# Mixed state branching evolution for cell division models

Shukai Chen,[1] Lina Ji[2] and Jie Xiong[3]

*School of Mathematics and Statistics & Key Laboratory of Analytical Mathematics and Applications (Ministry of Education) &*
*Fujian Provincial Key Laboratory of Statistics and Artificial Intelligence &*
*Fujian Key Laboratory of Analytical Mathematics and Applications (FJKLAMA) &*
*Center for Applied Mathematics of Fujian Province (FJNU), Fujian Normal University*
*Fuzhou 350007, People's Republic of China.*
*Faculty of Computational Mathematics and Cybernetics, Shenzhen MSU-BIT University*
*Shenzhen 518172, People's Republic of China.*
*Department of Mathematics and National Center for Applied Mathematics (Shenzhen)*
*Southern University of Science and Technology, Shenzhen 518055, China.*
*E-mail: skchen@mail.bnu.edu.cn*, *jiln@smbu.edu.cn* and *xiongj@sustech.edu.cn*



**Abstract**

We prove a scaling limit theorem for two-type Galton-Waston branching processes with interaction. The limit theorem gives rise to a class of mixed state branching processes with interaction using to simulate the evolution for cell division affected by parasites. Such process can also be obtained by the pathwise unique solution to a stochastic equation system. Moreover, we present sufficient conditions for extinction with probability one and the exponential ergodicity in the $L^1$-Wasserstein distance of such process in some cases.

**Keywords and phrases: mixed state branching process; stochastic integral equation; interaction.**

## 1 Introduction

Let $\mathbb{N} = \{0, 1, 2, ...\}$. We consider a continuous time model in $D = [0, \infty) \times \mathbb{N}$ for cells and parasites, where the behavior of cell division is infected by parasites. Informally, the quantity of parasites $(X(t))_{t\geq 0}$ in a cell evolves as a continuous state branching process. The cells divide in continuous time at a rate $h(x, y)$ which may depend on the quantity of parasites $x$ and cells $y$. This framework is general enough to be applied for the modelling of other structured populations, for instance, grass-rabbit models in [11].

Many studies have been conducted on branching within branching processes to study such population dynamics in continuous time. In [25], the evolution of parasites is modelled by a birth-death process, while the cells split according to a Yule process. [2] allows the quantity of parasites in a cell following a Feller diffusion. A continuous state branching process with jumps is considered to model the quantity of parasites in a cell in [24]. In particular, [25, 2, 24] describe cell populations in a tree structure, in this way, the population of cells at some time may be represented by a random point measure and associated martingale problems can be established by choosing test functions appropriately. Instead of [25, 2, 24], in this paper we ignore the tree

[1]Supported by National Key R&D Program of China (No. 2022YFA1006003) and China Postdoctoral Science Foundation (No. 2022M720735).

[2]Supported by NSFC grant (No. 12301167 and No. 12271029), Guangdong Basic and Applied Basic Research Foundation (No. 2022A1515110986) and Guangdong Young Innovative Talents Project (No. 2022KQNCX105).

[3]Supported by National Key R&D Program of China grant (No. 2022YFA1006102) and NSFC grant (No. 11831010).

structure and mainly focus on a parasite-cell model from a macro point of view. More precisely, we use a stochastic equation system to describe the sample path of such models,

$$\begin{cases} X(t) = X(0) - b\int_0^t X(s)\,\mathrm{d}s + \int_0^t \sqrt{2cX(s)}\,\mathrm{d}B(s) + \int_0^t \int_0^{X(s-)} \int_0^{\infty} \xi\,\tilde{M}(\mathrm{d}s,\mathrm{d}u,\mathrm{d}\xi), \\ Y(t) = Y(0) + \int_0^t \int_0^{Y(s-)} \int_0^{h(X(s-),Y(s-))} \int_{\mathbb{N}} (\xi-1)\,N(\mathrm{d}s,\mathrm{d}u,\mathrm{d}r,\mathrm{d}\xi), \end{cases}$$

where $b \in \mathbb{R}$ and $c \geq 0$ are constants, $(B(t))_{t\geq 0}$ is a standard Brownian motion, $h(\cdot,\cdot) \in C(\mathbb{R}_+^2)^+$, here $C(\mathbb{R}_+^2)^+$ is the collection of continuous positive functions defined on $\mathbb{R}_+^2$. Let $(\xi \wedge \xi^2)\,m(\mathrm{d}\xi)$ be a finite measure on $(0,\infty)$ and $(p_\xi : \xi \in \mathbb{N})$ be an offspring distribution satisfying $\sum_\xi \xi p_\xi < \infty$. Without losing generality, we assume $p_1 = 0$. The above $M(\mathrm{d}s,\mathrm{d}u,\mathrm{d}\xi)$ is a Poisson random measure on $(0,\infty)^3$ with intensity $\mathrm{d}s\mathrm{d}u m(\mathrm{d}\xi)$, and $\tilde{M}(\mathrm{d}s,\mathrm{d}u,\mathrm{d}\xi) = M(\mathrm{d}s,\mathrm{d}u,\mathrm{d}\xi) - \mathrm{d}s\mathrm{d}u m(\mathrm{d}\xi)$. The above $N$ is a Poisson random measure on $(0,\infty)^3 \times \mathbb{N}$ with intensity $\mathrm{d}s\mathrm{d}u\mathrm{d}r n(\mathrm{d}\xi)$, where $n(\mathrm{d}\xi) = p_\xi \sharp(\mathrm{d}\xi)$ and $\sharp(\cdot) = \sum_j \delta_j(\cdot)$ is the counting measure on $\mathbb{N}$. Those three random elements ($B(t)$, $M$ and $N$) are independent of each other. Apparently, $(X(t))_{t\geq 0}$ is indeed a *continuous-state branching process* (CB-process), see [6, 7]. In particular, when $h(\cdot,\cdot) \equiv r > 0$, $(Y(t))_{t\geq 0}$ is a standard continuous time Markov branching process with branching rate $r > 0$ and offspring $(p_\xi, \xi \in \mathbb{N})$, in this case, the system can be seen as a particular case of mixed state branching processes, which has been studied in [4].

For simplicity, we introduce another Poisson random measure and write it again by $N$ on $[0,\infty)^3 \times \mathbb{N}^{-1}$, $\mathbb{N}^{-1} = \mathbb{N} \cup \{-1\}$ with characteristic measure $n(\mathrm{d}\xi) = p'_\xi \sharp(\mathrm{d}\xi)$, $p'_\xi = p_{\xi+1}$. Then we can rewrite the system by

$$\begin{cases} X(t) = X(0) - b\displaystyle\int_0^t X(s)\,\mathrm{d}s + \int_0^t \sqrt{2cX(s)}\,\mathrm{d}B(s) + \int_0^t \int_0^{X(s-)} \int_0^{\infty} \xi\,\tilde{M}(\mathrm{d}s,\mathrm{d}u,\mathrm{d}\xi), & (1.1) \\ Y(t) = Y(0) + \displaystyle\int_0^t \int_0^{Y(s-)} \int_0^{h(X(s-),Y(s-))} \int_{\mathbb{N}^{-1}} \xi\,N(\mathrm{d}s,\mathrm{d}u,\mathrm{d}r,\mathrm{d}\xi). & (1.2) \end{cases}$$

In the rest of the paper, we use the stochastic equation system (1.1)–(1.2) to describe the parasite-cell model. In the literature on the theory of branching processes, the rescaling (in time or state) approach plays a valuable role in establishing the connection among those branching processes, see [12], [17, 18], [23], [4] and [21] and the references therein. To the best of our knowledge, limited work has been done in branching processes with interactions. This leads to the first purpose of this paper, and the establishment of strong uniqueness of solution to (1.1)–(1.2). For a sequence of two-type Galton–Watson processes with interactions $\{(x_k(n), y_k(n))_{n\in\mathbb{N}}\}_{k\geq 1}$, we prove that $(x_k(\lfloor\gamma_k t\rfloor)/k, y_k(\lfloor\gamma_k t\rfloor))_{t\geq 0}$ converges in distribution to the solution to (1.1)–(1.2) as $k \to \infty$ under suitable conditions. The pathwise uniqueness of solution to (1.1)–(1.2) is also given.

In addition, the second purpose of this paper is to study several long time behaviors of such process and we mainly obtain the extinction behavior and exponential ergodicity in the $L^1$-Wasserstein distance in some cases. The result of extinction behavior is inspired by [16]. Furthermore, ergodicity is the foundation for a wide class of limit theorems and long-time behavior for Markov processes. Due to the nonlinearity of function $h$, the semigroup transition of $(X, Y)$ is not explicit. We obtain the ergodic property by a coupling approach, which has been proved to be effective in the study of ergodicity of nonlinear case, see [3, 22] and the references therein.

We now introduce some notation. Let $\mathrm{e}_\lambda(z) = \mathrm{e}^{-\langle\lambda,z\rangle}$ for any $\lambda = (\lambda_1,\lambda_2) \in \mathbb{R}_+^2$ and $z = (x,y) \in D$, where $\langle\lambda,z\rangle = \lambda_1 x + \lambda_2 y$. We use $C_b(D)$ to denote the set of all bounded functions $(x,y) \mapsto f(x,y)$ on $D$ with $x \mapsto f(x,\cdot)$ continuous. Let $C_b^2(D)$ be the subset of $C_b(D)$ with continuous bounded derivatives up to 2nd order on $x$. Let $C_0^2(D)$ be the subset of $C_b^2(D)$ vanishing at infinity, and $C_c^2(D)$ be the subset of $C_0^2(D)$ with compact support. Define $C_b(\mathbb{R}_+^2)$ to be the collection of all bounded continuous functions on $\mathbb{R}_+^2$, which is a subset of $C_b(D)$. Let $C_b^{2,1}(\mathbb{R}_+^2)$ be the subset of $C_b(\mathbb{R}_+^2)$ with continuous bounded derivatives up to 2nd order on $x$ and continuous bounded derivatives up to first order on $y$. Then we have $C_b^{2,1}(\mathbb{R}_+^2) \subset C_b^2(D)$. Let $\mathbb{D}([0,\infty), D)$ denote the space of càdlàg paths from $[0,\infty)$ to $D$ furnished with the Skorokhod topology. In the integrals, we make the convention that, for $a \le b \in \mathbb{R}$,

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

This paper is structured as follows. The existence by a scaling limit of a sequence of two-type Galton-Watson processes with interaction and pathwise uniqueness of solution to (1.1)–(1.2) are given in Section 2. In Section 3 the extinction behavior of the system is studied. In Section 4, an exponential ergodic property is proved under some conditions.

## 2 Existence and pathwise uniqueness of solution

The generator $A$ of $(X(t), Y(t))_{t\ge0}$ satisfying (1.1)–(1.2) is determined by

$$\begin{aligned} Af(z) &= x\Big[-bf'_x + cf''_{xx} + \int_0^\infty \{f(x+\xi,y) - f(x,y) - \xi f'_x\} m(\mathrm{d}\xi)\Big] \\ &\quad + \gamma(x,y)\int_{\mathbb{N}^{-1}} \Big\{f(x,y+\xi) - f(x,y)\Big\} n(\mathrm{d}\xi) \end{aligned} \tag{2.3}$$

for any $f \in C_b^2(D)$, where $z = (x,y) \in D$ and $\gamma(x,y) = h(x,y)y$. Then

$$A\mathrm{e}_\lambda(z) = \mathrm{e}_\lambda(z)\Big[x\phi_1(\lambda_1) + \gamma(x,y)\phi_2(\lambda_2)\Big], \tag{2.4}$$

where

$$\phi_1(\lambda_1) = b\lambda_1 + c\lambda_1^2 + \int_0^\infty (\mathrm{e}^{-\lambda_1\xi} - 1 + \lambda_1\xi)\, m(\mathrm{d}\xi), \tag{2.5}$$

$$\phi_2(\lambda_2) = \int_{\mathbb{N}^{-1}} (\mathrm{e}^{-\lambda_2\xi} - 1)\, n(\mathrm{d}\xi). \tag{2.6}$$

We first consider the case of $h \in C_b(\mathbb{R}_+^2)^+$. Given the initial value $(x(0), y(0)) \in \mathbb{N}\times\mathbb{N}$, let $(x(n), y(n))_{n\ge0}$ be a two-dimensional process defined by

$$x(n) = \sum_{j=1}^{x(n-1)} \alpha_{n-1,j}, \qquad y(n) = \sum_{j=1}^{y(n-1)} \beta_{n-1,j,\theta(x(n-1),y(n-1))}, \quad n \ge 1, \tag{2.7}$$

where $\{\alpha_{n,j} : n \ge 0, j \ge 1\}$ are integer-valued i.i.d. random variables with offspring distribution $(w(i) : i \in \mathbb{N})$. Given $x, y \in \mathbb{N}$, the above $\{\beta_{n,j,\theta(x,y)} : n \ge 0, j \ge 1\}$ are i.i.d. integer-valued random variables with offspring distribution $(v^{\theta(x,y)}(i) : i \in \mathbb{N})$ depending on the function $\theta$. Let $g_1$ and $g_2^{\theta(x,y)}$ be the generating functions of $(w(i) : i \in \mathbb{N})$ and $(v^{\theta(x,y)}(i) : i \in \mathbb{N})$, respectively. It is known that $(x(n), y(n))_{n\ge0}$ is a Markov process and we call it *two-type Galton-Watson process with interaction.* Suppose that there exists a sequence of two-type Galton-Watson processes with interaction $\{(x_k(n), y_k(n))_{n\ge0}\}_{k\ge1}$ with parameters $(g_{k,1}, g_{k,2}^{\theta_k(x,y)})$. Let $\{\gamma_k\}_{k\ge1}$ be a sequence of positive numbers with $\gamma_k \to \infty$ as $k \to \infty$. For $(x,y) \in \mathbb{N}\times\mathbb{N}$, we introduce several functions on $\mathbb{R}_+$ as below:

$$\begin{aligned} &\bar{\Phi}_{k,1}(\lambda_1) = k\gamma_k \log\Big[1 - (k\gamma_k)^{-1}\Phi_{k,1}(\lambda_1)\mathrm{e}^{\lambda_1/k}\Big], \\ &\Phi_{k,1}(\lambda_1) = k\gamma_k\Big[\mathrm{e}^{-\lambda_1/k} - g_{k,1}(\mathrm{e}^{-\lambda_1/k})\Big], \\ &\bar{\Phi}_{k,2}^{\theta_k(x,y)}(\lambda_2) = \gamma_k \log\Big[1 - \gamma_k^{-1}\Phi_{k,2}^{\theta_k(x,y)}(\lambda_2)\mathrm{e}^{\lambda_2}\Big], \\ &\Phi_{k,2}^{\theta_k(x,y)}(\lambda_2) = \gamma_k\Big[\mathrm{e}^{-\lambda_2} - g_{k,2}^{\theta_k(x,y)}(\mathrm{e}^{-\lambda_2})\Big]. \end{aligned}$$

Let $E_k = \{0, k^{-1}, 2k^{-1}, \cdots\}$ for each $k \geq 1$. For any $x \in \mathbb{R}_+$, we take $x_k := \lfloor kx \rfloor / k$. Then $x_k \in E_k$ and $|x_k - x| \leq 1/k$. Let $D_k := E_k \times \mathbb{N}$. Then $D_k$ is a subset of $D$. We define a continuous-time stochastic process taking values on $D_k$ as $(X_k(t), Y_k(t))_{t \geq 0} := (x_k(\lfloor \gamma_k t \rfloor)/k, y_k(\lfloor \gamma_k t \rfloor))_{t \geq 0}$. Denote $Z_k(t) = (X_k(t), Y_k(t))$ to simplify the notation. In order to state our main results in this section, we first present the assumption taken throughout this section.

**Condition 2.1**

*(2.1.1) The sequence $\{\Phi_{k,1}(\lambda_1)\}_{k \geq 1}$ is uniformly Lipschitz in $\lambda_1$ on each bounded interval, and converges to a continuous function as $k \to \infty$;*

*(2.1.2) $\gamma_k[1 - v_k^{\theta_k(kx_k,y)}(1)] \to h(x, y)$ uniformly in $(x, y) \in \mathbb{R}_+ \times \mathbb{N}$ as $k \to \infty$;*

*(2.1.3) $\frac{v_k^{\theta_k(kx_k,y)}(\xi)}{1 - v_k^{\theta_k(kx_k,y)}(1)} \to p_\xi$ for $\xi \in \mathbb{N} \backslash \{1\}$ uniformly in $(x, y) \in \mathbb{R}_+ \times \mathbb{N}$ as $k \to \infty$.*

*(2.1.4) The sequence $\{\Phi_{k,2}^{\theta_k(kx_k,y)}(\lambda_2)\}_{k \geq 1}$ is uniformly Lipschitz in $\lambda_2$ on each bounded interval, where the Lipshcitz coefficient is independent from $x, y$.*

By [20, Proposition 2.5], under Condition (2.1.1), $\Phi_{k,1}(\lambda_1)$ converges to a function with representation (2.5) as $k \to \infty$, see also [17, 18]. Moreover, there exists a constant $K > 0$ such that

$$\sup_k \Phi'_{k,1}(0+) = \sup_k \gamma_k [g'_{k,1}(1-) - 1] \leq K. \tag{2.8}$$

**Example 2.1** *Let $\{p_\xi : \xi = 0, 1, \cdots\}$ be an offspring distribution with $p_1 = 0$. Let*

$$v_k^{\theta_k(kx_k,y)}(\xi) = p_\xi \gamma_k^{-1/2} \left(1 - e^{-\gamma_k^{-1/2} h(x_k,y)}\right)$$

*for any $\xi \in \mathbb{N} \backslash \{1\}$ and*

$$v_k^{\theta_k(kx_k,y)}(1) = 1 - \gamma_k^{-1/2} \left(1 - e^{-\gamma_k^{-1/2} h(x_k,y)}\right).$$

*Then we have*

$$\Phi_{k,2}^{\theta_k(kx_k,y)}(\lambda_2) = \gamma_k^{1/2} \left(1 - e^{-\gamma_k^{-1/2} h(x_k,y)}\right) \left(\mathrm{e}^{-\lambda_2} - g(\mathrm{e}^{-\lambda_2})\right)$$

*with $g(\mathrm{e}^{-\lambda_2}) = \sum_{\xi=0}^\infty p_\xi e^{-\lambda_2 \xi}$. It is easy to check that the above satisfies Condition (2.1.2)–(2.1.4).*

## 2.1 Main results

We have the following statements about a scaling limit theorem of mixed state branching processes with interactions, existence and pathwise uniqueness of solution to (1.1)–(1.2).

**Theorem 2.2** *Suppose that Condition 2.1 holds and $h \in C_b(\mathbb{R}_+^2)^+$. Let $Z_k(0)$ converge in distribution to $Z(0)$ as $k \to \infty$ with $\sup_k \mathbb{E}[X_k(0) + Y_k(0)] < \infty$. Then $\{(Z_k(t))_{t \geq 0}\}_{k \geq 1}$ converges in distribution on $\mathbb{D}(0, \infty), D)$ to $(Z(t))_{t \geq 0}$, which is a solution to (1.1)–(1.2).*

**Theorem 2.3** *Assume that $h \in C_b(\mathbb{R}_+^2)^+$. For any given initial value $(X(0), Y(0)) \in D$, the pathwise uniqueness holds for (1.1)–(1.2) on $D$.*

**Remark 2.4** *For the classical branching process, the transition semigroup can be determined uniquely by the log-Laplace transform since branching property. Then the linear hull of* $\{e^{-\lambda x} : \lambda \geq 0\}$ *is the core for generator of the process. Therefore, the scaling limit theorem can be obtained in the sense of convergence in distribution on the skorokhod space by [8, p.226 and pp.233-234]. We refer to [12, 17] and references therein for details. However, the branching property is invalid for our model since the interaction. It is not enough to give estimation only when* $f = e_\lambda$*. Therefore, in the following proof of Theorem 2.2, we give the existence of solution to the martingale problem by tightness for general* $f$*, which implies the existence of solution to* (1.1)-(1.2).

*As we know, there is a unique strong solution to* (1.1)*, which is a CB-process. Based on this, in the proof of Theorem 2.3, we then construct the pathwise unique solution to* (1.2) *by path stitching method.*

One sees that there is a unique positive strong solution to (1.1)-(1.2) by Theorems 2.2 and 2.3 when $h \in C_b(\mathbb{R}_+^2)^+$. However, the boundedness assumption of $h$ is removed in the following result.

**Theorem 2.5** *Suppose that* $h \in C(\mathbb{R}_+^2)^+$*. Then there exists a unique positive strong solution to* (1.1)–(1.2).

## 2.2 Proof of main results

**Proposition 2.6** *Under Conditions (2.1.2)–(2.1.3),* $e^{\lambda_2}\Phi_{k,2}^{\theta_k(kx_k,y)}(\lambda_2)$ *converges to* $-h(x,y)\phi_2(\lambda_2)$ *uniformly for* $(x,y,\lambda_2) \in \mathbb{R}_+ \times \mathbb{N} \times \mathbb{R}_+$ *as* $k \to \infty$*, where* $h \in C_b(\mathbb{R}_+^2)^+$ *and* $\phi_2(\lambda_2)$ *is given by* (2.6).

*Proof.* One can see that

$$
\begin{aligned}
&e^{\lambda_2}\Phi_{k,2}^{\theta_k(kx_k,y)}(\lambda_2)\\
&\quad = \gamma_k\Big[1 - e^{\lambda_2} g_{k,2}^{\theta_k(kx_k,y)}(e^{-\lambda_2})\Big]\\
&\quad = \gamma_k\Big[1 - e^{\lambda_2}\sum_{j=0}^{\infty} e^{-\lambda_2 j} v_k^{\theta_k(kx_k,y)}(j)\Big]\\
&\quad = \gamma_k \sum_{j=0}^{\infty}(1 - e^{-\lambda_2(j-1)}) v_k^{\theta_k(kx_k,y)}(j)\\
&\quad = \gamma_k\left[1 - v_k^{\theta_k(kx_k,y)}(1)\right]\int_{\mathbb{N}^{-1}\setminus\{0\}}(1 - e^{-\lambda_2\xi})\rho_k^{\theta_k(kx_k,y)}(d\xi),
\end{aligned}
$$

where

$$
\begin{aligned}
\rho_k^{\theta_k(kx_k,y)}(d\xi) &= \frac{1}{1 - v_k^{\theta_k(kx_k,y)}(1)}\sum_{j=0}^{\infty} v_k^{\theta_k(kx_k,y)}(j)\delta_{j-1}(d\xi)\\
&= \frac{v_k^{\theta_k(kx_k,y)}(\xi+1)}{1 - v_k^{\theta_k(kx_k,y)}(1)}\sharp(d\xi)
\end{aligned}
$$

for $\xi \in \mathbb{N}_{-1}\setminus\{0\}$ with $\sharp(d\xi)$ being the counting measure on $\mathbb{N}_{-1}$. The result follows from Conditions (2.1.2)–(2.1.3). □

For $\lambda = (\lambda_1, \lambda_2) \in \mathbb{R}_+^2$, we then have

$$
e_\lambda(Z_k(t)) = e_\lambda(Z_k(0)) + \sum_{i=1}^{\lfloor \gamma_k t\rfloor}\left[e_\lambda\left(Z_k\left(\frac{i}{\gamma_k}\right)\right) - e_\lambda\left(Z_k\left(\frac{i-1}{\gamma_k}\right)\right)\right]
$$

$$
\begin{aligned}
&= \mathrm{e}_\lambda(Z_k(0)) + \sum_{i=1}^{\lfloor\gamma_k t\rfloor} \gamma_k^{-1} A_k \mathrm{e}_\lambda\left(Z_k\left(\frac{i-1}{\gamma_k}\right)\right) + M_{k,\lambda}(t) \\
&= \mathrm{e}_\lambda(Z_k(0)) + \int_0^{\lfloor\gamma_k t\rfloor/\gamma_k} A_k e_\lambda(Z_k(s))\mathrm{d}s + M_{k,\lambda}(t), \qquad (2.9)
\end{aligned}
$$

where

$$
\begin{aligned}
M_{k,\lambda}(t) \;=\; \sum_{i=1}^{\lfloor\gamma_k t\rfloor} \Bigg\{ &\left[\mathrm{e}_\lambda\left(Z_k\left(\frac{i}{\gamma_k}\right)\right) - \mathrm{e}_\lambda\left(Z_k\left(\frac{i-1}{\gamma_k}\right)\right)\right] \\
&-\mathbb{E}\left[\mathrm{e}_\lambda\left(Z_k\left(\frac{i}{\gamma_k}\right)\right) - \mathrm{e}_\lambda\left(Z_k\left(\frac{i-1}{\gamma_k}\right)\right)\middle|\mathscr{F}_{\frac{i-1}{\gamma_k}}\right]\Bigg\} \qquad (2.10)
\end{aligned}
$$

is a martingale and for $z = (x, y) \in D$,

$$
A_k \mathrm{e}_\lambda(z) = \gamma_k\left[(g_{k,1}(\mathrm{e}^{-\lambda_1/k}))^{kx_k} \cdot (g_{k,2}^{\theta_k(kx_k,y)}(\mathrm{e}^{-\lambda_2}))^y - \mathrm{e}_\lambda(z)\right].
$$

One can check that

$$
A_k \mathrm{e}_\lambda(z) = \mathrm{e}_\lambda(z)\left[x\bar{\Phi}_{k,1}(\lambda_1) + y\bar{\Phi}_{k,2}^{\theta_k(kx_k,y)}(\lambda_2)\right] + o(1).
$$

By the above, [20, Proposition 2.5] and Proposition 2.6, we have the following estimation.

**Theorem 2.7** *Suppose that Condition 2.1 holds. Then for any $\lambda > 0$, we have*

$$
\lim_{k\to\infty} \sup_{z\in D_k} |A_k \mathrm{e}_\lambda(z) - A\mathrm{e}_\lambda(z)| = 0,
$$

*where $A$ is the generator defined by* (2.3)*.*

**Proposition 2.8** *Suppose that Condition 2.1 holds. Let $T > 0$ be a fixed constant and*

$$
\sup_k \mathbb{E}[X_k(0) + Y_k(0)] < \infty.
$$

*Then we have*

$$
\sup_k \sup_{0\le t\le T} \mathbb{E}[X_k(t) + Y_k(t)] < \infty.
$$

*Proof.* By (2.8) one sees that $0 \le g'_{k,1}(1-) \le K/\gamma_k + 1$. Then for $t \in [\frac{i}{\gamma_k}, \frac{i+1}{\gamma_k})$, we have

$$
\begin{aligned}
\mathbb{E}[X_k(t)] &= k^{-1}\mathbb{E}[x_k(\lfloor\gamma_k t\rfloor)] \\
&= g'_{k,1}(1-)k^{-1}\mathbb{E}[x_k(\lfloor\gamma_k t\rfloor - 1)] \\
&\le (K/\gamma_k + 1)k^{-1}\mathbb{E}[x_k(\lfloor\gamma_k t\rfloor - 1)].
\end{aligned}
$$

By induction, we have $\mathbb{E}[X_k(t)] \le (K/\gamma_k + 1)^{\lfloor\gamma_k t\rfloor}\mathbb{E}[X_k(0)]$. Moreover, by Condition (2.1.4), we have

$$
\sup_k \left|\frac{\partial}{\partial\lambda_2}\Phi_{k,2}^{\theta_k(kx_k,y)}(\lambda_2)\Big|_{\lambda_2=0}\right| = \sup_k \gamma_k\left|\frac{\partial}{\partial z}g_{k,2}^{\theta_k(kx_k,y)}(z)\Big|_{z=1} - 1\right| \le K.
$$

Similarly, for $t \in [\frac{i}{\gamma_k}, \frac{i+1}{\gamma_k})$, one sees that

$$\begin{aligned}
\mathbb{E}[Y_k(t)] &= \mathbb{E}[y_k(\lfloor \gamma_k t \rfloor)] = \mathbb{E}\left[ \sum_{k=1}^{y_k(n-1)} \mathbb{E}\left[ \beta_{n-1,k,\theta_k(x_k(n-1),y_k(n-1))} \Big| x_k(n-1), y_k(n-1) \right] \right] \Bigg|_{n=\lfloor \gamma_k t \rfloor} \\
&= \mathbb{E}\left[ y_k(n-1) \cdot \frac{\partial}{\partial z} g_{k,2}^{\theta_k(x_k(n-1),y_k(n-1))}(z) \Bigg|_{z=1} \right] \Bigg|_{n=\lfloor \gamma_k t \rfloor} \\
&\leq (1 + K/\gamma_k)\mathbb{E}[y_k(\lfloor \gamma_k t \rfloor - 1)].
\end{aligned}$$

Then we get $\mathbb{E}[Y_k(t)] \leq (K/\gamma_k + 1)^{\lfloor \gamma_k t \rfloor}\mathbb{E}[Y_k(0)]$ by induction. The result follows. □

Let $\{\tau_k : k \geq 1\}$ be a sequence of bounded stopping times, and $\{\delta_k : k \geq 1\}$ be a sequence of positive constants with $\delta_k \to 0$ as $k \to \infty$. For a fixed constant $T > 0$, we assume that

$$0 \leq \tau_k \leq \tau_k + \delta_k \leq T.$$

**Proposition 2.9** *Suppose that Condition 2.1 holds and $h \in C_b(\mathbb{R}_+^2)^+$. Then for any $\lambda \in \mathbb{R}_+^2$, we have*

$$\lim_{k \to \infty} \mathbb{E}\left[ |\mathrm{e}_\lambda(Z_k(\tau_k + \delta_k)) - \mathrm{e}_\lambda(Z_k(\tau_k))|^2 \right] = 0.$$

*Proof.* For any $\lambda \in \mathbb{R}_+^2$, by (2.9) we have

$$\begin{aligned}
&\mathbb{E}\left[ |\mathrm{e}_\lambda(Z_k(\tau_k + \delta_k)) - \mathrm{e}_\lambda(Z_k(\tau_k))|^2 \right] \\
&\quad \leq |\mathbb{E}\left[ \mathrm{e}_{2\lambda}(Z_k(\tau_k + \delta_k)) - \mathrm{e}_{2\lambda}(Z_k(\tau_k)) \right]| \\
&\qquad + |\mathbb{E}\left[ 2\mathrm{e}_\lambda(Z_k(\tau_k))[\mathrm{e}_\lambda(Z_k(\tau_k + \delta_k)) - \mathrm{e}_\lambda(Z_k(\tau_k))] \right]| \\
&\quad \leq I_1 + I_2 + I_3,
\end{aligned}$$

where

$$\begin{aligned}
I_1 &= \left| \mathbb{E}\left[ \int_{\lfloor \gamma_k \tau_k \rfloor/\gamma_k}^{\lfloor \gamma_k(\tau_k + \delta_k) \rfloor/\gamma_k} A_k \mathrm{e}_{2\lambda}(Z_k(s))\mathrm{d}s \right] \right|, \\
I_2 &= \left| \mathbb{E}\left[ 2\mathrm{e}_\lambda(Z_k(\tau_k)) \int_{\lfloor \gamma_k \tau_k \rfloor/\gamma_k}^{\lfloor \gamma_k(\tau_k + \delta_k) \rfloor/\gamma_k} A_k \mathrm{e}_\lambda(Z_k(s))\mathrm{d}s \right] \right|, \\
I_3 &= \left| \mathbb{E}\left[ 2\mathrm{e}_\lambda(Z_k(\tau_k))\left( M_{k,\lambda}(\tau_k + \delta_k) - M_{k,\lambda}(\tau_k) \right) \right] \right|.
\end{aligned}$$

Then by (2.4), Theorem 2.7 and Proposition 2.8, one can see that

$$\begin{aligned}
I_1 &\leq \mathbb{E}\left[ \int_{\lfloor \gamma_k \tau_k \rfloor/\gamma_k}^{\lfloor \gamma_k(\tau_k + \delta_k) \rfloor/\gamma_k} |A_k \mathrm{e}_{2\lambda}(Z_k(s)) - A\mathrm{e}_{2\lambda}(Z_k(s))|\,\mathrm{d}s \right] \\
&\quad + \mathbb{E}\left[ \int_{\lfloor \gamma_k \tau_k \rfloor/\gamma_k}^{\lfloor \gamma_k(\tau_k + \delta_k) \rfloor/\gamma_k} |A\mathrm{e}_{2\lambda}(Z_k(s))|\,\mathrm{d}s \right] \leq K\delta_k.
\end{aligned}$$

Similarly,

$$\begin{aligned}
I_2 &\leq 2\mathbb{E}\left[ \int_{\lfloor \gamma_k \tau_k \rfloor/\gamma_k}^{\lfloor \gamma_k(\tau_k + \delta_k) \rfloor/\gamma_k} |A_k \mathrm{e}_\lambda(Z_k(s))|\,\mathrm{d}s \right] \\
&\leq 2\mathbb{E}\left[ \int_{\lfloor \gamma_k \tau_k \rfloor/\gamma_k}^{\lfloor \gamma_k(\tau_k + \delta_k) \rfloor/\gamma_k} |A_k \mathrm{e}_\lambda(Z_k(s)) - A\mathrm{e}_\lambda(Z_k(s))|\,\mathrm{d}s \right]
\end{aligned}$$

$$+2\mathbb{E}\left[\int_{\lfloor\gamma_k\tau_k\rfloor/\gamma_k}^{\lfloor\gamma_k(\tau_k+\delta_k)\rfloor/\gamma_k}|A\mathrm{e}_\lambda(Z_k(s))|\,\mathrm{d}s\right]\le K\delta_k.$$

Moreover, recall that $h$ is bounded, it follows from (2.10) and Doob's stopping theorem that

$$\begin{aligned}
&\mathbb{E}\left[\mathrm{e}_\lambda(Z_k(\tau_k))\left(M_{k,\lambda}(\tau_k+\delta_k)-M_{k,\lambda}(\tau_k)\right)\right]\\
&\quad=\mathbb{E}\left[\mathbb{E}\left[\mathrm{e}_\lambda(Z_k(\tau_k))\left(M_{k,\lambda}(\tau_k+\delta_k)-M_{k,\lambda}(\tau_k)\right)|\mathscr{F}_{\lfloor\gamma_k\tau_k\rfloor}\right]\right]\\
&\quad=\mathbb{E}\left[\mathrm{e}_\lambda(Z_k(\tau_k))\left[\mathbb{E}\left[M_{k,\lambda}(\tau_k+\delta_k)\middle|\mathscr{F}_{\lfloor\gamma_k\tau_k\rfloor}\right]-M_{k,\lambda}(\tau_k)\right]\right]\\
&\quad=0,
\end{aligned}$$

which implies that $I_3=0$. The result holds. □

**Corollary 2.10** *Suppose that Condition 2.1 holds and $h\in C_b(\mathbb{R}_+^2)^+$. Then for any $\lambda:=(\lambda_1,\lambda_2)\in\mathbb{R}_+^2$, we have*

$$\lim_{k\to\infty}L^\lambda_{\tau_k,\delta_k}(Z_k)=0,$$

*where* $L^\lambda_{\tau_k,\delta_k}(Z_k):=\mathbb{E}\left[\left|e^{-\lambda_1X_k(\tau_k+\delta_k)}-e^{-\lambda_1X_k(\tau_k)}\right|^2+\left|e^{-\lambda_2Y_k(\tau_k+\delta_k)}-e^{-\lambda_2Y_k(\tau_k)}\right|^2\right]$.

*Proof.* The result follows by taking $\lambda=(\lambda_1,0)$ and $\lambda=(0,\lambda_2)$ in Proposition 2.9. □

Similar to the proof of [20, Theorem 3.6], we get the following result.

**Proposition 2.11** *Suppose that Condition 2.1 holds and $h\in C_b(\mathbb{R}_+^2)^+$. Let $Z_k(0)=(X_k(0),Y_k(0))$ be the initial value satisfying $\sup_k\mathbb{E}[X_k(0)+Y_k(0)]<\infty$. Then the process $\{(Z_k(t))_{t\ge0}\}_{k\ge1}=\{(X_k(t),Y_k(t))_{t\ge0}\}_{k\ge1}$ is tight on $\mathbb{D}([0,\infty),D)$.*

*Proof.* By Aldous's criterion, it suffices to show that, for any $\epsilon>0$,

$$\lim_{k\to\infty}\mathbb{P}\left[\|Z_k(\tau_k+\delta_k)-Z_k(\tau_k)\|_2>\epsilon\right]=0,\tag{2.11}$$

where $\|\cdot\|_2$ is the $L^2$ norm on $D$. For any $a:=(a_1,a_2),\ b:=(b_1,b_2)\in D$ satisfying $\|a-b\|_2>\epsilon$, we have $|a_1-b_1|\wedge|a_2-b_2|>\epsilon/2$. Then for a fixed constant $M>0$, by taking $0\le\|a\|_2,\|b\|_2\le M$, one sees that

$$|\mathrm{e}^{-\lambda_1a_1}-\mathrm{e}^{-\lambda_1b_1}|^2+|\mathrm{e}^{-\lambda_2a_2}-\mathrm{e}^{-\lambda_2b_2}|^2\ge\left(\frac{1}{2}(\lambda_1\wedge\lambda_2)\epsilon e^{-(\lambda_1+\lambda_2)M}\right)^2.$$

By Proposition 2.9, it is easy to see that

$$\begin{aligned}
&\mathbb{P}\left\{\|Z_k(\tau_k+\delta_k)-Z_k(\tau_k)\|_2>\epsilon;\|Z_k(\tau_k)\|_2\vee\|Z_k(\tau_k+\delta_k)\|_2\le M\right\}\\
&\quad\le\left(\frac{1}{2}(\lambda_1\wedge\lambda_2)\epsilon e^{-(\lambda_1+\lambda_2)M}\right)^{-2}L^\lambda_{\tau_k,\delta_k}(Z_k)\to0
\end{aligned}$$

as $k\to\infty$. Further, by Proposition 2.8, we have

$$\begin{aligned}
\mathbb{P}\left[\|Z_k(\tau_k+\delta_k)\|_2\ge M\right]&\le\mathbb{P}\left[X_k(\tau_k+\delta_k)\ge\frac{M}{2}\right]+\mathbb{P}\left[Y_k(\tau_k+\delta_k)\ge\frac{M}{2}\right]\\
&\le2\frac{\sup_{0\le t\le T}\mathbb{E}[X_k(t)+Y_k(t)]}{M}\le\frac{K}{M}.
\end{aligned}$$

Similarly, we get

$$\mathbb{P}\left[\|Z_k(\tau_k)\|_2 \geq M\right] \leq \frac{K}{M}.$$

As a result,

$$\begin{aligned}
&\mathbb{P}\left[\|Z_k(\tau_k+\delta_k) - Z_k(\tau_k)\|_2 > \epsilon\right] \\
&\quad \leq \mathbb{P}\left[\|Z_k(\tau_k+\delta_k) - Z_k(\tau_k)\|_2 > \epsilon; \|Z_k(\tau_k)\|_2 \vee \|Z_k(\tau_k+\delta_k)\|_2 \leq M\right] \\
&\qquad + \mathbb{P}\left[\|Z_k(\tau_k+\delta_k) - Z_k(\tau_k)\|_2 > \epsilon; \|Z_k(\tau_k+\delta_k)\|_2 \geq M\right] \\
&\qquad + \mathbb{P}\left[\|Z_k(\tau_k+\delta_k) - Z_k(\tau_k)\|_2 > \epsilon; \|Z_k(\tau_k)\|_2 \geq M\right] \\
&\quad \leq \mathbb{P}\left[\|Z_k(\tau_k+\delta_k) - Z_k(\tau_k)\|_2 > \epsilon; \|Z_k(\tau_k)\|_2 \vee \|Z_k(\tau_k+\delta_k)\|_2 \leq M\right] \\
&\qquad + \mathbb{P}\left[\|Z_k(\tau_k+\delta_k)\|_2 \geq M\right] + \mathbb{P}\left[\|Z_k(\tau_k)\|_2 \geq M\right]
\end{aligned}$$

goes to 0 as $k \to \infty$ and $M \to \infty$, which implies (2.11). The result follows. $\square$

**Lemma 2.12** *For any $f \in C_b^2(D)$, there exists a sequence of functions $f^{m,n} \in C_0^2(D)$ such that $f^{m,n} \to f$, $f_1^{m,n} \to f_1$ and $f_{11}^{m,n} \to f_{11}$ uniformly on any bounded subset of $D$ as $m,n \to \infty$, where $f_1^{m,n} := \frac{\partial f^{m,n}(x,y)}{\partial x}$, $f_1 := \frac{\partial f(x,y)}{\partial x}$, $f_{11}^{m,n} := \frac{\partial^2 f^{m,n}(x,y)}{\partial x^2}$ and $f_{11} := \frac{\partial^2 f(x,y)}{\partial x^2}$.*

*Proof.* For any nonnegative function $f \in C_b^2(D)$, we define

$$f^{m,n}(x,y) = \begin{cases} f(x,y), & (x,y) \in [0,m] \times [0,n] \cap D, \\ f(x,y)\left[1 - 2\int_m^x \rho(2(z-m)-1)dz\right], & (x,y) \in [m,m+1] \times [0,n] \cap D; \\ 0, & \text{others,} \end{cases}$$

where $\rho$ is the mollifier defined by

$$\rho(x) = \Lambda \exp\{-1/(1-x^2)\} 1_{\{|x|<1\}}$$

with $\Lambda$ being the constant such that $\int_{\mathbb{R}} \rho(x)dx = 1$. It is easy to see that $f^{m,n} \in C_0^2(D)$. Notice that, for $(x,y) \in [m,m+1] \times [0,n] \cap D$,

$$f_1^{m,n}(x,y) = f_1(x,y) - \frac{d}{dx}\left[2f(x,y)\int_m^x \rho(2(z-m)-1)dz\right]$$

and

$$f_{11}^{m,n}(x,y) = f_{11}(x,y) - \frac{d^2}{dx^2}\left[2f(x,y)\int_m^x \rho(2(z-m)-1)dz\right].$$

Let $D^b$ be a fixed bounded subset of $D$. Then we have

$$\sup_{(x,y)\in D^b} \left[|f^{m,n}(x,y) - f(x,y)| + |f_1^{m,n}(x,y) - f_1(x,y)| + |f_{11}^{m,n}(x,y) - f_{11}(x,y)|\right] \to 0$$

as $m,n \to \infty$. The result follows. $\square$

Now we are ready to give the existence of the solution to (1.1)–(1.2) for the case of $h \in C_b(D)^+$.

**Proof of Theorem 2.2** Let $P^{(k)}$ be the distributions of $Z_k$ on $\mathbb{D}([0,\infty), D)$. By Proposition 2.11, the sequence of processes $\{Z_k\}_{k\geq 1}$ is relatively compact. Then there are a probability measure $Q$ and a subsequence $P^{(k_i)}$ on $\mathbb{D}([0,\infty), D)$ such that $Q = \lim_{i\to\infty} P^{(k_i)}$. By Skorokhod representative theorem, there exists a probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathbb{P}})$ on which are defined càdlàg processes $(\tilde{Z}(t))_{t\geq 0}$ and $(\tilde{Z}_{k_i}(t))_{t\geq 0}$ such that the distribution of $\tilde{Z}$ and $\tilde{Z}_{k_i}$ on $\mathbb{D}([0,\infty), D)$ are $Q$ and $P^{(k_i)}$, respectively, and $\lim_{i\to\infty} \tilde{Z}_{k_i} = \tilde{Z}$, $\tilde{\mathbb{P}}$-almost surely.

Now it suffices to show that $(\tilde{Z}(t))_{t\geq 0}$ satisfies the following martingale problem: for any $f \in C_b^2(D)$, we have

$$f(\tilde{Z}(t)) = f(\tilde{Z}(0)) + \int_0^t Af(\tilde{Z}(s))ds + \text{local mart.} \tag{2.12}$$

Let $f(z) = \mathrm{e}_\lambda(z)$ for any $z \in D$. By (2.9), we get

$$\mathrm{e}_\lambda(\tilde{Z}_{k_i}(t)) = \mathrm{e}_\lambda(\tilde{Z}_{k_i}(0)) + \int_0^{\lfloor \gamma_{k_i} t\rfloor/\gamma_{k_i}} A_{k_i}\mathrm{e}_\lambda(\tilde{Z}_{k_i}(s))ds + M_{k_i,\lambda}(t).$$

One sees that

$$\begin{aligned}
&\int_0^{\lfloor \gamma_{k_i} t\rfloor/\gamma_{k_i}} \left| A_{k_i}\mathrm{e}_\lambda(\tilde{Z}_{k_i}(s)) - A\mathrm{e}_\lambda(\tilde{Z}(s))\right| ds \\
&\quad \leq \int_0^{\lfloor \gamma_{k_i} t\rfloor/\gamma_{k_i}} \left| A_{k_i}\mathrm{e}_\lambda(\tilde{Z}_{k_i}(s)) - A\mathrm{e}_\lambda(\tilde{Z}_{k_i}(s))\right| ds \\
&\qquad + \int_0^{\lfloor \gamma_{k_i} t\rfloor/\gamma_{k_i}} \left| A\mathrm{e}_\lambda(\tilde{Z}_{k_i}(s)) - A\mathrm{e}_\lambda(\tilde{Z}(s))\right| ds =: I^1_{k_i} + I^2_{k_i}.
\end{aligned}$$

Then $I^1_{k_i} \to 0$ as $i \to \infty$ by Theorem 2.7. On the other hand, let

$$C_X := \{t > 0 : \tilde{P}(\tilde{Z}(t-) = \tilde{Z}(t)) = 1\}.$$

Then the set $\mathbb{R}_+ \backslash C_X$ is at most countable. Then we have $I^2_{k_i} \to 0$ as $i \to \infty$. Consequently, the process $(\tilde{Z}(t))_{t\geq 0}$ satisfies the martingale problem (2.12) when $f(z) = \mathrm{e}_\lambda(z)$.

Let $f \in C_0^2(D)$ be fixed, and $E_0$ be the linear hull of $\{\mathrm{e}_\lambda(z) : \lambda \in \mathbb{R}_+^2\}$. By Stone-Weierstrass Theorem and (2.3), there exists a sequence of functions $f_n \in E_0$ such that $Af_n(z) \to Af(z)$ uniformly on each bounded subset of $D$ as $n \to \infty$. As a linear span of $\{\mathrm{e}_\lambda(z)\}$, we have

$$f_n(\tilde{Z}(t)) = f_n(\tilde{Z}(0)) + \int_0^t Af_n(\tilde{Z}(s))ds + \text{local mart.} \tag{2.13}$$

Let $\tilde{\tau}_N := \inf\{t > 0 : \tilde{X}(t) \geq N \text{ or } \tilde{Y}(t) \geq N\}$. Then $\tilde{\tau}_N \to \infty$ almost surely as $N \to \infty$ by Proposition 2.8 and Fatou's lemma. Replacing $t$ with $t \wedge \tilde{\tau}_N$, and taking limits as $n \to \infty$ on both sides of (2.13), we then have

$$f(\tilde{Z}(t \wedge \tilde{\tau}_N)) = f(\tilde{Z}(0)) + \int_0^t Af(\tilde{Z}(s \wedge \tilde{\tau}_N -))ds + \text{mart.} \tag{2.14}$$

Next, for the general function $f \in C_b^2(D)$, by Lemma 2.12, there exists a sequence functions $f^{m,n} \in C_0^2(D)$ such that $Af^{m,n}(z) \to Af(z)$ uniformly on each bounded subset of $D$ as $m, n \to \infty$. Similar to the above, (2.14) holds for any $f \in C_b^2(D)$. Letting $N \to \infty$, one can see that $(\tilde{Z}(t))_{t\geq 0}$ satisfies the martingale problem (2.12), which implies that $(\tilde{Z}(t))_{t\geq 0}$ is a weak solution to (1.1)–(1.2). The result follows. □

**Lemma 2.13** *Assume that* $h \in C_b(\mathbb{R}_+^2)^+$. *Let* $(X(t), Y(t))_{t\geq 0}$ *be the solution to* (1.1)-(1.2) *with* $\mathbb{E}[X(0) + Y(0)] < \infty$. *Then for any* $T > 0$ *we have*

$$\sup_{0\leq t\leq T} \mathbb{E}[X(t) + Y(t)] < \infty.$$

*Proof.* Taking $f(x,y)=x+y$ and $Z(t)=(X(t),Y(t))$, by (2.14) we have

$$\mathbb{E}[X(t\wedge\tilde{\tau}_N)+Y(t\wedge\tilde{\tau}_N)]=\mathbb{E}[X(0)+Y(0)]+\int_0^t\mathbb{E}\left[Af(Z(s\wedge\tilde{\tau}_{N-}))\right]\mathrm{d}s.$$

By (2.3) and $h\in C_b(\mathbb{R}_+^2)^+$, one sees that

$$\begin{aligned}\mathbb{E}[X(t\wedge\tilde{\tau}_N)+Y(t\wedge\tilde{\tau}_N)]\le\ &\mathbb{E}[X(0)+Y(0)]\\&+\int_0^t\mathbb{E}\left[-bX(s\wedge\tilde{\tau}_N)+\left(\sup_{(x,y)\in\mathbb{R}_+^2}h(x,y)\int_{\mathbb{N}^{-1}}\xi n(\mathrm{d}\xi)\right)Y(s\wedge\tilde{\tau}_N)\right]\mathrm{d}s\\\le\ &\mathbb{E}[X(0)+Y(0)]+K\int_0^t\mathbb{E}[X(s\wedge\tilde{\tau}_N)+Y(s\wedge\tilde{\tau}_N)]\mathrm{d}s.\end{aligned}$$

Then the result follows by Gronwall's inequality and letting $N\to\infty$. □

**Proof of Theorem 2.3** By [6, Theorems 5.1 and 5.2] and [10, Corollary 5.2], there is a unique positive strong solution to (1.1), which is a CB-process. Let $(X(t))_{t\ge0}$ be the unique positive strong solution to (1.1). The pathwise uniqueness of solution to (1.2) can be constructed by path stitching method. Let $\kappa_0=0$. Then $Y(\kappa_0)=Y(0)$. Given $\kappa_{k-1}\ge0$, $Y(\kappa_{k-1})\ge0$ and the process $(X(t))_{t\ge0}$, we first define

$$\kappa_k=\kappa_{k-1}+\inf\left\{t>0:\int_{\kappa_{k-1}}^{t+\kappa_{k-1}}\int_0^{Y(\kappa_{k-1})}\int_0^{h(X(s-),Y(\kappa_{k-1}))}\int_{\mathbb{N}^{-1}}N(\mathrm{d}s,\mathrm{d}u,\mathrm{d}r,\mathrm{d}\xi)=1\right\}.$$

Then we define $Y(t)=Y(0)$ for any $t\in[\kappa_0,\kappa_1)$ and

$$\Delta_k:=\int_{\kappa_{k-1}}^{\kappa_k}\int_0^{Y(\kappa_{k-1})}\int_0^{h(X(s-),Y(\kappa_{k-1}))}\int_{\mathbb{N}^{-1}}\xi N(\mathrm{d}s,\mathrm{d}u,\mathrm{d}r,\mathrm{d}\xi).$$

For any $k=1,2,\cdots$, let $Y(t):=Y(\kappa_{k-1})+\Delta_k$ for any $t\in[\kappa_k,\kappa_{k+1})$, which uniquely determine the behavior of the trajectory $t\to Y(t)$ on the time interval $[\kappa_k,\kappa_{k+1})$, $k=1,2,\cdots$. Let $\kappa:=\lim_{k\to\infty}\kappa_k$. Then $(Y(t))_{t\ge0}$ is the pathwise-unique solution to (1.2) up to time $\kappa$. Notice that $\kappa_k$ is the time of $k$-th jump of $Y$. By Lemma 2.13, one sees that

$$\mathbb{E}\left[\int_0^{t\wedge\kappa_k}\int_0^{Y(s-)}\int_0^{h(X(s-),Y(s-))}\int_{\mathbb{N}^{-1}}N(\mathrm{d}s,\mathrm{d}u,\mathrm{d}r,\mathrm{d}\xi)\right]\le Kt\sup_{0\le s\le t}\mathbb{E}[Y(s)]<\infty.$$

It follows that $\mathbb{P}(\kappa>t)=1$ for any $t\ge0$, which implies that $\mathbb{P}(\kappa=\infty)=1$. Then $(X(t),Y(t))_{t\ge0}$ is the pathwise-unique solution to (1.1)-(1.2). The result follows. □

**Proof of Theorem 2.5** Let $h_m(x,y):=h(x\wedge m,y\wedge m)$. Then $h_m$ is bounded for any $m\ge1$ and $h_m\to h$ as $m\to\infty$. By Theorems 2.2 and 2.3, there exists a unique strong solution $(X_m(t),Y_m(t))_{t\ge0}$ to the following stochastic integral equation system:

$$\begin{cases}X(t)=X(0)-b\int_0^tX(s)\,\mathrm{d}s+\int_0^t\sqrt{2cX(s)}\,\mathrm{d}B(s)+\int_0^t\int_0^{X(s-)}\int_0^\infty\xi\,\tilde{M}(\mathrm{d}s,\mathrm{d}u,\mathrm{d}\xi),\\Y(t)=Y(0)+\int_0^t\int_0^{Y(s-)}\int_0^{h_m(X(s-),Y(s-))}\int_{\mathbb{N}^{-1}}\xi\,N(\mathrm{d}s,\mathrm{d}u,\mathrm{d}r,\mathrm{d}\xi).\end{cases}\tag{2.15}$$

In fact, $(X_m(t))_{t\ge0}$ is the unique strong solution to (1.1) independent with $m$, which is written as $(X(t))_{t\ge0}$ in the following. Let $\tau_m^X:=\inf\{t>0:X(t)\ge m\}$, $\tau_m^Y:=\inf\{t>0:Y_m(t)\ge m\}$ and $\tau_m=\tau_m^X\wedge\tau_m^Y$. Then $0\le X(t)<m$ and $0\le Y_m(t)<m$ for $0\le t<\tau_m$, and $(X(t),Y_m(t))$ satisfies (1.1)-(1.2) for $0\le t<\tau_m$. For $n\ge m\ge1$, let

$$Y_m(\tau_m)=Y_m(\tau_m-)+\int_{\{\tau_m\}}\int_0^{Y_m(\tau_m-)}\int_0^{h_n(X(\tau_m-),Y_m(\tau_m-))}\int_{\mathbb{N}^{-1}}\xi N(\mathrm{d}s,\mathrm{d}u,\mathrm{d}r,\mathrm{d}\xi).$$

There exists a unique strong solution $(X(t), \tilde{Y}(t))_{t\geq\tau_m}$ to (1.1) and

$$Y(t) = Y_m(\tau_m) + \int_{\tau_m}^{t}\int_0^{Y(s-)}\int_{\tau_m}^{h_n(X(s-),Y(s-))}\int_{\mathbb{N}^{-1}} \xi\, N(\mathrm{d}s, \mathrm{d}u, \mathrm{d}r, \mathrm{d}\xi).$$

Let $Y'(t) = Y_m(t)$ for $0 \leq t < \tau_m$ and $Y'(t) = \tilde{Y}(t)$ for $t \geq \tau_m$. Then it is a solution to (2.15) by changing $m$ to $n$. By the strong uniqueness we get $(X(t), Y'(t))_{t\geq 0} = (X(t), Y_n(t))_{t\geq 0}$ almost surely. In particular, we infer $Y_n(t) = Y_m(t) < m$ for $0 \leq t < \tau_m$. Consequently, the sequence $\{\tau_m\}$ is non-decreasing. On the other hand, by (2.15) it is easy to check that $\mathbb{E}[X(t\wedge\tau_m^X)] \leq \mathbb{E}[X(0)]e^{Kt}$, where $K$ is a constant independent with $m$. Then we have $\tau_m^X \to \infty$ almost surely as $m \to \infty$. Let $\tau = \lim_{m\to\infty} \tau_m = \lim_{m\to\infty} \tau_m^Y$. Let $Y(t) = Y_m(t)$ for all $0 \leq t < \tau_m$ and $m \geq 1$. It is easily seen that $(X(t), Y(t))_{t\in[0,\tau)}$ is a unique strong solution to (1.1)–(1.2) up to $\tau$. For $t \geq \tau$, let $(X(t), Y(t)) = (X(t), \infty)$. The result follows. $\square$

# 3 Large time behaviors

## 3.1 Foster-Lyapunov criteria for extinction

In this subsection, we mainly discuss the extinction behavior of such processes under $b \geq 0$. Define $\tau_0 = \inf\{t > 0 : X(t) = 0 \text{ and } Y(t) = 0\}$. Moreover, we separately define the extinction time of $X, Y$ as $\tau_0^X := \inf\{t > 0 : X(t) = 0\}$ and $\tau_0^Y := \{t > 0 : Y(t) = 0\}$. Then we have $\tau_0 = \tau_0^X \vee \tau_0^Y$. For the extinction behavior of the process $X$, we introduce the so called Grey's condition:

**Condition 3.1** *There is some constant $\theta > 0$ so that $\phi_1(z) > 0$ for $z \geq \theta$ and $\int_\theta^\infty \phi_1^{-1}(z)\mathrm{d}z < \infty$, where $\phi_1$ is given by (2.5).*

Since $b \geq 0$, under Condition 3.1, one can see that $\mathbb{P}_x(\tau_0^X < \infty) = 1$ for all $x > 0$; see, e.g., [20, Corollary 3.8]. In the following, we present a Foster-Lyapunov criteria-type result for the process $(X, Y)$. For $z_1 = (x_1, y_2), z_2 = (x_2, y_2) \in D$, we say $z_1 \succeq z_2$ if $x_1 \geq x_2$ and $y_1 \geq y_2$. Let $\tilde{z} := (\tilde{x}, \tilde{y}) \succeq z_0$ and $(X(t), Y(t))_{t\geq 0}$ be the mixed state branching process satisfying (1.1)–(1.2) with initial value $z_0$. We define stopping time $\sigma_{\tilde{z}} = \inf\{t > 0 : X(t) \geq \tilde{x} \text{ or } Y(t) \geq \tilde{y}\}$. It is easy to see that $X_{t\wedge\sigma_{\tilde{z}}-} \leq \tilde{x}$ and $Y_{t\wedge\sigma_{\tilde{z}}-} \leq \tilde{y}$.

**Theorem 3.2** *Let $(Z(t))_{t\geq 0} = (X(t), Y(t))_{t\geq 0}$ be the mixed state branching process satisfying (1.1)–(1.2) with initial value $z_0 = (x_0, y_0) \in D$. Suppose that $\phi_1(\lambda_1) > 0$ and $\phi_2(\lambda_2) > 0$ for any $\lambda := (\lambda_1, \lambda_2) \in (0, \infty)^2$. Then we have $\mathbb{P}_{z_0}\{\tau_0 < \infty\} = 1$.*

*Proof.* It suffices to prove the case of $z_0 \in D\backslash(0, 0)$. The proof is inspired by [16, Lemma 4.1]. By Itô's formula, we have

$$\mathrm{e}_\lambda(Z(t\wedge\tau_0\wedge\sigma_{\tilde{z}})) = \mathrm{e}_\lambda(z_0) + \int_0^{t\wedge\tau_0\wedge\sigma_{\tilde{z}}} A\mathrm{e}_\lambda(Z(s-))\mathrm{d}s + mart. \tag{3.16}$$

Taking expectations on both sides, we have

$$\mathbb{E}_{z_0}\left[\mathrm{e}_\lambda(Z(t\wedge\tau_0\wedge\sigma_{\tilde{z}}))\right] = \mathrm{e}_\lambda(z_0) + \int_0^t \mathbb{E}_{z_0}\left[A\mathrm{e}_\lambda(Z(s-))1_{\{s<\tau_0\wedge\sigma_{\tilde{z}}\}}\right]\mathrm{d}s,$$

which implies that

$$\mathrm{d}(\mathbb{E}_{z_0}\left[\mathrm{e}_\lambda(Z(t\wedge\tau_0\wedge\sigma_{\tilde{z}}))\right]) = \mathbb{E}_{z_0}\left[A\mathrm{e}_\lambda(Z(t-))1_{\{t<\tau_0\wedge\sigma_{\tilde{z}}\}}\right]\mathrm{d}t.$$

Recall that $\phi_1(\lambda_1) > 0$ and $\phi_2(\lambda_2) > 0$ for all $\lambda \in (0,\infty)^2$. Then for all $\tilde{z} = (\tilde{x},\tilde{y}) \in D$ with $\tilde{z} \succeq z_0$ and $\lambda \in (0,\infty)^2$, there exists a constant $d_{z_0,\tilde{z},\lambda} > 0$ such that for all $z = (x,y) \in D$ with $z_0 \preceq z \preceq \tilde{z}$,

$$x\phi_1(\lambda_1) + h(x,y)y\phi_2(\lambda_2) \geq d_{z_0,\tilde{z},\lambda}. \tag{3.17}$$

Then by integration by parts,

$$\begin{aligned}
&\int_0^\infty \mathrm{e}^{-d_{z_0,\tilde{z},\lambda}t}\mathbb{E}_{z_0}\left[A\mathrm{e}_\lambda(Z(t))1_{\{t<\tau_0\wedge\sigma_{\tilde{z}}\}}\right]\mathrm{d}t \\
&\quad = \int_0^\infty \mathrm{e}^{-d_{z_0,\tilde{z},\lambda}t}\mathrm{d}(\mathbb{E}_{z_0}\left[\mathrm{e}_\lambda(Z(t\wedge\tau_0\wedge\sigma_{\tilde{z}}))\right]) \\
&\quad = d_{z_0,\tilde{z},\lambda}\int_0^\infty \mathrm{e}^{-d_{z_0,\tilde{z},\lambda}t}\mathbb{E}_{z_0}\left[\mathrm{e}_\lambda(Z(t\wedge\tau_0\wedge\sigma_{\tilde{z}}))\right]\mathrm{d}t - \mathrm{e}_\lambda(z_0).
\end{aligned}$$

Moreover, by (2.4) and (3.17) we have

$$\begin{aligned}
&\int_0^\infty \mathrm{e}^{-d_{z_0,\tilde{z},\lambda}t}\mathbb{E}_{z_0}\left[A\mathrm{e}_\lambda(Z(t))1_{\{t<\tau_0\wedge\sigma_{\tilde{z}}\}}\right]\mathrm{d}t \\
&\qquad \geq d_{z_0,\tilde{z},\lambda}\int_0^\infty \mathrm{e}^{-d_{z_0,\tilde{z},\lambda}t}\mathbb{E}_{z_0}\left[\mathrm{e}_\lambda(Z(t))1_{\{t<\tau_0\wedge\sigma_{\tilde{z}}\}}\right]\mathrm{d}t.
\end{aligned}$$

It follows that

$$\begin{aligned}
\mathrm{e}_\lambda(z_0) &\leq d_{z_0,\tilde{z},\lambda}\int_0^\infty e^{-d_{z_0,\tilde{z},\lambda}t}\mathbb{E}_{z_0}\left[\mathrm{e}_\lambda(Z(\tau_0\wedge\sigma_{\tilde{z}}))1_{\{t\geq\tau_0\wedge\sigma_{\tilde{z}}\}}\right]\mathrm{d}t \\
&\leq \mathbb{P}_{z_0}\{\tau_0 \leq \sigma_{\tilde{z}}\} + \sup_{z\succeq\tilde{z}}[\mathrm{e}^{-\lambda_1 x} + \mathrm{e}^{-\lambda_2 y}] \\
&\leq \mathbb{P}_{z_0}\{\tau_0 < \infty\} + [\mathrm{e}^{-\lambda_1\tilde{x}} + \mathrm{e}^{-\lambda_2\tilde{y}}].
\end{aligned}$$

Taking $\tilde{x},\tilde{y} \to \infty$, we get $\mathbb{P}_{z_0}\{\tau_0 < \infty\} \geq \mathrm{e}_\lambda(z_0)$, which holds for any $\lambda \in (0,\infty)^2$. The result follows by letting $\lambda \to (0,0)$. □

**Remark 3.3** *The processes $(X(t))_{t\geq0}$ and $(Y(t))_{t\geq0}$ are independent when $h$ is a positive constant. In this case, one can check that $\mathbb{P}(\tau_0^X < \infty) = 1$ when $\phi_1(\lambda_1) > 0$ for any $\lambda_1 > 0$, and $\mathbb{P}(\tau_0^Y < \infty) = 1$ if $\phi_2(\lambda_2) > 0$ for any $\lambda_2 > 0$.*

**Corollary 3.4** *Assume that $b \geq 0$, $R_1 := \int_{\mathbb{N}^{-1}} \xi n(\mathrm{d}\xi) < 0$ and Condition 3.1 holds. Then we have $\mathbb{P}_{z_0}\{\tau_0 < \infty\} = 1$.*

*Proof.* By Condition 3.1 and $b \geq 0$, one sees that $\phi_1(\lambda_1) > 0$ for any $\lambda_1 > 0$. Moreover, by the inequality $1 - e^{-\lambda_2\xi} \leq \lambda_2\xi$, we have $\phi_2(\lambda_2) \geq -R_1\lambda_2 > 0$. The result follows by Theorem 3.2. □

## 3.2 Exponential ergodicity in the $L^1$-Wasserstein distance

The coupling method is a powerful tool in the study of ergodicity of Markov processes. We refer the reader to [3, 13] for the systematical study on this topic.

We say a couple process $(Z_1(t), Z_2(t))_{t\geq0}$ is called a *coupling* of $(Z(t))_{t\geq0}$ with transition semigroup $(P_t)_{t\geq0}$ if both $(Z_1(t))_{t\geq0}$ and $(Z_2(t))_{t\geq0}$ are Markov processes with transition semigroup $(P_t)_{t\geq0}$ (possibly with different initial distributions) and $Z_1(t+\tau) = Z_2(t+\tau)$ for every $t \geq 0$, where

$$\tau = \inf\{t \geq 0 : Z_1(t) = Z_2(t)\}.$$

In this case, $(Z_1(t))_{t\geq 0}$ and $(Z_2(t))_{t\geq 0}$ are called the *marginal processes* of the coupling. Let $A$ and $\tilde{A}$ be infinitesimal generators of $(Z(t))_{t\geq 0}$ and $(Z_1(t), Z_2(t))_{t\geq 0}$, respectively. Then $\tilde{A}$ is called the *coupling generator* and satisfies the marginal property, i.e., for any $f, g \in \mathcal{D}(A)$,

$$\tilde{A}u(z, \tilde{z}) = Af(z) + Ag(\tilde{z}) \tag{3.18}$$

with $u(z, \tilde{z}) = f(z) + g(\tilde{z})$.

A coupling $(Z_1(t), Z_2(t))_{t\geq 0}$ is called *successful* if $\tau < \infty$ almost surely. A Markov process is said to have a *coupling property* if, for any initial distributions $\mu_1$ and $\mu_2$, there exists a successful coupling with marginal processes starting from $\mu_1$ and $\mu_2$, respectively. For any initial distribution $\mu$, let $\mathbb{P}^{\mu}$ be the distribution of this process with initial distribution $\mu$, and let $\mu P_t$ be the marginal distribution of $\mathbb{P}^{\mu}$ . It is known, see [5, 13], that the coupling property of the process is equivalent to the following statement that

$$\text{For any initial distributions } \ \mu_1, \mu_2, \ \lim_{t\to\infty} \|\mu_1 P_t - \mu_2 P_t\|_{\mathrm{var}} = 0,$$

where $\|\cdot\|_{\mathrm{var}}$ is the total variational norm in the sense of $\|\mu - \nu\|_{\mathrm{var}} := \sup\{|\mu(A) - \nu(A)| : A :$ Borel set$\}$.

The total variational norm is a special case of *Wasserstein distances*. By $\mathcal{P}(D)$ we denote the space of all Borel probability measures over $D$. Given $\mu, \nu \in \mathcal{P}(D)$, a coupling $H$ of $(\mu, \nu)$ is a Borel probability measure on $D \times D$ which has marginals $\mu$ and $\nu$, respectively. We write $\mathcal{H}(\mu, \nu)$ for the collection of all such couplings. Let $d$ be a metric on $D$ such that $(D, d)$ is a complete separable metric space and define

$$\mathcal{P}_d(D) = \Big\{\rho \in \mathcal{P}(D) : \int_D d(z, 0)\, \rho(\mathrm{d}z) < \infty\Big\}.$$

Then the Wasserstein distance on $\mathcal{P}_d(D)$ is defined by

$$W_d(\mu, \nu) = \inf\Big\{\int_{D\times D} d(z, \tilde{z})\, H(\mathrm{d}z, \mathrm{d}\tilde{z}) : H \in \mathcal{H}(\mu, \nu)\Big\}.$$

Moreover, it can be shown that this infimum is attained; see, e.g., [26, Theorem 6.16]. More precisely, there exists $\tilde{H} \in \mathcal{H}(\mu, \nu)$ such that

$$W_d(\mu, \nu) = \int_{D\times D} d(z, \tilde{z})\, H(\mathrm{d}z, \mathrm{d}\tilde{z}).$$

Taking $d(z, \tilde{z}) = 1_{\{z\neq\tilde{z}\}}$, then we have $\mathcal{P}_d(D) = \mathcal{P}(D)$ and $W_d(\mu, \nu) = \|\mu - \nu\|_{\mathrm{var}}$; see, e.g., [3, pp.18]. By taking $d(z, \tilde{z}) = \|z - \tilde{z}\|_1$, $\mathcal{P}_d(D) := \mathcal{P}_1(D)$ defined by

$$\mathcal{P}_1(D) := \Big\{\rho \in \mathcal{P}(D) : \int_D \|z\|_1\, \rho(\mathrm{d}z) < \infty\Big\},$$

where $\|\cdot\|_1$ is the $L^1$ norm on $D$. We say the corresponding Wasserstein distance $W_1$ is the *$L^1$-Wasserstein distance.*

**Definition 3.5** *We say $(Z(t))_{t\geq 0}$ on $D$ or its transition semigroup $(P_t)_{t\geq 0}$ is exponential ergodic in the $L^1$-Wasserstein distance with rate $\lambda_0 > 0$ if its possesses a unique stationary distribution $\mu$ and there is a nonnegative function $\nu \mapsto C(\nu)$ on $\mathcal{P}(D)$ such that*

$$W_1(\nu P_t, \mu) \leq C(\nu)\mathrm{e}^{-\lambda_0 t}, \quad t \geq 0, \nu \in \mathcal{P}(D). \tag{3.19}$$

By standard arguments, (3.19) follows if $(P_t(z,\cdot) := \delta_z P_t)$

$$W_1(P_t(z,\cdot), P_t(\tilde z,\cdot)) \le C_0(z,\tilde z)\mathrm{e}^{-\lambda_0 t}, \quad t \ge 0 \tag{3.20}$$

for $C_0(z,\tilde z) > 0$ depending on $z$ and $\tilde z$; see, e.g., the proof of [9, Theorem 3.5].

In the literature of the exponential ergodicity of branching processes, [19, 20, 9, 4] obtained the results by making full use of the branching property. If the property fails, it has been shown that coupling methods are very effective; see, e.g., [14, 15]. In this subsection, we mainly consider that the rate function $h(\cdot,\cdot)$ satisfies for any $z = (x, y) \in D$ that

$$h(x, y) = r + xm(y) \tag{3.21}$$

with $r > 0$ and $m(\cdot) \in C(\mathbb{R}_+)$ is nonnegative.

This subsection consists two parts. First, we need to construct a coupling for the mixed state branching process with interactions $(Z(t))_{t\ge0} = (X(t), Y(t))_{t\ge0}$. Second, we construct a proper function $F(x, y, \tilde x, \tilde y) = F(|x-\tilde x| + |y-\tilde y|)$ for any $z = (x, y), \tilde z = (\tilde x, \tilde z) \in D$ satisfying two properties:

(i) the exponential contraction property: for any $z = (x, y), \tilde z = (\tilde x, \tilde y) \in D$,

$$\tilde A F(|x-\tilde x| + |y-\tilde y|) \le -\lambda F(|x-\tilde x| + |y-\tilde y|)$$

with some $\lambda > 0$, where $\tilde A$ denotes the coupling generator.

(ii) control the $L^1$-Wasserstein distance in the sense that

$$F(|x-\tilde x| + |y-\tilde y|) \asymp |x-\tilde x| + |y-\tilde y|,$$

here, $f \asymp g$ means that there is a constant $C_1 \ge 1$ such that,

$$C_1^{-1} f(\cdot) \le g(\cdot) \le C_1 f(\cdot).$$

Then we can deduce (3.19) from (i)–(ii).

Recalling that the generator $A$ of $(Z(t))_{t\ge0} = (X(t), Y(t))_{t\ge0}$ is given by (2.3) for any $f \in C_b^{2,1}(\mathbb{R}_+^2)$. Let $\mathcal{D}(A)$ denote the linear space consisting of functions $f \in C_b^{2,1}(\mathbb{R}_+^2)$ such that the two integrals on the right-hand side of (2.3) are convergent and define continuous functions on $D$. To study the coupling and ergodicity of the process, we begin with the construction of a new coupling operator. we will combine the coupling by reflection for the Brownian motion and the synchronous coupling for Poisson random measures. Here the coupling by reflection for Brownian motion means that we will take $(-B(t))_{t\ge0}$ (which is regarded as a reflection of $(B(t))_{t\ge0}$) for the process $(X(t))_{t\ge0}$ before two marginal processes meet. To explain the meaning of the synchronous coupling for Poisson random measures, we will use the viewpoint from the coupling operator. Set $z = (x, y)$, $\tilde z = (\tilde x, \tilde y) \in D$ with $x \ge \tilde x$. Recall that $\gamma(x, y) = yh(x, y)$. The jump system corresponding to the synchronous coupling for the operator $A$ is given by

$$(x, \tilde x) \to \begin{cases} (x+\xi, \tilde x+\xi), & \tilde x m(\mathrm{d}\xi), \\ (x+\xi, \tilde x), & (x-\tilde x)m(\mathrm{d}\xi) \end{cases}$$

and

$$(y, \tilde y) \to \begin{cases} (y+\xi, \tilde y+\xi), & [\gamma(x,y) \wedge \gamma(\tilde x,\tilde y)]n(\mathrm{d}\xi), \\ (y+\xi, \tilde y), & [\gamma(x,y) - \gamma(\tilde x,\tilde y)]^+ n(\mathrm{d}\xi), \\ (y, \tilde y+\xi), & [\gamma(x,y) - \gamma(\tilde x,\tilde y)]^- n(\mathrm{d}\xi), \end{cases}$$

where $f^+$ ($f^-$) denotes the positive (negative) part of function $f$. Similarly, we can define the case that $x < \tilde{x}$. We refer to [3] for the details of such coupling and other couplings for jump systems. We also refer to [15, remark 2.4] for similar discussions in the setting of nonlinear continuous state branching processes. Let $C^2(D \times D)$ be the set of continuous function $(x, y, \tilde{x}, \tilde{y}) \mapsto F(x, y, \tilde{x}, \tilde{y})$ on $D \times D$ with continuous derivatives uo to 2nd order on $x$ and $\tilde{x}$. With the idea above in mind, we then define for any $F \in C^2(D \times D)$ and $x \geq \tilde{x}$ that

$$\begin{aligned}
\tilde{A}F(x, y, \tilde{x}, \tilde{y}) = & -bxF'_x - b\tilde{x}F'_{\tilde{x}} + cxF''_{xx} + c\tilde{x}F''_{\tilde{x}\tilde{x}} - 2c\sqrt{x\tilde{x}}F''_{x\tilde{x}} \\
& + \tilde{x}\int_0^\infty [F(x+\xi, y, \tilde{x}+\xi, \tilde{y}) - F(x, y, \tilde{x}, \tilde{y}) - \xi(F'_x + F'_{\tilde{x}})]\, m(\mathrm{d}\xi) \\
& + (x - \tilde{x})\int_0^\infty [F(x+\xi, y, \tilde{x}, \tilde{y}) - F(x, y, \tilde{x}, \tilde{y}) - \xi F'_x]\, m(\mathrm{d}\xi) \\
& + [\gamma(x, y) \wedge \gamma(\tilde{x}, \tilde{y})]\int_{\mathbb{N}^{-1}} [F(x, y+\xi, \tilde{x}, \tilde{y}+\xi) - F(x, y, \tilde{x}, \tilde{y})]\, n(\mathrm{d}\xi) \\
& + [\gamma(x, y) - \gamma(\tilde{x}, \tilde{y})]^+\int_{\mathbb{N}^{-1}} [F(x, y+\xi, \tilde{x}, \tilde{y}) - F(x, y, \tilde{x}, \tilde{y})]\, n(\mathrm{d}\xi) \\
& + [\gamma(x, y) - \gamma(\tilde{x}, \tilde{y})]^-\int_{\mathbb{N}^{-1}} [F(x, y, \tilde{x}, \tilde{y}+\xi) - F(x, y, \tilde{x}, \tilde{y})]\, n(\mathrm{d}\xi).
\end{aligned} \tag{3.22}$$

Here and in what follows, $F'_x = \frac{\partial F(x,y,\tilde{x},\tilde{y})}{\partial x}, F''_{xx} = \frac{\partial^2 F(x,y,\tilde{x},\tilde{y})}{\partial x^2}$ and so on. Similarly, we can define $\tilde{A}F(x, y, \tilde{x}, \tilde{y})$ for the case that $x < \tilde{x}$. By (3.18), it is not hard to see that $\tilde{A}$ is indeed a coupling generator of $A$ defined by (2.3).

**Theorem 3.6** *There exists a coupling $(Z(t), \tilde{Z}(t))_{t\geq 0} = (X(t), Y(t), \tilde{X}(t), \tilde{Y}(t))_{t\geq 0}$ whose generator is $\tilde{A}$.*

*Proof.* We consider the following SDE:

$$\begin{cases}
X(t) = X(0) - b\displaystyle\int_0^t X(s)\,\mathrm{d}s + \int_0^t \sqrt{2cX(s)}\,\mathrm{d}B(s) + \int_0^t\int_0^{X(s-)}\int_0^\infty \xi\,\tilde{M}(\mathrm{d}s, \mathrm{d}u, \mathrm{d}\xi), \\
Y(t) = Y(0) + \displaystyle\int_0^t\int_0^{Y(s-)}\int_0^{h(X(s-),Y(s-))}\int_{\mathbb{N}^{-1}} \xi\, N(\mathrm{d}s, \mathrm{d}u, \mathrm{d}r, \mathrm{d}\xi), \\
\tilde{X}(t) = \tilde{X}(0) - b\displaystyle\int_0^t \tilde{X}(s)\,\mathrm{d}s + \int_0^t \sqrt{2c\tilde{X}(s)}\,\mathrm{d}B^*(s) + \int_0^t\int_0^{\tilde{X}(s-)}\int_0^\infty \xi\,\tilde{M}(\mathrm{d}s, \mathrm{d}u, \mathrm{d}\xi), \\
\tilde{Y}(t) = \tilde{Y}(0) + \displaystyle\int_0^t\int_0^{\tilde{Y}(s-)}\int_0^{h(\tilde{X}(s-),\tilde{Y}(s-))}\int_{\mathbb{N}^{-1}} \xi\, N(\mathrm{d}s, \mathrm{d}u, \mathrm{d}r, \mathrm{d}\xi),
\end{cases} \tag{3.23}$$

where

$$B^*(t) = \begin{cases} -B(t), & t \leq T, \\ -2B(T) + B(t), & t > T \end{cases}$$

and $T = \inf\{t > 0 : X(t) = \tilde{X}(t)\}$. Clearly, $(B^*(t))_{t\geq 0}$ is still a standard Brownian motion. By the results in Section 2, we can determine the unique strong solution $(Z(t), \tilde{Z}(t))_{t\geq 0} := (X(t), Y(t), \tilde{X}(t), \tilde{Y}(t))_{t\geq 0}$ to (3.23). On the other hand, we can apply the Itô's formula to the SDE (3.23) to see that the infinitesimal generator of the process $(Z(t), \tilde{Z}(t))_{t\geq 0}$ is indeed the coupling generator defined by (3.22). □

Now let us define a function $F$ on $D \times D$ such that

$$F(x, y, \tilde{x}, \tilde{y}) = F(|x - \tilde{x}| + |y - \tilde{y}|) = |x - \tilde{x}| + \theta|y - \tilde{y}| \tag{3.24}$$

with $\theta > 0$, the exact value will be determined later. It is easy to see that

$$F(|x-\tilde{x}|+|y-\tilde{y}|) \asymp |x-\tilde{x}|+|y-\tilde{y}|.$$

For our main result in this subsection, we require the following assumptions.

**Condition 3.7** *$b > 0$ and $R_1 := \int_{\mathbb{N}^{-1}} \xi\, n(\mathrm{d}\xi) \in (-\infty, 0)$.*

**Condition 3.8** *For any $y \in \mathbb{N}_+$, $y \mapsto m(y)y$ is non-decreasing and bounded in the sense of $R_2 := \sup_{y\in\mathbb{N}_+} m(y)y < \infty$.*

**Remark 3.9** *(1). $b > 0$ means that $X$ is a subcritical CB-process; see, e.g., [18, 20]. There is not any restrictions on Lévy noises in this paper. It is reasonable since we mainly focus on the exponential ergodicity in the $L^1$-Wasserstein distance. When considering the exponential ergodicity in the total variation distance for CB-processes, some additional assumptions on Lévy noises are needed; see [19, 9]. In the setting of state-dependent branching cases, one has to make some restrictions on Lévy noises even in the $L^1$-Wasserstein distance; see [15]. Furthermore, we mention that either in other distances $W_d$ or state-dependent cases, more complicated functions are needed.*

*(2). $R_1 < 0$ of Condition 3.7 actually means that the associated first moment of offerspring of each individual strictly less than 1, i.e. $\sum_j jp_j < 1$. In this case, the process $Y$ is called the subcritical case of continuous-time Markov branching processes when $h$ is a positive constant; see, e.g., [1, pp.112].*

*(3). Condition 3.8 holds when $m(\cdot) \equiv 0$. In this case, $(Y(t))_{t\geq 0}$ is a standard continuous time branching process with branching rate $r > 0$ and offspring $(p_\xi, \xi \in \mathbb{N})$. Condition 3.8 also holds when $m(y) = \frac{1}{y+1}$. In this case, as the number of cells increasing, the rate of cell division is getting slower.*

We now present the main result.

**Theorem 3.10** *Suppose that Conditions 3.7–3.8 are satisfied. Then there are constants $\lambda_0 > 0$ such that for any $(x, y), (\tilde{x}, \tilde{y}) \in D$, (3.20) holds.*

*Proof.* We shall first give some estimates of $\tilde{A}F(|x-\tilde{x}|+|y-\tilde{y}|)$. By (3.22) and Taylor's formula, we have

$$\begin{aligned}
&\tilde{A}F(|x-\tilde{x}|+|y-\tilde{y}|)\\
&\quad\leq -bx\frac{x-\tilde{x}}{|x-\tilde{x}|} + b\tilde{x}\frac{x-\tilde{x}}{|x-\tilde{x}|}\\
&\qquad+\theta[\gamma(x,y)-\gamma(\tilde{x},\tilde{y})]^+\int_{\mathbb{N}^{-1}}\Big[|y-\tilde{y}+\xi|-|y-\tilde{y}|\Big]\,n(\mathrm{d}\xi)\\
&\qquad+\theta[\gamma(x,y)-\gamma(\tilde{x},\tilde{y})]^-\int_{\mathbb{N}^{-1}}\Big[|y-\tilde{y}-\xi|-|y-\tilde{y}|\Big]\,n(\mathrm{d}\xi)\\
&\quad\leq -b|x-\tilde{x}| + \theta[\gamma(x,y)-\gamma(\tilde{x},\tilde{y})]^+\mathbf{1}_{\{y>\tilde{y}\}}\int_{\mathbb{N}^{-1}}\xi\,n(\mathrm{d}\xi)\\
&\qquad+\theta[\gamma(x,y)-\gamma(\tilde{x},\tilde{y})]^+\mathbf{1}_{\{y\leq\tilde{y}\}}\int_{\mathbb{N}^{-1}}|\xi|\,n(\mathrm{d}\xi)+\theta[\gamma(x,y)-\gamma(\tilde{x},\tilde{y})]^-\mathbf{1}_{\{y>\tilde{y}\}}\int_{\mathbb{N}^{-1}}|\xi|\,n(\mathrm{d}\xi)\\
&\qquad+\theta[\gamma(x,y)-\gamma(\tilde{x},\tilde{y})]^-\mathbf{1}_{\{y\leq\tilde{y}\}}\int_{\mathbb{N}^{-1}}\xi\,n(\mathrm{d}\xi).
\end{aligned}$$

Since $\gamma(x,y) = ry + xm(y)y$, we arrive at

$$\begin{aligned}\tilde{A}F(|x-\tilde{x}|+|y-\tilde{y}|) &\le -b|x-\tilde{x}| - \theta n(\{-1\})(\gamma(x,y)-\gamma(\tilde{x},\tilde{y}))\frac{y-\tilde{y}}{|y-\tilde{y}|} \\ &\quad + \theta \int_{\mathbb{N}} \xi\, n(\mathrm{d}\xi)|\gamma(x,y)-\gamma(\tilde{x},\tilde{y})| \\ &= -b|x-\tilde{x}| - \theta n(\{-1\})(ry - r\tilde{y} + xm(y)y - \tilde{x}m(\tilde{y})y)\frac{y-\tilde{y}}{|y-\tilde{y}|} \\ &\quad + \theta \int_{\mathbb{N}} \xi\, n(\mathrm{d}\xi)|ry - r\tilde{y} + xm(y)y - \tilde{x}m(\tilde{y})y|.\end{aligned}$$

Notice that $\int_{\mathbb{N}^{-1}} |\xi|\, n(\mathrm{d}\xi) \in (0,\infty)$ by $R_1 \in (-\infty, 0)$ of Condition 3.7. For the case of $x > \tilde{x}$, it follows from Condition 3.8 that

$$\begin{aligned}\tilde{A}F(|x-\tilde{x}|+|y-\tilde{y}|) \;&\le\; -b|x-\tilde{x}| - \theta n(\{-1\})\Big(r|y-\tilde{y}| + (x-\tilde{x})m(y)y\frac{y-\tilde{y}}{|y-\tilde{y}|} \\ &\qquad + \tilde{x}(m(y)y - m(\tilde{y})\tilde{y})\frac{y-\tilde{y}}{|y-\tilde{y}|}\Big) \\ &\quad + \theta \int_{\mathbb{N}} \xi\, n(\mathrm{d}\xi)\Big(r|y-\tilde{y}| + (x-\tilde{x})m(y)y + \tilde{x}|m(y)y - m(\tilde{y})\tilde{y}|\Big) \\ &\le -\Big(b - \theta \int_{\mathbb{N}^{-1}} |\xi|\, n(\mathrm{d}\xi) m(y)y\Big)|x-\tilde{x}| + \theta r R_1 |y-\tilde{y}| + \tilde{x}\theta R_1 |m(y)y - m(\tilde{y})\tilde{y}| \\ &\le -\Big(b - \theta R_2 \int_{\mathbb{N}^{-1}} |\xi|\, n(\mathrm{d}\xi)\Big)|x-\tilde{x}| + \theta r R_1 |y-\tilde{y}| \\ &\le -\lambda_1 F(|x-\tilde{x}|+|y-\tilde{y}|)\end{aligned}$$

for some $\lambda_1 > 0$ by setting $\theta := \theta_1 = \frac{b}{2R_2 \int_{\mathbb{N}^{-1}} |\xi|\, n(\mathrm{d}\xi)}$. When $x \le \tilde{x}$, similarly we have

$$\begin{aligned}\tilde{A}F(|x-\tilde{x}|+|y-\tilde{y}|) &\le -(b + \theta R_1 R_2)|x-\tilde{x}| + \theta R_1 \Big(r|y-\tilde{y}| + x|m(\tilde{y})\tilde{y} - m(y)y|\Big) \\ &\le -\lambda_2 F(|x-\tilde{x}|+|y-\tilde{y}|).\end{aligned}$$

for some $\lambda_2 > 0$ by setting $\theta := \theta_2 = \frac{-b}{2R_1R_2} > 0$. In conclusion, let $\theta = \theta_1 \wedge \theta_2$ and $\lambda = \lambda_1 \wedge \lambda_2$. Following similar arguments in step 2 of the proof for [22, Theorem 3.1], we obtain the desired result. □

## References


[1] Athreya, K.B. and Ney, P.E. (1972): *Branching Processes.* New York, Heidelberg, Berlin.

[2] Bansaye, V. and Tran V.C. (2011): Branching Feller diffusion for cell division with parasite infection. *ALEA, Lat. Am. J. Probab. Math. Stat.* **8**: 95–127.

[3] Chen, M. (2005): *Eigenvalues, Inequalities, and Ergodic Theory*, Springer, New York.

[4] Chen, S. and Li, Z. (2021): Continuous time mixed state branching processes and stochastic equations. *Acta Math. Sci. (Engl. Ed.)* **41(5)**: 1445–1473.

[5] Cranston, M. and Greven, A. (1995): Coupling and harmonic functions in the case of continuous time Markov processes. *Stoch. Process. Appl.* **60**: 261–286.

[6] Dawson, D.A. and Li, Z. (2006): Skew convolution semigroups and affine Markov processes. *Ann. Probab.* **34**: 1103–1142.

[7] Dawson, D.A. and Li, Z. (2012): Stochastic equations, flows and measure-valued processes. *Ann. Probab.* **40**: 813–857.

[8] Ethier, S.N. and Kurtz, T.G. (1986): *Markov Processes: Characterization and Convergence*. John Wiley and Sons, New York.

[9] Friesen, M., Jin, P., Kremer, J. and Rüdiger, B. (2023): Exponential ergodicity for stochastic equations of nonnegative processes with jumps. *ALEA, Lat. Am. J. Probab. Math. Stat.* **20**: 593–627.

[10] Fu, Z. and Li, Z. (2010): Stochastic equations of non-negative processes with jumps. *Stoch. Process. Appl.* **120**: 306–330.

[11] Ji, L. and Xiong, J. (2022): Superprocesses for the population of rabbits on grassland. *Proc. Steklov Inst. Math.* **316**(1): 195–208.

[12] Kawazu, K. and Watanabe, S. (1971): Branching processes with immigration and related limit theorems. *Theory Probab. Appl.* **16**: 36–54.

[13] Lindvall, T. (1992): *Lectures on the coupling method.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York.

[14] Li, P., Li, Z., Wang, J. and Zhou, X. (2023+): Exponential ergodicity of branching processes with immigration and competition. To appear in *Annales de l'Institute Henri Poincaré-Probabilités et Statistiques*, available at arxiv: 2205.15499v1.

[15] Li, P. and Wang, J. (2020): Exponential ergodicity for general continuous-state nonlinear branching processes. *Electron. J. Probab.* **25**: article no. 125, 1–25.

[16] Li, P., Yang, X. and Zhou, X. (2019): A general continuous-state nonlinear branching process. *Ann. Appl. Probab.* **29**(4): 2523–2555.

[17] Li, Z. (2006): A limit theorem for discrete Galton-Watson branching processes with immigration. *J. Appl. Probab.* **43**: 289–295.

[18] Li, Z. (2011): *Measure-Valued Branching Markov Processes.* Probab. and Its Appl. Springer.

[19] Li, Z. and Ma, C. (2015): Asymptotic properties of estimators in a Cox-Ingersoll-Ross model. *Stoch. Process. Appl.* **125**: 3196–3233.

[20] Li, Z. (2020): *Continuous-State Branching Processes with Immigration.* In: Jiao, Y. (eds) From Probability to Finance. Mathematical Lectures from Peking University. Springer, Singapore.

[21] Liu, J. (2022+): Scaling Limits of Controlled Branching Processes. Available at arXiv: 2204.06796.

[22] Luo, D. and Wang, J. (2019): Refined basic couplings and Wasserstein-type distances for SDEs with Lévy noises. *Stoc. Process. Appl.* **129**: 3129–3173.

[23] Ma, C. (2009): A limit theorem of two-type Galton-Watson branching processes with immigration. *Stat. Prob. Lett.* **79**: 1710–1716.

[24] Marguet, A. and Smadi, C. (2020+): Parasite infection in a cell population with deaths. Available at arXiv: 2010. 16070v1.

[25] Osorio, L. and Winter, A. (2020+): Two level branching model for virus population under cell division. Available at arXiv: 2004.14352.

[26] Villani, C. (2009): *Optimal Transport, Old and New.* Springer, Berlin.