2}\sigma_0\mathbb{E}\Big(\Big|\int_0^t \mathrm{e}^{-\beta_{22}s}\mathrm{d}W_0(s)\Big|\Big)\\
&+|b_2|\int_0^t \mathrm{e}^{-\beta_{22}s}\mathrm{d}s+\sqrt{2}\sigma_{21}\Big[\mathbb{E}\Big(\Big|\int_0^t\int_0^{Y_s} \mathrm{e}^{-\beta_{22}s}W_1(\mathrm{d}s,\mathrm{d}u)\Big|^2\Big)\Big]^{1/2}\\
&+\sqrt{2}\sigma_{22}\Big[\mathbb{E}\Big(\Big|\int_0^t\int_0^{Y_s} \mathrm{e}^{-\beta_{22}s}W_2(\mathrm{d}s,\mathrm{d}u)\Big|^2\Big)\Big]^{1/2}\\
&+\Big[\mathbb{E}\Big(\Big|\int_0^t\int_0^{Y_{s-}}\int_{D_1} \mathrm{e}^{-\beta_{22}s}v_2\tilde{M}(\mathrm{d}s,\mathrm{d}u,\mathrm{d}v)\Big|^2\Big)\Big]^{1/2}\\
&+\mathbb{E}\Big(\Big|\int_0^t\int_0^{Y_{s-}}\int_{D_1^c} \mathrm{e}^{-\beta_{22}s}|v_2|\tilde{M}(\mathrm{d}s,\mathrm{d}u,\mathrm{d}v)\Big|\Big)\\
&+\Big[\mathbb{E}\Big(\Big|\int_0^t\int_{D_1} \mathrm{e}^{-\beta_{22}s}v_2\tilde{N}(\mathrm{d}s,\mathrm{d}v)\Big|^2\Big)\Big]^{1/2}\\
&+\mathbb{E}\Big(\Big|\int_0^t\int_{D_1^c} \mathrm{e}^{-\beta_{22}s}|v_2|\tilde{N}(\mathrm{d}s,\mathrm{d}v)\Big|\Big)\\
\le{}& |z|+\Big(|b_2|+2\int_{D_1^c}|v_2|\nu(\mathrm{d}v)\Big)\int_0^t \mathrm{e}^{-\beta_{22}s}\mathrm{d}s\\
&+\Big(|\beta_{21}|+2\int_{D_1^c}|v_2|\mu(\mathrm{d}v)\Big)\int_0^t \mathrm{e}^{-\beta_{22}s}\mathbb{E}(Y_s)\mathrm{d}s\\
&+\Big[\sqrt{2}\sigma_0+\Big(\int_{D_1} v_2^2\nu(\mathrm{d}v)\Big)^{1/2}\Big]\Big(\int_0^t \mathrm{e}^{-2\beta_{22}s}\mathrm{d}s\Big)^{1/2}\\
&+\sqrt{2}(\sigma_{21}+\sigma_{22})\Big(\int_0^t \mathrm{e}^{-2\beta_{22}s}\mathbb{E}(Y_s)\mathrm{d}s\Big)^{1/2}\\
&+\Big(\int_{D_1} v_2^2\mu(\mathrm{d}v)\Big)^{1/2}\Big(\int_0^t \mathrm{e}^{-2\beta_{22}s}\mathbb{E}(Y_s)\mathrm{d}s\Big)^{1/2}.
\end{aligned}
$$

Then (2.18) follows. □

Proposition 2.5. *Suppose that $\beta_{11} < 0$ and $\beta_{22} < 0$. Then the transition semigroup $(P_t)_{t\ge0}$ defined by (2.3) and (2.4) has a unique stationary distribution π, which is given by*

$$
\int_D \mathrm{e}^{\langle u,v\rangle}\pi(\mathrm{d}v)=\exp\Big\{\int_0^\infty F(\psi(t,u))\mathrm{d}s\Big\},\qquad u\in U. \tag{2.19}
$$

https://doi.org/10.1017/jpr.2022.100 Published online by Cambridge University Press

Moreover, the distribution π has finite first moment; that is,

$$\int_D (v_1 + |v_2|)\pi(\mathrm{d}v) < \infty.$$

Proof. By Theorem 2.7 of Jin *et al.* [11], the affine process has a unique stationary distribution π given by (2.19). In particular, we have

$$\int_D \mathrm{e}^{-\lambda v_1}\pi(\mathrm{d}v) = \exp\left\{\int_0^\infty F(\psi_1(t, -\lambda, 0), 0)\mathrm{d}s\right\}, \quad \lambda \geq 0.$$

By differentiating both sides of the above equality at $\lambda = 0+$ and using (2.1), one may see that

$$m_1(\pi) := \int_D v_1\pi(\mathrm{d}v) = [b_1 + m_1(\nu)]\int_0^\infty \mathrm{e}^{-|\beta_{22}|s}\mathrm{d}s < \infty.$$

By (2.8) and (2.18) there is a constant $C \geq 0$ such that

$$\int_D |v_2|P_t(x, \mathrm{d}v) \leq \mathrm{e}^{-|\beta_{22}|t}|z| + C(1 + y), \quad t \geq 0, \; x = (y, z) \in D.$$

Since π is a stationary distribution for $(P_t)_{t\geq 0}$, it follows that

$$\begin{aligned}\int_D (|v_2| \wedge k)\pi(\mathrm{d}v) &= \int_D \pi(\mathrm{d}u)\int_D (|v_2| \wedge k)P_t(u, \mathrm{d}v)\\ &\leq \int_D \big[(\mathrm{e}^{-|\beta_{22}|t}|u_2| \wedge k) + C(1 + u_1)\big]\pi(\mathrm{d}u)\\ &\leq \int_D (\mathrm{e}^{-|\beta_{22}|t}|u_2| \wedge k)\pi(\mathrm{d}u) + C[1 + m_1(\pi)].\end{aligned}$$

Then, letting $t \to \infty$ and $k \to \infty$, we obtain

$$\int_D |v_2|\pi(\mathrm{d}v) \leq C[1 + m_1(\pi)] < \infty.$$

This proves the result. □

Proposition 2.6. *For $i = 1, 2$, let $x_i = (y_i, z_i) \in D$, and let $\{X_i(t) : t \geq 0\} = \{(Y_i(t), Z_i(t)) : t \geq 0\}$ be defined by (2.14) and (2.15) with $(Y_i(0), Z_i(0)) = x_i$. Then, for any $t \geq 0$,*

$$\begin{aligned}&\mathbb{E}\Big(\sup_{0\leq s\leq t}|Z_1(s) - Z_2(s)|\Big)\\ &\quad\leq \mathrm{e}^{\beta_{22}t}|z_1 - z_2| + |\beta_{21}||y_1 - y_2|\int_0^t \mathrm{e}^{\beta_{11}s}\mathrm{e}^{\beta_{22}(t-s)}\mathrm{d}s\\ &\qquad + 2|y_1 - y_2|\int_{D_1^c}|v_2|\mu(\mathrm{d}v)\int_0^t \mathrm{e}^{\beta_{11}s}\mathrm{e}^{\beta_{22}(t-s)}\mathrm{d}s\\ &\qquad + 2\sqrt{2}(\sigma_{21} + \sigma_{22})|y_1 - y_2|^{1/2}\Big(\int_0^t \mathrm{e}^{\beta_{11}s}\mathrm{e}^{2\beta_{22}(t-s)}\mathrm{d}s\Big)^{1/2}\\ &\qquad + 2|y_1 - y_2|^{1/2}\Big(\int_{D_1} v_2^2\mu(\mathrm{d}v)\Big)^{1/2}\Big(\int_0^t \mathrm{e}^{\beta_{11}s}\mathrm{e}^{2\beta_{22}(t-s)}\mathrm{d}s\Big)^{1/2}.\end{aligned}$$

https://doi.org/10.1017/jpr.2022.100 Published online by Cambridge University Press

Proof. Without loss of generality, we may assume $y_1 \geq y_2$. By Theorem 10.1 in Li [17] one can see that $\{Y_1(t) \geq Y_2(t)$ for every $t \geq 0\} = 1$ and $\{Y_1(t) - Y_2(t) : t \geq 0\}$ is a CB process with transition semigroup $(Q_t)_{t\geq 0}$. Then we apply (2.16) to $Z_1(t)$ and $Z_2(t)$ and take the difference to see

$$\begin{aligned}
Z_1(t) - Z_2(t) = \mathrm{e}^{\beta_{22}t}\Big\{&(z_1 - z_2) + \beta_{21}\int_0^t \mathrm{e}^{-\beta_{22}s}[Y_1(s) - Y_2(s)]\mathrm{d}s \\
&+ \sqrt{2}\sigma_{21}\int_0^t \int_{Y_1(s)}^{Y_2(s)} \mathrm{e}^{-\beta_{22}s} W_1(\mathrm{d}s, \mathrm{d}u) \\
&+ \sqrt{2}\sigma_{22}\int_0^t \int_{Y_1(s)}^{Y_2(s)} \mathrm{e}^{-\beta_{22}s} W_2(\mathrm{d}s, \mathrm{d}u) \\
&+ \int_0^t \int_{Y_1(s-)}^{Y_2(s-)} \int_D \mathrm{e}^{-\beta_{22}s} v_2 \tilde{M}(\mathrm{d}s, \mathrm{d}u, \mathrm{d}v)\Big\}.
\end{aligned}$$

It follows that

$$\begin{aligned}
&\mathbb{E}\Big(\sup_{0\leq s\leq t} |Z_1(s) - Z_2(s)|\Big) \\
&\leq \mathrm{e}^{\beta_{22}t}\Big\{|z_1 - z_2| + |\beta_{21}|\int_0^t \mathrm{e}^{-\beta_{22}s}\mathbb{E}[Y_1(s) - Y_2(s)]\mathrm{d}s \\
&\quad + 2\sqrt{2}\sigma_{21}\Big[\mathbb{E}\Big(\Big|\int_0^t \int_{Y_1(s)}^{Y_2(s)} \mathrm{e}^{-\beta_{22}s} W_1(\mathrm{d}s, \mathrm{d}u)\Big|^2\Big)\Big]^{1/2} \\
&\quad + 2\sqrt{2}\sigma_{22}\Big[\mathbb{E}\Big(\Big|\int_0^t \int_{Y_1(s)}^{Y_2(s)} \mathrm{e}^{-\beta_{22}s} W_2(\mathrm{d}s, \mathrm{d}u)\Big|^2\Big)\Big]^{1/2} \\
&\quad + 2\Big[\mathbb{E}\Big(\Big|\int_0^t \int_{Y_1(s-)}^{Y_2(s-)} \int_{D_1} \mathrm{e}^{-\beta_{22}s} v_2 \tilde{M}(\mathrm{d}s, \mathrm{d}u, \mathrm{d}v)\Big|^2\Big)\Big]^{1/2} \\
&\quad + \mathbb{E}\Big(\Big|\int_0^t \int_{Y_1(s-)}^{Y_2(s-)} \int_{D_1^c} \mathrm{e}^{-\beta_{22}s} |v_2| \tilde{M}(\mathrm{d}s, \mathrm{d}u, \mathrm{d}v)\Big|\Big)\Big\} \\
&\leq \mathrm{e}^{\beta_{22}t}\Big\{|z_1 - z_2| + |\beta_{21}|\int_0^t \mathrm{e}^{-\beta_{22}s}\mathbb{E}[Y_1(s) - Y_2(s)]\mathrm{d}s \\
&\quad + 2\sqrt{2}(\sigma_{21} + \sigma_{22})\Big(\int_0^t \mathrm{e}^{-2\beta_{22}s}\mathbb{E}[Y_1(s) - Y_2(s)]\mathrm{d}s\Big)^{1/2} \\
&\quad + 2\Big(\int_{D_1} v_2^2 \mu(\mathrm{d}v)\Big)^{1/2}\Big(\int_0^t \mathrm{e}^{-2\beta_{22}s}\mathbb{E}[Y_1(s) - Y_2(s)]\mathrm{d}s\Big)^{1/2} \\
&\quad + 2\int_{D_1^c} |v_2|\mu(\mathrm{d}v)\int_0^t \mathrm{e}^{-\beta_{22}s}\mathbb{E}[Y_1(s) - Y_2(s)]\mathrm{d}s\Big\},
\end{aligned}$$

where we have used Doob's L^2 inequality for the square-integrable martingales; see, e.g., Revuz and Yor [25, Theorem 1.7, p. 54]. Then the desired estimate follows by (2.10). □

https://doi.org/10.1017/jpr.2022.100 Published online by Cambridge University Press

A Markov process $\{(X_1(t), X_2(t)) : t \geq 0\}$ with state space D^2 is called a *Markov coupling* of the (1+1)-affine process with transition semigroup $(P_t)_{t\geq 0}$ defined by (2.3) and (2.4) with *coupling time* $\tau := \inf\{t \geq 0 : X_1(t) = X_2(t)\}$ if both $\{X_1(t) : t \geq 0\}$ and $\{X_2(t) : t \geq 0\}$ are Markov processes with transition semigroup $(P_t)_{t\geq 0}$ and $X_1(\tau + t) = X_2(\tau + t)$ for every $t \geq 0$.

The method of couplings provides an efficient way to estimate the variations of the transition probabilities of the affine process. Let $\|\cdot\|_{\mathrm{var}}$ denote the total variation norm of signed measures. Let $\mathscr{B}_1$ be the set of Borel functions f on D satisfying $|f| \leq 1$. Then we have

$$\big\|P_t(x_1, \cdot) - P_t(x_2, \cdot)\big\|_{\mathrm{var}} = \sup_{f\in\mathscr{B}_1} \big[P_t f(x_1) - P_t f(x_2)\big]. \tag{2.20}$$

Let $\{(X_1(t), X_2(t)) : t \geq 0\}$ be a Markov coupling of the affine process with initial state $(X_1(0), X_2(0)) = (x_1, x_2)$ and coupling time τ. From (2.20) it follows that

$$\big\|P_t(x_1, \cdot) - P_t(x_2, \cdot)\big\|_{\mathrm{var}} = \sup_{f\in\mathscr{B}_1} \mathbb{E}\big[f(X_1(t)) - f(X_2(t))\big] \leq 2\mathbb{P}(\tau > t). \qquad (2.21$$

By the pathwise uniqueness for (2.14)–(2.15), the process $\{(X_1(t), X_2(t)) : t \geq 0\}$ defined in Proposition 2.6 is a Markov coupling of the affine process with transition semigroup $(P_t)_{t\geq 0}$. Based on this coupling, several different couplings of the affine process will be given in the next two sections. We shall see that the stochastic equations (2.14)–(2.15) and (2.16) provide an efficient method for the pathwise construction and analysis of those couplings. The approach of stochastic equations has also played an important role in other recent developments concerning branching processes in continuous state spaces; see, e.g., Bansaye and Méléard [1], Li [17], Pardoux [22], and the references therein.

3. Estimates for variations of probabilities

In this section, we study the strong Feller property and the exponential ergodicity of the total variation distance of the affine process. Let $\|\cdot\|_{\mathrm{var}}$ denote the total variation norm of signed measures. Our strategy is to establish some estimates for the differences of the transition probabilities in the form (2.20). The proofs of the estimates are based on couplings of the affine process constructed in terms of strong solutions of stochastic equations. From those estimates we deduce the strong Feller property and the exponential ergodicity under natural conditions.

We first consider the case $\sigma_0 > 0$. Write $x_1 = (y_1, z_1)$ and $x_2 = (y_2, z_2)$, where $y_1, y_2 \in \mathbb{R}_+$ and $z_1, z_2 \in \mathbb{R}$. For $i = 1, 2$ let $(Y_i(t), Z_i(t))$ be defined by (2.14)–(2.16) with $(Y_i(0), Z_i(0)) = (y_i, z_i)$, and write $X_i(t) = (Y_i(t), Z_i(t))$. Then $\{(X_1(t), X_2(t)) : t \geq 0\}$ is a Markov coupling of the affine process. The pathwise uniqueness of the solution for (2.14) implies $Y_1(\tau_0 + t) = Y_2(\tau_0 + t)$ for $t \geq 0$. In fact, by Theorem 10.1 in Li [17] it is easy to see that $\{|Y_1(t) - Y_2(t)| : t \geq 0\}$ is a CB process with transition semigroup $(Q_t)_{t\geq 0}$. Let $\tau_0 = \inf\{t \geq 0 : Y_1(t) = Y_2(t)\}$ be the extinction time of the process. By Proposition 2.3 we have

$$\mathbb{P}(\tau_0 > t) = 1 - \mathrm{e}^{-|y_1 - y_2|\bar{v}_t} \leq |y_1 - y_2|\bar{v}_t, \qquad t \geq 0. \tag{3.1}$$

Let $a(\tau_0) = [Z_1(\tau_0) - Z_2(\tau_0)]/2\sqrt{2}\sigma_0$ and

$$\tau = \inf\Big\{t \geq 0 : \int_0^t \mathrm{e}^{-\beta_{22}s}\mathrm{d}W_0(\tau_0 + s) = -a(\tau_0)\Big\}.$$

https://doi.org/10.1017/jpr.2022.100 Published online by Cambridge University Press

Then we define the process $\{Z_2{}'(t) : t \geq 0\}$ by

$$\begin{aligned}
\mathrm{e}^{-\beta_{22}t}Z_2{}'(t) = z_1 &+ \int_0^t \mathrm{e}^{-\beta_{22}s}[b_2 + \beta_{21}Y_2(s)]\mathrm{d}s + \sqrt{2}\sigma_{21}\int_0^t\int_0^{Y_2(s)} \mathrm{e}^{-\beta_{22}s}W_1(\mathrm{d}s, \mathrm{d}u)\\
&+ \sqrt{2}\sigma_{22}\int_0^t\int_0^{Y_2(s)} \mathrm{e}^{-\beta_{22}s}W_2(\mathrm{d}s, \mathrm{d}u) + \sqrt{2}\sigma_0\Big[\int_0^{t\wedge\tau_0} \mathrm{e}^{-\beta_{22}s}\mathrm{d}W_0(s)\\
&- \int_{t\wedge\tau_0}^{t\wedge(\tau_0+\tau)} \mathrm{e}^{-\beta_{22}s}\mathrm{d}W_0(s) + \int_{t\wedge(\tau_0+\tau)}^{t} \mathrm{e}^{-\beta_{22}s}\mathrm{d}W_0(s)\Big]\\
&+ \int_0^t\int_0^{Y_2(s-)}\int_D \mathrm{e}^{-\beta_{22}s}v_2\tilde{M}(\mathrm{d}s, \mathrm{d}u, \mathrm{d}v) + \int_0^t\int_D \mathrm{e}^{-\beta_{22}s}v_2\tilde{N}(\mathrm{d}s, \mathrm{d}v).
\end{aligned}$$

It is clear that $Z_2{}'(t\wedge\tau_0) = Z_2(t\wedge\tau_0)$ for $t \geq 0$. Write $X_2{}'(t) = (Y_2(t), Z_2{}'(t))$. Then $\{(X_1(t), X_2{}'(t)) : t \geq 0\}$ is also a Markov coupling of the affine process. For $t \geq 0$ let

$$\zeta(t) = Z_1(\tau_0 + t) - Z_2{}'(\tau_0 + t).$$

Since $Y_1(\tau_0 + t) = Y_2(\tau_0 + t)$ for $t \geq 0$, by the construction of $Z_1(t)$ and $Z_2{}'(t)$ we have

$$\zeta(t) = 2\sqrt{2}\sigma_0\mathrm{e}^{\beta_{22}t}\Big[a(\tau_0) + \int_0^{t\wedge\tau} \mathrm{e}^{-\beta_{22}s}\mathrm{d}W_0(\tau_0 + s)\Big]. \tag{3.2}$$

It follows that \P$\tau = \inf\{t \geq 0 : \zeta(t) = 0\}$, and so

$$\tau_0 + \tau = \inf\{t \geq \tau_0 : Z_1(t) = Z_2{}'(t)\} = \inf\{t \geq \tau_0 : X_1(t) = X_2{}'(t)\}.$$

Then $\tau_0 + \tau$ is the coupling time of $\{(X_1(t), X_2{}'(t)) : t \geq 0\}$.

Theorem 3.1. *Suppose that* $\sigma_0 > 0$. *Then there is a constant* $C \geq 0$ *such that*

$$\begin{aligned}
&\|P_t(x_1, \cdot) - P_t(x_2, \cdot)\|_{\mathrm{var}}\\
&\quad\leq 2|y_1 - y_2|\bar{v}_{t/2} + C\Big\{\mathrm{e}^{\beta_{22}t}|z_1 - z_2| + |\beta_{21}||y_1 - y_2|\int_0^{t/2} \mathrm{e}^{\beta_{11}s}\mathrm{e}^{\beta_{22}(t/2-s)}\mathrm{d}s\\
&\quad\quad+ 2\sqrt{2}(\sigma_{21} + \sigma_{22})|y_1 - y_2|^{1/2}\Big(\int_0^{t/2} \mathrm{e}^{\beta_{11}s}\mathrm{e}^{2\beta_{22}(t/2-s)}\mathrm{d}s\Big)^{1/2}\\
&\quad\quad+ 2\Big[\int_{D_1} v_2^2\mu(\mathrm{d}v)\Big]^{1/2}|y_1 - y_2|^{1/2}\Big(\int_0^{t/2} \mathrm{e}^{\beta_{11}s}\mathrm{e}^{2\beta_{22}(t/2-s)}\mathrm{d}s\Big)^{1/2}\\
&\quad\quad+ 2\int_{D_1^c} |v_2|\mu(\mathrm{d}v)|y_1 - y_2|\int_0^{t/2} \mathrm{e}^{\beta_{11}s}\mathrm{e}^{\beta_{22}(t/2-s)}\mathrm{d}s\Big\}\Big(\int_0^{t/2} \mathrm{e}^{-2\beta_{22}s}\mathrm{d}s\Big)^{-1/2},\\
&\qquad\qquad t > 0,\ x_i = (y_i, z_i) \in D,\ i = 1, 2.
\end{aligned}$$

Proof. Let $\{(X_1(t), X_2{}'(t)) : t \geq 0\}$ be the coupling of the affine process constructed as above. Then we have

$$\mathbb{P}(\tau_0 + \tau > t) \leq \mathbb{P}(\tau_0 > t/2) + \mathbb{P}(\tau_0 \leq t/2, \tau_0 + \tau > t),$$

https://doi.org/10.1017/jpr.2022.100 Published online by Cambridge University Press

where $\mathbb{P}(\tau_0 > t/2) \le |y_1 - y_2|\bar{v}_{t/2}$ by (3.1). In view of (3.2), there is a standard Brownian motion $\{B(t) : t \ge 0\}$ independent of $\mathscr{F}_{\tau_0}$ such that

$$\zeta(t) = 2\sqrt{2}\sigma_0 e^{\beta_{22}t}\big[a(\tau_0) + B(\rho(t \wedge \tau))\big],$$

where

$$\rho(t) = \int_0^t e^{-2\beta_{22}s} ds, \qquad t \ge 0.$$

Since $t - \tau_0$ is measurable relative to $\mathscr{F}_{\tau_0}$, by the reflection principle for the Brownian motion we get

$$\begin{aligned}
\mathbb{P}(\tau_0 \le t/2, \tau_0 + \tau > t) &= \mathbb{E}\big[1_{\{\tau_0 \le t/2\}}\mathbb{P}(\tau_0 + \tau > t|\mathscr{F}_{\tau_0})\big] \\
&\le \mathbb{E}\big[1_{\{\tau_0 \le t/2\}}\mathbb{P}(|B(\rho(t - \tau_0))| < |a(\tau_0)|)\big] \\
&\le \mathbb{E}\Big[1_{\{\tau_0 \le t/2\}}\frac{2|a(\tau_0)|}{\sqrt{2\pi\rho(t - \tau_0)}}\Big] \\
&\le \frac{1}{2\sigma_0\sqrt{\pi\rho(t/2)}}\mathbb{E}\big[1_{\{\tau_0 \le t/2\}}|Z_1(\tau_0) - Z_2(\tau_0)|\big] \\
&\le \frac{1}{2\sigma_0\sqrt{\pi\rho(t/2)}}\mathbb{E}\Big(\sup_{0\le s\le t/2}|Z_1(s) - Z_2(s)|\Big).
\end{aligned}$$

Then the result follows by Proposition 2.6 and (2.21). □

Corollary 3.2. *Suppose that $\sigma_0 > 0$. Then $(P_t)_{t\ge0}$ is a strong Feller transition semigroup.*

Corollary 3.3. *Suppose that $\beta_{11} < 0$, $\beta_{22} < 0$, and $\sigma_0 > 0$. Let π be the unique stationary distribution for $(P_t)_{t\ge0}$. Then for every $\delta > 0$ there is a constant $C_\delta \ge 0$ such that*

$$\|P_t(x, \cdot) - \pi\|_{\text{var}} \le C_\delta(1 + |x|)e^{-\kappa t/2}, \qquad t \ge \delta, \; x \in D,$$

where $\kappa = |\beta_{11}| \wedge |\beta_{22}|$.

Proof. By Proposition 2.5, the stationary distribution π possesses a finite first moment. It is well known that

$$\|P_t(x, \cdot) - \pi\|_{\text{var}} \le \int_D \|P_t(x, \cdot) - P_t(x_2, \cdot)\|_{\text{var}}\pi(dx_2). \tag{3.3}$$

By Theorem 3.1, there is a constant $C \ge 0$ such that, for $x_i = (y_i, z_i) \in D$, $i = 1, 2$,

$$\begin{aligned}
&\|P_t(x_1, \cdot) - P_t(x_2, \cdot)\|_{\text{var}} \\
&\qquad \le C\big(|x_1 - x_2| + |y_1 - y_2|^{1/2}\big)\big[\bar{v}_{t/2} \vee e^{-|\beta_{22}|t/2}(1 - e^{-|\beta_{22}|t})^{-1/2}\big].
\end{aligned}$$

Then the desired estimate follows by (2.13). □

Now let us consider the case where $\nu(D) > 0$. For $\varepsilon > 0$ let $D_\varepsilon = \mathbb{R}_+ \times [-\varepsilon, \varepsilon]$ and $D_\varepsilon^c = D \setminus D_\varepsilon$. By choosing sufficiently small $\varepsilon \in (0, 1]$ we have $0 < \nu(D_\varepsilon^c) < \infty$. Let ν_ε be the finite measure on D defined by

$$\nu_\varepsilon(A) = \begin{cases} \nu(A) & \text{if } \nu(D) < \infty, \\ \nu(A \cap D_\varepsilon^c) & \text{if } \nu(D) = \infty, \end{cases} \tag{3.4}$$

where $A \in \mathscr{B}(D)$. Let $\hat{\nu}_\varepsilon = \nu_\varepsilon(D)^{-1}\nu_\varepsilon$.

https://doi.org/10.1017/jpr.2022.100 Published online by Cambridge University Press

Condition 3.4. *There exists* $\varepsilon \in (0, 1]$ *such that*

$$\limsup_{|z|\to 0} |z|^{-1} \|\hat{\nu}_\varepsilon - \delta_{(0,z)} * \hat{\nu}_\varepsilon\|_{\mathrm{var}} < \infty.$$

The above condition is a slight modification of (10) in Wang [29] for OU-type processes. As in Wang [29, p. 996], one may see that the above condition implies

$$K_\varepsilon := \sup_{z\in\mathbb{R}} |z|^{-1} \|\hat{\nu}_\varepsilon - \delta_{(0,z)} * \hat{\nu}_\varepsilon\|_{\mathrm{var}} < \infty. \tag{3.5}$$

Theorem 3.5. *Suppose that Condition 3.4 is satisfied for some* $\varepsilon \in (0, 1]$. *Then we have*

$$\begin{aligned}
&\|P_t(x_1, \cdot) - P_t(x_2, \cdot)\|_{\mathrm{var}} \\
&\quad \le 2\big(|y_1 - y_2|\bar{v}_{t/3} + \mathrm{e}^{-\nu(D_\varepsilon^c)t/3}\big) + K_\varepsilon \mathrm{e}^{\beta_{22}t/3}\Big\{\mathrm{e}^{\beta_{22}t/3}|z_1 - z_2| \\
&\qquad + \Big(\beta_{21}| + 2\int_{D_1^c} |v_2|\mu(\mathrm{d}v)\Big)|y_1 - y_2| \int_0^{t/3} \mathrm{e}^{\beta_{11}s}\mathrm{e}^{\beta_{22}(t/3-s)}\mathrm{d}s \\
&\qquad + 2\sqrt{2}(\sigma_{21} + \sigma_{22})|y_1 - y_2|^{1/2}\Big(\int_0^{t/3} \mathrm{e}^{\beta_{11}s}\mathrm{e}^{2\beta_{22}(t/3-s)}\mathrm{d}s\Big)^{1/2} \\
&\qquad + 2\Big(\int_{D_1} v_2^2\mu(\mathrm{d}v)\Big)^{1/2}|y_1 - y_2|^{1/2}\Big(\int_0^{t/3} \mathrm{e}^{\beta_{11}s}\mathrm{e}^{2\beta_{22}(t/3-s)}\mathrm{d}s\Big)^{1/2}\Big\}, \\
&\qquad\qquad t \ge 0,\ x_i = (y_i, z_i) \in D,\ i = 1, 2.
\end{aligned}$$

Proof. Step 1. Consider the case where $y_1 = y_2 = y \in \mathbb{R}_+$. Let $\{Y_t : t \ge 0\}$ be the solution of (2.14) with $Y_0 = y$. Let $z_0 = 0$. For $i = 0, 1, 2$ let $\{Z_i(t) : t \ge 0\}$ be defined by (2.16), with $Z_i(0) = z_i$. It is easy to see that

$$Z_i(t) = \mathrm{e}^{\beta_{22}t}z_i + Z_0(t), \qquad t \ge 0,\ i = 1, 2. \tag{3.6}$$

Let $\{\eta_\varepsilon(t) : t \ge 0\}$ be the compensated compound Poisson process defined by

$$\eta_\varepsilon(t) = \int_0^t \int_{D_\varepsilon^c} v_2\tilde{N}(\mathrm{d}s, \mathrm{d}v).$$

Let $\tau_1 = \inf\{t \ge 0 : \eta_\varepsilon(t) \ne \eta_\varepsilon(t-)\}$ be the first jump time of this process. For any $f \in \mathscr{B}_1$ we have

$$\begin{aligned}
\big|P_tf(y, z_1) - P_tf(y, z_2)\big| &= \big|\mathbb{E}\big[f(Y_t, Z_1(t)) - f(Y_t, Z_2(t))\big]\big| \\
&\le 2\mathbb{P}(\tau_1 > t) + p_\varepsilon(t),
\end{aligned}$$

where $\mathbb{P}(\tau_1 > t) = \mathrm{e}^{-\nu(D_\varepsilon^c)t}$ and

$$p_\varepsilon(t) = \Big|\mathbb{E}\Big\{\big[f(Y_t, Z_1(t)) - f(Y_t, Z_2(t))\big]1_{\{\tau_1 \le t\}}\Big\}\Big|.$$

https://doi.org/10.1017/jpr.2022.100 Published online by Cambridge University Press

By the strong Markov property and (2.16),

$$\begin{aligned}
p_\varepsilon(t) &= \bigg|\mathbb{E}\bigg\{\int_0^t \nu(D_\varepsilon^c)\mathrm{e}^{-\nu(D_\varepsilon^c)s}\bigg[\int_D P_{t-s}f(Y_s, Z_1(s-)+r)\hat{\nu}_\varepsilon(\mathrm{d}r) \\
&\qquad - \int_D P_{t-s}f(Y_s, Z_2(s-)+r)\hat{\nu}_\varepsilon(\mathrm{d}r)\bigg]\mathrm{d}s\bigg\}\bigg| \\
&= \bigg|\mathbb{E}\bigg\{\int_0^t \nu(D_\varepsilon^c)\mathrm{e}^{-\nu(D_\varepsilon^c)s}\bigg[\int_D P_{t-s}f(Y_s, Z_1(s)+r)\hat{\nu}_\varepsilon(\mathrm{d}r) \\
&\qquad - \int_D P_{t-s}f(Y_s, Z_1(s)+r)\delta_{(0,\mathrm{e}^{\beta_{22}s}(z_2-z_1))}*\hat{\nu}_\varepsilon(\mathrm{d}r)\bigg]\mathrm{d}s\bigg\}\bigg| \\
&\le \int_0^t \nu(D_\varepsilon^c)\mathrm{e}^{-\nu(D_\varepsilon^c)s}\|\hat{\nu}_\varepsilon - \delta_{(0,\mathrm{e}^{\beta_{22}s}(z_2-z_1))}*\hat{\nu}_\varepsilon\|\mathrm{d}s \\
&\le K_\varepsilon \nu(D_\varepsilon^c)|z_2-z_1|\int_0^t \mathrm{e}^{-\nu(D_\varepsilon^c)s}\mathrm{d}s \le K_\varepsilon|z_2-z_1|.
\end{aligned}$$

It follows that

$$\big|P_tf(y,z_1) - P_tf(y,z_2)\big| \le 2\mathrm{e}^{-\nu(D_\varepsilon^c)t} + K_\varepsilon|z_1-z_2|. \tag{3.7}$$

Then we can use the Markov property and the representation (3.6) to get

$$\begin{aligned}
&\big|P_tf(y,x_1) - P_tf(y,x_2)\big| \\
&\quad = \Big|\mathbb{E}\big[f(Y_t, Z_1(t)) - f(Y_t, Z_2(t))\big]\Big| \\
&\quad = \Big|\mathbb{E}\big[P_{t/2}f(Y_{t/2}, Z_1(t/2)) - P_{t/2}f(Y_{t/2}, Z_2(t/2))\big]\Big| \\
&\quad \le 2\mathrm{e}^{-\nu(D_\varepsilon^c)t/2} + K_\varepsilon\mathbb{E}\big[|Z_1(t/2) - Z_2(t/2)|\big] \\
&\quad \le 2\mathrm{e}^{-\nu(D_\varepsilon^c)t/2} + K_\varepsilon|z_1-z_2|\mathrm{e}^{\beta_{22}t/2}.
\end{aligned}$$

Step 2. In the general case, we have $x_1 = (y_1, z_1)$ and $x_2 = (y_2, z_2)$, where $y_1, y_2 \in \mathbb{R}_+$ and $z_1, z_2 \in \mathbb{R}_+$. It suffices to consider the case of $y_1 \ge y_2$. Let $\{(Y_i(t), Z_i(t)) : t \ge 0\}$ be defined by (2.14)–(2.15) with $(Y_i(0), Z_i(0)) = x_i$, $i = 1, 2$. Then

$$\begin{aligned}
\big|P_tf(x_1) - P_tf(x_2)\big| &= \Big|\mathbb{E}\big[f(Y_1(t), Z_1(t)) - f(Y_2(t), Z_2(t))\big]\Big| \\
&\le 2\mathbb{P}(\tau_0 > t/3) + q_\varepsilon(t),
\end{aligned}$$

with $\mathbb{P}(\tau_0 > t/3) \le |y_1-y_2|\bar{v}_{t/3}$ and

$$\begin{aligned}
q_\varepsilon(t) &= \Big|\mathbb{E}\big\{\big[f(Y_1(t), Z_1(t)) - f(Y_1(t), Z_2(t))\big]1_{\{\tau_0\le t/3\}}\big\}\Big| \\
&= \Big|\mathbb{E}\big\{1_{\{\tau_0\le t/3\}}\mathbb{E}\big[f(Y_1(t), Z_1(t)) - f(Y_1(t), Z_2(t))\big|\mathscr{F}_{\tau_0}\big]\big\}\Big| \\
&= \Big|\mathbb{E}\big\{1_{\{\tau_0\le t/3\}}\big[P_{t-\tau_0}f(Y_1(\tau_0), Z_1(\tau_0)) - P_{t-\tau_0}f(Y_1(\tau_0), Z_2(\tau_0))\big]\big\}\Big| \\
&\le 2\mathrm{e}^{-\nu(D_\varepsilon^c)t/3} + K_\varepsilon\mathbb{E}\big(1_{\{\tau_0\le t/3\}}|Z_1(\tau_0) - Z_2(\tau_0)|\big)\mathrm{e}^{\beta_{22}t/3} \\
&\le 2\mathrm{e}^{-\nu(D_\varepsilon^c)t/3} + K_\varepsilon\mathbb{E}\Big(\sup_{s\le t/3}|Z_1(s) - Z_2(s)|\Big)\mathrm{e}^{\beta_{22}t/3},
\end{aligned}$$

https://doi.org/10.1017/jpr.2022.100 Published online by Cambridge University Press

where we have used (3.7) for the first inequality. Then the desired estimate follows by (2.20) and Proposition 2.6. □

Corollary 3.6. *Suppose that Condition 3.4 is satisfied for a sequence* $\{\varepsilon_n\} \subset (0, 1]$ *and* $\lim_{n\to\infty} \nu(D^c_{\varepsilon_n}) = \infty$. *Then* $(P_t)_{t\geq 0}$ *is a strong Feller transition semigroup.*

Proof. Suppose that $\{x_k\} \in D$ is a sequence such that $\lim_{k\to\infty} x_k = x_0 \in D$. By Theorem 3.5, for $t > 0$ and $n \geq 1$ we have

$$\limsup_{k\to\infty} \|P_t(x_k, \cdot) - P_t(x_0, \cdot)\|_{\mathrm{var}} \leq 2\mathrm{e}^{-\nu(D^c_{\varepsilon_n})t/3}.$$

The left-hand side vanishes since $\lim_{n\to\infty} \nu(D^c_{\varepsilon_n}) = \infty$. □

Corollary 3.7. *Suppose that* $\beta_{11} < 0$, $\beta_{22} < 0$, *and Condition 3.4 is satisfied. Then there is a constant* $C_\varepsilon \geq 0$ *such that*

$$\|P_t(x, \cdot) - \pi\|_{\mathrm{var}} \leq C_\varepsilon(1 + |x|)\mathrm{e}^{-\kappa_\varepsilon t/3}, \qquad t \geq 0, \ x \in D, \tag{3.8}$$

where $\kappa_\varepsilon = |\beta_{11}| \wedge |\beta_{22}| \wedge \nu(D^c_\varepsilon)$.

Proof. By Theorem 3.5 there is a constant $C_\varepsilon \geq 0$ such that, for $t > 0$ and $x_i = (y_i, z_i) \in D$, $i = 1, 2$,

$$\begin{aligned}&\|P_t(x_1, \cdot) - P_t(x_2, \cdot)\|_{\mathrm{var}} \\ &\quad \leq 2\mathrm{e}^{-\nu(D^c_\varepsilon)t/3} + C_\varepsilon\big(|x_1 - x_2| + |y_1 - y_2|^{1/2}\big)(\bar{v}_{t/3} \vee \mathrm{e}^{-|\beta_{22}|t/3}).\end{aligned}$$

Then the result follows as in the proof of Corollary 3.3. □

Theorems 3.1 and 3.5 and their corollaries are natural extensions of the existing results on CBI and OU-type processes in the literature. In fact, one may see that some parts of the proofs given above essentially follow the ideas of Li and Ma [16] and Wang [29]; see also Wang [28]. For general affine processes on cones, Mayerhofer *et al.* [19] studied the exponential ergodicity in the total variation distance under certain irreducibility, aperiodicity, and finite second moment assumptions. Their techniques were based on the theory of stochastic stability of Markov processes; see Meyn and Tweedie [20, 21] and the references therein. While the results of Mayerhofer *et al.* [19] were formulated in an abstract framework, it seems a delicate task to check their conditions for the process discussed here. Moreover, the finite second moment condition of Mayerhofer *et al.* [19] rules out some natural examples.

4. A weaker condition for ergodicity

Throughout this section, we assume $\beta_{11} < 0$ and $\beta_{22} < 0$. We shall establish the ergodicity of the affine process under a condition on the Lévy measure ν weaker than Condition 3.4. The proof of the result is based on a coupling similar to those used in the last section. Suppose that $\nu(D) > 0$ and choose $0 < \varepsilon < 1$ so that $0 < \nu(D_\varepsilon) < \infty$, where $D_\varepsilon = \mathbb{R}_+ \times [-\varepsilon, \varepsilon]$. Let ν_ε and $\hat{\nu}_\varepsilon$ be defined as in the last section. Let $\gamma^\varepsilon_z = \hat{\nu}_\varepsilon \wedge (\delta_{(0,z)} * \hat{\nu}_\varepsilon)$ for $z \in \mathbb{R}$.

Condition 4.1. *There are constants* $\varepsilon \in (0, 1]$ *and* $\delta > 0$ *such that*

$$q := \inf_{|z|\leq\delta} \gamma^\varepsilon_z(D) = \inf_{|z|\leq\delta} \hat{\nu}_\varepsilon \wedge (\delta_{(0,z)} * \hat{\nu}_\varepsilon)(D) > 0.$$

https://doi.org/10.1017/jpr.2022.100 Published online by Cambridge University Press

The above condition was introduced for general finite-dimensional OU-type processes by Schilling and Wang [27] and Wang [29]; see also Wang [28]. As observed in Wang [29, p. 992], the condition is weaker than Condition 3.4.

Lemma 4.2. *Suppose that Condition 4.1 is satisfied. For $|z| \leq \delta$, let $\hat{\gamma}_z^\varepsilon = \gamma_z^\varepsilon(D)^{-1}\gamma_z^\varepsilon$, and let (η, ρ_1, ζ) be a random vector such that, for $A \in \mathscr{B}(D)$,*

$$\mathbb{P}\{(\eta, \rho_1, \zeta) \in A \times B\} = \begin{cases} q\hat{\gamma}_{-z}^\varepsilon(A)/2, & B = \{z\}, \\ q\hat{\gamma}_z^\varepsilon(A)/2, & B = \{-z\}, \\ [\hat{\nu}_\varepsilon - q(\hat{\gamma}_{-z}^\varepsilon + \hat{\gamma}_z^\varepsilon)/2](A), & B = \{0\}. \end{cases} \tag{4.1}$$

Let $\rho_2 = \rho_1 + \zeta$. Then we have

$$\mathbb{P}(\zeta = z) = \mathbb{P}(\zeta = -z) = q/2, \quad \mathbb{P}(\zeta = 0) = 1 - q \tag{4.2}$$

and

$$\mathbb{P}\{(\eta, \rho_1) \in A\} = \mathbb{P}\{(\eta, \rho_2) \in A\} = \hat{\nu}_\varepsilon(A), \quad A \in \mathscr{B}(D). \tag{4.3}$$

Proof. From (4.1) it is easy to see that ζ has distribution given by (4.2). Moreover, for any $A \in \mathscr{B}(D)$ we have

$$\begin{aligned} \mathbb{P}\{(\eta, \rho_1) \in A\} &= \mathbb{P}\{(\eta, \rho_1) \in A, \zeta = z\} + \mathbb{P}\{(\eta, \rho_1) \in A, \zeta = -z\} \\ &\quad + \mathbb{P}\{(\eta, \rho_1) \in A, \zeta = 0\} \\ &= q\hat{\gamma}_{-z}^\varepsilon(A)/2 + q\hat{\gamma}_z^\varepsilon(A)/2 + [\hat{\nu}_\varepsilon - q(\hat{\gamma}_{-z}^\varepsilon + \hat{\gamma}_z^\varepsilon)/2](A) \\ &= \hat{\nu}_\varepsilon(A) \end{aligned}$$

and

$$\begin{aligned} \mathbb{P}\{(\eta, \rho_2) \in A\} &= \mathbb{P}\{(\eta, \rho_2) \in A, \zeta = z\} + \mathbb{P}\{(\eta, \rho_2) \in A, \zeta = -z\} \\ &\quad + \mathbb{P}\{(\eta, \rho_2) \in A, \zeta = 0\} \\ &= \mathbb{P}\{(\eta, \rho_1) \in A - (0, z), \zeta = z\} + \mathbb{P}\{(\eta, \rho_1) \in A + (0, z), \zeta = -z\} \\ &\quad + \mathbb{P}\{(\eta, \rho_1) \in A, \zeta = 0\} \\ &= q\hat{\gamma}_{-z}^\varepsilon(A - (0, z))/2 + q\hat{\gamma}_z^\varepsilon(A + (0, z))/2 + [\hat{\nu}_\varepsilon - q(\hat{\gamma}_{-z}^\varepsilon + \hat{\gamma}_z^\varepsilon)/2](A) \\ &= q\hat{\gamma}_z^\varepsilon(A)/2 + q\hat{\gamma}_{-z}^\varepsilon(A)/2 + [\hat{\nu}_\varepsilon - q(\hat{\gamma}_{-z}^\varepsilon + \hat{\gamma}_z^\varepsilon)/2](A) \\ &= \hat{\nu}_\varepsilon(A). \end{aligned}$$

Then (4.3) holds. □

Lemma 4.3. *Suppose that Condition 4.1 is satisfied. Let $Q(z, \cdot)$ denote the joint distribution of (η, ρ_1, ρ_2). Then $Q(z, \cdot)$ is a probability kernel from $[-\delta, \delta]$ to $\mathbb{R}_+ \times \mathbb{R}^2$.*

Proof. It is easy to see that $z \mapsto \delta_{(0,z)} * \hat{\nu}_\varepsilon$ is a Borel probability kernel from $[-\delta, \delta]$ to D. Let

$$m^\varepsilon(z, \cdot) = (\hat{\nu}_\varepsilon - \delta_{(0,z)} * \hat{\nu}_\varepsilon)^+ + (\hat{\nu}_\varepsilon - \delta_{(0,z)} * \hat{\nu}_\varepsilon)^-$$

https://doi.org/10.1017/jpr.2022.100 Published online by Cambridge University Press

denote the total variation of the signed measure $\hat{\nu}_\varepsilon - \delta_{(0,z)} * \hat{\nu}_\varepsilon$. By the regularity of the measures, for any bounded positive continuous function f on D we have

$$\int_D f(v) m^\varepsilon(z, \mathrm{d}v) = \sup_{g \in \mathscr{C}_1} \int_D f(v) g(v) \big[\hat{\nu}_\varepsilon(\mathrm{d}v) - (\delta_z * \hat{\nu}_\varepsilon)(\mathrm{d}v)\big],$$

where $\mathscr{C}_1$ is the set of continuous functions g on D satisfying $|g| \leq 1$. Then the mapping

$$z \mapsto \int_D f(v) m^\varepsilon(z, \mathrm{d}v)$$

is lower semicontinuous, so it is a Borel function on $[-\delta, \delta]$. It follows that $m^\varepsilon(z, \cdot)$ and $\gamma_z^\varepsilon = \hat{\nu}_\varepsilon + \delta_z * \hat{\nu}_\varepsilon - m^\varepsilon(z, \cdot)$ are kernels from $[-\delta, \delta]$ to D. From (4.1) we see that $Q(z, \cdot)$ is a Borel probability kernel from $[-\delta, \delta] \setminus \{0\}$ to $\mathbb{R}_+ \times \mathbb{R}^2$. □

Now let us define the first coupling in this section. The basic idea follows that of Schilling and Wang [27] and Wang [29]. Here we give a pathwise construction of the coupling in terms of stochastic integrals. Let $x_1 = (y, z_1) \in D$ and $x_2 = (y, z_2) \in D$, where $y \in \mathbb{R}_+$ and $z_1, z_2 \in \mathbb{R}$ satisfy $z_1 \neq z_2$. Let $k = \lfloor \delta^{-1} |z_1 - z_2| \rfloor + 1$. Then $k^{-1}|z_1 - z_2| \leq \delta$. Let $G = \mathbb{R}_+ \times \mathbb{R}^2$ be endowed with its Borel σ-algebra. In addition to the noises in (2.14) and (2.15), let $N_0(\mathrm{d}s, \mathrm{d}u, \mathrm{d}v_1, \mathrm{d}v_2)$ be an $(\mathscr{F}_t)$-Poisson random measure on $(0, \infty) \times G$ with intensity

$$\nu_\varepsilon(D)\mathrm{d}s Q(k^{-1}(z_1 - z_2)\mathrm{e}^{\beta_{22}s}, \mathrm{d}u, \mathrm{d}v_1, \mathrm{d}v_2). \tag{4.4}$$

We assume all of these noises are independent of each other. For $t \geq 0$ let

$$\xi(t) = (z_1 - z_2) + \int_0^t \int_G (v_1 - v_2)\mathrm{e}^{-\beta_{22}s} N_0(\mathrm{d}s, \mathrm{d}u, \mathrm{d}v_1, \mathrm{d}v_2). \tag{4.5}$$

Then we have

$$\xi(t) = (z_1 - z_2)[1 + L(t)], \qquad t \geq 0, \tag{4.6}$$

where

$$L(t) = (z_1 - z_2)^{-1} \int_0^t \int_G (v_1 - v_2)\mathrm{e}^{-\beta_{22}s} N_0(\mathrm{d}s, \mathrm{d}u, \mathrm{d}v_1, \mathrm{d}v_2).$$

Let $N_0(\mathrm{d}s, \mathrm{d}r)$ be the image of $N_0(\mathrm{d}s, \mathrm{d}u, \mathrm{d}v_1, \mathrm{d}v_2)$ under the mapping

$$(s, u, v_1, v_2) \mapsto (s, r) = (s, (z_1 - z_2)^{-1}(v_1 - v_2)\mathrm{e}^{\beta_{11}s}).$$

We easily see that $N_0(\mathrm{d}s, \mathrm{d}r)$ is a Poisson random measure on $(0, \infty) \times \mathbb{R}$ with intensity $\nu_\varepsilon(D)\mathrm{d}s\pi(s, \mathrm{d}r)$, where $\pi(s, \mathrm{d}r)$ is the probability measure on $\mathbb{R}$ defined by

$$\pi(s, \{1/k\}) = \pi(s, \{-1/k\}) = q/2, \qquad \pi(s, \{0\}) = 1 - q.$$

It follows that

$$N(t) := \int_0^t \int_{\mathbb{R}_+} \int_{\mathbb{R}^2} 1_{\{v_1 \neq v_2\}} N_0(\mathrm{d}s, \mathrm{d}u, \mathrm{d}v_1, \mathrm{d}v_2)$$

https://doi.org/10.1017/jpr.2022.100 Published online by Cambridge University Press

is a Poisson process with parameter $q\nu_\varepsilon(D)$. For $i \geq 1$, let ξ_i denote the size of the ith jump of the process $\{L(t) : t \geq 0\}$. Then $\{\xi_i : i \geq 1\}$ are independent and identically distributed (i.i.d.) random variables with

$$\mathbb{P}(\xi_i = 1/k) = \mathbb{P}(\xi_i = -1/k) = 1/2. \tag{4.7}$$

Let $S(n) = \sum_{i=1}^n \xi_i$. Then the processes $\{N(t) : t \geq 0\}$ and $\{S(n) : n \geq 0\}$ are independent and

$$L(t) = S(N(t)), \qquad t \geq 0.$$

This proves the following.

Lemma 4.4. *The process* $\{L(t) : t \geq 0\}$ *is a continuous-time simple random walk with i.i.d. jumps* $\{\xi_i : i \geq 1\}$ *satisfying* (4.7).

Let $\tau = \inf\{t \geq 0 : \xi(t) = 0\} = \inf\{t \geq 0 : L(t) = -1\}$, and let $\{Y'(t) : t \geq 0\}$ be the solution of

$$\begin{aligned} Y'(t) = y &+ \int_0^t [b_1 + \beta_{11} Y'(s)]\mathrm{d}s + \sqrt{2}\sigma_{11} \int_0^t \int_0^{Y'(s)} W_1(\mathrm{d}s, \mathrm{d}u) \\ &+ \sqrt{2}\sigma_{12} \int_0^t \int_0^{Y'(s)} W_2(\mathrm{d}s, \mathrm{d}u) + \int_0^t \int_0^{Y'(s-)} \int_D v\tilde{M}(\mathrm{d}s, \mathrm{d}u, \mathrm{d}v, \mathrm{d}r) \\ &+ \int_{t\wedge\tau}^t \int_D v_1 N(\mathrm{d}s, \mathrm{d}v) + \int_0^{t\wedge\tau} \int_G u\tilde{N}_0(\mathrm{d}s, \mathrm{d}u, \mathrm{d}v_1, \mathrm{d}v_2). \end{aligned} \tag{4.8}$$

Let $\{Z_1{}'(t) : t \geq 0\}$ and $\{Z_2{}'(t) : t \geq 0\}$ be defined by

$$\begin{aligned} Z_1{}'(t) = \mathrm{e}^{\beta_{22}t}\Big[z_1 &+ \int_0^t \mathrm{e}^{-\beta_{22}s}[b_2 + \beta_{21} Y'(s)]\mathrm{d}s + \sqrt{2}\sigma_0 \int_0^t \mathrm{e}^{-\beta_{22}s}\mathrm{d}W_0(s) \\ &+ \sqrt{2}\sigma_{21} \int_0^t \int_0^{Y'(s)} \mathrm{e}^{-\beta_{22}s} W_1(\mathrm{d}s, \mathrm{d}u) + \sqrt{2}\sigma_{22} \int_0^t \int_0^{Y'(s)} \mathrm{e}^{-\beta_{22}s} W_2(\mathrm{d}s, \mathrm{d}u) \\ &+ \int_0^t \int_0^{Y'(s-)} \int_D \mathrm{e}^{-\beta_{22}s} r\tilde{M}(\mathrm{d}s, \mathrm{d}u, \mathrm{d}v, \mathrm{d}r) + \int_{t\wedge\tau}^t \int_D \mathrm{e}^{-\beta_{22}s} v_2 \tilde{N}(\mathrm{d}s, \mathrm{d}v) \\ &+ \int_0^{t\wedge\tau} \int_G \mathrm{e}^{-\beta_{22}s} v_1 \tilde{N}_0(\mathrm{d}s, \mathrm{d}u, \mathrm{d}v_1, \mathrm{d}v_2)\Big] \end{aligned} \tag{4.9}$$

and

$$\begin{aligned} Z_2{}'(t) = \mathrm{e}^{\beta_{22}t}\Big[z_2 &+ \int_0^t \mathrm{e}^{-\beta_{22}s}[b_2 + \beta_{21} Y'(s)]\mathrm{d}s + \sqrt{2}\sigma_0 \int_0^t \mathrm{e}^{-\beta_{22}s}\mathrm{d}W_0(s) \\ &+ \sqrt{2}\sigma_{21} \int_0^t \int_0^{Y'(s)} \mathrm{e}^{-\beta_{22}s} W_1(\mathrm{d}s, \mathrm{d}u) + \sqrt{2}\sigma_{22} \int_0^t \int_0^{Y'(s)} \mathrm{e}^{-\beta_{22}s} W_2(\mathrm{d}s, \mathrm{d}u) \\ &+ \int_0^t \int_0^{Y'(s-)} \int_D \mathrm{e}^{-\beta_{22}s} r\tilde{M}(\mathrm{d}s, \mathrm{d}u, \mathrm{d}v, \mathrm{d}r) + \int_{t\wedge\tau}^t \int_D \mathrm{e}^{-\beta_{22}s} v_2 \tilde{N}(\mathrm{d}s, \mathrm{d}v) \\ &+ \int_0^{t\wedge\tau} \int_G \mathrm{e}^{-\beta_{22}s} v_2 \tilde{N}_0(\mathrm{d}s, \mathrm{d}u, \mathrm{d}v_1, \mathrm{d}v_2)\Big]. \end{aligned} \tag{4.10}$$

Let $N_1(\mathrm{d}s, \mathrm{d}u, \mathrm{d}v_1)$ and $N_2(\mathrm{d}s, \mathrm{d}u, \mathrm{d}v_2)$ respectively denote the images of the random measure $N_0(\mathrm{d}s, \mathrm{d}u, \mathrm{d}v_1, \mathrm{d}v_2)$ under the mappings

$$(s, u, v_1, v_2) \mapsto (s, u, v_1), \quad (s, u, v_1, v_2) \mapsto (s, u, v_2).$$

Clearly, both $N_1(\mathrm{d}s, \mathrm{d}u, \mathrm{d}w)$ and $N_2(\mathrm{d}s, \mathrm{d}u, \mathrm{d}w)$ are Poisson random measures on $(0, \infty) \times D$ with intensity $\nu_\varepsilon(D)\mathrm{d}s\hat{\nu}_\varepsilon(\mathrm{d}u, \mathrm{d}w) = \mathrm{d}s\nu_\varepsilon(\mathrm{d}u, \mathrm{d}w)$. It follows that both $\{(Y'(t), {Z_1}'(t)) : t \geq 0\}$ and $\{(Y'(t), {Z_2}'(t)) : t \geq 0\}$ are affine processes with transition semigroup $(P_t)_{t\geq 0}$. From (4.8), (4.9), and (4.10) we see that

$$\zeta(t) := {Z_1}'(t) - {Z_2}'(t) = \mathrm{e}^{\beta_{22}t}\xi(t \wedge \tau). \tag{4.11}$$

Then $\{(Y'(t), {Z_1}'(t), Y'(t), {Z_2}'(t)) : t \geq 0\}$ is a coupling of the affine process with coupling time

$$\tau = \inf\{t \geq 0 : \zeta(t) = 0\} = \inf\{t \geq 0 : {X_1}'(t) = {X_2}'(t)\}.$$

Remark 4.5. The intensity (4.4) of the Poisson random measure $N_0(\mathrm{d}s, \mathrm{d}u, \mathrm{d}v_1, \mathrm{d}v_2)$ in (4.8) and (4.9)–(4.10) depends on the difference ${Z_1}'(0) - {Z_2}'(0) = z_1 - z_2$.

Lemma 4.6. *Suppose that Condition 4.1 is satisfied. Then there exists a constant $C_\varepsilon > 0$ such that*

$$\mathbb{P}\{{Z_1}'(t) \neq {Z_2}'(t)\} = \mathbb{P}(\tau > t) \leq C_\varepsilon\big(1 + |x_2 - x_1|\big)\frac{1}{\sqrt{t}}, \quad t > 0. \tag{4.12}$$

Proof. By the reflection principle for the symmetric simple random walk, we have

$$\mathbb{P}\Big(\min_{k\leq n} S(k) > -1\Big) \leq \mathbb{P}\big(|S(n)| \leq 1\big) = \mathbb{P}\Big(\frac{k|S(n)|}{\sqrt{n}} \leq \frac{k}{\sqrt{n}}\Big);$$

see, e.g., Lemma 2.3 in Schilling and Wang [27]. By the Berry–Esseen inequality, there is a universal constant $C_0 > 0$ such that

$$\Big|\mathbb{P}\Big(x \leq \frac{kS(n)}{\sqrt{n}} \leq y\Big) - \frac{1}{\sqrt{2\pi}}\int_x^y \mathrm{e}^{-z^2/2}\mathrm{d}z\Big| \leq \frac{C_0}{\sqrt{n}}, \quad x \leq y \in \mathbb{R}.$$

Let $T = \inf\{n \geq 0 : S(n) = -1\}$. Then we have

$$\begin{aligned}\mathbb{P}(T > n) &= \mathbb{P}\Big(\min_{k\leq n} S(k) > -1\Big)\\ &\leq \frac{C_0}{\sqrt{n}} + \frac{1}{\sqrt{2\pi}}\int_{-k/\sqrt{n}}^{k/\sqrt{n}} \mathrm{e}^{-z^2/2}\mathrm{d}z\\ &\leq \frac{C_0}{\sqrt{n}} + \frac{\sqrt{2}k}{\sqrt{n\pi}} \leq C_1(k+1)\frac{1}{\sqrt{n}},\end{aligned}$$

https://doi.org/10.1017/jpr.2022.100 Published online by Cambridge University Press

where $C_1 = 1 \vee C_0$. By the total probability formula and the independence of $\{S(n) : n \geq 0\}$ and $\{N(t) : t \geq 0\}$, it follows that

$$
\begin{aligned}
\mathbb{P}(\tau > t) &= \sum_{n=0}^{\infty} \mathbb{P}(N(t) = n)\mathbb{P}(T > n | N(t) = n) \\
&= \mathrm{e}^{-q\nu_\varepsilon(D)t}\Big[1 + \sum_{n=1}^{\infty} \frac{(q\nu_\varepsilon(D)t)^n}{n!}\mathbb{P}(T > n)\Big] \\
&\leq \mathrm{e}^{-q\nu_\varepsilon(D)t}\Big[1 + C_1(k+1)\sum_{n=1}^{\infty} \frac{(q\nu_\varepsilon(D)t)^n}{n!}\frac{1}{\sqrt{n}}\Big] \\
&\leq \mathrm{e}^{-q\nu_\varepsilon(D)t}\Big[1 + C_1(k+1)(\mathrm{e}^{q\nu_\varepsilon(D)t} - 1)^{1/2}\Big(\sum_{n=1}^{\infty} \frac{(q\nu_\varepsilon(D)t)^n}{n \cdot n!}\Big)^{1/2}\Big] \\
&\leq \mathrm{e}^{-q\nu_\varepsilon(D)t}\Big[1 + C_1(k+1)\Big(\frac{2(\mathrm{e}^{q\nu_\varepsilon(D)t} - 1)}{q\nu_\varepsilon(D)t}\sum_{n=1}^{\infty} \frac{(q\nu_\varepsilon(D)t)^{n+1}}{(n+1) \cdot n!}\Big)^{1/2}\Big] \\
&\leq \mathrm{e}^{-q\nu_\varepsilon(D)t}\Big[1 + \sqrt{2}C_1(k+1)(\mathrm{e}^{q\nu_\varepsilon(D)t} - 1)\frac{1}{\sqrt{q\nu_\varepsilon(D)t}}\Big] \\
&\leq \mathrm{e}^{-q\nu_\varepsilon(D)t} + \sqrt{2}C_1(|x_1 - x_2| + 2)(1 - \mathrm{e}^{-q\nu_\varepsilon(D)t})\frac{1}{\sqrt{q\nu_\varepsilon(D)t}}.
\end{aligned}
$$

Then (4.12) holds for some constant $C_\varepsilon \geq 0$. □

We next construct the main coupling of this section, through a concatenation of two couplings. Let $D([0, \infty), D^2)$ denote the space of càdlàg paths from $[0, \infty)$ to D. Let $\{w(t) : t \geq 0\} = \{(w_1(t), w_2(t), w_3(t), w_4(t)) : t \geq 0\}$ denote the coordinate process of this space, and let $(\mathscr{F}_t : t \geq 0)$ be its natural filtration generated by the coordinate process. Let $\tau_0^w = \{t \geq 0 : w_1(t) = w_3(t)\}$ and $\tau^w = \{t \geq \tau_0(w) : w_2(t) = w_4(t)\}$. For $s \geq 0$, let θ_s be the *shifting operator* on $D([0, \infty), D^2)$ defined by $\theta_s w(t) = w(s + t)$, $t \geq 0$. For $s \geq 0$ and $w \in D([0, \infty), D^2)$, the *stopped path* $w^s \in D([0, \infty), D^2)$ is defined by $w^s(t) = w(s \wedge t)$, $t \geq 0$.

Let $x_1 = (y_1, z_1) \in D$ and $x_2 = (y_2, z_2) \in D$, where $y_1, y_2 \in \mathbb{R}_+$ and $z_1, z_2 \in \mathbb{R}$. For $i = 1, 2$, let $\{Y_i(t) : t \geq 0\}$ be the solution of (2.14) with $Y_i(0) = y_i$, and let $\{Z_i(t) : t \geq 0\}$ be defined by (2.16) with $Z_i(0) = z_i$. Let $\mathbb{P}^{2,2}_{(x_1,x_2)}$ be the distribution on $D([0, \infty), D^2)$ of $\{(Y_1(t), Z_1(t), Y_2(t), Z_2(t)) : t \geq 0\}$. Let $\mathbb{P}^{1,2}_{(y,z_1,z_2)}$ be the distribution on $D([0, \infty), D^2)$ of the process $\{(Y'(t), {Z_1}'(t), Y'(t), {Z_2}'(t)) : t \geq 0\}$ defined by (4.8)–(4.10). Let $\mathbb{P}_{(x_1,x_2)}$ be the probability measure on $D([0, \infty), D^2)$ defined by

$$
\begin{aligned}
&\mathbb{P}_{(x_1,x_2)}\big[F\big((w_1, w_2, w_3, w_4)^{\tau_0^w}\big)G\big(\theta_{\tau_0^w}(w_1, w_2, w_3, w_4)\big)\big] \\
&\quad = \mathbb{P}^{2,2}_{(x_1,x_2)}\big[F\big((w_1, w_2, w_3, w_4)^{\tau_0^w}\big)\mathbb{P}^{1,2}_{(w_1(\tau_0^w), w_3(\tau_0^w), w_4(\tau_0^w))}G(w_1, w_2, w_3, w_4)\big],
\end{aligned}
$$

where F and G are Borel functions on $D([0, \infty), D^2)$, and the probability symbols are used to denote the corresponding expectations. Then under $\mathbb{P}_{(x_1,x_2)}$ the coordinate process $\{(w_1(t), w_2(t), w_3(t), w_4(t)) : t \geq 0\}$ evolves according to the transition law of $\{(Y_1(t), Z_1(t), Y_2(t), Z_2(t)) : t \geq 0\}$ up to time τ_0^w, after which it evolves according to the transition law of $\{(Y'(t), {Z_1}'(t), Y'(t), {Z_2}'(t)) : t \geq 0\}$. It is clear that both $\{(w_1(t), w_2(t)) : t \geq 0\}$ and

https://doi.org/10.1017/jpr.2022.100 Published online by Cambridge University Press

$\{(w_3(t), w_4(t)) : t \geq 0\}$ are affine processes with transition semigroup $(P_t)_{t\geq 0}$. Thus they form a coupling of the affine process with coupling time τ^w.

Remark 4.7. One might wish to construct a coupling through stochastic equations by a direct concatenation of the two sets of stochastic equations. As above, one first constructs the process $\{(Y_1(t), Z_1(t), Y_2(t), Z_2(t)) : t \geq 0\}$ by (2.14)–(2.16) and defines the stopping time $\tau_0 = \{t \geq 0 : Y_1(t) = Y_2(t)\}$. By Remark 4.5 one would need a Poisson random measure with intensity depending on the random variable $Z_1(\tau_0) - Z_2(\tau_0)$ to define the coupling process on the time interval $[\tau_0, \infty)$ by (4.8)–(4.10). We leave the details to the interested reader.

Theorem 4.8. *Suppose that Condition 4.1 is satisfied. Then there is a constant $C_\varepsilon \geq 0$ such that*

$$\|P_t(x_1, \cdot) - P_t(x_2, \cdot)\|_{\mathrm{var}} \leq C_\varepsilon(1 + |x_1 - x_2|)t^{-1/2}, \quad t > 0,\ x_1, x_2 \in D. \tag{4.13}$$

Proof. There is no loss of generality in assuming $y_1 \geq y_2$. Using the coupling of the affine process constructed above, we have

$$\mathbb{P}_{(x_1,x_2)}(\tau_1^w > t) \leq \mathbb{P}_{(x_1,x_2)}(\tau_0^w > t/2) + \mathbb{P}_{(x_1,x_2)}(\tau_0^w \leq t/2,\ \tau_1^w > t),$$

where $\mathbb{P}_{(x_1,x_2)}(\tau_0^w > t/2) \leq |y_1 - y_2|\bar{v}_{t/2}$ by (3.1). Since $t - \tau_0^w$ is measurable relative to $\mathscr{F}_{\tau_0^w}$, we can use Lemma 4.6 to see

$$\begin{aligned}
&\mathbb{P}_{(x_1,x_2)}\big(\tau_0^w \leq t/2,\ \tau_1^w > t\big) \\
&\quad = \mathbb{E}_{(x_1,x_2)}\big[1_{\{\tau_0^w \leq t/2\}}\mathbb{P}_{(x_1,x_2)}\big(\tau_1^w > t|\mathscr{F}_{\tau_0^w}\big)\big] \\
&\quad \leq C_\varepsilon\mathbb{E}_{(x_1,x_2)}\big[1_{\{\tau_0^w \leq t/2\}}\big(1 + |w_3(\tau_0^w) - w_4(\tau_0^w)|\big)(t - \tau_0^w)^{-1/2}\big] \\
&\quad \leq C_\varepsilon t^{-1/2}\mathbb{E}\Big(1 + \sup_{0\leq s\leq t/2}|Z_1(s) - Z_2(s)|\Big).
\end{aligned}$$

Then the result follows by Proposition 2.6 and (2.21). □

Corollary 4.9. *Suppose that Condition 4.1 is satisfied. Then there is a constant $C_\varepsilon \geq 0$ such that*

$$\|P_t(x, \cdot) - \pi\|_{\mathrm{var}} \leq C_\varepsilon(1 + |x|)t^{-1/2}, \quad t > 0,\ x \in D. \tag{4.14}$$

The above corollary gives an extension of Theorem 1.1 of Schilling and Wang [27]; see also Theorem 2(i) of Wang [29].

Acknowledgements

We would like to express our gratitude to two anonymous referees for their very helpful comments and suggestions on the paper and on the literature. We are grateful to Peisen Li and Jian Wang for their helpful comments.

Funding information

This research was supported by the National Key R&D Program of China (No. 2020YFA0712900).

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] BANSAYE, V. AND MÉLÉARD, S. (2015). *Stochastic Models for Structured Populations: Scaling Limits and Long Time Behavior*. Springer, Cham.

[2] BAO, J. H. AND WANG, J. (2020). *Coupling methods and exponential ergodicity for two factor affine processes*. Preprint. Available at https://arxiv.org/abs/2004.10384.

[3] BARCZY, M., DÖRING, L., LI, Z. AND PAP, G. (2014). Stationarity and ergodicity for an affine two-factor model. *Adv. Appl. Prob.* **46**, 878–898.

[4] DAWSON, D. A. AND LI, Z. (2006). Skew convolution semigroups and affine Markov processes. *Ann. Prob.* **34**, 1103–1142.

[5] DAWSON, D. A. AND LI, Z. (2012). Stochastic equations, flows and measure-valued processes. *Ann. Prob.* **40**, 813–857.

[6] DUFFIE, D., FILIPOVIĆ, D. AND SCHACHERMAYER, W. (2003). Affine processes and applications in finance. *Ann. Appl. Prob.* **13**, 984–1053.

[7] FRIESEN, M. AND JIN, P. (2020). On the anisotropic stable JCIR process. *ALEA Lat. Amer. J. Prob. Math. Statist.* **2**, 643–674.

[8] FRIESEN, M., JIN, P. AND RÜDIGER, B. (2020). Stochastic equation and exponential ergodicity in Wasserstein distances for affine processes. *Ann. Appl. Prob.* **30**, 2165–2195.

[9] JIN, P., KREMER, J. AND RÜDIGER, B. (2017). Exponential ergodicity of an affine two-factor model based on the α-root process. *Adv. Appl. Prob.* **49**, 1144–1169.

[10] JIN, P., KREMER, J. AND RÜDIGER, B. (2019). Moments and ergodicity of the jump-diffusion CIR process. *Stochastics* **91**, 974–997.

[11] JIN, P., KREMER, J. AND RÜDIGER, B. (2020). Existence of limiting distribution for affine processes. *J. Math. Anal. Appl.* **486**, article no. 123912, 31 pp.

[12] KAWAZU, K. AND WATANABE, S. (1971). Branching processes with immigration and related limit theorems. *Theory Prob. Appl.* **16**, 36–54.

[13] KELLER-RESSEL, M., SCHACHERMAYER, W. AND TEICHMANN, J. (2011). Affine processes are regular. *Prob. Theory Relat. Fields* **151**, 591–611.

[14] KYPRIANOU, A. E. (2014). *Fluctuations of Lévy Processes with Applications*, 2nd edn. Springer, Heidelberg.

[15] LI, Z. (2011). *Measure-Valued Branching Markov Processes*. Springer, Heidelberg.

[16] LI, Z. AND MA, C. (2015). Asymptotic properties of estimators in a stable Cox–Ingersoll–Ross model. *Stoch. Process. Appl.* **125**, 3196–3233.

[17] LI, Z. (2020). Continuous-state branching processes with immigration. In *From Probability to Finance: Lecture Notes of BICMR Summer School on Financial Mathematics*, ed. Y. Jiao, Springer, Singapore, pp. 1–69.

[18] LI, Z. (2021). Ergodicities and exponential ergodicities of Dawson–Watanabe type processes. *Theory Prob. Appl.* **66**, 276–298.

[19] MAYERHOFER, E., STELZER, R. AND VESTWEBER, J. (2020). Geometric ergodicity of affine processes on cones. *Stoch. Process. Appl.* **130**, 4141–4173.

[20] MEYN, S. P. AND TWEEDIE, R. L. (1993). Stability of Markovian processes III: Foster–Lyapunov criteria for continuous-time processes. *Adv. Appl. Prob.* **25**, 518–548.

[21] MEYN, S. P. AND TWEEDIE, R. L. (2009). *Markov Chains and Stochastic Stability*, 2nd edn. Cambridge University Press.

[22] PARDOUX, E. (2016). *Probabilistic Models of Population Evolution: Scaling Limits, Genealogies and Interactions*. Springer, Cham.

[23] PINSKY, M. A. (1972). Limit theorems for continuous state branching processes with immigration. *Bull. Amer. Math. Soc.* **78**, 242–244.

[24] PRIOLA, E. AND ZABCZYK, J. (2009). Densities for Ornstein–Uhlenbeck processes with jumps. *Bull. London Math. Soc.* **41**, 41–50.

[25] REVUZ, D. AND YOR, M. (1999). *Continuous Martingales and Brownian Motion*, 3rd edn. Springer, Berlin.

[26] SATO, K. AND YAMAZATO, M. (1984). Operator-self-decomposable distributions as limit distributions of processes of Ornstein–Uhlenbeck type. *Stoch. Process. Appl.* **17**, 73–100.

[27] SCHILLING, R. L. AND WANG, J. (2012). On the coupling property and the Liouville theorem for Ornstein–Uhlenbeck processes. *J. Evol. Equat.* **12**, 119–140.

[28] WANG, F. (2011). Coupling for Ornstein–Uhlenbeck jump processes with jumps. *Bernoulli* **17**, 1136–1158.
[29] WANG, J. (2012). On the exponential ergodicity of Lévy-driven Ornstein–Uhlenbeck processes. *J. Appl. Prob.* **49**, 990–1004.
[30] ZHANG, X. AND GLYNN, P. (2018). *Affine jump-diffusions: stochastic stability and limit theorems*. Preprint. Available at https://arxiv.org/abs/1811.00122.