J. Appl. Probab. **60,** 465–478 (2023)
*doi:*10.1017/jpr.2022.56

EXPONENTIAL ERGODICITY FOR A CLASS OF MARKOV PROCESSES WITH INTERACTIONS

JIANHAI BAO,* *Tianjin University*
JIAN WANG,** *Fujian Normal University*

Abstract

We establish exponential ergodicity for a class of Markov processes with interactions, including two-factor type processes and Gruschin type processes. The proof is elementary and direct via the Markov coupling technique.

Keywords: Exponential ergodicity; two-factor type process; Gruschin type process; coupling

2020 Mathematics Subject Classification: Primary 60G51
Secondary 60G52; 60J25

1. Introduction and main results

Let $D \subset \mathbb{R}^m$. For smooth functions $b : D \to \mathbb{R}^m$ and $\gamma : D \times \mathbb{R}^n \to \mathbb{R}^n$, consider the deterministic dynamical system

$$\begin{cases} \mathrm{d}X_t = b(X_t)\,\mathrm{d}t, \\ \mathrm{d}Y_t = \gamma(X_t, Y_t)\,\mathrm{d}t. \end{cases} \tag{1.1}$$

The feature of this system is that although it consists of two components, the first component is determined completely by itself. The interpretations of the model (1.1) can be explained in at least the following two aspects. The first one arises from the point of view of ecological systems. Regard X_t as the quantity of grass and Y_t as the number of vegetarians at time t. Then, (1.1) demonstrates that the quantity of grass is sufficient and does not change as the number of vegetarians varies; however, the quantity of grass influences the increase in vegetarians. The other aspect comes from the mathematical modelling of infectious diseases. For numerous infections, e.g. measles, many newborn individuals might be immune to the disease since they are protected by maternal antibodies (in particular, via the placenta and the colostrum). This is termed 'passive immunity' in the literature. To describe this phenomenon in a mathematical framework, an extra M class is taken into consideration at the beginning of the classical

Received 25 March 2021; revision received 11 May 2022.
* Postal address: Center for Applied Mathematics, Tianjin University, 300072 Tianjin, P.R. China. Email: jianhaibao@tju.edu.cn
** Postal address: School of Mathematics and Statistics & Fujian Key Laboratory of Mathematical Analysis and Applications (FJKLMAA) & Center for Applied Mathematics of Fujian Province (FJNU), Fujian Normal University, 350007 Fuzhou, P.R. China. Email: jianwang@fjnu.edu.cn

SIR (susceptible–infectious–recovered) dynamics (see, e.g., [4, Chapter 2]), which leads to the MSIR model (M = maternally derived immunity; see, e.g., [3]):

$$\begin{cases} \mathrm{d}M_t = \{p\Lambda - (\mu+\eta)M_t\}\,\mathrm{d}t, \\ \mathrm{d}S_t = \{(1-p)\Lambda + \eta M_t - \beta S_t I_t + \gamma I_t - \mu S_t\}\,\mathrm{d}t, \\ \mathrm{d}I_t = \{\beta S_t I_t - (\mu+\delta+\nu+\gamma)I_t\}\,\mathrm{d}t, \\ \mathrm{d}R_t = \{\nu I_t - \mu R_t\}\,\mathrm{d}t, \end{cases}$$

where $0 \le p < 1$ and $\Lambda, \mu, \eta, \beta, \gamma, \delta, \nu > 0$. It is apparent that this MSIR model is a typical example of (1.1).

Since the model (1.1) is often subject to environmental noise, it is natural to consider the corresponding stochastic version,

$$\begin{cases} \mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma^{(1)}(X_t)\,\mathrm{d}L_t^{(1)}, \\ \mathrm{d}Y_t = \gamma(X_t, Y_t)\,\mathrm{d}t + \sigma^{(2)}(X_t)\,\mathrm{d}L_t^{(2)}, \end{cases} \tag{1.2}$$

where $\big(L_t^{(1)}\big)_{t\ge0}$ (resp. $\big(L_t^{(2)}\big)_{t\ge0}$) is a k_1-dimensional (resp. k_2-dimensional) Lévy noise (including the standard Brownian motion), $\sigma^{(1)} : D \to \mathbb{R}^m \otimes \mathbb{R}^{k_1}$ and $\sigma^{(2)} : D \to \mathbb{R}^n \otimes \mathbb{R}^{k_2}$. Intuitively speaking, (1.2) can be regarded as a class of stochastic differential equation (SDE) with simple interactions.

Before we explain the focus of the current article, let us introduce two motivating examples related to the SDE (1.2).

Example 1.1 (*Two-factor type process.*) Let $\big(B_t^{(1)}, B_t^{(2)}\big)_{t\ge0}$ be a standard two-dimensional Brownian motion. For $i = 1, 2$ and $\gamma \in \{\alpha, \beta\} \subset (1, 2)$, let $\big(L_t^{(i,\gamma)}\big)_{t\ge0}$ be a spectrally positive γ-stable process with the Lévy measure $\nu_\gamma(\mathrm{d}z) = C_\gamma z^{-(1+\gamma)}\mathbf{1}_{\{z>0\}}\mathrm{d}z$, where $C_\gamma := (\gamma\Gamma(-\gamma))^{-1}$ and Γ is the Gamma function. We suppose that $\big(B_t^{(1)}, B_t^{(2)}\big)_{t\ge0}$, $\big(L_t^{(1,\alpha)}\big)_{t\ge0}$, and $\big(L_t^{(2,\beta)}\big)_{t\ge0}$ are mutually independent. Set $\mathbb{R}_+ := [0, \infty)$. We consider the following process $(X_t, Y_t)_{t\ge0}$ on $\mathbb{R}_+ \times \mathbb{R}$ solving

$$\begin{cases} \mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \theta_1 X_t^{1/2}\,\mathrm{d}B_t^{(1)} + \theta_2 X_{t-}^{1/\alpha}\,\mathrm{d}L_t^{(1,\alpha)}, & X_0 = x \in \mathbb{R}_+, \\ \mathrm{d}Y_t = \gamma(X_t, Y_t)\,\mathrm{d}t + \delta_1 X_t^{1/2}\,\mathrm{d}B_t^{(2)} + \delta_2 X_{t-}^{1/\beta}\,\mathrm{d}L_t^{(2,\beta)}, & Y_0 = y \in \mathbb{R}, \end{cases} \tag{1.3}$$

where $\theta_1, \theta_2, \delta_1, \delta_2 \ge 0$ with $\theta_1 + \theta_2 > 0$ and $\delta_1 + \delta_2 > 0$.

In particular, when $b(x) = a - bx$ and $\gamma(x, y) = c - dy$, the corresponding SDE (1.3) has a unique strong solution, and the associated process becomes a two-factor affine model; see, e.g., [2].

Example 1.2 (*Gruschin type process*) Let $(X_t, Y_t)_{t\ge0}$ be the process on $\mathbb{R}^{m+n} := \mathbb{R}^m \times \mathbb{R}^n$ solving the following SDE:

$$\begin{cases} \mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \mathrm{d}B_t^{(1)}, & X_0 = x \in \mathbb{R}^m, \\ \mathrm{d}Y_t = \gamma(X_t, Y_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}B_t^{(2)}, & Y_0 = y \in \mathbb{R}^n, \end{cases} \tag{1.4}$$

where $(B_t)_{t\ge0} := \big(B_t^{(1)}, B_t^{(2)}\big)_{t\ge0}$ is an $(m+n)$-dimensional Brownian motion.

https://doi.org/10.1017/jpr.2022.56 Published online by Cambridge University Press

In particular, when $m=n=1$, $b(x)=0$, $\gamma(x, y)=0$, and $\sigma(x)=|x|^l$ for $l>0$, the SDE (1.4) is reduced into

$$\begin{cases} \mathrm{d}X_t = \mathrm{d}B_t^{(1)}, \\ \mathrm{d}Y_t = |X_t|^l \, \mathrm{d}B_t^{(2)}, \end{cases}$$

where $\big(B_t^{(1)}, B_t^{(2)}\big)_{t\geq 0}$ is a two-dimensional Brownian motion. The generator of the process $(X_t, Y_t)_{t\geq 0}$ solving this SDE is given by $L=\frac{1}{2}\big(\partial_{11}+|x^{(1)}|^{2l}\partial_{22}\big)$ with $x=\big(x^{(1)}, x^{(2)}\big)$, where ∂_{ii}, $i=1, 2$, denotes the second-order gradient operator with respect to the ith component. The associated semigroup is called the classical Gruschin semigroup on $\mathbb{R}^2$ with order $l>0$; see, e.g., [19].

The objective of this article is to seek verifiable conditions under which SDEs such as (1.3) and (1.4) are exponentially ergodic. Ergodicity is the foundation for a wide class of limit theorems and long-time behavior for Markov processes. In this setting, we will provide a (more direct) strategy via the Markov coupling technique to prove the exponential ergodicity for a class of Markov processes with interactions.

In the following two subsections, we state our main results for the exponential ergodicity for the SDEs (1.3) and (1.4), respectively. For a more general result concerning the exponential ergodicity for a class of Markov processes with interactions, see Section 2.1. For simplicity, we call processes solving (1.3) and (1.4) two-factor type processes and Gruschin type processes, respectively.

Given a distance-like function $\rho:\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}_+$ (i.e. $\rho(x, y)$ is non-negative, symmetric and lower semicontinuous on $\mathbb{R}^d\times\mathbb{R}^d$, and $\rho(x, y)=0$ if only if $x=y$), the Wasserstein-type quasi-metric on $\mathscr{P}\big(\mathbb{R}^d\big)$, the set of probability measures on $\mathbb{R}^d$, is defined by

$$\mathbb{W}_\rho(\mu, \nu)=\inf_{\pi\in\mathscr{C}(\mu,\nu)}\int_{\mathbb{R}^d\times\mathbb{R}^d}\rho(x, y)\,\pi(\mathrm{d}x, \mathrm{d}y), \qquad \mu, \nu\in\mathscr{P}\big(\mathbb{R}^d\big),$$

where $\mathscr{C}(\mu, \nu)$ stands for the couplings of μ and ν. If the distance-like function ρ satisfies the triangle inequality in the weak sense, i.e. there exists a constant $K>0$ such that

$$\rho(x, y)\leq K\big(\rho(x, z)+\rho(z, y)\big), \qquad x, y, z\in\mathbb{R}^d, \tag{1.5}$$

then $\big(\mathscr{P}\big(\mathbb{R}^d\big), \mathbb{W}_\rho\big)$ is a complete space; see the proof of [6, Theorem 5.4] for more details. Let $P(t, (x, y), \cdot)$ be the transition probability kernel of the process (X_t, Y_t) with the initial value (x, y).

1.1. Two-factor type processes

We make the following assumptions.

Assumption 1.1. *There are constants $\lambda_1, \lambda_2>0$ and $r_0\geq 0$ such that, for all $x, \widetilde{x}\in\mathbb{R}_+$, $(b(x)-b(\widetilde{x}))(x-\widetilde{x})\leq -\lambda_1(x-\widetilde{x})^2\mathbf{1}_{\{|x-\widetilde{x}|\geq r_0\}}+\lambda_2(x-\widetilde{x})^2\mathbf{1}_{\{|x-\widetilde{x}|<r_0\}}$.*

Assumption 1.2. *There are constants $\kappa_1, \kappa_2>0$ and $\kappa_3\geq 0$ such that, for all $x, \widetilde{x}\in\mathbb{R}_+$ and $y, \widetilde{y}\in\mathbb{R}$,*

$$(\gamma(x, y)-\gamma(\widetilde{x}, \widetilde{y}))(y-\widetilde{y})\leq -\kappa_1(y-\widetilde{y})^2+\kappa_2|(x-\widetilde{x})(y-\widetilde{y})|+\kappa_3(y-\widetilde{y})^2\mathbf{1}_{\{|y-\widetilde{y}|\leq\kappa_3\}}. \tag{1.6}$$

https://doi.org/10.1017/jpr.2022.56 Published online by Cambridge University Press

Theorem 1.1. *Under Assumptions 1.1 and 1.2, the SDE given by (1.3) has a unique strong solution $(X_t, Y_t)_{t\geq 0}$, and the process $(X_t, Y_t)_{t\geq 0}$ is exponentially ergodic in the sense that there exist a unique invariant probability measure μ on $\mathbb{R}_+\times\mathbb{R}$ and a constant $\eta>0$ such that, for all $x\in\mathbb{R}_+$, $y\in\mathbb{R}$, and $t>0$, $\mathbb{W}_V\big(P(t,(x,y),\cdot),\mu\big)\leq C_0(x,y)\,\mathrm{e}^{-\eta t}$, where $C_0(x,y)>0$ is independent of t, and*

$$V(x,\widetilde{x},y,\widetilde{y}):=\mathbf{1}_{\{\kappa_3>0,x\neq\widetilde{x}\}}+|x-\widetilde{x}|+|y-\widetilde{y}|\wedge|y-\widetilde{y}|^2,\qquad x,\widetilde{x}\in\mathbb{R}_+,\ y,\widetilde{y}\in\mathbb{R},$$

with $\kappa_3\geq 0$ given in (1.6).

When $b(x)=a-bx$ with $a\geq 0$ and $b>0$, and $\gamma(x,y)=c-dy$ with $c\in\mathbb{R}$ and $d>0$, the exponential ergodicity in terms of the total variation norm for $(X_t, Y_t)_{t\geq 0}$ solving (1.3) (in this case the process $(X_t, Y_t)_{t\geq 0}$ is the so-called two-factor affine process) was investigated in [2, 11] via the Meyn–Tweedie approach, and in terms of the L^1-Wasserstein-type distance was studied in [1] by employing the coupling argument (see [8] for the corresponding results for general affine processes). The novelty of Theorem 1.1 is twofold.

First, (1.3) is universal and incorporates all the frameworks in [1, 2, 11]. Under Assumption 1.1, the drift term $b(x)$ satisfies the dissipative condition only for large distances. Meanwhile, under Assumption 1.2, the mapping $\gamma(x,y)$ may also be dissipative for large distances (when $\kappa_3>0$) with respect to the second variable. Hence, all the frameworks in [1, 2, 11] satisfy Assumptions 1.1 and 1.2. Therefore, to some extent, Theorem 1.1 improves the corresponding results in [1, 2, 11]. In particular, Theorem 1.1 can yield the exponential ergodicity of the process $(X_t, Y_t)_{t\geq 0}$ both in terms of the total variation norm (with respect to the first marginal) and the L^1-Wasserstein-type distance (when the distance between the second components is large). Note that, according to Theorem 1.1, the process $(X_t, Y_t)_{t\geq 0}$ determined by

$$\begin{cases}\mathrm{d}X_t=(aX_t-bX_t^3)\,\mathrm{d}t+\theta_1X_t^{1/2}\,\mathrm{d}B_t^{(1)}+\theta_2X_{t-}^{1/\alpha}\,\mathrm{d}L_t^{(1,\alpha)},\\ \mathrm{d}Y_t=(cX_t+\gamma_0(Y_t)-dY_t)\,\mathrm{d}t+\delta_1X_t^{1/2}\,\mathrm{d}B_t^{(2)}+\delta_2X_{t-}^{1/\beta}\,\mathrm{d}L_t^{(2,\beta)},\end{cases}\tag{1.7}$$

with $X_0=x\in\mathbb{R}_+$, $Y_0=y\in\mathbb{R}$, $\gamma_0(y)$ being a bounded and Lipschitz continuous function on $\mathbb{R}$, $a, b, c, d>0$, and $\theta_1, \theta_2, \delta_1, \delta_2\geq 0$ satisfying $\theta_1+\theta_2>0$ and $\delta_1+\delta_2>0$, is exponentially ergodic.

Second, the proofs of [1, 8] are based on the same coupling idea as the present paper. However, the details for the construction of the coupling are quite different. In detail, reflection or synchronous coupling for the Brownian motion part and synchronous coupling for the pure jump counterpart are employed in [1, 8], but here we apply reflection coupling for the Brownian motion part and exploit the refined basic coupling for the pure jump part. The advantage of our approach lies in that it does not depend on the order-preserving property enjoyed by two-factor affine processes. Note that the arguments in [1, 8] cannot handle exponential ergodicity for the SDE (1.3) (even for (1.7)) since the order-preserving property plays a crucial role in [1, 8]. Besides, the exponential ergodicity of $(X_t, Y_t)_{t\geq 0}$ is addressed in [1, 8] under the Wasserstein-type distance with the cost function $((x,y),(\widetilde{x},\widetilde{y}))\mapsto|x-\widetilde{x}|+|x-\widetilde{x}|^\theta+|y-\widetilde{y}|$ for $\theta\in(0,1)$ (in particular $\theta=\frac{1}{2}$ is taken in [8]), while in Theorem 1.1 we take the cost function $((x,y),(\widetilde{x},\widetilde{y}))\mapsto\mathbf{1}_{\{\kappa_3>0,x\neq\widetilde{x}\}}+|x-\widetilde{x}|+|y-\widetilde{y}|^2\wedge|y-\widetilde{y}|$. Since the cost functions involved are different, the corresponding results in [1, 8] and Theorem 1.1 cannot be compared.

https://doi.org/10.1017/jpr.2022.56 Published online by Cambridge University Press

1.2. Gruschin type processes

Suppose that the coefficients in (1.4) satisfy the following assumptions.

Assumption 1.3. *There are constants* $\lambda_1 > 0$ *and* $\lambda_2, \lambda_3 \geq 0$ *such that, for all* $x, \widetilde{x} \in \mathbb{R}^m$, $\langle b(x) - b(\widetilde{x}), x - \widetilde{x}\rangle \leq \lambda_2 |x - \widetilde{x}|^2$ *and* $\langle b(x), x\rangle \leq -\lambda_1 |x|^2 + \lambda_3$.

Assumption 1.4. *There are constants* $\kappa_1, \kappa_2 > 0$, $l \geq 1$, *and* $\kappa_3 \geq 0$ *such that, for all* $x, \widetilde{x} \in \mathbb{R}^m$ *and* $y, \widetilde{y} \in \mathbb{R}^n$,

$$
\begin{aligned}
&\langle \gamma(x, y) - \gamma(\widetilde{x}, \widetilde{y}), y - \widetilde{y}\rangle \\
&\qquad \leq -\kappa_1 |y - \widetilde{y}|^2 + \kappa_2 |y - \widetilde{y}|(1 \wedge |x - \widetilde{x}|)\big(1 + |x|^l + |\widetilde{x}|^l\big) + \kappa_3 |y - \widetilde{y}|^2 \mathbf{1}_{\{|y - \widetilde{y}| \leq \kappa_3\}}
\end{aligned}
$$

and $\|\sigma(x) - \sigma(\widetilde{x})\|^2 \leq \kappa_2 (1 \wedge |x - \widetilde{x}|)\big(1 + |x|^l + |\widetilde{x}|^l\big) + \kappa_3$.

Theorem 1.2. *Under Assumptions 1.3 and 1.4, the SDE given by (1.4) has a unique strong solution* $(X_t, Y_t)_{t\geq 0}$, *and the process* $(X_t, Y_t)_{t\geq 0}$ *is exponentially ergodic in the sense that there exist a unique invariant probability measure* μ *on* $\mathbb{R}^m \times \mathbb{R}^n$ *and a constant* $\eta > 0$ *such that, for all* $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$, *and* $t > 0$, $\mathbb{W}_V(P(t, (x, y), \cdot), \mu) \leq C_0(x, y)\,\mathrm{e}^{-\eta t}$, *where* $C_0(x, y) > 0$ *is independent of t, and*

$$
V(x, \widetilde{x}, y, \widetilde{y}) := \mathbf{1}_{\{\kappa_3 > 0, x \neq \widetilde{x}\}} + (1 \wedge |x - \widetilde{x}|)\big(1 + |x|^l + |\widetilde{x}|^l\big) + |y - \widetilde{y}|^2 \wedge |y - \widetilde{y}|,
$$

with κ_3 *as in Assumption 1.4.*

Note that the Gruschin semigroup is a typical example of hypoelliptic semigroups. There are a few regularity results concerning the Gruschin semigroups; see, for instance, [18, 19]. However, the long-time behavior of diffusions associated with Gruschin type semigroups is unavailable.

Assumption 1.3 guarantees the existence of the Lyapunov function for the first component $(X_t)_{t\geq 0}$ and ensures that $(X_t)_{t\geq 0}$ is exponentially ergodic, while Assumption 1.4 shows that the function $\gamma(x, y)$ may be dissipative for large distances with respect to the second variable and of polynomial growth with respect to the first variable, and that the diffusion term $\sigma(x)$ is allowed to be of polynomial growth. We also note that under Assumption 1.3 the process $(X_t)_{t\geq 0}$ need not preserve the order property as the first component in the two-factor affine processes.

2. Proofs of main results

We present the proofs of Theorems 1.1 and 1.2. For this, we first provide the general framework to handle the exponential ergodicity for a class of Markov processes with interactions, and then provide sufficient conditions for this general result. Some comments on extensions of Theorems 1.1 and 1.2 are given at the end of the section.

2.1. A general result

The setting below is highly motivated by the SDE (1.2), but from the point of view of the generator. Let $(X_t, Y_t)_{t\geq 0}$ be a strong Markov process on $D \times \mathbb{R}^n \subset \mathbb{R}^{m+n} := \mathbb{R}^m \times \mathbb{R}^n$ such that $(X_t)_{t\geq 0}$ itself is a strong Markov process on D, and, given $(X_t)_{t\geq 0}$, $(Y_t)_{t\geq 0}$ is also a strong Markov process on $\mathbb{R}^n$. Let L be the generator of the process $(X_t, Y_t)_{t\geq 0}$. Then we have $Lf(x, y) = L_X f(\cdot, y)(x) + L_{x,Y} f(x, \cdot)(y)$, $f \in \mathscr{D}(L)$, where L_X is the infinitesimal generator

https://doi.org/10.1017/jpr.2022.56 Published online by Cambridge University Press

of the process $(X_t)_{t\geq 0}$ and, for any fixed $x\in D$, $L_{x,Y}$ is the generator of the process $(Y_t)_{t\geq 0}$ for given $(X_t)_{t\geq 0}$.

A measurable function $f: D\times\mathbb{R}^n\to\mathbb{R}$ belongs to the domain of the extended generator L of the Markov process $(X_t, Y_t)_{t\geq 0}$ if there exists a measurable function $F: D\times\mathbb{R}^n\to\mathbb{R}$ such that the process $f(X_t, Y_t)-\int_0^t F(X_s, Y_s)\,\mathrm{d}s$, $t\geq 0$, is well defined, and is an $\mathscr{F}^{X,Y}$-martingale with respect to $\mathbb{P}_{(x,y)}$ for any $(x, y)\in D\times\mathbb{R}^n$. For any such pair (f, F), we write $f\in\mathscr{D}(L)$ and $Lf=F$.

Suppose that $\big((X_t, Y_t), (\widetilde{X}_t, \widetilde{Y}_t)\big)_{t\geq 0}$ is a strong Markov coupling process of the process $(X_t, Y_t)_{t\geq 0}$ such that $\big(X_t, \widetilde{X}_t\big)_{t\geq 0}$ itself is a strong Markov process on $D\times D$, and, for given $\big(X_t, \widetilde{X}_t\big)_{t\geq 0}$, $\big(Y_t, \widetilde{Y}_t\big)_{t\geq 0}$ is also a strong Markov process on $\mathbb{R}^n\times\mathbb{R}^n$. In particular, $\big(X_t, \widetilde{X}_t\big)_{t\geq 0}$ is a Markov coupling of the process $(X_t)_{t\geq 0}$. Let $\mathscr{L}$ be the generator of the process $\big((X_t, Y_t), (\widetilde{X}_t, \widetilde{Y}_t)\big)_{t\geq 0}$. Then, for any $f\in\mathscr{D}(\mathscr{L})$ (that is, f belongs to the domain of the extended generator $\mathscr{L}$, which is defined as $\mathscr{D}(L)$ above),

$$\mathscr{L}f(x, \widetilde{x}, y, \widetilde{y})=\mathscr{L}_X f(\cdot, \cdot, y, \widetilde{y})(x, \widetilde{x})+\mathscr{L}_{x,\widetilde{x},Y} f(x, \widetilde{x}, \cdot, \cdot)(y, \widetilde{y}),$$

where $\mathscr{L}_X$ is the generator of the coupling process $\big(X_t, \widetilde{X}_t\big)_{t\geq 0}$ and $\mathscr{L}_{x,\widetilde{x},Y}$ is the generator of the process $\big(Y_t, \widetilde{Y}_t\big)_{t\geq 0}$ for given $\big(X_t, \widetilde{X}_t\big)_{t\geq 0}$. Since $\big((X_t, Y_t), (\widetilde{X}_t, \widetilde{Y}_t)\big)_{t\geq 0}$ is a strong Markov coupling process of the process $(X_t, Y_t)_{t\geq 0}$, and $\big(X_t, \widetilde{X}_t\big)_{t\geq 0}$ is a Markov coupling of the process $(X_t)_{t\geq 0}$, $\mathscr{L}$ is a coupling operator of L (that is, for any $h((x, y), (\widetilde{x}, \widetilde{y}))=f(x, y)+g(\widetilde{x}, \widetilde{y})$ with $f, g\in\mathscr{D}(L)$, $\widetilde{L}h((x, y), (\widetilde{x}, \widetilde{y}))=Lf(x, y)+Lg(\widetilde{x}, \widetilde{y})$) and $\mathscr{L}_X$ is a coupling operator of L_X. Hence, $\mathscr{L}_{x,\widetilde{x},Y}$ is also a coupling operator of the operator $L_{x,Y}$ and $L_{\widetilde{x},Y}$.

We impose the following assumptions on $\mathscr{L}_X$ and $\mathscr{L}_{x,\widetilde{x},Y}$, respectively.

Assumption 2.1. *There are a constant $\lambda_1>0$ and a sequence of distance-like functions $(f_k)_{k\geq 1}$ such that, for all $k\geq 1$, $f_k\in\mathscr{D}(\mathscr{L}_X)$ and*

$$\mathscr{L}_X f_k(x, \widetilde{x})\leq -\lambda_1 f_k(x, \widetilde{x}), \qquad x, \widetilde{x}\in D \text{ with } |x-\widetilde{x}|\geq 1/k. \tag{2.1}$$

Assumption 2.2. *There are a distance-like function $g:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}_+$ with $g\in\mathscr{D}(\mathscr{L}_{x,\widetilde{x},Y})$ for all $x, \widetilde{x}\in D$, and constants $\lambda_2, \lambda_3>0$ such that, for any $x, \widetilde{x}\in D$ with $|x-\widetilde{x}|\geq 1/k$ and any $y, \widetilde{y}\in\mathbb{R}^n$,*

$$\mathscr{L}_{x,\widetilde{x},Y}\, g(y, \widetilde{y})\leq -\lambda_2 g(y, \widetilde{y})+\lambda_3 f_k(x, \widetilde{x}). \tag{2.2}$$

Theorem 2.1. *Under Assumptions 2.1 and 2.2, there are constants $\theta, \eta>0$ such that, for all $x, \widetilde{x}\in D$, $y, \widetilde{y}\in\mathbb{R}^n$, and $t>0$,*

$$\mathbb{E}^{((x,y),(\widetilde{x},\widetilde{y}))} V\big(X_t, \widetilde{X}_t, Y_t, \widetilde{Y}_t\big)\leq \max\{\theta, \theta^{-1}\}\mathrm{e}^{-\eta t} V(x, \widetilde{x}, y, \widetilde{y}), \tag{2.3}$$

where $\mathbb{E}^{((x,y),(\widetilde{x},\widetilde{y}))}$ is the expectation of the process $\big((X_t, Y_t), (\widetilde{X}_t, \widetilde{Y}_t)\big)_{t\geq 0}$ starting from $((x, y), (\widetilde{x}, \widetilde{y}))$, and $V(x, \widetilde{x}, y, \widetilde{y}):=f(x, \widetilde{x})+g(y, \widetilde{y})$, $x, \widetilde{x}\in D$, $y, \widetilde{y}\in\mathbb{R}^n$ with $f(x, \widetilde{x})=\liminf_{k\to\infty} f_k(x, \widetilde{x})$. Moreover, if both $(x, \widetilde{x})\mapsto f(x, \widetilde{x})$ and $(y, \widetilde{y})\mapsto g(y, \widetilde{y})$ satisfy the triangle inequality in the weak sense (see (1.5) for more details), and, for any $t>0$, $x, x_0\in D$, and $y\in\mathbb{R}^n$,

$$\mathbb{E}^x f(X_t, x_0)+\mathbb{E}^y g(Y_t, 0)<\infty, \tag{2.4}$$

then the process $(X_t, Y_t)_{t\geq 0}$ is exponentially ergodic under the quasi-distance $\mathbb{W}_V$, i.e. there exist a unique invariant probability measure μ on $D\times\mathbb{R}^n$ such that, for all $x\in D$, $y\in\mathbb{R}^n$, and $t>0$, $\mathbb{W}_V(P(t, (x, y), \cdot), \mu)\leq C_0(x, y)\,\mathrm{e}^{-\eta t}$, where $C_0(x, y)>0$ is independent of t.

https://doi.org/10.1017/jpr.2022.56 Published online by Cambridge University Press

Proof. For any $k \ge 1$ and $\theta > 0$, define $V_{k,\theta}(x, \widetilde{x}, y, \widetilde{y}) = f_k(x, \widetilde{x}) + \theta g(y, \widetilde{y})$ and $V_\theta(x, \widetilde{x}, y, \widetilde{y}) = f(x, \widetilde{x}) + \theta g(y, \widetilde{y})$ for $x, \widetilde{x} \in D$, $y, \widetilde{y} \in \mathbb{R}^n$. We claim that there exist constants $\theta, \eta > 0$ such that, for all $x, \widetilde{x} \in D$, $y, \widetilde{y} \in \mathbb{R}^n$, and $t > 0$,

$$\mathbb{E}^{((x,y),(\widetilde{x},\widetilde{y}))} V_\theta\big(X_t, \widetilde{X}_t, Y_t, \widetilde{Y}_t\big) \le \mathrm{e}^{-\eta t} V_\theta(x, \widetilde{x}, y, \widetilde{y}). \tag{2.5}$$

If (2.3) holds, then the first desired assertion (2.5) follows by the basic fact that

$$(1 \wedge \theta) V(x, \widetilde{x}, y, \widetilde{y}) \le V_\theta(x, \widetilde{x}, y, \widetilde{y}) \le (1 \vee \theta) V(x, \widetilde{x}, y, \widetilde{y}), \qquad x, \widetilde{x} \in D,\ y, \widetilde{y} \in \mathbb{R}^n.$$

Now, since $(x, \widetilde{x}) \mapsto f(x, \widetilde{x})$ and $(y, \widetilde{y}) \mapsto g(y, \widetilde{y})$ satisfy the weak form of the triangle inequality in the sense of (1.5), $((x, y), (\widetilde{x}, \widetilde{y})) \mapsto V(x, \widetilde{x}, y, \widetilde{y})$ satisfies the weak form of the triangle inequality, and so $(\mathscr{P}(D \times \mathbb{R}^n), \mathbb{W}_V)$ is a complete space. Thus, the second assertion is a consequence of (2.5) by following the argument of [15, Corollary 1.8] and by taking the condition (2.4) into consideration.

We first suppose that $x = \widetilde{x}$. Since $\big(X_t, \widetilde{X}_t\big)_{t \ge 0}$ is a Markov coupling of the process $(X_t)_{t \ge 0}$ and f is a distance-like function, for any $t > 0$, $x \in D$, and $y, \widetilde{y} \in \mathbb{R}^n$,

$$\begin{aligned}\mathbb{E}^{((x,y),(x,\widetilde{y}))} V_\theta\big(X_t, \widetilde{X}_t, Y_t, \widetilde{Y}_t\big) &= \theta \mathbb{E}^{((x,y),(x,\widetilde{y}))} g(Y_t, \widetilde{Y}_t)\\ &\le \theta \mathrm{e}^{-\lambda_2 t} g(y, \widetilde{y}) = \mathrm{e}^{-\lambda_2 t} V_\theta(x, x, y, \widetilde{y}),\end{aligned} \tag{2.6}$$

where the inequality follows from (2.2) with $x = \widetilde{x}$, i.e. $\mathscr{L}_{x,x,Y} g(y, \widetilde{y}) \le -\lambda_2 g(y, \widetilde{y})$.

In the following, we show that (2.5) is still valid for the case $x \ne \widetilde{x}$. Define the stopping time $T_X = \inf\{t > 0 : X_t = \widetilde{X}_t\}$. By the strong Markov property and (2.6), we have

$$\begin{aligned}&\mathbb{E}^{((x,y),(\widetilde{x},\widetilde{y}))}\big(V_\theta\big(X_t, \widetilde{X}_t, Y_t, \widetilde{Y}_t\big)\mathbf{1}_{\{T_X \le t\}}\big)\\ &\quad = \mathbb{E}^{((x,y),(\widetilde{x},\widetilde{y}))}\Big(\mathbf{1}_{\{T_X \le t\}} \mathbb{E}^{\left(\left(X_{T_X}, y\right),\left(X_{T_X}, \widetilde{y}\right)\right)} V_\theta\big(X_{t-T_X}, \widetilde{X}_{t-T_X}, Y_{t-T_X}, \widetilde{Y}_{t-T_X}\big)\Big)\\ &\quad \le \mathbb{E}^{((x,y),(\widetilde{x},\widetilde{y}))}\Big(\mathbf{1}_{\{T_X \le t\}} \mathrm{e}^{-\lambda_2 (t-T_X)} V_\theta\big(X_{T_X}, \widetilde{X}_{T_X}, Y_{T_X}, \widetilde{Y}_{T_X}\big)\Big).\end{aligned}$$

This yields that, for any $t > 0$, $x, \widetilde{x} \in D$ with $x \ne \widetilde{x}$, and $\lambda_0 \in (0, \lambda_2]$,

$$\begin{aligned}&\mathrm{e}^{\lambda_0 t} \mathbb{E}^{((x,y),(\widetilde{x},\widetilde{y}))} V_\theta\big(X_t, \widetilde{X}_t, Y_t, \widetilde{Y}_t\big)\\ &\quad = \mathbb{E}^{((x,y),(\widetilde{x},\widetilde{y}))}\Big(\mathrm{e}^{\lambda_0 (t \wedge T_X)} V_\theta\big(X_{t \wedge T_X}, \widetilde{X}_{t \wedge T_X}, Y_{t \wedge T_X}, \widetilde{Y}_{t \wedge T_X}\big)\mathbf{1}_{\{T_X > t\}}\Big)\\ &\quad\quad + \mathrm{e}^{\lambda_0 t} \mathbb{E}^{((x,y),(\widetilde{x},\widetilde{y}))}\Big(V_\theta\big(X_t, \widetilde{X}_t, Y_t, \widetilde{Y}_t\big)\mathbf{1}_{\{T_X \le t\}}\Big)\\ &\quad \le \mathbb{E}^{((x,y),(\widetilde{x},\widetilde{y}))}\Big(\mathrm{e}^{\lambda_0 \left(t \wedge T_X\right)} V_\theta\big(X_{t \wedge T_X}, \widetilde{X}_{t \wedge T_X}, Y_{t \wedge T_X}, \widetilde{Y}_{t \wedge T_X}\big)\mathbf{1}_{\{T_X > t\}}\Big)\\ &\quad\quad + \mathbb{E}^{((x,y),(\widetilde{x},\widetilde{y}))}\Big(\mathbf{1}_{\{T_X \le t\}} \mathrm{e}^{\lambda_0 T_X} V_\theta\big(X_{T_X}, \widetilde{X}_{T_X}, Y_{T_X}, \widetilde{Y}_{T_X}\big)\Big)\\ &\quad = \mathbb{E}^{((x,y),(\widetilde{x},\widetilde{y}))}\Big(\mathrm{e}^{\lambda_0 (t \wedge T_X)} V_\theta\big(X_{t \wedge T_X}, \widetilde{X}_{t \wedge T_X}, Y_{t \wedge T_X}, \widetilde{Y}_{t \wedge T_X}\big)\Big).\end{aligned}$$

Thus, (2.5) follows for the case $x \ne \widetilde{x}$ if we can verify that there is a constant $\lambda_0 \in (0, \lambda_2]$ such that, for all $t > 0$ and $y, \widetilde{y} \in \mathbb{R}^n$,

$$\mathbb{E}\Big(\mathrm{e}^{\lambda_0 (t \wedge T_X)} V_\theta\big(X_{t \wedge T_X}, \widetilde{X}_{t \wedge T_X}, Y_{t \wedge T_X}, \widetilde{Y}_{t \wedge T_X}\big)\Big) \le V_\theta(x, \widetilde{x}, y, \widetilde{y}). \tag{2.7}$$

Define the stopping time $T_{X,k} := \inf\{t > 0 : |X_t - \widetilde{X}_t| \le 1/k\}$, $k \ge 1$. Since $x \ne \widetilde{x}$, there exists a constant $k_0 \ge 1$ sufficiently large such that $|x - \widetilde{x}| > 1/k_0$. On the other hand, choosing $\theta \in (0, 1 \wedge (\lambda_1/\lambda_3))$ and using (2.1) and (2.2), we have, for all $k \ge k_0$, $x, \widetilde{x} \in D$ with $|x - \widetilde{x}| \ge 1/k$, and $y, \widetilde{y} \in \mathbb{R}^n$,

$$\mathscr{L}V_{k,\theta}(x, \widetilde{x}, y, \widetilde{y}) \le -\big((\lambda_1 - \theta\lambda_3) \wedge \theta\lambda_2\big) V_{k,\theta}(x, \widetilde{x}, y, \widetilde{y}).$$

Now let $\lambda_0 = \eta := (\lambda_1 - \theta\lambda_3) \wedge \theta\lambda_2$. Then, for any $k \ge k_0$, $x, \widetilde{x} \in D$ with $|x - \widetilde{x}| \ge 1/k_0$, $y, \widetilde{y} \in \mathbb{R}^n$, and $t > 0$,

$$\begin{aligned}
&\mathbb{E}^{((x,y),(\widetilde{x},\widetilde{y}))}\big(e^{\eta(t\wedge T_{X,k})} V_{k,\theta}\big(X_{t\wedge T_{Y,n}}, \widetilde{X}_{t\wedge T_{X,k}}, Y_{t\wedge T_{X,k}}, \widetilde{Y}_{t\wedge T_{X,k}}\big)\big)\\
&= V_{k,\theta}(x, \widetilde{x}, y, \widetilde{y}) + \mathbb{E}^{((x,y),(\widetilde{x},\widetilde{y}))}\int_0^{t\wedge T_{X,k}} e^{\eta s}\big(\eta V_{k,\theta}\big(X_s, \widetilde{X}_s, Y_s, \widetilde{Y}_s\big) + \mathscr{L}V_{k,\theta}\big(X_s, \widetilde{X}_s, Y_s, \widetilde{Y}_s\big)\big)\, ds\\
&\le V_{k,\theta}(x, \widetilde{x}, y, \widetilde{y}).
\end{aligned}$$

Consequently, (2.7) holds from the inequality above and Fatou's lemma. The proof is therefore completed. □

2.2. Proofs

The proofs of Theorems 1.1 and 1.2 are based on Theorem 2.1. For this, we need to provide some explicit sufficient conditions to guarantee that Assumption 2.2 holds. First, we assume that, for any $x \in D$ and $f \in C_b^2(\mathbb{R}^n)$,

$$\begin{aligned}
L_{x,Y} f(y) &= \langle \nabla f(y), \gamma(x, y)\rangle + \tfrac{1}{2}\big\langle \nabla^2 f(y), (\sigma\sigma^*)(x)\big\rangle_{\mathrm{HS}}\\
&\quad + \int_{\mathbb{R}^n} \big(f(y+z) - f(y) - \langle \nabla f(y), z\rangle \mathbf{1}_{\{|z|\le 1\}}\big)\, \nu(x, dz),
\end{aligned} \tag{2.8}$$

where $\gamma : D \times \mathbb{R}^n \to \mathbb{R}^n$, $\sigma : D \to \mathbb{R}^n \otimes \mathbb{R}^n$, $\nu(x, dz)$ is a Lévy kernel on $D \times \mathscr{B}(\mathbb{R}^n)$ such that

$$\int_{\mathbb{R}^n} \big(|z|^{\alpha_0} \wedge |z|^2\big)\, \nu(x, dz) < \infty, \qquad x \in D, \tag{2.9}$$

for some $\alpha_0 \in (0, 1]$, ∇ (resp. ∇^2) denotes the gradient operator (resp. the Hessian operator) on $\mathbb{R}^n$, and $\langle \cdot, \cdot\rangle$ means the Hilbert–Schmidt inner product on $\mathbb{R}^n \otimes \mathbb{R}^n$.

Given the coupling $\big(X_t, \widetilde{X}_t\big)_{t\ge 0}$, we adopt simultaneously the synchronous couplings for the diffusion term and the pure jump term in the second component process $(Y_t)_{t\ge 0}$. For the synchronous coupling of multidimensional diffusion processes, we refer to, e.g., [5, 7, 17] for more details. With regard to the pure jump counterpart, the synchronous coupling is constructed via the following relationship (see, e.g., [6, Theorem 0.21] and [20, Section 3.1]): for given $x, \widetilde{x} \in D$,

$$(y, \widetilde{y}) \to \begin{cases} (y+z, \widetilde{y}+z), & (\nu(x,\cdot) \wedge \nu(\widetilde{x},\cdot))(dz),\\ (y+z, \widetilde{y}), & \nu(x, dz) - (\nu(x,\cdot) \wedge \nu(\widetilde{x},\cdot))(dz),\\ (z, \widetilde{y}+z), & \nu(\widetilde{x}, dz) - (\nu(x,\cdot) \wedge \nu(\widetilde{x},\cdot))(dz). \end{cases}$$

Thus, the generator of the coupling above is given by

$$\begin{aligned}
\mathscr{L}_{x,\widetilde{x},Y}f(y,\widetilde{y}) &= \langle \partial_1 f(y,\widetilde{y}), \gamma(x,y)\rangle + \langle \partial_2 f(y,\widetilde{y}), \gamma(\widetilde{x},\widetilde{y})\rangle \\
&\quad + \tfrac{1}{2}\langle \partial_{11} f(y,\widetilde{y}), (\sigma\sigma^*)(x)\rangle_{\mathrm{HS}} + \tfrac{1}{2}\langle \partial_{22} f(y,\widetilde{y}), (\sigma\sigma^*)(\widetilde{x})\rangle_{\mathrm{HS}} \\
&\quad + \langle \partial_{12} f(y,\widetilde{y}), (\sigma\sigma^*)(\widetilde{x})\rangle_{\mathrm{HS}} \\
&\quad + \int_{\mathbb{R}^n} \big(f(y+z,\widetilde{y}+z) - f(y,\widetilde{y}) - \langle \partial_1 f(y,\widetilde{y}), z\rangle \mathbf{1}_{\{|z|\le 1\}} \\
&\qquad\quad - \langle \partial_2 f(y,\widetilde{y}), z\rangle \mathbf{1}_{\{|z|\le 1\}}\big)(\nu(x,\cdot)\wedge\nu(\widetilde{x},\cdot))(\mathrm{d}z) \\
&\quad + \int_{\mathbb{R}^n} \big(f(y+z,\widetilde{y}) - f(y,\widetilde{y}) - \langle \partial_1 f(y,\widetilde{y}), z\rangle \mathbf{1}_{\{|z|\le 1\}}\big) \\
&\qquad\quad \times \big(\nu(x,\mathrm{d}z) - (\nu(x,\cdot)\wedge\nu(\widetilde{x},\cdot))(\mathrm{d}z)\big) \\
&\quad + \int_{\mathbb{R}^n} \big(f(y,\widetilde{y}+z) - f(y,\widetilde{y}) - \langle \partial_2 f(y,\widetilde{y}), z\rangle \mathbf{1}_{\{|z|\le 1\}}\big) \\
&\qquad\quad \times \big(\nu(\widetilde{x},\mathrm{d}z) - (\nu(x,\cdot)\wedge\nu(\widetilde{x},\cdot))(\mathrm{d}z)\big),
\end{aligned} \tag{2.10}$$

where ∂_{11} (resp. ∂_{22}) denotes the second-order gradient operator with respect to the first (resp. second) variable, and ∂_{12} stands for the second-order gradient operator by taking gradient with respect to the first component followed by doing so with respect to the second component.

Let

$$g_0(r) = \eta(r)r^2 + (1-\eta(r))r^{\alpha_0}, \qquad r \ge 0, \tag{2.11}$$

where $\alpha_0 \in (0, 1]$ is introduced in (2.9) and $\eta : [0, \infty) \to [0, 1]$ is defined as

$$\eta(r) = \begin{cases} 1, & r \in \left[0, \tfrac{5}{8}\right], \\ -30\cdot 4^5 \cdot \left[\tfrac{1}{5}\left(r-\tfrac{7}{8}\right)^5 + \tfrac{1}{8}\left(r-\tfrac{7}{8}\right)^4 + \tfrac{1}{48}\left(r-\tfrac{7}{8}\right)^3\right], & \tfrac{5}{8} < r < \tfrac{7}{8}, \\ 0, & r \ge \tfrac{7}{8}. \end{cases}$$

A direct calculation shows that $\eta \in C^2(\mathbb{R}_+)$ and is decreasing. Moreover, it is easy to see that $g_0 \in C^2(\mathbb{R}_+)$ is an increasing function satisfying $g_0(r) \le r^{\alpha_0}$ for all $r \ge 0$.

Proposition 2.1. *Assume that Assumption 2.1 holds. Suppose further that there are constants $c_1^* > 0$ and $c_2^*, c_3^* \ge 0$ such that, for all $k \ge 1$, $x, \widetilde{x} \in D$ with $|x-\widetilde{x}| \ge 1/k$, and $y, \widetilde{y} \in \mathbb{R}^n$,*

$$\langle y-\widetilde{y}, \gamma(x,y) - \gamma(\widetilde{x},\widetilde{y})\rangle \le -c_1^*|y-\widetilde{y}|^2 + c_2^*|y-\widetilde{y}|f_k(x,\widetilde{x}) \tag{2.12}$$

and

$$\|\sigma(x) - \sigma(\widetilde{x})\|^2 + \int_{\mathbb{R}^n} \big(|z|^2 \wedge |z|^{\alpha_0}\big)|\nu(x,\cdot) - \nu(\widetilde{x},\cdot)|(\mathrm{d}z) \le c_3^* f_k(x,\widetilde{x}), \tag{2.13}$$

where $\{f_k\}_{k\ge 1}$ is a sequence of distance-like functions given in Assumption 2.1. Then, Assupmtion 2.2 holds, i.e. there are constants $c_1, c_2 > 0$ such that, for any $k \ge 1$, $x, \widetilde{x} \in D$ with $|x-\widetilde{x}| \ge 1/k$, and $y, \widetilde{y} \in \mathbb{R}^n$, $\mathscr{L}_{x,\widetilde{x},Y}g(y,\widetilde{y}) \le -c_1 g(y,\widetilde{y}) + c_2 f_k(x,\widetilde{x})$, where $g(y,\widetilde{y}) := g_0(|y-\widetilde{y}|)$ with g_0 given by (2.11).

Proof. According to the definition of the function g_0, there exist constants $c_3 > 0$ and $c_4 \ge 1$ such that

$$\|g_0'\|_\infty + \|g_0''\|_\infty \le c_3, \quad c_4^{-1}g_0'(r)r \le g_0(r) \le c_4 g_0'(r)r, \quad g_0'(r)/r \le c_4, \quad r \ge 0. \tag{2.14}$$

https://doi.org/10.1017/jpr.2022.56 Published online by Cambridge University Press

From (2.10), we deduce that, for any $k \geq 1$, $x, \widetilde{x} \in D$ with $|x-\widetilde{x}| \geq 1/k$, and $y, \widetilde{y} \in \mathbb{R}^n$,

$$\begin{aligned}
\mathscr{L}_{x,\widetilde{x},Y} g_0(|y-\widetilde{y}|) &= \frac{g_0'(|y-\widetilde{y}|)}{|y-\widetilde{y}|}\langle y-\widetilde{y}, \gamma(x,y)-\gamma(\widetilde{x},\widetilde{y})\rangle \\
&\quad + \Bigg[\frac{g_0''(|y-\widetilde{y}|)}{2|y-\widetilde{y}|^2}\langle (y-\widetilde{y})\otimes(y-\widetilde{y}), \Gamma(x,\widetilde{x})\rangle_{\mathrm{HS}} \\
&\qquad + \frac{g_0'(|y-\widetilde{y}|)}{2|y-\widetilde{y}|}\big(\|\sigma(x)-\sigma(\widetilde{x})\|^2 \\
&\qquad\qquad - \langle (y-\widetilde{y})/|y-\widetilde{y}|, \Gamma(x,\widetilde{x})(y-\widetilde{y})/|y-\widetilde{y}|\rangle_{\mathrm{HS}}\big)\Bigg] \\
&\quad + \int_{\mathbb{R}^n}\bigg(g_0(|y+z-\widetilde{y}|)-g_0(|y-\widetilde{y}|)-\frac{g_0'(|y-\widetilde{y}|)}{|y-\widetilde{y}|}\langle y-\widetilde{y}, z\rangle \mathbf{1}_{\{|z|\leq 1\}}\bigg) \\
&\qquad \times \big(\nu(x,\mathrm{d}z)-(\nu(x,\cdot)\wedge\nu(\widetilde{x},\cdot))(\mathrm{d}z)\big) \\
&\quad + \int_{\mathbb{R}^n}\bigg(g_0(|y-\widetilde{y}-z|)-g_0(|y-\widetilde{y}|)-\frac{g_0'(|y-\widetilde{y}|)}{|y-\widetilde{y}|}\langle \widetilde{y}-y, z\rangle \mathbf{1}_{\{|z|\leq 1\}}\bigg) \\
&\qquad \times \big(\nu(\widetilde{x},\mathrm{d}z)-(\nu(x,\cdot)\wedge\nu(\widetilde{x},\cdot))(\mathrm{d}z)\big) \\
&=: \Lambda_1+\Lambda_2+\Lambda_3+\Lambda_4,
\end{aligned}$$

where $\Gamma(x,\widetilde{x}) := (\sigma(x)-\sigma(\widetilde{x}))(\sigma(x)-\sigma(\widetilde{x}))^* \in \mathbb{R}^n\otimes\mathbb{R}^n$.

Due to (2.12) and (2.14), for any $k \geq 1$, $x, \widetilde{x} \in D$ with $|x-\widetilde{x}| \geq 1/k$, and $y, \widetilde{y} \in \mathbb{R}^n$,

$$\Lambda_1 \leq g_0'(|y-\widetilde{y}|)\big(-c_1^*|y-\widetilde{y}|+c_2^* f_k(x,\widetilde{x})\big) \leq -c_5 g_0(|y-\widetilde{y}|)+c_6 f_k(x,\widetilde{x})$$

holds with some constants $c_5, c_6 > 0$. Using the fact that the matrix Γ is non-negative definite, and (2.13) as well as (2.14), we find that there is a constant $c_7 > 0$ such that, for any $k \geq 1$, $x, \widetilde{x} \in D$ with $|x-\widetilde{x}| \geq 1/k$, and $y, \widetilde{y} \in \mathbb{R}^n$, $\Lambda_2 \leq c_7 f_k(x,\widetilde{x})$. Furthermore, we can derive that, for any $k \geq 1$, $x, \widetilde{x} \in D$ with $|x-\widetilde{x}| \geq 1/k$, and $y, \widetilde{y} \in \mathbb{R}^n$,

$$\Lambda_3 \leq c_8 f_k(x,\widetilde{x}) \tag{2.15}$$

holds with some constant $c_8 > 0$. Indeed, by using $g_0(r) \leq r^{\alpha_0}$ for $r \geq 0$, $g_0(r) = r^{\alpha_0}$ for $r \geq 1$, and the fact that $(a+b)^{\alpha_0} \leq a^{\alpha_0}+b^{\alpha_0}$ for all $a, b \geq 0$, we find that, for any $y, \widetilde{y} \in \mathbb{R}^n$ with $|y-\widetilde{y}| \geq 2$ and any $z \in \mathbb{R}^n$,

$$g_0(|y+z-\widetilde{y}|)-g_0(|y-\widetilde{y}|) \leq |y+z-\widetilde{y}|^{\alpha_0}-|y-\widetilde{y}|^{\alpha_0} \leq |z|^{\alpha_0}.$$

Then, for any $k \geq 1$, $x, \widetilde{x} \in D$ with $|x-\widetilde{x}| \geq 1/k$, and any $y, \widetilde{y} \in \mathbb{R}^n$ with $|y-\widetilde{y}| \geq 2$, according to the mean value theorem, we have

$$\begin{aligned}
\Lambda_3 &= \int_{\{|z|\leq 1\}}\bigg(g_0(|y+z-\widetilde{y}|)-g_0(|y-\widetilde{y}|)-\frac{g_0'(|y-\widetilde{y}|)}{|y-\widetilde{y}|}\langle y-\widetilde{y}, z\rangle\bigg) \\
&\qquad \times \big(\nu(x,\mathrm{d}z)-(\nu(x,\cdot)\wedge\nu(\widetilde{x},\cdot))(\mathrm{d}z)\big) \\
&\quad + \int_{\{|z|>1\}}\big(g_0(|y+z-\widetilde{y}|)-g_0(|y-\widetilde{y}|)\big)\big(\nu(x,\mathrm{d}z)-(\nu(x,\cdot)\wedge\nu(\widetilde{x},\cdot))(\mathrm{d}z)\big)
\end{aligned}$$

https://doi.org/10.1017/jpr.2022.56 Published online by Cambridge University Press

$$
\begin{aligned}
&\leq \frac{\|g_0''\|_\infty}{2} \int_{\{|z|\leq 1\}} |z|^2 \big(\nu(x, \mathrm{d}z) - (\nu(x, \cdot) \wedge \nu(\widetilde{x}, \cdot))(\mathrm{d}z)\big) \\
&\quad + \int_{\{|z|>1\}} |z|^{\alpha_0} \big(\nu(x, \mathrm{d}z) - (\nu(x, \cdot) \wedge \nu(\widetilde{x}, \cdot))(\mathrm{d}z)\big) \\
&\leq c_9 \int_{\mathbb{R}^n} \big(|z|^2 \wedge |z|^{\alpha_0}\big) \big(\nu(x, \mathrm{d}z) - (\nu(x, \cdot) \wedge \nu(\widetilde{x}, \cdot))(\mathrm{d}z)\big)
\end{aligned}
$$

for some constant $c_9 > 0$. On the other hand, for any $k \geq 1$, $x, \widetilde{x} \in D$ with $|x - \widetilde{x}| \geq 1/k$, and any $y, \widetilde{y} \in \mathbb{R}^n$ with $|y - \widetilde{y}| \leq 2$, by the mean value theorem again, we have

$$
\begin{aligned}
\Lambda_3 &= \int_{\{|z|\leq 1\}} \left(g_0(|y + z - \widetilde{y}|) - g_0(|y - \widetilde{y}|) - \frac{g_0'(|y - \widetilde{y}|)}{|y - \widetilde{y}|} \langle y - \widetilde{y}, z\rangle \right) \\
&\qquad \times \big(\nu(x, \mathrm{d}z) - (\nu(x, \cdot) \wedge \nu(\widetilde{x}, \cdot))(\mathrm{d}z)\big) \\
&\quad + \int_{\{|z|>1\}} \big(g_0(|y + z - \widetilde{y}|) - g_0(|y - \widetilde{y}|)\big)\big(\nu(x, \mathrm{d}z) - (\nu(x, \cdot) \wedge \nu(\widetilde{x}, \cdot))(\mathrm{d}z)\big) \\
&\leq \frac{\|g_0''\|_\infty}{2} \int_{\{|z|\leq 1\}} |z|^2 \big(\nu(x, \mathrm{d}z) - (\nu(x, \cdot) \wedge \nu(\widetilde{x}, \cdot))(\mathrm{d}z)\big) \\
&\quad + \int_{\{|z|>1\}} (2 + |z|)^{\alpha_0} \big(\nu(x, \mathrm{d}z) - (\nu(x, \cdot) \wedge \nu(\widetilde{x}, \cdot))(\mathrm{d}z)\big) \\
&\leq c_{10} \int_{\mathbb{R}^n} \big(|z|^2 \wedge |z|^{\alpha_0}\big) \big(\nu(x, \mathrm{d}z) - (\nu(x, \cdot) \wedge \nu(\widetilde{x}, \cdot))(\mathrm{d}z)\big)
\end{aligned}
$$

for some constant $c_{10} > 0$. Combining both estimates above with (2.13) yields (2.15). Similarly, there exists a constant $c_{11} > 0$ such that $\Lambda_4 \leq c_{11} f_k(x, \widetilde{x})$.

Therefore, the desired assertion follows from all the estimates for Λ_i $(1 \leq i \leq 4)$. □

Now, we are in a position to present the proofs of Theorems 1.1 and 1.2.

Proof of Theorem 1.1. According to [9, Theorem 5.6], under Assumption 1.1 the first SDE in (1.3) has a unique strong solution $(X_t)_{t\geq 0}$ taking values in $\mathbb{R}_+$. Once $(X_t)_{t\geq 0}$ is available, the second SDE in (1.3) also has a unique strong solution $(Y_t)_{t\geq 0}$ since Assumption 1.2 yields that the coefficient $\gamma(x, \cdot)$ satisfies the one-sided Lipschitz condition; see, e.g., [10, Theorem 1] or [16, Theorem 1.1]. Therefore, the SDE (1.3) has a unique strong solution $(X_t, Y_t)_{t\geq 0}$. We can also follow the proof of [2, Theorem 2.2] (which only considers the affine two-factor models but works for general cases) to directly construct the strong solution to the SDE (1.3) with the aid of the Itô formula.

We split the proof into two cases.

Case (i): $\kappa_3 = 0$. For the first component process $(X_t)_{t\geq 0}$, we consider the Markov coupling $\big(X_t, \widetilde{X}_t\big)_{t\geq 0}$, where $(\widetilde{X}_t)_{t\geq 0}$ is constructed by applying the coupling by reflection for the Brownian motion and the refined basic coupling (see, e.g., [14]) for the spectrally positive stable process. Roughly speaking, the refined basic coupling is a revised version of the classical maximal coupling which will couple two marginal processes to meet together with half of the biggest jump intensity but fully adopt for the characterization of Lévy jumps. Then, it is seen from the proof of [12, Theorem 3.1] that, under Assumption 1.1, Assumption 2.1 holds for $f_k(x, \widetilde{x}) = f_0(|x - \widetilde{x}|)$ for all $x, \widetilde{x} \in \mathbb{R}_+$ with $|x - \widetilde{x}| \geq 1/k$, where the explicit expression for f_0, satisfying $c_1 r \leq f_0(r) \leq c_2 r$ for some $c_2 \geq c_1 > 0$, is given in [12, (4.2)]. (That is,

https://doi.org/10.1017/jpr.2022.56 Published online by Cambridge University Press

in this case we can choose $f_k(x,\widetilde{x})=f_0(|x-\widetilde{x}|)$ independent of k.) For the second component process $(Y_t)_{t\geq 0}$, we adopt simultaneously the synchronous couplings for the diffusion term and the pure jump term. Then, we can see that the coupling process $\big((X_t,Y_t),(\widetilde{X}_t,\widetilde{Y}_t)\big)_{t\geq 0}$ has the strong Markov property. Moreover, in this case, $\sigma(x)=\delta_1 x^{1/2}$ and $\nu(x,\mathrm{d}z)=\delta_2^{\beta} x\nu_\beta(\mathrm{d}z)$ in (2.8), and (2.9) is satisfied with $\alpha_0=1$. In particular, by Assumption 1.2, (2.12) and (2.13) hold with $f_k(x,\widetilde{x})=f_0(|x-\widetilde{x}|)$ as mentioned above. Indeed, when $\kappa_3=0$, by (1.6) we know that, for all $x,\widetilde{x}\in\mathbb{R}_+$ and $y,\widetilde{y}\in\mathbb{R}$,

$$\begin{aligned}(\gamma(x,y)-\gamma(\widetilde{x},\widetilde{y}))(y-\widetilde{y})&\leq -\kappa_1(y-\widetilde{y})^2+\kappa_2|(x-\widetilde{x})(y-\widetilde{y})|\\&\leq -\kappa_1(y-\widetilde{y})^2+\kappa_2c_1^{-1}|y-\widetilde{y}|f_0(|x-\widetilde{x}|)\\&=-\kappa_1(y-\widetilde{y})^2+\kappa_2c_1^{-1}|y-\widetilde{y}|f_k(x,\widetilde{x}).\end{aligned}$$

Therefore, according to Proposition 2.1, Assumption 2.2 is satisfied with $f_k(x,\widetilde{x})=f_0(|x-\widetilde{x}|)$ and $g(y,\widetilde{y})=g_0(|y-\widetilde{y}|)$, with g_0 given by (2.11).

Now let $V(x,\widetilde{x},y,\widetilde{y})=|x-\widetilde{x}|+|y-\widetilde{y}|\wedge|y-\widetilde{y}|^2$, $x,\widetilde{x}\in\mathbb{R}_+$, $y,\widetilde{y}\in\mathbb{R}$. It is clear that there are constants $0<c_3\leq c_4$ such that, for all $x,\widetilde{x}\in\mathbb{R}_+$ and $y,\widetilde{y}\in\mathbb{R}$,

$$c_3(f_0(|x-\widetilde{x}|)+g_0(|y-\widetilde{y}|))\leq V(x,\widetilde{x},y,\widetilde{y})\leq c_4(f_0(|x-\widetilde{x}|)+g_0(|y-\widetilde{y}|)),$$

and $(y,\widetilde{y})\mapsto|y-\widetilde{y}|\wedge|y-\widetilde{y}|^2$ satisfies the weak form of the triangle inequality. On the other hand, due to $\alpha,\beta\in(1,2)$, we can see from the Itô formula that, for all $t>0$ and $x,y\in\mathbb{R}_+$, $\mathbb{E}^{(x,y)}(|X_t|+|Y_t|)<\infty$; see, e.g., the proof of [9, Proposition 2.3]. Therefore, the desired assertion follows from Theorem 2.1.

Case (ii): $\kappa_3>0$. For the first component process $(X_t)_{t\geq 0}$, we still apply the coupling by reflection for the Brownian motion and the refined basic coupling for the spectrally positive stable process. Suppose that Assumption 1.1 holds. According to the proof of [12, Theorem 3.2], there are a constant $\lambda_1>0$ and a bounded, nondecreasing, and concave continuous function f_0^* such that Assumption 2.1 holds for $f_k(x,\widetilde{x})$, where $f_k(x,x)=0$ and $f_k(x,\widetilde{x})=1+f_0^*(|x-\widetilde{x}|)+f_0(|x-\widetilde{x}|)$ for all $x,\widetilde{x}\in\mathbb{R}_+$ with $|x-\widetilde{x}|\geq 1/k$. Here, f_0 is the same function taken in the case $\kappa_3=0$ (i.e. from [12, (4.2)]). We consider the same coupling for the second component process $(Y_t)_{t\geq 0}$ as in the case $\kappa_3=0$. Recall that $\sigma(x)=\delta_1x^{1/2}$ and $\nu(x,\mathrm{d}z)=\delta_2^{\beta}x\nu_\beta(\mathrm{d}z)$ in (2.8), and that (2.9) is satisfied with $\alpha_0=1$. Then, by Assumption 1.2, for $x,\widetilde{x}\in\mathbb{R}_+$ with $|x-\widetilde{x}|\geq 1/k$ and $y,\widetilde{y}\in\mathbb{R}$,

$$\begin{aligned}(\gamma(x,y)-\gamma(\widetilde{x},\widetilde{y}))(y-\widetilde{y})&\leq -\kappa_1(y-\widetilde{y})^2+\kappa_2|(x-\widetilde{x})(y-\widetilde{y})|+\kappa_3|y-\widetilde{y}|^2\mathbf{1}_{\{|y-\widetilde{y}|\leq\kappa_3\}}\\&\leq -\kappa_1(y-\widetilde{y})^2+c_5|y-\widetilde{y}|f_k(x,\widetilde{x}),\end{aligned}$$

and (2.13) holds as well since $f_k(x,\widetilde{x})\geq f_0(|x-\widetilde{x}|)$ for all $x,\widetilde{x}\in\mathbb{R}_+$ with $|x-\widetilde{x}|\geq 1/k$. Hence, by Proposition 2.1, Assumption 2.2 is satisfied with $f_k(x,\widetilde{x})$ and $g(y,\widetilde{y})=g_0(|y-\widetilde{y}|)$, with g_0 given by (2.11).

Note that there is a constant $0<c_6\leq c_7$ such that, for all $x,\widetilde{x}\in\mathbb{R}_+$,

$$c_6\mathbf{1}_{\{x\neq\widetilde{x}\}}(1+|x-\widetilde{x}|)\leq\lim_{k\to\infty}f_k(x,\widetilde{x})\leq c_7\mathbf{1}_{\{x\neq\widetilde{x}\}}(1+|x-\widetilde{x}|).$$

With this at hand, the desired assertion follows by similar arguments to Case (i). □

Proof of Theorem 1.2. Following the argument at the beginning of the proof of Theorem 1.1, we can see that the SDE (1.3) has a unique strong solution $(X_t,Y_t)_{t\geq 0}$ under Assumptions 1.3 and 1.4.

https://doi.org/10.1017/jpr.2022.56 Published online by Cambridge University Press

In the first place, we consider the case $\kappa_3 = 0$. For the first component process $(X_t)_{t\ge0}$, we will directly apply the reflection coupling. Then, according to Assumption 1.3 and [7, Theorem 2.2], we know that Assumption 2.1 holds for $f_k(x, \widetilde{x}) = f(x, \widetilde{x})$ for all $x, \widetilde{x} \in \mathbb{R}_+$ with $|x - \widetilde{x}| \ge 1/k$, where $f(x, \widetilde{x})$ satisfies

$$c_1(1 \wedge |x - \widetilde{x}|)\big(1 + |x|^l + |\widetilde{x}|^l\big) \le f(x, \widetilde{x}) \le c_2(1 \wedge |x - \widetilde{x}|)\big(1 + |x|^l + |\widetilde{x}|^l\big)$$

for any $l > 0$ with $c_1, c_2 > 0$; see [7, Remark 2.1 and Assumption 2.4] for more details. For the second component process $(Y_t)_{t\ge0}$, we adopt the synchronous coupling. It is clear that the coupling process $((X_t, Y_t), (\widetilde{X}_t, \widetilde{Y}_t))_{t\ge0}$ enjoys the strong Markov property, and that (2.12) and (2.13) hold under Assumption 1.4. Note that $(x, \widetilde{x}) \mapsto (1 \wedge |x - \widetilde{x}|)\big(1 + |x|^l + |\widetilde{x}|^l\big)$ satisfies the weak form of the triangle inequality. Then, following the argument at the end of the proof of Theorem 1.1, we can obtain the desired assertion by invoking Proposition 2.1 and Theorem 2.1.

Next, we turn to the case $\kappa_3 > 0$. Then, we can follow the arguments in Case (ii) of the proof of Theorem 1.1 and apply [7, Theorem 2.1] (instead of [7, Theorem 2.2]). Since the details are similar to those for the proof of Theorem 1.1, we omit them here. □

Finally, we give a remark on the extension of Theorems 1.1 and 1.2.

Remark 2.1. According to [12, Theorem 3.1] and the proof of Theorem 1.1, we can consider more general two-factor type processes by replacing the first component process in the SDE (1.3) by a class of continuous-state nonlinear branching processes as studied in [12]. Similarly, by [13, Sections 2.2 and 2.3] and the proof of Theorem 1.1, we can consider more general Gruschin type processes by replacing the Brownian motion with general Lévy noise.

Acknowledgements

The authors thank two referees for their helpful comments and corrections.

Funding information

The research of Jianhai Bao is supported by the National Natural Science Foundation of China (Nos. 11831014 and 12071340). The research of Jian Wang is supported by the National Natural Science Foundation of China (Nos. 11831014, 12071076, 12225104, and 2022YFA1000033) and the Education and Research Support Program for Fujian Provincial Agencies.

Competing interests

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

References

[1] Bao, J. and Wang, J. Coupling methods and exponential ergodicity for two-factor affine processes. Submitted. Preprint, arXiv:2004.10384.

[2] Barczy, M., Döring, L., Li, Z. and Pap, G. (2014). Stationarity and ergodicity for an affine two-factor model. *Adv. Appl. Prob.* **46**, 878–898.

[3] Bichara, D., Iggidr, A. and Sallet, G. (2014). Global analysis of multi-strains SIS, SIR and MSIR epidemic models. *J. Appl. Math. Comput.* **46**, 273–292.

[4] Capasso, V. (1993). *Mathematical Structure of Epidemic Systems*. Springer, Berlin.

[5] Chen, M.-F. and Li, S.-F. (1989). Coupling methods for multi-dimensional diffusion process. *Ann. Prob.* **17**, 151–177.

[6] CHEN, M.-F. (2004). *From Markov Chains to Non-Equilibrium Particle Systems*, 2nd edn. World Scientific, River Edge, NJ.
[7] EBERLE, A., GUILLIN, A. AND ZIMMER, R. (2019). Quantitative Harris type theorems for diffusions and McKean–Vlasov processes. *Trans. Amer. Math. Soc.* **371**, 7135–7173.
[8] FRIESEN, M., JIN, P. AND RÜDIGER, B. (2020). Stochastic equation and exponential ergodicity in Wasserstein distances for affine processes. *Ann. Appl. Prob.* **30**, 2165–2195.
[9] FU, Z. AND LI, Z. (2010). Stochastic equations of non-negative processes with jumps. *Stoch. Process. Appl.* **120**, 306–330.
[10] GYÖNGY, I. AND KRYLOV, N.V. (1980). On stochastic equations with respect to semimartingales I. *Stochastics* **4**, 1–21.
[11] JIN, P., KREMER, J. AND RÜDIGER, B. (2017). Exponential ergodicity of an affine two-factor model based on the α-root process. *Adv. Appl. Prob.* **49**, 1144–1169.
[12] LI, P.-S. AND WANG, J. (2020). Exponential ergodicity for general continuous-state nonlinear branching processes. *Electron. J. Prob.* **25**, 125.
[13] LIANG, M., MAJKA, M. AND WANG, J. (2021). Exponential ergodicity for SDEs and McKean–Vlasov processes with Lévy noise. *Ann. Inst. H. Poincaré Prob. Statist.* **57**, 1665–1701.
[14] LUO, D. AND WANG, J. (2019). Refined couplings and Wasserstein-type distances for SDEs with Lévy noises. *Stoch. Process. Appl.* **129**, 3129–3173.
[15] MAJKA, M. B. (2017). Coupling and exponential ergodicity for stochastic differential equations driven by Lévy processes. *Stoch. Process. Appl.* **127**, 4083–4125.
[16] MAJKA, M. B. (2020). A note on existence of global solutions and invariant measures for jump SDEs with locally one-sided Lipschitz drift. *Prob. Math. Statist.* **40**, 37–55.
[17] PRIOLA, E. AND WANG, F.-Y. (2006). Gradient estimates for diffusion semigroups with singular coefficients. *J. Funct. Anal.* **236**, 244–264.
[18] WANG, F.-Y. (2014). Derivative formula and gradient estimates for Gruschin type semigroups. *J. Theoret. Prob.* **27**, 30–49.
[19] WANG, F.-Y. AND XU, L. (2014). Log-Harnack inequality for Gruschin type semigroups. *Rev. Mat. Iberoam.* **30**, 405–418.
[20] WANG, J. (2010). Regularity of semigroups generated by Lévy type operators via coupling. *Stoch. Process. Appl.* **120**, 1680–1700.