Distribution Dependent Birth-Death Processes: $\mathbb{W}_p$-Estimate, Ergodicity and Propagation of Chaos*

Feng-Yu Wang, Yi Zhao

Center for Applied Mathematics and KL-AAGDM, Tianjin University, Tianjin 300072, China

wangfy@tju.edu.cn, zhaoyibnu@mail.bnu.edu.cn

Abstract

For a class of time inhomogenous distribution dependent birth-death processes, we derive the well-posedness, $\mathbb{W}_p$-estimate, exponential ergodicity, and uniform in time propagation of chaos. These extend the corresponding results derived for distribution dependent SDEs and mean field particle systems. As preparation, a criterion on the well-posedness of inhomogenous jump process is presented in the end of the paper, which should be interesting by itself.

AMS Subject Classification: 60J27; 60J74 .
Keywords: Distribution dependent birth-death processes; $\mathbb{W}_p$-estimate; exponential ergodicity; uniform in time propagation of chaos.

1 Introduction

Due to its solid background and wide applications, the distribution dependent SDE (DDSDE for short) initiated by McKean [13] have been intensively investigated, and plentiful results have been derived for the well-posedness, regularities, ergodicity, propagation of chaos, see the monographs [18, 2, 23] and references therein. Since the solution of DDSDE is Markov in time but nonlinear in initial distribution, it is called nonlinear Markov process, see [11] for a literature of nonlinear Markov processes and applications.

To describe stochastic systems with discontinuous noise, the theory of Markov jump processes has been developed. This type processes are known as Q-processes or continuous

*Supported in part by the National Key R&D Program of China (2022YFA1006000), NSFC(12531007) and the State Key Laboratory of Synthetic Biology in Tianjin University.

time Markov chains when the state space is discrete, and are called q-processes in general, see [3] and references therein. In particular, for three specific Markov chains, namely birth-death processes, single birth processes and single death processes, sharp criteria have been derived for different type ergodicity properties, see [3, 4, 24].

Comparing with DDSDEs, the study on distribution dependent Markov chains is rather limited. Existing results focus on expectation dependent Markov chains, i.e. the transition rate depends on the state and expectation of the process, first introduced by Nicolis and Prigogine [15] to describe a class of nonlinear pure jump processes arising from chemistry, physics and biology, and further studied in [17, 7, 5, 6, 14] for the well-posedness, stationary distributions and large deviations. However, to our best knowledge, rare is known on general distribution dependent Markov chains, and the propagation of chaos is open even in the expectation dependent case.

In this paper, we study the well-posedness, exponential ergodicity and propagation of chaos for distribution dependent birth-death processes, and the argument we develop might be extended to general distribution dependent jump processes in the future.

Let $\mathbb{Z}_+$ be the set of nonnegative integer numbers, let $\mathscr{B}(\mathbb{Z}_+)$ ($\mathscr{B}_b(\mathbb{Z}_+)$) be the space of all real (bounded) functions on $\mathbb{Z}_+$, and let $\mathscr{P}$ be the space of probability measures on $\mathbb{Z}_+$ equipped with the weak topology, which is equivalent to the strong topology as all measurable functions on $\mathbb{Z}_+$ are continuous. For a random variable ξ on $\mathbb{Z}_+$, let $\mathscr{L}_\xi \in \mathscr{P}$ denote its distribution.

For any $p \in [1, \infty)$, let

$$\mathscr{P}_p := \Big\{\mu \in \mathscr{P} : \ \|\mu\|_p := \mu(|\cdot|^p)^{\frac{1}{p}} < \infty\Big\}.$$

It is a Polish space under the p-Wasserstein distance

$$\mathbb{W}_p(\mu,\nu) := \inf_{\pi\in\mathscr{C}(\mu,\nu)} \left(\int_{\mathbb{Z}_+\times\mathbb{Z}_+} |x-y|^p \pi(\mathrm{d}x,\mathrm{d}y)\right)^{\frac{1}{p}},$$

where $\mathscr{C}(\mu,\nu)$ is the set of all couplings of μ and ν.

Consider the following measurable maps

$$a, b : [0,\infty)\times\mathbb{Z}_+\times\mathscr{P}_1 \to [0,\infty), \quad a_t(0,\mu) = 0.$$

Let

$$\mathbf{L}^{a,b} := \big\{L^{a,b}_{t,\mu} : t \ge 0, \mu \in \mathscr{P}_1\big\}$$

be a family of linear operators on $\mathscr{B}(\mathbb{Z}_+)$ defined by

$$L^{a,b}_{t,\mu} f(i) := a_t(i,\mu)\big[f(i-1) - f(i)\big] + b_t(i,\mu)\big[f(i+1) - f(i)\big], \quad i \in \mathbb{Z}_+, \ f \in \mathscr{B}(\mathbb{Z}_+),$$

where $b_t(i,\mu)$ and $a_t(i,\mu)$ stand for the birth and death rates of the system at time t with state i and distribution μ. The distribution dependent birth-death process generated by $\mathbf{L}^{a,b}$ is defined as follows.

Definition 1.1. Let $\gamma \in \mathscr{P}$. We call $\mathbb{P}^\gamma$ a distribution dependent birth-death process generated by $\mathbf{L}^{a,b}$, denoted by $\mathbf{L}^{a,b}$-process for simplicity, if it is a probability measure on the space D consisting of càdlàg paths from $[0,\infty)$ to $\mathbb{Z}_+$, such that the time-marginal distributions

$$P_t^*\gamma(\cdot) := \mathbb{P}^\gamma(\{\omega \in D: \ \omega_t \in \cdot\}), \quad t \geq 0$$

satisfy the following conditions:

(1) $P_0^*\gamma = \gamma$;

(2) The time-dependent operator $L_{t,P_t^*\gamma}$ generates a unique inhomogenous jump process $\{P_{s,t}^\gamma(i,\cdot)\}_{0\leq s\leq t, i\in\mathbb{Z}_+}$ on $E = \mathbb{Z}_+$ in the sense of Theorem 5.1, such that

$$P_t^*\gamma = \sum_{i=0}^\infty \gamma(i) P_{0,t}^\gamma(i,\cdot), \quad t \geq 0.$$

Let $\mathbb{P}^\gamma$ be a distribution dependent birth-death process generated by $\mathbf{L}^{a,b}$, and let X_t be the coordinate process on the probability space $(\Omega, \mathscr{F}, \mathbb{P}) := (D, \mathscr{B}(D), \mathbb{P}^\gamma)$, where $\mathscr{B}(D)$ is the Borel σ-filed on D. Then it is standard to conclude that for any $t \geq 0$ and $i, j \in \mathbb{Z}_+$ with $\mathbb{P}(X_t = i) > 0$,

$$\lim_{\varepsilon\downarrow 0} \frac{1}{\varepsilon}\mathbb{P}(X_{t+\varepsilon} = j | X_t = i) = \begin{cases} b_t(i, \mathscr{L}_{X_t}), & \text{if } j = i+1, \\ a_t(i, \mathscr{L}_{X_t}), & \text{if } j = i-1, \\ -a_t(i, \mathscr{L}_{X_t}) - b_t(i, \mathscr{L}_{X_t}), & \text{if } j = i, \\ 0, & \text{otherwise,} \end{cases}$$

where $\mathscr{L}_{X_t} = P_t^*\gamma$ is the distribution of X_t.

In applications, we study the $\mathbf{L}^{a,b}$-process for distributions in a specific sub-space $\hat{\mathscr{P}}$ of $\mathscr{P}$, for instance $\hat{\mathscr{P}} = \mathscr{P}_p$ for some $p \geq 1$, which leads to the following definition.

Definition 1.2. We call the $\mathbf{L}^{a,b}$-problem well-posed for distributions in $\hat{\mathscr{P}}$, if for any $\gamma \in \hat{\mathscr{P}}$, there is a unique distribution dependent birth-death process generated by $\mathbf{L}^{a,b}$ such that $P_t^*\gamma \in \hat{\mathscr{P}}$ for $t \geq 0$.

The remainder of the paper is organized as follows. In Section 2, we study the well-posedness and exponential ergodicity of the $\mathbf{L}^{a,b}$-problem for distributions in $\mathscr{P}_1$. In Section 3, we estimate the $\mathbb{W}_p$-Lipschitz constant of P_t^* for $p \in [1,\infty)$, which are then applied in Section 4 to characterize the propagation of chaos for the associated mean field birth-death systems. Finally, a result on the well-posedness of inhomogenous jump processes is presented in Section 5 as appendix, which is used in Section 2 to construct the $\mathbf{L}^{a,b}$-process and its coupling processes.

2 Well-posedness and exponential ergodicity in $\mathbb{W}_1$

Let δ_i be the Dirac measure at $i \in \mathbb{Z}_+$, and let $C^w([0,\infty);\mathscr{P}_1)$ be the set of all weakly continuous maps $\mu : [0,\infty) \mapsto \mathscr{P}_1$ with

$$\sup_{t\in[0,T]} \|\mu_t\|_1 < \infty, \quad T \in (0,\infty).$$

Let

$$\mathrm{sgn}(r) := \begin{cases} 1, & \text{if } r > 0, \\ -1, & \text{if } r < 0, \\ 0, & \text{if } r = 0. \end{cases}$$

To prove the existence and uniqueness of the $\mathbf{L}^{a,b}$-process, we make the following assumption.

(H_1) For any $i \in \mathbb{Z}_+$ and $\mu \in C^w([0,\infty);\mathscr{P}_1)$, $a_t(i,\mu_t)$ and $b_t(i,\mu_t)$ are continuous in $t \in [0,\infty)$ locally uniform in $i \in \mathbb{Z}_+$. Moreover, there exist $K_1 \in C([0,\infty);\mathbb{R})$ and $K_2 \in C([0,\infty);[0,\infty))$ such that

$$\begin{aligned} &\big[b_t(i,\mu) - b_t(j,\nu) + a_t(j,\nu) - a_t(i,\mu)\big]\mathrm{sgn}(i-j) \\ &\quad + \big(|a_t(i,\mu) - a_t(i,\nu)| + |b_t(i,\mu) - b_t(i,\nu)|\big)1_{\{i=j\}} \\ &\le K_1(t)|i-j| + K_2(t)\mathbb{W}_1(\mu,\nu), \quad t \ge 0,\ i,j \in \mathbb{Z}_+,\ \mu,\nu \in \mathscr{P}_1, \end{aligned} \tag{2.1}$$

(H_2) There exist $\theta \in (1,\infty)$, $K_3 \in C([0,\infty);[0,\infty))$, $c_0 > 1$, and an increasing function $V : \mathbb{Z}_+ \to [1,\infty)$ with set $\{x : V(x) \le c\}$ being compact for any $c > 0$ such that

$$V(i+1) \le c_0 V(i), \quad i \in \mathbb{Z}_+, \tag{2.2}$$

$$a_t(i,\delta_0) + b_t(i,\delta_0) \le K_3(t)V(i), \quad t \ge 0,\ i \in \mathbb{Z}_+, \tag{2.3}$$

$$\begin{aligned} &\big(b_t(i,\delta_0) - a_t(i,\delta_0)\big)\big(V^\theta(i+1) - V^\theta(i)\big) \\ &\quad + a_t(i,\delta_0)\big(V^\theta(i+1) - 2V^\theta(i) + V^\theta(i-1)\big) \\ &\le K_3(t)V^\theta(i), \quad t \ge 0,\ i \in \mathbb{Z}_+. \end{aligned} \tag{2.4}$$

The assumption (H_1) is a standard monotone condition, while (H_2) contains growth and Lyapunov condition on the birth-death rate at δ_0. Before moving on, let us present two examples to illustrate the assumption (H_2).

Example 2.1. *Assume* (H_1).

(a) *If there exists* $K \in C([0,\infty);(0,\infty))$ *such that*

$$a_t(i,\delta_0) \le K(t) i^2, \quad t \ge 0,\ i \in \mathbb{Z}_+, \tag{2.5}$$

then (H_2) *with* $\theta = 2$ *holds for some* $K_3 \in C([0,\infty);(0,\infty))$ *and* $V(i) := (1+i)^2$.

(b) *If there exist constants* $q \ge 1, \varepsilon \in (0,1)$ *and* $C_1, C_2, K \in C([0,\infty);(0,\infty))$ *such that*

$$\begin{aligned} &b_t(i,\delta_0) - a_t(i,\delta_0) \le C_1(t) - C_2(t) i^q, \\ &a_t(i,\delta_0) \le K(t) i^{q+\varepsilon}, \quad t \ge 0,\ i \in \mathbb{Z}_+, \end{aligned} \tag{2.6}$$

with $K(t) \le C_2(t)$ *for any* $t \ge 0$, *then* (H_2) *with* $\theta = 2$ *holds for some* $K_3 \in C([0,\infty);(0,\infty))$ *and* $V(i) := (1+i)^{q+\varepsilon}$.

Proof. Obviously, for any $p \in (0,\infty)$, $V(i) := (1+i)^p$ satisfies (2.2) for some constant $c_0 > 1$. By (2.1) with $\mu = \nu = \delta_0$ and $j = 0$, we obtain

$$b_t(i,\delta_0) - a_t(i,\delta_0) \le b_t(0,\delta_0) + K_1(t) i \le C_1(t)(1+i), \quad i \in \mathbb{Z}_+,\ t \in [0,T] \tag{2.7}$$

for $C_1(t) := b_t(0,\delta_0) + K_1(t)$.

(a) If (2.5) holds, then (2.7) implies the second condition in (H_2) for $K_3(t) = 2K(t) + C_1(t)$ and $V(i) := (1+i)^2$. Moreover, for any $p \in [1,\infty)$, we find a constant $c_1 > 1$ such that $V(i) := (1+i)^p$ satisfies

$$\begin{aligned} &V^2(i+1) - V^2(i) \\ &\quad = (2+i)^{2p} - (1+i)^{2p} \le c_1 (1+i)^{2p-1} \\ &V^2(i+1) - 2V^2(i) + V^2(i-1) \\ &\quad = (2+i)^{2p} - 2(1+i)^{2p} + i^{2p} \le c_1(1+i)^{2p-2}, \quad i \in \mathbb{Z}_+. \end{aligned} \tag{2.8}$$

Letting $p = 2$ and combining this with (2.5) and (2.7), we derive

$$\begin{aligned} &\big(b_t(i,\delta_0) - a_t(i,\delta_0)\big)\big(V^2(i+1) - V^2(i)\big) + a_t(i,\delta_0)\big(V^2(i+1) - 2V^2(i) + V^2(i-1)\big) \\ &\le C_1(t) c_1 (1+i)^{2p} + K(t) i^2 c_1 (1+i)^{2p-2} \le \big(C_1(t) + K(t)\big) c_1 V^2(i), \quad t \ge 0,\ i \in \mathbb{Z}_+. \end{aligned}$$

So, the third condition in (H_2) holds for some K_3 and $V(i) = (1+i)^2$.

(b) Let $V(i) = (1+i)^{q+\varepsilon}$. By the second inequality in (2.6) and (2.7), we obtain the second condition in (H_2) for some K_3. Next, by (2.7) and (2.8) with $p = q+\varepsilon$ for the present V, we obtain

$$\begin{aligned} &\big(b_t(i,\delta_0) - a_t(i,\delta_0)\big)\big(V^2(i+1) - V^2(i)\big) + a_t(i,\delta_0)\big(V^2(i+1) - 2V^2(i) + V^2(i-1)\big) \\ &\le C_1(t) c_1 (1+i)^{2(q+\varepsilon)-1} - C_2(t) c_1 (1+i)^{2(q+\varepsilon)-1} i^q + K(t) i^{q+\varepsilon} c_1 (1+i)^{2(q+\varepsilon)-2} \\ &\le C_1(t) c_1 (1+i)^{2(q+\varepsilon)}, \quad t \ge 0,\ i \in \mathbb{Z}_+. \end{aligned}$$

Here we use the fact that $K(t) \le C_2(t)$ and $\varepsilon \in (0,1)$. This implies the third condition in (H_2) for $\theta = 2$ and $V(i) = (1+i)^{q+\varepsilon}$.

□

Theorem 2.2. *Assume (H_1) and (H_2). Then the following assertions hold.*

(1) *The $\mathbf{L}^{a,b}$-problem is well-posed for distributions in $\mathscr{P}_1$, and*

$$\|P_t^*\mu\|_1 \le \mathrm{e}^{\int_0^t (K_1+K_2)(s)\mathrm{d}s}\|\mu\|_1 + \int_0^t b_s(0,\delta_0)\mathrm{e}^{\int_s^t (K_1+K_2)(r)\mathrm{d}r}\mathrm{d}s, \quad t \ge 0, \tag{2.9}$$

$$\mathbb{W}_1(P_t^*\mu, P_t^*\nu) \le \mathrm{e}^{\int_0^t (K_1(s)+K_2(s))\mathrm{d}s}\mathbb{W}_1(\mu,\nu), \quad t \ge 0, \ \mu,\nu \in \mathscr{P}_1. \tag{2.10}$$

(2) *If $(a_t(i,\mu), b_t(i,\mu)) = (a(i,\mu), b(i,\mu))$ does not depend on t such that (H_1) holds for some constants $K_1 \in \mathbb{R}$ and $K_2 \ge 0$ with $K_1 + K_2 < 0$, then P_t^* has a unique invariant probability measure $\bar\mu \in \mathscr{P}_1$, i.e. $P_t^*\bar\mu = \bar\mu$ for all $t \ge 0$, and*

$$\mathbb{W}_1(P_t^*\mu, \bar\mu) \le \mathrm{e}^{(K_1+K_2)t}\mathbb{W}_1(\mu,\bar\mu), \quad t \ge 0, \ \mu \in \mathscr{P}_1. \tag{2.11}$$

Proof. (1) It suffices to consider the $\mathbf{L}^{a,b}$-problem up to an arbitrarily finite time. So, in the following, we fix $T \in (0,\infty)$ and an initial distribution $\mu \in \mathscr{P}_1$.

(a) For $\mu \in \mathscr{P}_1$ and $T \in (0,\infty)$, let

$$\mathscr{C}_T^\mu := \left\{\gamma \in C^w([0,T];\mathscr{P}) : \ \gamma_0 = \mu, \ H_T(\gamma) := \sup_{t\in[0,T]} \|\gamma_t\|_1 < \infty\right\}.$$

For fixed $\gamma \in \mathscr{C}_T^\mu$, (2.1) with $\mu = \nu = \gamma_t$ and $j = 0$ implies

$$\begin{aligned} b_t(i,\gamma_t) - a_t(i,\gamma_t) &\le b_t(0,\gamma_t) + K_1(t)i \\ &\le b_t(0,\delta_0) + K_2(t)\|\gamma_t\|_1 + K_1(t)i \le c_1 + c_1 i, \quad i \in \mathbb{Z}_+, \ t \in [0,T] \end{aligned} \tag{2.12}$$

for some constant $c_1 > 0$, since $b_t(0,\gamma_t)$ and $K_1(t)$ are bounded in $t \in [0,T]$. Let

$$L_t^\gamma f(i) := a_t(i,\gamma_t)(f(i-1) - f(i)) + b_t(i,\gamma_t)(f(i+1) - f(i)), \quad f \in \mathscr{B}_b(\mathbb{Z}_+). \tag{2.13}$$

Next, taking $i = j$ in (2.1) we derive

$$|a_t(i,\mu_1) - a_t(i,\mu_2)| + |b_t(i,\mu_1) - b_t(i,\mu_2)| \le K_2(t)\mathbb{W}_1(\mu_1,\mu_2), \quad \mu_1,\mu_2 \in \mathscr{P}_1. \tag{2.14}$$

This together with (2.3) yields

$$q_t(i) := a_t(i,\gamma_t) + b_t(i,\gamma_t) \le c_2(\gamma)V(i), \quad i \in \mathbb{Z}_+, \ t \in [0,T] \tag{2.15}$$

for some constant $c_2(\gamma)$ depending on γ.

Next, by (2.2), (2.4) and (2.14), we find a constant $c_3(\gamma) > 0$ such that

$$
\begin{aligned}
L_t^\gamma V^\theta(i) &= L_t^{\delta_0} V^\theta(i) + (a_t(i,\gamma_t) - a_t(i,\delta_0))(V^\theta(i-1) - V^\theta(i)) \\
&\quad + (b_t(i,\gamma_t) - b_t(i,\delta_0))(V^\theta(i+1) - V^\theta(i)) \\
&\le K_3(t) V^\theta(i) + (c_0^\theta - 1) K_2(t) \|\gamma_t\|_1 V^\theta(i) \le c_3(\gamma) V^\theta(i), \quad t \in [0,T],\ i \in \mathbb{Z}_+.
\end{aligned}
\tag{2.16}
$$

So, by Theorem 5.1, there is a unique inhomogenous birth-death process

$$\{P_{s,t}^\gamma(i,\cdot)\}_{i\in\mathbb{Z}_+, 0\le s\le t\le T}$$

generated by $\{L_t^\gamma\}_{t\in[0,T]}$. Let

$$(\Phi\gamma)_t = \Phi_t\gamma := \sum_{i=0}^\infty \mu(i) P_{s,t}^\gamma(i,\cdot), \quad t \in [0,T].$$

If we can verify that $\Phi : \mathscr{C}_T^\mu \to \mathscr{C}_T^\mu$ has a unique fixed point $\bar\gamma$, then the inhomogenous birth-death process generated by $\{L_t^{\bar\gamma}\}_{t\in[0,T]}$ is the unique $\mathbf{L}^{a,b}$-process up to time T with initial distribution μ.

(b) We first show $\Phi\gamma \in \mathscr{C}_T^\mu$ for $\gamma \in \mathscr{C}_T^\mu$. The weak continuity of $t \mapsto (\Phi\gamma)_t$ is standard for the birth-death process. By (2.12) we have

$$\sum_{j=0}^\infty (1+j) P_{s,t}^\gamma(i,\{j\}) \le (1+i) \mathrm{e}^{c_1(t-s)}, \quad 0 \le s \le t \le T,\ i \in \mathbb{Z}_+,$$

so that

$$\|\Phi_t\gamma\|_1 = \sum_{i,j=0}^\infty \mu(i) j P_{s,t}^\gamma(i,\{j\}) \le (1 + \|\mu\|_1) \mathrm{e}^{c_1 T} < \infty, \quad t \in [0,T].$$

Then $\Phi\gamma \in \mathscr{C}_T^\mu$.

To see that Φ has a unique fixed point, we need only to show that Φ is contractive under the complete metric

$$\rho_\lambda(\gamma,\tilde\gamma) := \sup_{t\in[0,T]} \mathrm{e}^{-\lambda t} \mathbb{W}_1(\gamma_t,\tilde\gamma_t), \quad \gamma,\tilde\gamma \in \mathscr{C}_T^\mu$$

for some constant $\lambda > 0$. To this end, we consider the following synchronized coupling operator for L_t^γ and $L_t^{\tilde\gamma}$: for $t \ge 0, i,j \in \mathbb{Z}_+$ and $f \in \mathscr{B}(\mathbb{Z}_+^2)$,

$$
\begin{aligned}
L_t^{\gamma,\tilde\gamma} f(i,j) &:= \big[b_t(i,\gamma_t) \wedge b_t(j,\tilde\gamma_t)\big]\big(f(i+1,j+1) - f(i,j)\big) \\
&\quad + \big[a_t(i,\gamma_t) \wedge a_t(j,\tilde\gamma_t)\big]\big(f(i-1,j-1) - f(i,j)\big) \\
&\quad + \big[b_t(i,\gamma_t) - b_t(j,\tilde\gamma_t)\big]^+\big(f(i+1,j) - f(i,j)\big) \\
&\quad + \big[b_t(i,\gamma_t) - b_t(j,\tilde\gamma_t)\big]^-\big(f(i,j+1) - f(i,j)\big) \\
&\quad + \big[a_t(i,\gamma_t) - a_t(j,\tilde\gamma_t)\big]^+\big(f(i-1,j) - f(i,j)\big) \\
&\quad + \big[a_t(i,\gamma_t) - a_t(j,\tilde\gamma_t)\big]^-\big(f(i,j-1) - f(i,j)\big).
\end{aligned}
\tag{2.17}
$$

By taking

$$\bar{V}(i,j) := \big(V(i)^\theta + V(j)^\theta\big)^{\frac{1}{\theta}},$$

the marginal property of the coupling operator and (2.16) imply

$$L_t^{\gamma,\tilde{\gamma}} \bar{V}^\theta(i,j) \le (c_3(\gamma) + c_3(\tilde{\gamma}))\bar{V}^\theta(i,j), \quad t \in [0,T],\ i,j \in \mathbb{Z}_+.$$

Moreover, by (2.15), q-pair of the coupling operator satisfies

$$\bar{q}_t(i,j) := a_t(i,\gamma_t) + b_t(i,\gamma_t) + a_t(j,\tilde{\gamma}_t) + b_t(j,\tilde{\gamma}_t) \le \big(c_2(\gamma) + c_2(\tilde{\gamma})\big)\bar{V}(i,j).$$

So, by Theorem 5.1, there exists a unique coupling process up to time T generated by $\{L_t^{\gamma,\tilde{\gamma}}\}_{t\in[0,T]}$. Let $(X_t^\gamma, Y_t^{\tilde{\gamma}})$ be the coordinate process on $(\Omega, \mathscr{F}, \mathbb{P})$, where

$$\Omega := D([0,T]; \mathbb{Z}_+ \times \mathbb{Z}_+)$$

is the space of c$\grave{a}$dl$\grave{a}$g paths from $[0,T]$ to $\mathbb{Z}_+ \times \mathbb{Z}_+$, $\mathscr{F}$ is the Borel σ-field of Ω, and $\mathbb{P}$ is the distribution of the coupling process with initial distribution

$$\pi_{ij} := \mu(i) 1_{\{i=j\}}, \quad i,j \in \mathbb{Z}_+. \tag{2.18}$$

By (2.1), for $\rho(i,j) := |i-j|$ we obtain

$$\begin{aligned} L_t^{\gamma,\tilde{\gamma}}\rho(i,j) &= \big[b_t(i,\gamma_t) - b_t(j,\tilde{\gamma}_t)\big]^+ \big(\rho(i+1,j) - \rho(i,j)\big) \\ &\quad + \big[b_t(i,\gamma_t) - b_t(j,\tilde{\gamma}_t)\big]^- \big(\rho(i,j+1) - \rho(i,j)\big) \\ &\quad + \big[a_t(i,\gamma_t) - a_t(j,\tilde{\gamma}_t)\big]^+ \big(\rho(i-1,j) - \rho(i,j)\big) \\ &\quad + \big[a_t(i,\gamma_t) - a_t(j,\tilde{\gamma}_t)\big]^- \big(\rho(i,j-1) - \rho(i,j)\big) \\ &= \operatorname{sgn}(i-j)\big[b_t(i,\gamma_t) - b_t(j,\tilde{\gamma}_t) - a_t(i,\gamma_t) + a_t(j,\tilde{\gamma}_t)\big] \\ &\quad + 1_{\{i=j\}}\big[|b_t(i,\gamma_t) - b_t(i,\tilde{\gamma}_t)| + |a_t(i,\gamma_t) - a_t(i,\tilde{\gamma}_t)|\big] \\ &\le K_1(t)|i-j| + K_2(t)\mathbb{W}_1(\gamma_t,\tilde{\gamma}_t), \quad t \ge 0,\ i,j \in \mathbb{Z}_+. \end{aligned} \tag{2.19}$$

This implies

$$\mathbb{E}[|X_t^\gamma - Y_t^{\tilde{\gamma}}|] - \mathbb{E}[|X_s^\gamma - Y_s^{\tilde{\gamma}}|] \le \int_s^t \Big(K_1(r)\mathbb{E}[|X_r^\gamma - Y_r^{\tilde{\gamma}}|] + K_2(r)\mathbb{W}_1(\gamma_r,\tilde{\gamma}_r)\Big)\mathrm{d}r, \quad t \ge s \ge 0.$$

By Gronwall's inequality and (2.18), we derive

$$\mathbb{W}_1(\Phi_t\gamma, \Phi_t\tilde{\gamma}) \le \mathbb{E}[|X_t^\gamma - Y_t^{\tilde{\gamma}}|] \le \int_0^t \mathrm{e}^{\int_0^t K_1(r)\mathrm{d}r} K_2(r)\mathbb{W}_1(\gamma_r,\tilde{\gamma}_r)\mathrm{d}r, \quad t \in [0,T].$$

So, when $\lambda > 0$ is large enough,

$$\rho_\lambda(\Phi\gamma, \Phi\tilde{\gamma}) := \sup_{t\in[0,T]} \mathrm{e}^{-\lambda t}\mathbb{W}_1(\Phi_t\gamma, \Phi_t\tilde{\gamma})$$

$$\le \Big(\sup_{t\in[0,T]} K_2(t)\Big)\mathrm{e}^{\int_0^T |K_1(r)|\mathrm{d}r}\rho_\lambda(\gamma,\tilde\gamma)\int_0^T \mathrm{e}^{-\lambda r}\mathrm{d}r$$
$$\le \frac{1}{2}\rho_\lambda(\gamma,\tilde\gamma), \quad \gamma,\tilde\gamma\in\mathscr{C}_T^\mu.$$

Thus, when $\lambda>0$ is large enough, Φ is contractive in ρ_λ, so that it has a unique fixed point in $\mathscr{C}_T^\mu$. Since $T\in(0,\infty)$ is arbitrary, there is a unique $\mathbf{L}^{a,b}$-process with initial distribution μ.

(c) To prove (2.9), let $\gamma_t=P_t^*\mu$, so that (2.12) implies

$$\|\gamma_t\|_1-\|\gamma_s\|_1\le\int_s^t \big[b_r(0,\delta_0)+(K_1+K_2)(r)\|\gamma_r\|_1\big]\mathrm{d}r, \quad 0\le s\le t<\infty.$$

By Gronwall's inequality, this implies (2.9).

Let $\gamma_t=P_t^*\mu,\tilde\gamma_t=P_t^*\nu$, and let $(X_t^\gamma,Y_t^{\tilde\gamma})$ be the coupling process constructed as in above, but with initial distribution $\pi=(\pi_{ij})_{i,j\in\mathbb{Z}_+}\in\mathscr{C}(\mu,\nu)$ such that

$$\mathbb{W}_1(\mu,\nu)=\sum_{i,j=0}^\infty \pi_{ij}|i-j|=\mathbb{E}[|X_0^\gamma-Y_0^{\tilde\gamma}|]. \tag{2.20}$$

On the other hand, we have

$$\mathbb{W}_1(P_t^*\mu,P_t^*\nu)\le\mathbb{E}[|X_t^\gamma-Y_t^{\tilde\gamma}|]=:h_t.$$

Then (2.19) implies

$$h_t-h_s\le\int_s^t \big[K_1(r)h_r+K_2(r)\mathbb{W}_1(\gamma_r,\tilde\gamma_r)\big]\mathrm{d}r\le\int_s^t (K_1+K_2)(r)h_r\mathrm{d}r, \quad 0\le s\le t<\infty.$$

Combining this with (2.20) and Gronwall's inequality, we derive (2.10).

(2) When $(a_t(i,\mu),b_t(i,\mu))=(a(i,\mu),b(i,\mu))$ such that K_1 and K_2 are constant with $\lambda:=-(K_1+K_2)>0$, then (2.10) becomes the exponential contraction of P_t^*:

$$\mathbb{W}_1(P_t^*\mu,P_t^*\nu)\le\mathrm{e}^{-\lambda t}\mathbb{W}_1(\mu,\nu), \quad t\ge0,\mu,\nu\in\mathscr{P}_1.$$

According to [23, Theorem 5.1.1], this together with (2.9) implies that P_t^* has a unique invariant probability measure $\bar\mu\in\mathscr{P}_1$ and (2.11) holds. □

Remark 2.1. In the context of DDSDEs, the exponential ergodicity has been proved for the partially dissipative case and non-dissipative case, see [21, 22]. In the same spirit, it should be possibly to derive the exponential ergodicity for the present model by replacing (2.1) with $K_1+K_2<0$ by the weaker condition

$$\begin{aligned}&\big[b(i,\mu)-b(j,\nu)+a(j,\nu)-a(i,\mu)\big]\mathrm{sgn}(i-j)\\ &\quad+\big(|a(i,\mu)-a(i,\nu)|+|b(i,\mu)-b(i,\nu)|\big)1_{\{i=j\}}\\ &\le K+K_1|i-j|1_{\{|i-j|\ge N\}}+K_2\mathbb{W}_1(\mu,\nu), \quad t\ge0,\ i,j\in\mathbb{Z}_+,\ \mu,\nu\in\mathscr{P}_1,\end{aligned} \tag{2.21}$$

for some constants $K>0,N\in\mathbb{N},K_1<0$ and small enough $K_2>0$. To this end, one should adopt the coupling by reflection rather than the synchronized coupling in (2.17).

3 Lipschitz continuity in $\mathbb{W}_p$

In this section, for fixed $p \in (1, \infty)$, we study the the Lipschitz continuity of the maps

$$
\begin{aligned}
&P_t^* : \mathscr{P}_p \to \mathscr{P}_p,\\
&\mathscr{P}_p \ni \mu \mapsto P_t f(\mu) := \int_{\mathbb{Z}_+} f \mathrm{d}(P_t^*\mu), \quad f \in \mathscr{B}_b(\mathbb{Z}_+)
\end{aligned}
$$

for $t > 0$. To this end, we make the following assumption.

(H_3) Let $p \in (1, \infty)$. There exist continuous functions $\beta, \beta_1 \in C([0, \infty); (0, \infty)), \beta_k \in C([0, \infty); \mathbb{R}), 2 \le k \le 3$, such that

$$
\begin{aligned}
&p\big[(i+1)^{p-1} b_t(i,\mu) - (i-1)^{p-1} a_t(i,\mu)\big] \le \beta_1(t) + \beta_2(t) i^p + \beta_3(t)\|\mu\|_1^p,\\
&|a_t(i,\mu) - a_t(j,\nu)| + |b_t(i,\mu) - b_t(j,\nu)| \le \beta(t)\big(|i-j| + \mathbb{W}_1(\mu,\nu)\big),\\
&\quad i, j \in \mathbb{Z}_+, \quad t \ge 0, \quad \mu, \nu \in \mathscr{P}_1.
\end{aligned}
$$

We note that if (H_1) and one of (2.5) and (2.6) hold, then (H_2) and (H_3) are satisfied, where (H_2) is already verified in Example 2.1 and (H_3) can be deduced similarly.

Theorem 3.1. *Let $p \in (1, \infty)$ and assume (H_1)-(H_3). Then*

$$
\|P_t^*\mu\|_p^p \le \mathrm{e}^{\int_0^t (\beta_2+\beta_3)(s)\mathrm{d}s}\|\mu\|_p^p + \int_0^t \beta_1(s)\mathrm{e}^{\int_s^t (\beta_2+\beta_3)(r)\mathrm{d}r}\mathrm{d}s, \quad t \ge 0, \ \mu \in \mathscr{P}_p. \tag{3.1}
$$

$$
\mathbb{W}_p(P_t^*\mu, P_t^*\nu) \le \mathrm{e}^{2^p \int_0^t \beta(s)\mathrm{d}s}\mathbb{W}_p(\mu,\nu), \quad t \ge 0, \ \mu, \nu \in \mathscr{P}_p. \tag{3.2}
$$

Consequently, letting

$$
|\nabla f(i)| := \sup_{j \ne i} \frac{|f(j) - f(i)|}{|i-j|}, \quad i \in \mathbb{Z}_+, \ f \in \mathscr{B}_b(\mathbb{Z}_+),
$$

for any $p > 1$, we have

$$
\sup_{\nu \ne \mu} \frac{|P_t f(\mu) - P_t f(\nu)|}{\mathbb{W}_p(\mu,\nu)} \le \mathrm{e}^{2^p \int_0^t \beta(s)\mathrm{d}s}\big(P_t|\nabla f|^{\frac{p}{p-1}}(\mu)\big)^{\frac{p-1}{p}}, \quad t \ge 0, \ \mu \in \mathscr{P}_p. \tag{3.3}
$$

Proof. Let $\mu, \nu \in \mathscr{P}_1$ and $\gamma_t = P_t^*\mu, \tilde\gamma_t = P_t^*\nu, \ t \ge 0$. For fixed constant $p > 1$, let $g_p(i) := i^p, i \in \mathbb{Z}_+$. By $(H3)$, the operator L_t^γ defined in (2.13) satisfies

$$
\begin{aligned}
&L_t^\gamma g_p(i) = b_t(i,\gamma_t)\big[(i+1)^p - i^p\big] + a_t(i,\gamma_t)\big[(i-1)^p - i^p\big]\\
&\le p\big[(i+1)^{p-1} b_t(i,\gamma_t) - (i-1)^{p-1} a_t(i,\gamma_t)\big]\\
&\le \beta_1(t) + \beta_2(t) i^p + \beta_3(t)\|\gamma_t\|_1^p, \quad t \ge 0, \ i \in \mathbb{Z}_+.
\end{aligned}
$$

This and (2.9) imply that $\sup_{t\in[0,T]}\|\gamma_t\|_p<\infty$ for $T\in(0,\infty)$ and

$$\|\gamma_t\|_p^p \le \|\mu\|_p^p + \int_0^t \big[\beta_1(s)+(\beta_2+\beta_3)(s)\|\gamma_s\|_p^p\big]\mathrm{d}s,\quad t\ge 0,$$

which implies (3.1) by Gronwall's inequality.

Next, let $f(i,j)=|i-j|^p$. By (H_3) and

$$\begin{aligned}&\max\Big\{\big||i-j+1|^p-|i-j|^p\big|,\ \big||i-j-1|^p-|i-j|^p\big|\Big\}\\ &\le p(|i-j|+1)^{p-1}\le p2^{p-1}(1\vee|i-j|)^{p-1},\end{aligned}$$

the coupling operator $L_t^{\gamma,\tilde\gamma}$ defined in (2.17) satisfies

$$\begin{aligned}L_t^{\gamma,\tilde\gamma}f(i,j) &= \big[b_t(i,\gamma_t)-b_t(j,\tilde\gamma_t)\big]^+\big(|i-j+1|^p-|i-j|^p\big)\\ &\quad+\big[b_t(i,\gamma_t)-b_t(j,\tilde\gamma_t)\big]^-\big(|i-j-1|^p-|i-j|^p\big)\\ &\quad+\big[a_t(i,\gamma_t)-a_t(j,\tilde\gamma_t)\big]^+\big(|i-j-1|^p-|i-j|^p\big)\\ &\quad+\big[a_t(i,\gamma_t)-a_t(j,\tilde\gamma_t)\big]^-\big(|i-j+1|^p-|i-j|^p\big)\\ &\le p2^{p-1}(1\vee|i-j|)^{p-1}\big(|b_t(i,\gamma_t)-b_t(j,\tilde\gamma_t)|+|a_t(i,\gamma_t)-a_t(j,\tilde\gamma_t)|\big)\\ &\le p2^{p-1}\beta(t)(1\vee|i-j|)^{p-1}\big(|i-j|+\mathbb{W}_1(\gamma_t,\tilde\gamma_t)\big),\quad t\ge0,\ i,j\in\mathbb{Z}_+.\end{aligned}\tag{3.4}$$

By Theorem 5.1, let $(X_t^\gamma,Y_t^{\tilde\gamma})$ be generated by $L_t^{\gamma,\tilde\gamma}$ with initial distribution $\pi\in\mathscr{C}(\mu,\nu)$ such that

$$\mathbb{W}_p(\mu,\nu)^p=\sum_{i,j=0}^\infty \pi_{ij}|i-j|^p=\mathbb{E}[|X_0^\gamma-Y_0^{\tilde\gamma}|^p].$$

Noting that

$$|i-j|(1\vee|i-j|)^{p-1}=|i-j|^p,\quad i,j\in\mathbb{Z}_+,$$

which together with $\mathbb{W}_1(\gamma_t,\tilde\gamma_t)\le\mathbb{E}[|X_t^\gamma-Y_t^{\tilde\gamma}|]$ and the FKG inequality implies

$$\begin{aligned}&\mathbb{E}[(1\vee|X_t^\gamma-Y_t^{\tilde\gamma}|)^{p-1}]\mathbb{W}_1(\gamma_t,\tilde\gamma_t)\le\mathbb{E}[(1\vee|X_t^\gamma-Y_t^{\tilde\gamma}|)^{p-1}]\mathbb{E}[|X_t^\gamma-Y_t^{\tilde\gamma}|]\\ &\le\mathbb{E}[(1\vee|X_t^\gamma-Y_t^{\tilde\gamma}|)^{p-1}|X_t^\gamma-Y_t^{\tilde\gamma}|]=\mathbb{E}[|X_t^\gamma-Y_t^{\tilde\gamma}|^p].\end{aligned}$$

We deduce from (3.4) that

$$\begin{aligned}&\mathbb{E}[|X_t^\gamma-Y_t^{\tilde\gamma}|^p]\le\mathbb{W}_p(\mu,\nu)^p\\ &+p2^{p-1}\mathbb{E}\int_0^t\beta(s)\big(1\vee|X_s^\gamma-Y_s^{\tilde\gamma}|\big)^{p-1}\big(|X_s^\gamma-Y_s^{\tilde\gamma}|+\mathbb{W}_1(\gamma_s,\tilde\gamma_s)\big)\mathrm{d}s\\ &\le\mathbb{W}_p(\mu,\nu)^p+p2^p\int_0^t\beta(s)\mathbb{E}[|X_s^\gamma-Y_s^{\tilde\gamma}|^p]\mathrm{d}s,\quad t\ge0.\end{aligned}$$

Since $\mathbb{E}[|X_t^\gamma - Y_t^{\tilde\gamma}|^p] < \infty$ is ensured by (3.1), by Gronwall's inequality we obtain

$$\Big(\mathbb{E}[|X_t^\gamma - Y_t^{\tilde\gamma}|^p]\Big)^{\frac{1}{p}} \le \mathrm{e}^{2^p \int_0^t \beta(s)\mathrm{d}s} \mathbb{W}_p(\mu,\nu). \tag{3.5}$$

This implies (3.2) since $\mathbb{W}_p(P_t^*\mu, P_t^*\nu)^p \le \mathbb{E}[|X_t^\gamma - Y_t^{\tilde\gamma}|^p]$.

Finally, by Hölder's inequality, we obtain

$$\begin{aligned}
&\frac{|P_t f(\mu) - P_t f(\nu)|}{\mathbb{W}_p(\mu,\nu)} \le \frac{\mathbb{E}[|f(X_t^\gamma) - f(Y_t^{\tilde\gamma})|]}{\mathbb{W}_p(\mu,\nu)} \\
&\le \mathbb{E}\bigg[|\nabla f|(X_t^\gamma)\cdot \frac{|X_t^\gamma - Y_t^{\tilde\gamma}|}{\mathbb{W}_p(\mu,\nu)}\bigg] \\
&\le \big(P_t|\nabla f|^{\frac{p}{p-1}}\big)^{\frac{p-1}{p}} \frac{(\mathbb{E}[|X_t^\gamma - Y_t^{\tilde\gamma}|^p])^{\frac{1}{p}}}{\mathbb{W}_p(\mu,\nu)}.
\end{aligned}$$

Combining this with (3.5) we derive (3.3) . □

4 Propagation of chaos

Let $N \ge 2$, and $\mathscr{P}_1(N)$ be the space of all probability measures ν on $\mathbb{Z}_+^N$ with $\nu(\rho_N(x,o)) < \infty$, where $o = (0,0,\cdots,0)$ and

$$\rho_N(x,y) = |x-y| := \sum_{l=1}^N |x_l - y_l|, \quad x,y \in \mathbb{Z}_+^N.$$

A measure $\nu \in \mathscr{P}_1(N)$ is called symmetric, if for any permutation $l := (l_1,\cdots,l_N)$ of $(1,\cdots,N)$, ν is invariant under the map

$$\mathbb{Z}_+^N \ni x = (x_1,\cdots,x_N) \mapsto \pi_l x := (x_{l_1},\cdots,x_{l_N}) \in \mathbb{Z}_+^N.$$

For any $x \in \mathbb{Z}_+^N$, let

$$\mu^N(x) := \frac{1}{N}\sum_{l=1}^N \delta_{x_l},$$

where δ_{x_l} is the Dirac measure at $x_l \in \mathbb{Z}_+$. Let $e_l \in \mathbb{Z}_+^N$ with the l-th component 1 and others 0.

We consider the mean field particle systems $X_t^N = (X_t^{N,1},\cdots,X_t^{N,N})$ on $\mathbb{Z}_+^N$ generated by

$$L_t^N f(x) = \sum_{l=1}^N \Big[b_t\big(x_l,\mu^N(x)\big)\big(f(x+e_l)-f(x)\big) + a_t\big(x_l,\mu^N(x)\big)\big(f(x-e_l)-f(x)\big)\Big]$$

for $f\in \mathscr{B}(\mathbb{Z}_+^N),\ x\in\mathbb{Z}_+^N$. For any $1\le k\le N$, let

$$P_t^{N*k}\nu=\mathscr{L}_{(X_t^{N,1},\cdots,X_t^{N,k})}\text{ for }\mathscr{L}_{X_0^N}=\nu,\quad t\ge 0,\ \nu\in\mathscr{P}_1(N).$$

When $k=N$ we simply denote $P_t^{N*}=P_t^{N*N}$.

For any $\mu_0\in\mathscr{P}_2$, let $\mu_0^{\otimes N}$ be the N times product measure of μ_0. Under (H_1)-$(H3)$ for $p=2$, let

$$\begin{aligned}
h_t(\mu_0)&:=\bigg(\mathrm{e}^{\int_0^t(\beta_2+\beta_3)(s)\mathrm{d}s}\|\mu_0\|_2^2+\int_0^t\beta_1(s)\mathrm{e}^{\int_s^t(\beta_2+\beta_3)(r)\mathrm{d}r}\mathrm{d}s\bigg)^{\frac12},\\
H_t(\mu_0)&:=1+h_t(\mu_0)+\int_0^t\big(1+h_s(\mu_0)\big)K_2(s)\mathrm{e}^{\int_s^tK_1(r)\mathrm{d}r}\mathrm{d}s,\quad t\ge 0.
\end{aligned}$$

Let $\mathbb{W}_1$ be the 1-Wasserstein distance on $\mathscr{P}_1(N)$ induced by ρ_N.

Theorem 4.1. *Assume (H_1)-(H_3) for $p=2$.*

(1) *There exists a constant $c>0$ such that for any $N\ge 2$, $\mu_0\in\mathscr{P}_2$ and $\nu\in\mathscr{P}_1(N)$, we have*

$$\begin{aligned}
&\mathbb{W}_1\big(P_t^{N*}\nu,(P_t^*\mu_0)^{\otimes N}\big)\\
&\le \mathrm{e}^{\int_0^t(K_1+K_2)(s)\mathrm{d}s}\mathbb{W}_1(\nu,\mu_0^{\otimes N})+c\sqrt{N}\int_0^t\mathrm{e}^{\int_s^tK_1(r)\mathrm{d}r}K_2(s)H_s(\mu_0)\mathrm{d}s.
\end{aligned}\tag{4.1}$$

Consequently, if ν is symmetric, then for any $N\ge 2$ and $1\le k\le N$,

$$\begin{aligned}
&\mathbb{W}_1\big(P_t^{N*k}\nu,(P_t^*\mu_0)^{\otimes k}\big)\\
&\le \frac{k}{N}\mathrm{e}^{\int_0^t(K_1+K_2)(s)\mathrm{d}s}\mathbb{W}_1(\nu,\mu_0^{\otimes N})+\frac{ck}{\sqrt{N}}\int_0^t\mathrm{e}^{\int_s^tK_1(r)\mathrm{d}r}K_2(s)H_s(\mu_0)\mathrm{d}s.
\end{aligned}\tag{4.2}$$

(2) *If $(a_t(i,\mu),b_t(i,\mu))=(a(i,\mu),b(i,\mu))$ does not depend on t such that (H_1) holds with $K_1+K_2<0$ and $\beta_2+\beta_3<0$, then for any $N\ge 2$, $1\le k\le N$, $\mu_0\in\mathscr{P}_2$ and $\nu\in\mathscr{P}_1(N)$ being symmetric, there exists a constant $c>0$ such that*

$$\mathbb{W}_1\big(P_t^{N*k}\nu,(P_t^*\mu_0)^{\otimes k}\big)\le\frac{k}{N}\mathrm{e}^{(K_1+K_2)t}\mathbb{W}_1(\nu,\mu_0^{\otimes N})+c\frac{k}{\sqrt{N}}(1+\|\mu_0\|_2).\tag{4.3}$$

Proof. It is standard that (4.1) implies (4.2). By the triangle inequality, we have

$$\mathbb{W}_1\big(P_t^{N*}\nu,(P_t^*\mu_0)^{\otimes N}\big)\le\mathbb{W}_1\big(P_t^{N*}\nu,P_t^{N*}\mu_0^{\otimes N}\big)+\mathbb{W}_1\big(P_t^{N*}\mu_0^{\otimes N},(P_t^*\mu_0)^{\otimes N}\big).$$

Then (4.1) follows from (4.5) and (4.8) for $\tilde\nu=\mu_0^{\otimes N}$, which are proved in the following two lemmas. (4.3) is easily followed by (4.2) by direct calculations. □

Lemma 4.2. *Under the assumption of Theorem 4.1, for any $\mu_0 \in \mathscr{P}_2$, let $\mu_t := P_t^*\mu_0$ for $t \geq 0$, and moreover, for any $N \geq 2$, let X_t^N have initial distribution $\mu_0^{\otimes N}$. Then there exists a constant $c > 0$ such that for any $t \geq 0$ and X_t^N having initial distribution $\mu_0^{\otimes N}$, we have*

$$\mathbb{E}\big[\mathbb{W}_1(\mu^N(X_t^N), \mu_t)\big] \leq \frac{c}{\sqrt{N}} H_t(\mu_0), \tag{4.4}$$

$$\mathbb{W}_1\big(P_t^{N*}\mu_0^{\otimes N}, (P_t^*\mu_0)^{\otimes N}\big) \leq c\sqrt{N} \int_0^t \mathrm{e}^{\int_s^t K_1(r)\mathrm{d}r} K_2(s) H_s(\mu_0)\mathrm{d}s. \tag{4.5}$$

Proof. Given the initial value $X_0^N = Y_0^N$ with $\mathscr{L}_{X_0^N} = \mu_0^{\otimes N}$, by Theorem 5.1 there is a unique inhomogenous Markov process (X_t^N, Y_t^N) on $\mathbb{Z}_+^N \times \mathbb{Z}_+^N$ generated by $\{\tilde{L}_t^{\mu,N} : t \geq 0\}$, where for $t \geq 0, f \in \mathscr{B}(\mathbb{Z}_+^N)$ and $x = (x_1, \cdots, x_N), y = (y_1, \cdots, y_N) \in \mathbb{Z}_+^N$,

$$\begin{aligned}
\tilde{L}_t^{\mu,N} f(x,y) := \sum_{l=1}^N \Big[&\big(b_t(x_l, \mu^N(x)) \wedge b_t(y_l, \mu_t)\big)\big(f(x+e_l, y+e_l) - f(x,y)\big) \\
&+ \big(a_t(x_l, \mu^N(x)) \wedge a_t(y_l, \mu_t)\big)\big(f(x-e_l, y-e_l) - f(x,y)\big) \\
&+ \big(b_t(x_l, \mu^N(x)) - b_t(y_l, \mu_t)\big)^+\big(f(x+e_l, y) - f(x,y)\big) \\
&+ \big(b_t(x_l, \mu^N(x)) - b_t(y_l, \mu_t)\big)^-\big(f(x, y+e_l) - f(x,y)\big) \\
&+ \big(a_t(x_l, \mu^N(x)) - a_t(y_l, \mu_t)\big)^+\big(f(x-e_l, y) - f(x,y)\big) \\
&+ \big(a_t(x_l, \mu^N(x)) - a_t(y_l, \mu_t)\big)^-\big(f(x, y-e_l) - f(x,y)\big)\Big].
\end{aligned}$$

Then

$$\mathscr{L}_{X_t^N} = P_t^{N*}\mu_0^{\otimes N}, \quad \mathscr{L}_{Y_t^N} = (P_t^*\mu_0)^{\otimes N}, \quad t \geq 0.$$

Then as in (2.19), we have

$$\begin{aligned}
\tilde{L}_t^{\mu,N} \rho_N(x,y) &\leq \sum_{l=1}^N \big[K_1(t)|x_l - y_l| + K_2(t)\mathbb{W}_1(\mu^N(x), \mu_t)\big] \\
&= K_1(t)\rho_N(x,y) + NK_2(t)\mathbb{W}_1(\mu^N(x), \mu_t), \quad t \geq 0, \ x, y \in \mathbb{Z}_+^N.
\end{aligned}$$

Noting that $X_0^N = Y_0^N$, this implies

$$\mathbb{E}[|X_t^N - Y_t^N|] \leq N \int_0^t \mathrm{e}^{\int_s^t K_1(r)\mathrm{d}r} K_2(s)\mathbb{E}[\mathbb{W}_1(\mu^N(X_s^N), \mu_s)]\mathrm{d}s. \tag{4.6}$$

Consequently,

$$\begin{aligned}
\mathbb{E}[\mathbb{W}_1(\mu^N(X_t^N), \mu^N(Y_t^N))] &\leq \frac{1}{N}\mathbb{E}[|X_t^N - Y_t^N|] \\
&\leq \int_0^t \mathrm{e}^{\int_s^t K_1(r)\mathrm{d}r} K_2(s)\mathbb{E}[\mathbb{W}_1(\mu^N(X_s^N), \mu_s)]\mathrm{d}s, \quad t \geq 0.
\end{aligned} \tag{4.7}$$

On the other hand, noting that the components $\{Y_t^i\}_{1\le i\le N}$ of Y_t^N are i.i.d. with distribution μ_t, by [9, Lemma 2.1], there exists a constant $c>0$ such that

$$\mathbb{E}[\mathbb{W}_1(\mu^N(Y_t^N),\mu_t)]\le \frac{c}{\sqrt{N}}\big(1+\|\mu_t\|_2\big),\quad t\ge 0,\ N\ge 2.$$

Combining this with (4.1) and the triangle inequality, we obtain

$$\begin{aligned}&\mathbb{E}[\mathbb{W}_1(\mu^N(X_t^N),\mu_t)]\le \mathbb{E}[\mathbb{W}_1(\mu^N(X_t^N),\mu^N(Y_t^N))]+\mathbb{E}[\mathbb{W}_1(\mu^N(Y_t^N),\mu_t)]\\ &\le \frac{c}{\sqrt{N}}\big(1+\|\mu_t\|_2\big)+\int_0^t \mathrm{e}^{\int_s^t K_1(r)\mathrm{d}r}K_2(s)\mathbb{E}[\mathbb{W}_1(\mu^N(X_s^N),\mu_s)]\mathrm{d}s,\quad t\ge 0.\end{aligned}$$

Combining this with $\|\mu_t\|_2\le h_t(\mu_0)$ due to (3.1) for $p=2$, we derive (4.4) from Gronwall's inequality, while (4.5) follows from (4.4) and (4.6).

□

Lemma 4.3. *Assume (H_1)-(H_2). Then for any $N\in\mathbb{N}$, $\nu,\tilde\nu\in\mathscr{P}_1(\mathbb{Z}_+^N)$ and $t>0$,*

$$\mathbb{W}_1(P_t^{N*}\nu,P_t^{N*}\tilde\nu)\le \mathrm{e}^{\int_0^t(K_1+K_2)(s)\mathrm{d}s}\mathbb{W}_1(\nu,\tilde\nu). \tag{4.8}$$

Proof. For the initial value $(X_0^N,\tilde X_0^N)$ with

$$\mathscr{L}_{X_0^N}=\nu,\quad \mathscr{L}_{\tilde X_0^N}=\tilde\nu,\quad \mathbb{E}[\rho_N(X_0^N,\tilde X_0^N)]=\mathbb{W}_1(\nu,\tilde\nu),$$

by Theorem 5.1 let $(X_t^N,\tilde X_t^N)$ be the inhomogenous Markov process on $\mathbb{Z}_+^N\times\mathbb{Z}_+^N$ generated by $\{\tilde L_t^N: t\ge 0\}$, where

$$\begin{aligned}\tilde L_t^N f(x,y):=\sum_{l=1}^N\Big[&\big(b_t(x_l,\mu^N(x))\wedge b_t(y_l,\mu^N(y))\big)\big(f(x+e_l,y+e_l)-f(x,y)\big)\\ &+\big(a_t(x_l,\mu^N(x))\wedge a_t(y_l,\mu^N(y))\big)\big(f(x-e_l,y-e_l)-f(x,y)\big)\\ &+\big(b_t(x_l,\mu^N(x))-b_t(y_l,\mu^N(y))\big)^+\big(f(x+e_l,y)-f(x,y)\big)\\ &+\big(b_t(x_l,\mu^N(x))-b_t(y_l,\mu^N(y))\big)^-\big(f(x,y+e_l)-f(x,y)\big)\\ &+\big(a_t(x_l,\mu^N(x))-a_t(y_l,\mu^N(y))\big)^+\big(f(x-e_l,y)-f(x,y)\big)\\ &+\big(a_t(x_l,\mu^N(x))-a_t(y_l,\mu^N(y))\big)^-\big(f(x,y-e_l)-f(x,y)\big)\Big].\end{aligned}$$

Then as in (2.19), we have

$$\begin{aligned}&\tilde L_t^N\rho_N(x,y)\le\sum_{l=1}^N\big[K_1(t)|x_l-y_l|+K_2(t)\mathbb{W}_1(\mu^N(x),\mu^N(y))\big]\\ &=K_1(t)\rho_N(x,y)+NK_2(t)\mathbb{W}_1(\mu^N(x),\mu^N(y)),\quad t\ge 0,\ x,y\in\mathbb{Z}_+^N.\end{aligned}$$

So,

$$
\begin{aligned}
&\mathbb{W}_1(P_t^{N*}\nu, P_t^{N*}\tilde{\nu}) \le \mathbb{E}[\rho_N(X_t^N, \tilde{X}_t^N)] \\
&\le \mathbb{E}[\rho_N(X_0^N, \tilde{X}_0^N)] + \int_0^t K_1(s)\mathbb{E}[\rho_N(X_s^N, \tilde{X}_s^N)]\mathrm{d}s + N\int_0^t K_2(s)\mathbb{E}[\mathbb{W}_1(\mu^N(X_s^N), \mu^N(\tilde{X}_s^N))]\mathrm{d}s \\
&\le \mathbb{E}[\rho_N(X_0^N, \tilde{X}_0^N)] + \int_0^t K_1(s)\mathbb{E}[\rho_N(X_s^N, \tilde{X}_s^N)]\mathrm{d}s + \int_0^t K_2(s)\mathbb{E}[\rho_N(X_s^N, \tilde{X}_s^N)]\mathrm{d}s \\
&\le \mathbb{E}[\rho_N(X_0^N, \tilde{X}_0^N)] + \int_0^t (K_1(s) + K_2(s))\mathbb{E}[\rho_N(X_s^N, \tilde{X}_s^N)]\mathrm{d}s, \quad t \ge 0.
\end{aligned}
$$

By Gronwall's inequality, we derive (4.8). □

5 Appendix: inhomogenous jump processes

Let E be a locally compact Polish space with Borel σ-algebra $\mathscr{E}$, and let

$$q : [0,\infty) \times E \times \mathscr{E} \to [0,\infty)$$

such that for each $t \ge 0$, q_t gives a stable conservative q-pair, i.e. for every $x \in E$, $q_t(x,\cdot)$ is a finite measure on E with $q_t(x,\{x\}) = 0$; and for each $A \in \mathscr{E}$, $q_t(\cdot, A)$ is a measurable function on E. Define $q_t(x) := q_t(x, E\backslash\{x\})$. Let $\mathscr{B}_b(E)$ be the class of bounded measurable functions on E.

(A_1) For each $t \ge 0$, there exists a unique homogenous jump process with transition probability kernel $\{P_s^{(t)}(x,\mathrm{d}y)\}_{s\ge0,x\in E}$ satisfying the Kolmogorov equations

$$\partial_s P_s^{(t)} f = L_t P_s^{(t)} f = P_s^{(t)} L_t f, \quad f \in \mathscr{B}_b(E),\ s \ge 0,$$

where $P_s^{(t)} f(x) := \int_E f(y) P_s^{(t)}(x,\mathrm{d}y)$ and

$$L_t f(x) := \int_E \big(f(y) - f(x)\big) q_t(x,\mathrm{d}y), \quad f \in C_0(E).$$

(A_2) The q-pairs $(q_t(x,\cdot))_{t\ge0,x\in E}$ satisfies

$$\lim_{\varepsilon\downarrow0} \sup_{s,t\le T,|t-s|\le\varepsilon} \|q_t(x,\cdot) - q_s(x,\cdot)\|_{var} = 0, \quad T > 0,$$

locally uniform in $x \in E$, where $\|\cdot\|_{var}$ is the total variation distance.

(A_3) There exist a measurable function $V \ge 1$ on E with set $\{x : V(x) \le c\}$ being compact for any $c > 0$, $K \in C([0,\infty);(0,\infty))$ and a constant $\theta > 1$ such that

$$q_t(x) \le K(t)V(x), \quad L_t V^\theta(x) \le K(t)V^\theta(x), \quad t \ge 0, x \in E.$$

Theorem 5.1. *Assume (A_1)-(A_3). There exists a unique inhomogeneous Markov process generated by $(L_t)_{t\geq 0}$, which is a family of transition probability kernels $\{P_{s,t}(x,\cdot):\ 0\leq s\leq t<\infty, x\in E\}$ on E, such that*

$$P_{s,t}f(x) := \int_E f(y)P_{s,t}(x,\mathrm{d}y),\quad t\geq s\geq 0,\ f\in\mathscr{B}_b(E)$$

satisfies the semigroup property $P_{s,t}=P_{s,r}P_{r,t}$ for $0\leq s\leq r\leq t$, and the Kolmogorov equations

$$\partial_s P_{s,t}f := \lim_{\varepsilon\downarrow 0}\frac{P_{s+\varepsilon,t}f-P_{s,t}f}{\varepsilon} = -L_sP_{s,t}f,\quad t>s\geq 0,\ f\in\mathscr{B}_b(E), \tag{5.1}$$

$$\partial_t P_{s,t}f := \lim_{\varepsilon\downarrow 0}\frac{P_{s,t+\varepsilon}f-P_{s,t}f}{\varepsilon} = P_{s,t}L_tf,\quad t>s\geq 0,\ f\in\mathscr{B}_b(E). \tag{5.2}$$

Proof. (a) By (A_1), for any $n\in\mathbb{N}$ and $k\in\mathbb{Z}_+$, let $\{P_t^{n,k}(x,\cdot)\}_{t\geq 0,x\in E}$ be the unique Markov transition kernel generated by $L_{k2^{-n}}$ satisfying the Kolmogorov equations. For any $t\geq 0$, let

$$k_n(t) := \sup\{k\in\mathbb{Z}_+:\ k2^{-n}\leq t\}.$$

We now construct the transition probability kernel $\{P_{s,t}^{(n)}(x,\mathrm{d}y)\}$ for any $0\leq s\leq t<\infty, x\in E$ and $A\in\mathscr{E}$ piecewisely in time intervals $\{[k2^{-n},(k+1)2^{-n}):k\in\mathbb{Z}_+\}$:

$$\begin{aligned}
P_{s,t}^{(n)}(x,A) &:= P_{t-s}^{n,k_n(s)}(x,A) \text{ if } k_n(t)-k_n(s)=0;\\
&:= \int_E P^{n,k_n(t)}_{t-k_n(t)2^{-n}}(y_1,A)P^{n,k_n(s)}_{(k_n(s)+1)2^{-n}-s}(x,\mathrm{d}y_1) \text{ if } k_n(t)-k_n(s)=1;\\
&:= \int_{E^l} P^{n,k_n(t)}_{t-k_n(t)2^{-n}}(y_l,A)P^{n,k_n(s)}_{(k_n(s)+1)2^{-n}-s}(x,\mathrm{d}y_1)\prod_{i=1}^{l-1}P^{n,k_n(s)+i}_{2^{-n}}(y_i,\mathrm{d}y_{i+1}) \text{ if } k_n(t)-k_n(s)=l\geq 2.
\end{aligned}$$

By (A_1), it is easy to check that $\{P_{s,t}^{(n)}(x,\cdot)\}_{0\leq s\leq t<\infty,x\in E}$ is an inhomogeneous Markov process generated by

$$L_t^{(n)} := L_{k_n(t)2^{-n}},\quad t\geq 0,$$

in the sense that

$$P_{s,t}^{(n)}f(x) := \int_E f(y)P_{s,t}^{(n)}(x,\mathrm{d}y),\quad t\geq s\geq 0,\ f\in\mathscr{B}_b(E)$$

satisfies the semigroup property $P_{s,t}^{(n)}=P_{s,r}^{(n)}P_{r,t}^{(n)}$ for $0\leq s\leq r\leq t$, and the Kolmogorov equations

$$\partial_s P_{s,t}^{(n)}f = -L_s^{(n)}P_{s,t}^{(n)}f,\quad t>s\geq 0,\ f\in\mathscr{B}_b(E), \tag{5.3}$$

$$\partial_t P^{(n)}_{s,t} f = P^{(n)}_{s,t} L^{(n)}_t f, \quad t > s \geq 0,\ f \in \mathscr{B}_b(E). \tag{5.4}$$

Below we show that when $n \to \infty$, $P_{s,t}(x, \mathrm{d}y) := \lim_{n\to\infty} P^{(n)}_{s,t}(x, \mathrm{d}y)$ exists and is the unique jump process generated by $(L_t)_{t\geq 0}$ as claimed.

(b) For any $f \in \mathscr{B}_b(E)$, let

$$P^{(n)}_{s,t} f(x) := \int_E f(y) P^{(n)}_{s,t}(x, \mathrm{d}y), \quad 0 \leq s \leq t < \infty, x \in E.$$

For any $x \in E$ and fixed $T > 0$, define

$$\varepsilon_m(x) := \sup_{n>m, t\in[0,T+1]} (|b_{k_n(t)2^{-n}}(x) - b_{k_n(t)2^{-m}}(x)| + |a_{k_n(t)2^{-n}}(x) - a_{k_n(t)2^{-m}}(x)|),\ m \geq 1.$$

By (A_2) we know that $0 \leq \varepsilon_n(x) \to 0$ as $n \to \infty$. By (A_3), for any $x \in E$, there exists a constant $c > 0$ such that

$$\sup_{t\in[0,T+1]} |(L^{(n)}_t - L^{(m)}_t) f(x)| \leq 2\|f\|_\infty \varepsilon_m(x) \leq c\|f\|_\infty V(x), \quad n \geq m \geq 1. \tag{5.5}$$

Next, by (5.3) and (5.4) we obtain the Duhamel formula

$$(P^{(n)}_{s,t} - P^{(m)}_{s,t}) f = \int_s^t P^{(n)}_{s,u} (L^{(n)}_u - L^{(m)}_u) P^{(m)}_{u,t} f \mathrm{d}u, \quad f \in \mathscr{B}_b(E).$$

So,

$$\sup_{0\leq s\leq t\leq T} |(P^{(n)}_{s,t} - P^{(m)}_{s,t}) f(x)| \leq 2T\|f\|_\infty \sup_{n\geq 1, 0\leq s\leq t\leq T} P^{(n)}_{s,t} \varepsilon_m(x). \tag{5.6}$$

By (A_3) we have

$$R := \sup_{n\geq 1, 0\leq s\leq t\leq T} P^{(n)}_{s,t} V^\theta(x) < \infty, \tag{5.7}$$

where $\theta > 1$. For any $\varepsilon > 0$, let M be the constant such that $cR/M^{\theta-1} < \varepsilon/2$. Since by (A_2), $\varepsilon_n(x) \to 0$ uniformly in $x \in \{V \leq M\}$ as $n \to \infty$, we know that there exists $m_0 = m_0(\varepsilon) > 0$ such that $\varepsilon_m(x) < \epsilon/2$ for $x \in \{V \leq M\}$. Then

$$\begin{aligned}
\sup_{n\geq 1, 0\leq s\leq t\leq T} P^{(n)}_{s,t} \varepsilon_m(x) &\leq \sup_{n\geq 1, 0\leq s\leq t\leq T} P^{(n)}_{s,t} (\varepsilon_m(x) 1_{\{V\geq M\}}) + \sup_{n\geq 1, 0\leq s\leq t\leq T} P^{(n)}_{s,t} (\varepsilon_m(x) 1_{\{V\leq M\}}) \\
&\leq cR/M^{\theta-1} + \sup_{n\geq 1, 0\leq s\leq t\leq T} P^{(n)}_{s,t} (\varepsilon_m(x) 1_{\{V\leq M\}}) \\
&\leq \varepsilon/2 + \varepsilon/2 = \varepsilon.
\end{aligned}$$

Combining this with (5.6), we derive

$$\lim_{n\to\infty} \sup_{0\leq s\leq t\leq T} \|P^{(n)}_{s,t}(x, \cdot) - P_{s,t}(x, \cdot)\|_{var} = 0, \quad x \in E, T > 0 \tag{5.8}$$

for some Markov transition probability kernel $\{P_{s,t}(x,\cdot)\}_{0\le s\le t<\infty, x\in E}$, which is an in-homogeneous Markov jump process generated by $(L_t)_{t\ge 0}$ as wanted. Indeed, for any $f\in\mathscr{B}_b(E), x\in E$ and $0\le s<t\le T$, we have

$$\frac{P^{(n)}_{s+\varepsilon,t}f(x)-P^{(n)}_{s,t}f(x)}{\varepsilon}=-\frac{1}{\varepsilon}\int_s^{s+\varepsilon}L^{(n)}_rP^{(n)}_{r,t}f(x)\mathrm{d}r,\quad \varepsilon\in(0,t-s).$$

By (A_1) and (5.8), when $n\to\infty$ this gives

$$\frac{P_{s+\varepsilon,t}f(x)-P_{s,t}f(x)}{\varepsilon}=-\frac{1}{\varepsilon}\int_s^{s+\varepsilon}L_rP_{r,t}f(x)\mathrm{d}r,\quad \varepsilon\in(0,t-s),$$

so that (5.1) holds. Similarly, by

$$\text{(5.9)}\qquad \frac{P^{(n)}_{s,t+\varepsilon}f(x)-P^{(n)}_{s,t}f(x)}{\varepsilon}=\frac{1}{\varepsilon}\int_t^{t+\varepsilon}P^{(n)}_{s,r}L^{(n)}_rf(x)\mathrm{d}r,\quad \varepsilon\in(0,T-t).$$

By (A_1) and (A_2) we have $L^{(n)}_rf\to L_rf$ as $n\to\infty$, and there exists a constant $c>0$ such that

$$\text{(5.10)}\qquad \sup_{r\in[0,T+1],n\ge 1}|L^{(n)}_rf|\le cV.$$

Combining this with (5.7) and (5.8), we may letting $n\to\infty$ in (5.9) to derive

$$\frac{P_{s,t+\varepsilon}f(x)-P_{s,t}f(x)}{\varepsilon}=\frac{1}{\varepsilon}\int_t^{t+\varepsilon}P_{s,r}L_rf(x)\mathrm{d}r,\quad \varepsilon\in(0,T-t).$$

So, (5.2) holds. Indeed, via (5.7) it suffices to show

$$\text{(5.11)}\qquad \lim_{n\to\infty}|P^{(n)}_{s,r}L^{(n)}_rf(x)-P_{s,r}L_rf(x)|=0.$$

To this end, we note that

$$\begin{aligned}&|P^{(n)}_{s,r}L^{(n)}_rf(x)-P_{s,r}L_rf(x)|\\ &\le|P^{(n)}_{s,r}L^{(n)}_rf(x)-P_{s,r}L^{(n)}_rf(x)|+|P_{s,r}L^{(n)}_rf(x)-P_{s,r}L_rf(x)|\\ &=: I_1(n)+I_2(n).\end{aligned}$$

For any $N\in\mathbb{Z}_+$, we have

$$\begin{aligned}I_1(n)&\le P^{(n)}_{s,r}(|L^{(n)}_rf|1_{\{V\ge N\}})(x)+P^{(n)}_{s,r}(|L^{(n)}_rf|1_{\{V\ge N\}})(x)\\ &\quad+|P^{(n)}_{s,r}(L^{(n)}_rf1_{\{V\le N\}})(x)-P_{s,r}(L^{(n)}_rf1_{\{V\le N\}})(x)|\\ &\le 2cP^{(n)}_{s,r}(V1_{\{V\ge N\}})(x)+|P^{(n)}_{s,r}(1_{\{V\le N\}}L^{(n)}_rf)(x)-P_{s,r}(1_{\{V\le N\}}L^{(n)}_rf)(x)|\end{aligned}$$

$$\leq \frac{2c(P^{(n)}_{s,t}V^\theta)(x)}{N^{\theta-1}} + |P^{(n)}_{s,r}(1_{\{V\leq N\}}L^{(n)}_r f)(x) - P_{s,r}(1_{\{V\leq N\}}L^{(n)}_r f)(x)|.$$

Since $\theta > 1$, by (5.7), (5.8), (5.10), we may first letting $n \to \infty$ then $N \to \infty$ to derive

$$\lim_{n\to\infty} I_1(n) = 0.$$

Moreover, by (5.10) and dominated convergence theorem we have $I_2(n) \to 0$ as $n \to \infty$. Therefore, (5.11) is verified.

(c) If there are two inhomogenous Markov processes generated by $(L_t)_{t\geq 0}$ with semigroups $(P^i_{s,t})_{i=1,2}$ satisfying the Kolmogorov equations, then follows from the Duhamel formula

$$P^1_{s,t}f = P^2_{s,t}f + \int_s^t P^1_{s,r}(L_r - L_r)P^2_{r,t}f\mathrm{d}r = 0, \quad f \in \mathscr{B}_b(E),$$

which is implied by the Kolmogorov equations. So, the uniqueness is confirmed. □

Declarations

Conflict of interest The authors declare that they have no conflict of interest that might influence the results reported in this paper.

Data Availability Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Author Contributions All authors contributed to this work.

Acknowledgement The authors would like to thank the referee for corrections and very helpful suggestions.

References

[1] O. A. Butkovsky. On ergodic properties of nonlinear Markov chains and stochastic McKean-Vlasov equations. *Theory of Probability and Its Applications*, 58(4):661–674, 2014.

[2] R. Carmona, F. Delarue, *Probabilistic Theory of Mean Field Games with Applications I,* Springer, 2019.

[3] M.-F. Chen. *From Markov chains to non-equilibrium particle systems.* World Scientific, Singapore, 2nd edition, 2004.

[4] M.-F. Chen and Y.-H. Zhang. Unified representation of formulas for single birth processes. *Front. Math. China*, 9:761–796, 2014.

[5] S. Feng. Large deviations for empirical process of mean-field interacting particle system with unbounded jumps. *Ann. Probab.*, 22(4):2122–2151, 1994.

[6] S. Feng. Large deviations for Markov processes with mean field interaction and unbounded jumps. *Probab. Theory Relat. Fields*, 100:227–252, 1994.

[7] S. Feng and X. Zheng. Solutions of a class of nonlinear master equations. *Stoc. Proc. Appl.*, 43:65–84, 1992.

[8] I. Karatzas and S. Shreve. *Brownian motion and stochastic calculus*, volume 113. Springer Science & Business Media, 2012.

[9] X. Huang. Uniform in time propagation of chaos for mean field particle system with interacting noise and partially dissipative drifts. *arXiv:2409.01606*, 2024.

[10] X. Huang, P. Ren, and F.-Y. Wang. Distribution dependent stochastic differential equations. *Front. Math. China*, 16:257–301, 2021.

[11] V. N. Kolokoltsov. *Nonlinear Markov processes and kinetic equations.* Cambridge University Press, Cambridge, UK, 2010.

[12] V. N. Kolokoltsov and O. A. Malafeyev. *Many Agent Games in Socio-economic Systems: Corruption, Inspection, Coalition Building, Network Growth, Security.* Springer, Cham, 2019.

[13] H. P. McKean. A class of Markov processes associated with nonlinear parabolic equations. *Proc. Natl. Acad. Sci.*, 56(6):1907–1911, 1966.

[14] B. A. Neumann. Nonlinear Markov chains with finite state space: invariant distributions and long-term behaviour. *J. Appl. Prob.*, 60:30–44, 2023.

[15] G. Nicolis and I. Prigogine. *Self-Organization in Nonequilibrium Systems.* Wiley, New York, 1982.

[16] M. Saburov. Ergodicity of nonlinear Markov operators on the finite dimensional space. *Nonlinear Anal. Theory Methods Appl.*, 143:105–119, 2016.

[17] T. Shiga and H. Tanaka. Central limit theorem for a system for a system of Markovian particle with mean field interactions. *Z. Wahrsch. Verw Gebiete*, 69:439–459, 1985.

[18] A.-S. Sznitman. *Topics in propagation of chaos.* in École d'Été de Probabilités de Saint-Flour XIX — 1989, Lecture Notes in Math 1464, Springer Berlin Heidelberg, 1991.

[19] C. Villani. *Optimal Transport, Old and New.* Springer, New York, 2009.

[20] F.-Y. Wang. Distribution-dependent SDEs for Landau type equations. *Stoc. Proc. Appl.*, 128:595–621, 2018.

[21] F.-Y. Wang. Exponential ergodicity for non-dissipative McKean–Vlasov SDEs. *Bernoulli*, 29:1035–1062, 2023.

[22] F.-Y. Wang. Exponential ergodicity for singular reflecting McKean–Vlasov SDEs. *Stoc. Proc. Appl.*, 160:265–293, 2023.

[23] F.-Y. Wang and P. Ren. Distribution Dependent Stochastic Differential Equations. World Scientifc, Singapore, 2025.

[24] Y.-H. Zhang. Criteria on ergodicity and strong ergodicity of single death processes. *Front. Math. China*, 13:1215–1243, 2018.

[25] R. Durrett. *Probability: Theory and Examples.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 5 edition, 2019.

[26] S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence.* Wiley Series in Probability and Statistics 1986-mar 21. John Wiley Sons, Inc., 1986.