Coupling methods and exponential ergodicity for two-factor affine processes

Jianhai Bao | Jian Wang

1Center for Applied Mathematics, Tianjin University, Tianjin, China
2School of Mathematics and Statistics & Fujian Key Laboratory of Mathematical Analysis and Applications (FJKLMAA) & Center for Applied Mathematics of Fujian Province (FJNU), Fujian Normal University, Fuzhou, China

Abstract
In this paper, by invoking the coupling approach, we establish exponential ergodicity under the $L^1$-Wasserstein distance for two-factor affine processes. The method employed herein is universal in a certain sense so that it is applicable to general two-factor affine processes, which allow that the first component solves a general Cox-Ingersoll-Ross (CIR) process, and that there are interactions in the second component, as well as that the Brownian noises are correlated; and even to some models beyond two-factor processes.

KEYWORDS
coupling by reflection, exponential ergodicity, synchronous coupling, two-factor affine process

MSC (2020)
60G51, 60G52, 60J25, 60J76

1 | INTRODUCTION

An affine process $(X_t)_{t \geq 0}$ is a time-homogeneous Markov process which satisfies that

$$
E^x e^{(u,X_t)} = \exp\{\phi(t,iu) + \langle x, \psi(t,iu) \rangle\}, \quad t > 0,
$$

where $\phi$ and $\psi$ satisfy the semi-flow property. See the pioneer work [13] upon the general theory of affine processes, and [12] on succinct mathematical foundations and complete characterizations of regular affine processes. In particular, when the state space $D = \mathbb{R}_+^m \times \mathbb{R}^n$, where $m, n \in \mathbb{N}_0 := \{0, 1, 2, \ldots\}$ with $m + n \geq 1$, the functions $\phi$ and $\psi$ satisfy certain generalized Riccati equations; see [12, Theorem 2.7]. Nowadays, the theory of affine processes has been developed in various directions; see, for example, [9, 10, 14, 24, 26, 27]. Meanwhile, affine processes have been applied considerably in mathematical finance due to their computational tractability and flexibility in capturing many empirical features of financial series; see the book [2], and references therein. The reader is also referred to more recent works on affine processes that are extended to infinite dimensions and beyond stochastic continuity (see, e.g., [19, 28, 39]) and the extension to Volterra processes (see, e.g., [1]).
The set of affine processes contains a large class of important Markov processes such as continuous-state branching processes with immigration with the state space \( D = \mathbb{R}^m_+ \), and Ornstein–Uhlenbeck-type processes with the state space \( D = \mathbb{R}^n \). The long time behaviors (e.g., existence and uniqueness of the invariant probability measure, ergodicity, exponential ergodicity, and so on) of those two processes have been studied extensively in the literature; see, for example [25, 30, 33, 36, 40]. In order to study long time behaviors for general affine processes on the canonical state space \( D = \mathbb{R}^m_+ \times \mathbb{R}^n \), it is natural to start from the following simplest (but interesting and important) two-factor affine process on \( \mathbb{R}_+ \times \mathbb{R} \) (i.e., \( m = n = 1 \)):

\[
\begin{align*}
    dY_t &= (a - bY_t) \, dt + Y_t^{1/\beta} \, dB_t, \quad t \geq 0, \, Y_0 = y \geq 0, \\
    dX_t &= (\lambda - \lambda X_t) \, dt + Y_t^{1/\alpha} \, dZ_t, \quad t \geq 0, \, X_0 = x \in \mathbb{R},
\end{align*}
\]  

(1.1)

where \( a \geq 0, b, \lambda \in \mathbb{R}, \beta, \alpha \in (1, 2) \), \((L_t)_{t \geq 0}\) is a spectrally positive \( \beta \)-stable process with the Lévy measure \( \nu_\beta(dz) := C_\beta z^{-(1+\beta)} 1_{[z>0]}(dz) \) with \( C_\beta = (\beta \Gamma(-\beta))^{-1} \) (where \( \Gamma \) denotes the Gamma function) in case of \( \beta \in (1, 2) \), a standard Brownian motion (which will be denoted by \((B_t)_{t \geq 0}\) later) in case of \( \beta = 2 \), and similarly \((Z_t)_{t \geq 0}\) is a spectrally positive \( \alpha \)-stable process in case of \( \alpha \in (1, 2) \), a standard Brownian motion in case of \( \alpha = 2 \). We further assume that \((L_t)_{t \geq 0}\) and \((Z_t)_{t \geq 0}\) are mutually independent.

The first coordinate process \((Y_t)_{t \geq 0}\) in Equation (1.1) is the so-called \( \beta \)-root process. In the literature, \((Y_t)_{t \geq 0}\) is also referred as the \( \beta \)-stable CIR process (see e.g. [11]), where the associated spectrally positive stable process \((L_t)_{t \geq 0}\) is interpreted as the branching mechanism; see, for example, [18]. In the sovereign bond market, financial data might experience large fluctuations and the interest rate might reach certain low level. In this setting, the short interest rate model (1.1) solving by the first component \((Y_t)_{t \geq 0}\) is one of the better candidates to reconcile large fluctuations with low rate; see, for example, [21, 30]. The second coordinate process is an Ornstein–Uhlenbeck-type process driven by the first coordinate process \((Y_t)_{t \geq 0}\). As we will see, the main purpose of this paper is concerned with the case that 

\[
\begin{align*}
    dX_t &= (\lambda - \lambda X_t) \, dt + Y_t^{1/\alpha} \, dZ_t, \quad t \geq 0, \, X_0 = x \in \mathbb{R},
\end{align*}
\]  

(1.2)

where \( a \geq 0, b, \lambda \in \mathbb{R}, (B_t)_{t \geq 0}\) is a standard Brownian motion, and \((Z_t)_{t \geq 0}\) is an independently spectrally positive \( \alpha \)-stable process with \( \alpha \in (1, 2) \). In terms of [4, Theorem 2.1], the SDE (1.2) has a unique strong solution \((Y_t, X_t)_{t \geq 0}\) in Equation (1.1) have already been received much attention in the literature; see [30] for the exponential ergodicity, [20] for the application on the sovereign bond market (indeed, the setting of [20] includes the first-coordinate process in [20]), and similar results for \( \alpha \in (1, 2) \) are still open. We will fill the gap in this paper. To taste our contribution, we herein state the result for the following SDE on \( \mathbb{R}_+ \times \mathbb{R} \):

\[
\begin{align*}
    dY_t &= (a - bY_t) \, dt + Y_t^{1/\beta} \, dB_t, \quad t \geq 0, \, Y_0 = y \geq 0, \\
    dX_t &= (\lambda - \lambda X_t) \, dt + Y_t^{1/\alpha} \, dZ_t, \quad t \geq 0, \, X_0 = x \in \mathbb{R},
\end{align*}
\]  

where \( a \geq 0, b, \lambda \in \mathbb{R}, (B_t)_{t \geq 0}\) is a standard Brownian motion, and \((Z_t)_{t \geq 0}\) is an independently spectrally positive \( \alpha \)-stable process with \( \alpha \in (1, 2) \). In terms of [4, Theorem 2.1], the SDE (1.2) has a unique strong solution \((Y_t, X_t)_{t \geq 0}\). Let \( P(t, (y, x), \cdot) \) be the transition probability kernel of the process \((Y_t, X_t)_{t \geq 0}\) with the initial value \((y, x)\).

For a strictly increasing function \( \psi \) on \( \mathbb{R}^2_+ \) and two probability measures \( \mu_1 \) and \( \mu_2 \) on \( \mathbb{R}_+ \times \mathbb{R} \), define

\[
W_\psi(\mu_1, \mu_2) = \inf_{\Pi \in \mathcal{E}(\mu_1, \mu_2)} \int_{\mathbb{R}^2_+ \times \mathbb{R}^2_+} \psi(y - y, x - x) \, \Pi(dy, dy, dx, dx),
\]
where \( \mathcal{C}(\mu_1, \mu_2) \) is the collection of all probability measures on \( \mathbb{R}^2_+ \times \mathbb{R}^2 \) with marginals \( \mu_1 \) and \( \mu_2 \). When \( \psi \) is concave, the above definition yields a Wasserstein distance \( W_\psi \) in the space of probability measures \( \mu \) on \( \mathbb{R}_+ \times \mathbb{R} \) such that \( \int_{\mathbb{R}_+ \times \mathbb{R}} \psi(y, |x|) \mu(dy, dx) < \infty \). If \( \psi(s, t) = (s^2 + t^2)^{1/2} \) for all \( s, t \geq 0 \), then \( W_\psi \) is the standard \( L^1 \)-Wasserstein distance, which will be denoted simply by \( W_1 \). Another well-known example for \( W_\psi \) is given by \( \psi(s, t) = 1_{\mathbb{R}^2_+ \setminus \{(0,0)\}}(s, t) \), which leads to the total variation distance

\[
W_\psi(\mu_1, \mu_2) = \frac{1}{2} \|\mu_1 - \mu_2\|_{\text{Var}} \coloneqq \frac{1}{2} [(\mu_1 - \mu_2)^+(\mathbb{R}_+ \times \mathbb{R}) + (\mu_1 - \mu_2)^-(\mathbb{R}_+ \times \mathbb{R})],
\]

where \( (\mu_1 - \mu_2)^+ \) and \( (\mu_1 - \mu_2)^- \) stand respectively for the positive part and the negative part for the Jordan–Hahn decomposition of the signed measure \( \mu_1 - \mu_2 \).

**Theorem 1.1.** Let \( (Y_t, X_t)_{t \geq 0} \) be the unique strong solution to Equation \( (1.2) \). If \( b > 0 \) and \( \lambda > 0 \), then the process \( (Y_t, X_t)_{t \geq 0} \) is exponentially ergodic with respect to the \( L^1 \)-Wasserstein distance \( W_1 \), that is, there exist a probability measure \( \mu \) on \( \mathbb{R}_+ \times \mathbb{R} \) and a constant \( \eta_0 > 0 \) such that for all \( y \geq 0, x \in \mathbb{R} \) and \( t > 0 \),

\[
W_1(P(t, (y, x), \cdot), \mu) \leq C_0(y, x)e^{-\eta_0 t},
\]

where \( C_0(y, x) > 0 \) is independent of \( t \).

We emphasize that, to investigate ergodicity of the two-factor affine processes, the Foster–Lyapunov criteria [35] developed by Meyn–Tweedie was adopted in [4, 5, 21], where the key ingredient is to examine the irreducibility of skeleton chains. Whereas, in general, it is a cumbersome task to check the irreducible property. Instead of the approach above, in this work we shall take advantage of the probabilistic coupling argument. Compared with [4, 5, 21], the coupling method enjoys several advantages. For instance, it avoids verification of the irreducible property of skeleton chains; on the other hand, it is universal in a certain sense that all the frameworks in [4, 5, 21] can be handled.

The approach of our paper is based on the probabilistic coupling, which is a well-known powerful tool in the study of Markov processes. Nowadays, there are considerable developments on discontinuous Lévy processes and related topics. For the initial investigations and systematic expositions on coupling of Lévy processes, we refer to [7, 37, 38] and [8, Chap. 6.2], respectively. Recently the coupling technique has been applied to study the exponential ergodicity under both the \( L^1 \)-Wasserstein distance and the total variation norm in [31, 32] for SDEs driven by Lévy noises, and in [29] for general continuous-state nonlinear branching processes, which in particular include continuous-state branching processes (i.e., typical class of affine processes on \( \mathbb{R}_+ \)). In contrast to these quoted papers, there are essential differences to realize the coupling approach for two-factor processes we are concerned with in the present paper. Roughly speaking, we do not consider directly the coupling of two-factor processes. Whereas we first couple the first component. For the two-factor process solved by Equation \( (1.2) \), the marginals of the coupling process for the first component will stay together once they meet at the first time. Also due to the structure of Equation \( (1.2) \), we then deal with the coupling concerning the second component as well. From the point of view above, the constructions of proper coupling processes and appropriate Lyapunov functions as well as their refined estimates, which are crucial to adopt the coupling approach, require much more effort than that in [29, 31, 32]. See Remarks 3.3 and 3.4 for more details. We emphasize that a two-factor process is a typical (and simple) example of Lévy-type processes with degenerate coefficients and that can describe interactions among the stochastic system. For example, the coefficient of jump noise of the \( x \)-component for the SDE \( (1.2) \) equals to \( y \), which can be zero and is determined by the \( y \)-component. It seems that our paper is the first one to construct a non-trivial coupling to such Lévy-type processes with two characteristics above.

Recently, there are a few of developments on the topics related with ergodicity of affine processes. For example, see [22] for the existence of limit distributions for affine processes, [17] for exponential ergodicity in Wasserstein distances for affine processes, [23] concerning positive Harris recurrence and exponential ergodicity for basic affine jump-diffusions, and [16, 34] and the references therein for exponential/geometric ergodicity of affine processes on cones, and so on. In particular, the exponential ergodicity of affine processes in terms of suitably chosen Wasserstein distances has been established in [17, Theorem 1.5] under the first moment condition on the state-dependent and log-moment conditions on the state-independent jump measures, respectively. Similar ideas have been used in [15] to study the exponential ergodicity for SDEs of nonnegative processes with jumps. Here, are two main differences between the approach of [17, Theorem 1.5] and the counterpart in our paper.
(i) Applying [17, Theorem 1.5(a)] to the setting of Theorem 1.1, one may get the exponential ergodicity of the two-factor process defined by (1.2) in terms of the $L^1$-Wasserstein distance $W_1$. The intermediate key step to yield [17, Theorem 1.5(a)] is [17, Proposition 7.3], which claims that there exist constants $K$ and $\delta > 0$ such that for all $y, \tilde{y} \in \mathbb{R}_+, x, \tilde{x} \in \mathbb{R}$ and $t > 0$,

$$W_{\psi^*}(P(t, (y, x), \cdot), P(t, (\tilde{y}, \tilde{x}), \cdot)) \leq Ke^{-\delta t}\psi^*(|y - \tilde{y}|, |x - \tilde{x}|),$$

where

$$\psi^*(u, v) := u + u^{1/2} + v, \quad u, v \geq 0;$$

see [17, (7.2)]. In this paper, we indeed can verify that for any $\delta \in (0, 1)$, there are positive constants $\eta := \eta(\delta)$ and $C := C(\delta)$ such that for all $y, \tilde{y} \in \mathbb{R}_+, x, \tilde{x} \in \mathbb{R}$ and $t > 0$,

$$W_{\psi^*}(P(t, (y, x), \cdot), P(t, (\tilde{y}, \tilde{x}), \cdot)) \leq Ce^{-\eta t}\psi(\eta(\delta), |y - \tilde{y}|, |x - \tilde{x}|),$$

where

$$\psi(\theta, u, v) := u + u^\theta + v, \quad u, v \geq 0;$$

see Equation (3.16).

(ii) The proof of [17] is based on the characterization of affine processes; see, for example, the proof of the crucial statement [17, Proposition 6.1]. However, our approach does not rely heavily on the structure of two-factor processes. Indeed, our argument still works for some models beyond two-factor processes; see Section 4.3 for further details.

The remainder of this paper is arranged as follows. In Section 2, we overview the existing result on existence and uniqueness of non-negative solutions to Equation (1.2), reveal that the first order moment of solutions to Equation (1.2) is finite, and construct the coupling operator by applying the coupling by reflection for a small distance and the synchronous coupling for a big distance to the first component and the synchronous coupling to the second component in Equation (1.2). Section 3 is devoted to the proof of Theorem 1.1 via the coupling approach and by constructing appropriate Lyapunov functions. In Section 4, we aim to apply the ideas adopted in Sections 2 and 3 to general two-factor affine models, (which allow that the first component solves a general CIR process, that there are interactions in the second component, and even that the Brownian noises are correlated), as well as some models beyond two-factor models.

## 2 COUPLING FOR TWO-FACTOR PROCESSES

### 2.1 Preliminary: Existence and uniqueness of strong solutions

In this part, we recall some known results on the existence and uniqueness of strong solutions to the SDE (1.2), and we also investigate moment estimates for the corresponding solution.

The existence and uniqueness of non-explosive strong solutions to the SDE (1.2) follows essentially from [4, Theorem 2.1]. Roughly speaking, the pathwise uniqueness of the non-negative strong solution $(Y_t)_{t \geq 0}$ of the first equation in Equation (1.2) is due to the well-known Yamada–Watanabe approximation approach. Once the first component $(Y_t)_{t \geq 0}$ is available, the second component $(X_t)_{t \geq 0}$ solved by the second equation in Equation (1.2) is indeed a one-dimensional Ornstein–Uhlenbeck-type process.

Throughout this paper, a spectrally positive $\alpha$-stable process $(L_t)_{t \geq 0}$ with $\alpha \in (1, 2)$ is a real-valued stochastic process with càdlàg paths, independent and stationary increments, and non-negative jumps; moreover, its Laplace exponent is given by

$$\log \mathbb{E}_0 e^{-\lambda L_1} = \int_0^\infty (e^{-\lambda z} - 1 + \lambda z) \nu_\alpha(dz), \quad \lambda > 0,$$
where \( \nu_\alpha(dz) := C_\alpha z^{-\alpha} \mathbb{1}_{\{z > 0\}} dz \) with \( C_\alpha = (\alpha \Gamma(-\alpha))^{-1} \). Then, according to the Itô formula, the infinitesimal generator of the process \((Y_t, X_t)_{t \geq 0}\) associated with the SDE (1.2) is given by

\[
(Lf)(y, x) = (a - bx) \partial_1 f(y, x) + \frac{y}{2} \partial_{11} f(y, x) + (\kappa - \lambda x) \partial_2 f(y, x)
+ y \int \mathbb{R} \left( f(y, x + z) - f(y, x) - \partial_2 f(y, x) z \right) \nu_\alpha(dz)
\]

(2.1)

for any \( f \in C^2_b(\mathbb{R}_+ \times \mathbb{R}) \). Here and below, \( \partial_i f \) stands for the first-order derivative w.r.t. the \( i \)th component, and \( \partial_{ij} f \) means the second-order derivative w.r.t. the \( i \)th component followed by the \( j \)th component. In particular, the operator \( L \) above is a special case of infinitesimal generators for affine processes given in [12, Equation (2.12) in Theorem 2.7] with \( m = n = 1 \) and the admissible parameters \( (a, \alpha, b, \beta, c, \gamma, m, \mu) \), where

\[
a = \begin{pmatrix} 0 & 0 \\ 0 & 1/2 \end{pmatrix}, \quad b = (x, a), \quad \beta = (\lambda, -b), \quad \alpha = c = \gamma = m = 0, \quad \mu = \nu_\alpha.
\]

Now, we take \( W(y, x) = 1 + y + h(x) \), where \( 0 \leq h \in C^2(\mathbb{R}) \) such that \( h(x) = |x| \) for all \( |x| \geq 2 \) and \( \|h'\|_\infty + \|h''\|_\infty < \infty \). Then, according to Equation (2.1), it follows that

\[
(LW)(y, x) = (a - bx) (x - \lambda x) h'(x) + y \int \mathbb{R} \left( h(x + z) - h(x) - h'(x) z \right) \nu_\alpha(dz)
\leq a + b|y| + \|h'\|_\infty (|x| + |\lambda| \cdot |x|) + \frac{1}{2} y \|h''\|_\infty \int_{|z|<1} z^2 \nu_\alpha(dz) + 2 y \|h''\|_\infty \int_{|z|\geq1} |z| \nu_\alpha(dz)
\leq C_0 (1 + y + h(x)) = C_0 W(y, x),
\]

where \( C_0 > 0 \) is independent of \( y \) and \( x \). Hence, for any \( (y, x) \in \mathbb{R}_+ \times \mathbb{R} \) and \( t > 0 \),

\[
\mathbb{E}(Y_t, X_t) W(Y_t, X_t) \leq W(y, x) + C_0 \int_0^t \mathbb{E}(Y_s, X_s) W(Y_s, X_s) \, ds.
\]

This, along with Gronwall’s inequality, yields that for any \( (y, x) \in \mathbb{R}_+ \times \mathbb{R} \) and \( t > 0 \),

\[
\mathbb{E}(Y_t, X_t) W(Y_t, X_t) \leq W(y, x) e^{C_0 t}.
\]

In particular, the first-order moment of \((Y_t, X_t)_{t \geq 0}\) is finite, that is, for any \( (y, x) \in \mathbb{R}_+ \times \mathbb{R} \) and \( t > 0 \),

\[
\mathbb{E}(Y_t, X_t) (Y_t + |X_t|) < \infty.
\]

(2.2)

### 2.2 | Markovian coupling for two-factor affine processes

Now, we consider the following SDE on \( \mathbb{R}_+ \times \mathbb{R} \):

\[
\begin{aligned}
\begin{cases}
\frac{d\tilde{Y}_t}{dt} = \begin{cases} (a - b \tilde{Y}_t) dt + \left( - \mathbb{1}_{\{0 < |Y_t - \tilde{Y}_t| < 1\}} + \mathbb{1}_{\{|Y_t - \tilde{Y}_t| \geq 1\}} \right) \tilde{Y}_t^{1/2} dB_t, & t < T_Y, \\
(a - b \tilde{Y}_t) dt + \tilde{Y}_t^{1/2} dB_t, & t \geq T_Y,
\end{cases}
\end{cases}
\end{aligned}
\]

\[
\begin{aligned}
\begin{cases}
\frac{d\tilde{X}_t}{dt} = (\kappa - \lambda \tilde{X}_t) dt + \tilde{X}_t^{1/\alpha} dL_t,
\end{cases}
\end{aligned}
\]

(2.3)
with the initial value \((\bar{Y}_0, \bar{X}_0) = (\bar{y}, \bar{x}) \in \mathbb{R}_+ \times \mathbb{R}\), where \(T_Y := \inf\{t \geq 0 : Y_t = \bar{Y}_t\}\). Define
\[
B^*_t = \begin{cases} 
B_0 + \int_0^t (-1_{\{0 < |Y_s - \bar{Y}_s| < 1\}} + 1_{\{|Y_s - \bar{Y}_s| \geq 1\}}) \, dB_s, & 0 \leq t \leq T_Y, \\
B^*_t + B_t - B_{T_Y}, & t \geq T_Y.
\end{cases}
\]

Then, we can write the first component process \((Y_t)_{t \geq 0}\) in Equation (2.3) as follows:
\[
d\bar{Y}_t = (a - b \bar{Y}_t) \, dt + \bar{Y}_t^{1/2} \, dB^*_t, \quad t > 0.
\]

Since \((B^*_t)_{t \geq 0}\) is still a standard Brownian motion and the SDE (1.2) has a unique strong (and so a unique weak) solution, the SDE (2.3) also admits a unique weak solution \((\bar{Y}_t, \bar{X}_t)_{t \geq 0}\) so that \(Y_t = \bar{Y}_t\) for all \(t \geq T_Y\) (thanks to the fact that after \(t \geq T_Y\) the process \((Y_t)_{t \geq 0}\) and \((\bar{Y}_t)_{t \geq 0}\) are driven by the same Brownian motion \((B_t)_{t \geq 0}\); moreover, the process \((\bar{Y}_t, \bar{X}_t)_{t \geq 0}\) enjoys the same law (i.e., the same transition probabilities) as that of \((Y_t, X_t)_{t \geq 0}\). In particular, \(((Y_t, X_t), (\bar{Y}_t, \bar{X}_t))_{t \geq 0}\) is a non-explosive coupling of the process \((Y_t, X_t)_{t \geq 0}\). Roughly speaking, before two marginal processes meet we will apply the coupling by reflection for the first component \((Y_t)_{t \geq 0}\) when the distance of them is less than 1 and the synchronous coupling when the distance of them is large or equal to 1, and once two marginal processes meet we will adopt the synchronous coupling; while we always take the synchronous coupling for the second component \((X_t)_{t \geq 0}\).

Furthermore, it is not hard to see that the generator (later we call it the coupling operator of the generator \(L\) given by Equation (2.1)) of the coupling process \(((Y_t, X_t), (\bar{Y}_t, \bar{X}_t))_{t \geq 0}\) is given by
\[
(L^* f)(y, \bar{y}, x, \bar{x}) = (a - b y) \partial_1 f(y, \bar{y}, x, \bar{x}) + \frac{y}{2} \partial_1 f(y, \bar{y}, x, \bar{x}) + (a - b \bar{y}) \partial_2 f(y, \bar{y}, x, \bar{x}) + \frac{\bar{y}}{2} \partial_2 f(y, \bar{y}, x, \bar{x})
\]
\[
- \sqrt{y\bar{y}} \partial_{12} f(y, \bar{y}, x, \bar{x}) 1_{\{0 < y - \bar{y} < 1\}} + \sqrt{y\bar{y}} \partial_{12} f(y, \bar{y}, x, \bar{x}) 1_{\{|y - \bar{y}| \geq 1\} \cup \{|y| = |\bar{y}|\}}
\]
\[
+ (x - \lambda x) \partial_3 f(y, \bar{y}, x, \bar{x}) + (x - \bar{\lambda} x) \partial_3 f(y, \bar{y}, x, \bar{x})
\]
\[
+ y \int_0^\infty (f(y, \bar{y}, x + z, \bar{x} + z) - f(y, \bar{y}, x, \bar{x}) - \partial_3 f(y, \bar{y}, x, \bar{x}) z - \partial_4 f(y, \bar{y}, x, \bar{x}) z) \, d\nu_{\alpha}(dz)
\]
\[
+ (y - \bar{y}) \int_0^\infty (f(y, \bar{y}, x + z, \bar{x}) - f(y, \bar{y}, x, \bar{x}) - \partial_3 f(y, \bar{y}, x, \bar{x}) z) \, d\nu_{\alpha}(dz)
\]
for any \(f \in C^2(\mathbb{R}_+^2 \times \mathbb{R}^2)\) and \((y, \bar{y}, x, \bar{x}) \in \mathbb{R}_+^2 \times \mathbb{R}^2\) with \(y \geq \bar{y}\). Similarly, we can write the expression of \((L^* f)(y, \bar{y}, x, \bar{x})\) for \(f \in C^2(\mathbb{R}_+^2 \times \mathbb{R}^2)\) and \((y, \bar{y}, x, \bar{x}) \in \mathbb{R}_+^2 \times \mathbb{R}^2\) with \(y \leq \bar{y}\).

## 3 Exponential Convergence in the Wasserstein Distance

This section is devoted to the proof of Theorem 1.1. Throughout this section, we shall fix \(\alpha \in (1, 2)\), which is the stability index for the spectrally positive \(\alpha\)-stable process \((Z_t)_{t \geq 0}\) in the second equation of Equation (1.2). Below, let \(\nu_{\alpha}(dz)\) be the associated Lévy measure of \((Z_t)_{t \geq 0}\).

We begin with the following simple lemma.

**Lemma 3.1.** Let \(g \in C^2_0(\mathbb{R}_+)\) such that \(0 \leq g \leq 1\). Then, for the function
\[
F(s, t) := (1 - g(t/s)) s + g(t/s) t, \quad s, t > 0,
\]
\[
\text{it holds that}
\]
\[
|\partial_i F(s, t)| \leq c_0, \quad |\partial_i^2 F(s, t)| \leq c_0 s^{-1}, \quad 1 \leq t/s \leq 2, \quad i = 1, 2,
\]
where \( c_0 > 0 \) is independent of \( s, t \), and \( \partial F \) (resp. \( \partial_i F \)) stands for the first (resp. second)-order derivative w.r.t. the \( i \)-th component of the function \( F \).

**Proof.** A straightforward calculation shows that for any \( s, t > 0 \),
\[
\partial_1 F(s, t) = g'(t/s)ts^{-1} + (1 - g(t/s)) - g'(t/s)t^2s^{-2},
\]
\[
\partial_{11} F(s, t) = -2g''(t/s)ts^{-3} + 2g'(t/s)t^2s^{-3} + g''(t/s)t^3s^{-4},
\]
\[
\partial_2 F(s, t) = -g'(t/s) + g'(t/s)ts^{-1} + g(t/s),
\]
\[
\partial_{22} F(s, t) = -g''(t/s)ts^{-1} + g''(t/s)ts^{-2} + 2g'(t/s)s^{-1}.
\]

Then, Equation (3.2) follows by \( g \in C^2_b(\mathbb{R}^+) \).

Next, we will take \( g \in C^2_b(\mathbb{R}^+) \) with \( g' \geq 0 \) such that
\[
g(r) = \begin{cases} 
0, & 0 \leq r < 1, \\
(r - 1)^{2+\delta}, & 1 < r < 3/2, \\
1, & r \geq 2 
\end{cases}
\] (3.3)

for some constant \( \delta > 0 \). With this choice, \( F(s, t) = s \) if \( t \leq s \); \( F(s, t) = t \) if \( t \geq 2s \).

Now, for any \( c > 0 \) and \( \theta \in (0,1) \) we define the function
\[
V_{c,\theta}(s, t) = c(s + s^{\theta}) + F(s, t), \ s, t \geq 0.
\] (3.4)

It is clear from \( g \in [0,1] \) that for any \( c > 0 \),
\[
V_{c,\theta}(s, t) \asymp (s \lor s^{\theta}) + t, \ s, t \geq 0.
\] (3.5)

Herein, we use the shorthand notation \( f \asymp g \) for two non-negative functions \( f \) and \( g \), which means that there exists a constant \( c \geq 1 \) such that \( c^{-1}f \leq g \leq cf \) on the domain.

Below, concerning \( V_{c,\theta}^*(y, \tilde{y}, x, \tilde{x}) := V_{c,\theta}(y - \tilde{y}, |x - \tilde{x}|) \) for any \((y, \tilde{y}, x, \tilde{x}) \in \mathbb{R}^2_+ \times \mathbb{R}^2 \) with \( y \geq \tilde{y} \), we will simply write \( (L^* V_{c,\theta})(y, \tilde{y}, x, \tilde{x}) \) := \( (L^* V_{c,\theta}^*)(y, \tilde{y}, x, \tilde{x}) \). We now have the following statement, which is crucial for the proof of Theorem 1.1.

**Proposition 3.2.** For any \( \theta \in (0,1) \), there exist constants \( c, \zeta > 0 \) such that for any \( y > \tilde{y} \geq 0 \) and \( x, \tilde{x} \in \mathbb{R} \),
\[
(L^* V_{c,\theta})(y - \tilde{y}, |x - \tilde{x}|) \leq -\zeta V_{c,\theta}(y - \tilde{y}, |x - \tilde{x}|).
\] (3.6)

**Proof.** (1) Let \( G(|y - \tilde{y}|, |x - \tilde{x}|) \in C^2((\mathbb{R}^2_+ \times \mathbb{R}^2) \setminus \{(y, \tilde{y}, x, \tilde{x}) : y = \tilde{y}\}) \). According to the definition of the coupling operator \( L^* \), we find that for any \( y \geq \tilde{y} \geq 0 \) and \( x, \tilde{x} \in \mathbb{R} \),
\[
(L^* G)(y - \tilde{y}, |x - \tilde{x}|) = -b \delta_1 G(y - \tilde{y}, |x - \tilde{x}|)(y - \tilde{y}) + \frac{1}{2} \left( \sqrt{y} + \sqrt{\tilde{y}} \right)^2 \delta_{11} G(y - \tilde{y}, |x - \tilde{x}|) 1_{\{0 < y - \tilde{y} < 1\}}
\]
\[
+ \frac{1}{2} \left( \sqrt{y} - \sqrt{\tilde{y}} \right)^2 \delta_{11} G(y - \tilde{y}, |x - \tilde{x}|) 1_{\{y - \tilde{y} \geq 1\}} - \lambda \delta_2 G(y - \tilde{y}, |x - \tilde{x}|)|x - \tilde{x}|
\]
\[
+ (y - \tilde{y}) \int_0^\infty (G(y - \tilde{y}, |x + z - \tilde{x}|) - G(y - \tilde{y}, |x - \tilde{x}|))
\]
\[
- \delta_2 G(y - \tilde{y}, |x - \tilde{x}|) \frac{(x - \tilde{x})z}{|x - \tilde{x}|}) v_2(dz),
\] (3.7)
where $\partial_i G$ (resp. $\partial_{i1} G$) stands for the first (resp. second) order derivative w.r.t. the $i$th component of the function $G$.

Below, let $y > \bar{y} \geq 0$ and $x, \bar{x} \in \mathbb{R}$ be arbitrary. For any $c > 0$ and $\theta \in (0, 1)$, let

$$U_{c, \theta}(s, t) = c(s + s^\theta), \quad s, t \geq 0.$$ 

Then, we get from Equation (3.7) and $\theta \in (0, 1)$ that

$$(L^* U_{c, \theta})(y - \bar{y}, |x - \bar{x}|) = -c b(1 + \theta(y - \bar{y})^{\theta - 1})(y - \bar{y}) - \frac{1}{2} c \theta(1 - \theta) \left(\sqrt{y} + \sqrt{\bar{y}}\right)^2 (y - \bar{y})^{\theta - 2} 1_{\{0 < y - \bar{y} < 1\}}$$

$$- \frac{1}{2} c \theta(1 - \theta) \left(\sqrt{y} - \sqrt{\bar{y}}\right)^2 (y - \bar{y})^{\theta - 2} 1_{\{y - \bar{y} \geq 1\}}$$

$$\leq -c b(y - \bar{y}) - c b \theta(y - \bar{y})^\theta - \frac{1}{2} c \theta(1 - \theta) \left(\sqrt{y} + \sqrt{\bar{y}}\right)^2 (y - \bar{y})^{\theta - 2} 1_{\{0 < y - \bar{y} < 1\}}.$$ 

Thanks to Equation (3.4), this yields Equation (3.6) provided that there is a constant $C > 0$ such that

$$(L^* F)(y - \bar{y}, |x - \bar{x}|) \leq -\lambda |x - \bar{x}| + 2(\lambda + C)(y - \bar{y}) + C \left(\sqrt{y} + \sqrt{\bar{y}}\right)^2 (y - \bar{y})^{-1} 1_{\{0 < y - \bar{y} < 1\}}$$

(3.8)

and by taking

$$c = \frac{4(\lambda + C)}{b} + \frac{4C}{\theta(1 - \theta)}.$$ 

Whereas, Equation (3.8) is available as long as there exists a constant $C > 0$ such that

$$(L^* F)(y - \bar{y}, |x - \bar{x}|) \leq -\lambda |x - \bar{x}| 1_{\{|x - \bar{x}| \geq 2(y - \bar{y})\}} + C(y - \bar{y}) + \frac{C}{2} (y - \bar{y})^{2-\alpha}$$

$$+ \frac{C}{2} \left(\sqrt{y} + \sqrt{\bar{y}}\right)^2 (y - \bar{y})^{-1} 1_{\{0 < y - \bar{y} < 1\}} + \frac{C}{2} \left(\sqrt{y} - \sqrt{\bar{y}}\right)^2 (y - \bar{y})^{-1} 1_{\{y - \bar{y} \geq 1\}}.$$ 

(3.9)

Indeed, Equation (3.8) holds true since Equation (3.9) implies

$$(L^* F)(y - \bar{y}, |x - \bar{x}|) \leq -\lambda |x - \bar{x}| + 2(\lambda + C)(y - \bar{y}) + C \left(\sqrt{y} + \sqrt{\bar{y}}\right)^2 (y - \bar{y})^{-1} 1_{\{0 < y - \bar{y} < 1\}},$$

where we used the facts that

$$(y - \bar{y})^{2-\alpha} \leq (y - \bar{y})_{\{y - \bar{y} \geq 1\}} + \left(\sqrt{y} + \sqrt{\bar{y}}\right)^2 (y - \bar{y})^{-1} 1_{\{0 < y - \bar{y} < 1\}}$$

(3.10)

and, for $y - \bar{y} \geq 1$,

$$(\sqrt{y} - \sqrt{\bar{y}})^2 (y - \bar{y})^{-1} = (\sqrt{y} - \sqrt{\bar{y}})/(\sqrt{y} + \sqrt{\bar{y}}) \leq 1 \leq y - \bar{y}.$$ 

So, to achieve the desired assertion (3.6), it is sufficient to show Equation (3.9) for the following three cases:

(i) $|x - \bar{x}| \geq 2(y - \bar{y})$;
(ii) $|x - \bar{x}| \leq y - \bar{y}$;
(iii) $y - \bar{y} \leq |x - \bar{x}| \leq 2(y - \bar{y})$.

**Proof of Equation (3.9) for the case (i).** In this case, $F(y - \bar{y}, |x - \bar{x}|) = |x - \bar{x}|$ so that, by Equation (3.7), one has

$$(L^* F)(y - \bar{y}, |x - \bar{x}|) = -\lambda |x - \bar{x}| + (y - \bar{y})(I_1 + I_2 + I_3),$$
where

\[
I_1 := \int_{\{x + z - \bar{x} < y - \bar{y}\}} \left( y - \bar{y} - |x - \bar{x}| - \frac{(x - \bar{x})z}{|x - \bar{x}|} \right) \nu_\alpha(dz),
\]

\[
I_2 := \int_{\{x + z - \bar{x} > 2(y - \bar{y})\}} \left( |x + z - \bar{x}| - |x - \bar{x}| - \frac{(x - \bar{x})z}{|x - \bar{x}|} \right) \nu_\alpha(dz),
\]

\[
I_3 := \int_{\{y - \bar{y} \leq x + z - \bar{x} \leq 2(y - \bar{y})\}} \left[ \left( 1 - g \left( \frac{|x + z - \bar{x}|}{y - \bar{y}} \right) \right) (y - \bar{y}) + g \left( \frac{|x + z - \bar{x}|}{y - \bar{y}} \right) \right] \nu_\alpha(dz).
\]

(3.11)

Note that, if \(|x + z - \bar{x}| < y - \bar{y}\) and \(|x - \bar{x}| \geq 2(y - \bar{y})\), then

\[
z \geq |x - \bar{x}| - |x + z - \bar{x}| > |x - \bar{x}| - (y - \bar{y}) > y - \bar{y}.
\]

Whence, there exists a constant \(c_1 > 0\) such that

\[
I_1 \leq \int_{\{z > y - \bar{y}\}} (y - \bar{y} + z) \nu_\alpha(dz) \leq c_1(y - \bar{y})^{1-\alpha}.
\]

(3.12)

A simple calculation shows that

\[
I_2 = \int_{\{x + z - \bar{x} > 2(y - \bar{y})\}} ((x + z - \bar{x}) - (x - \bar{x}) - z) 1_{\{x \geq \bar{x}\}} \nu_\alpha(dz)
+ \int_{\{x + z - \bar{x} > 2(y - \bar{y})\}} (-z + (x + z - \bar{x}) + (x - \bar{x}) + z) 1_{\{x < \bar{x}, x - \bar{x} + z \leq 0\}} \nu_\alpha(dz)
+ 2 \int_{\{z > x - \bar{x} + 2(y - \bar{y})\}} (x - \bar{x} + z) 1_{\{x < \bar{x}, x - \bar{x} + z > 0\}} \nu_\alpha(dz)
\leq 2 \int_{\{z \geq 3(y - \bar{y})\}} z 1_{\{x < \bar{x}, x - \bar{x} + z \geq 0\}} \nu_\alpha(dz) \leq c_2(y - \bar{y})^{1-\alpha}
\]

(3.13)

for some constant \(c_2 > 0\), where in the inequality we used \(x < \bar{x}\) and \(\bar{x} - x \geq 2(y - \bar{y})\).

Observe that

\[
I_3 = \int_{\{y - \bar{y} \leq x + z - \bar{x} \leq 2(y - \bar{y})\}} \left( 1 - g \left( \frac{|x + z - \bar{x}|}{y - \bar{y}} \right) \right) (y - \bar{y} - |x + z - \bar{x}|) \nu_\alpha(dz)
+ \int_{\{y - \bar{y} \leq x + z - \bar{x} \leq 2(y - \bar{y})\}} \left( |x + z - \bar{x}| - |x - \bar{x}| - \frac{(x - \bar{x})z}{|x - \bar{x}|} \right) \nu_\alpha(dz)
\leq 2 \int_{\{y - \bar{y} \leq x + z - \bar{x} \leq 2(y - \bar{y})\}} (x - \bar{x} + z) 1_{\{x < \bar{x}, x + z - \bar{x} > 0\}} \nu_\alpha(dz)
\leq 2 \int_{\{z \geq 3(y - \bar{y})\}} z \nu_\alpha(dz) = c_3(y - \bar{y})^{1-\alpha}
\]

(3.14)

for some constant \(c_3 > 0\), where in the first inequality we used \(g \in [0, 1]\) and \(y - \bar{y} \leq |x + z - \bar{x}|\), and the second inequality follows from \(x < \bar{x}\) and \(\bar{x} - x \geq 2(y - \bar{y})\). So, combining Equation (3.12) with Equations (3.13), (3.14) and (3.10) yields Equation (3.9) for the case \((i)\).
**Proof of Equation (3.9) for the case (ii).** In this case, \( F(y - \bar{y}, |x - \bar{x}|) = y - \bar{y} \). Then, according to Equation (3.7), we have

\[
(L^*F)(y - \bar{y}, |x - \bar{x}|) = -b(y - \bar{y}) + (y - \bar{y}) \left\{ \int_{|x-\bar{x}+z|>2(y-\bar{y})} (|x-\bar{x}+z|- (y-\bar{y})) \nu_\alpha(dz) 
\right.
\]

\[
+ \int_{|y-\bar{y}|<|x-\bar{x}+z|<2(y-\bar{y})} g \left( \frac{|x-\bar{x}+z|}{y-\bar{y}} \right) (|x-\bar{x}+z|- (y-\bar{y})) \nu_\alpha(dz) \right\} =: -b(y - \bar{y}) + (y - \bar{y})(J_1 + J_2).
\]

It is easy to get that

\[
J_1 \leq \int_{|x+z| > 2(y-\bar{y})} (|x| + z) \nu_\alpha(dz) \leq 2 \int_{|z-y|} z \nu_\alpha(dz) = c_1(y - \bar{y})^{1-\alpha}
\]

for some \( c_1 > 0 \), where in the second inequality we used \( y - \bar{y} > |x-\bar{x}| \) and

\[
z > |x-\bar{x} + z| - |x-\bar{x}| > 2(y-\bar{y}) - |x-\bar{x}| \geq y - \bar{y}.
\]

According to the definition of the function \( g(\cdot) \) given by Equation (3.3), without loss of generality, we can assume that there is a constant \( c_\ast > 0 \) such that \( g(r) \leq c_\ast(r-1)^2 + \delta \) for \( r \geq 1 \). Then, it holds

\[
J_2 \leq \frac{c_\ast}{(y - \bar{y})^{2+\delta}} \int_{|y| < |x-\bar{x}+z| < 2(y-\bar{y})} (|x-\bar{x}+z|- (y-\bar{y}))^{3+\delta} \nu_\alpha(dz)
\]

\[
\leq \frac{c_\ast}{(y - \bar{y})^{2+\delta}} \int_{|y| < |x-\bar{x}+z| < 2(y-\bar{y})} (|x-\bar{x}+z| - (y-\bar{y}))^{3+\delta} \nu_\alpha(dz)
\]

\[
= \frac{c_\ast}{(y - \bar{y})^{2+\delta}} \int_{|z| < 3(y-\bar{y})} z^{3+\delta} \nu_\alpha(dz) = c_2(y - \bar{y})^{1-\alpha}
\]

for some constant \( c_2 > 0 \), where the last inequality follows from the facts that \( |x-\bar{x}| < y-\bar{y} \) and \( z < |x-\bar{x} + z| + |x-\bar{x}| \leq 3(y-\bar{y}) \). Hence, combining the estimates above and taking Equation (3.10) into account gives Equation (3.9) for the case (ii).

**Proof of Equation (3.9) for the case (iii).** With regard to this case, we derive from Equation (3.2) and (3.7) that there exists a constant \( c_1 > 0 \) so that

\[
(L^*F)(y - \bar{y}, |x - \bar{x}|) \leq c_1(y - \bar{y}) + (y - \bar{y})(\Lambda_1 + \Lambda_2) + \frac{c_0}{2} \left( \sqrt{y} + \sqrt{\bar{y}} \right)^2 (y - \bar{y})^{-1} 1_{\{0 < y - \bar{y} < 1\}}
\]

\[
+ \frac{c_0}{2} \left( \sqrt{y} - \sqrt{\bar{y}} \right)^2 (y - \bar{y})^{-1} 1_{\{y - \bar{y} \geq 1\}}.
\]

where

\[
\Lambda_1 := \int_{|z| < (y-\bar{y})/2} \left( F(y - \bar{y}, |x + z - \bar{x}|) - F(y - \bar{y}, |x - \bar{x}|) - \partial_2 F(y - \bar{y}, |x - \bar{x}|) \right) \frac{(x-\bar{x})_z}{|x-\bar{x}|} \nu_\alpha(dz)
\]

and

\[
\Lambda_2 := \int_{|z| > (y-\bar{y})/2} \left( F(y - \bar{y}, |x + z - \bar{x}|) - F(y - \bar{y}, |x - \bar{x}|) - \partial_2 F(y - \bar{y}, |x - \bar{x}|) \right) \frac{(x-\bar{x})_z}{|x-\bar{x}|} \nu_\alpha(dz).
\]
Note that \( x + z - \bar{x} < 0 \) in case of \( x < \bar{x}, z \leq (y - \bar{y})/2 \) and \( y - \bar{y} \leq |x - \bar{x}| \leq 2(y - \bar{y}) \). Then, we have

\[
\Lambda_1 = \int_{z \leq (y - \bar{y})/2} \left( F(y - \bar{y}, x - \bar{x} + z) - F(y - \bar{y}, x - \bar{x}) - \partial_2 F(y - \bar{y}, x - \bar{x})z \right) \mathbb{1}_{\{|x\geq \bar{x}|} \nu_\alpha(dz)
\]

\[
+ \int_{\{|z \leq (y - \bar{y})/2, x + z - \bar{x} < 0\}} \left( F(y - \bar{y}, \bar{x} - x - z) - F(y - \bar{y}, \bar{x} - x) + \partial_2 F(y - \bar{y}, \bar{x} - x)z \right) \mathbb{1}_{\{|x< \bar{x}|} \nu_\alpha(dz).
\]

Applying the mean value theorem and taking Equation (3.2) into account, we find that there is a constant \( c_2 > 0 \) so that

\[
\Lambda_1 = \frac{1}{2} \int_{\{|z \leq (y - \bar{y})/2\}} \partial_2 F(y - \bar{y}, \xi_1)z^2 \mathbb{1}_{\{|x\geq \bar{x}|} \nu_\alpha(dz) + \frac{1}{2} \int_{\{|z \leq (y - \bar{y})/2, x + z - \bar{x} < 0\}} \partial_2 F(y - \bar{y}, \xi_2)z^2 \mathbb{1}_{\{|x< \bar{x}|} \nu_\alpha(dz)
\]

\[
\leq c_2(y - \bar{y})^{1-\alpha},
\]

where \( \xi_1 \in [y - \bar{y}, 5(y - \bar{y})/2] \) and \( \xi_2 \in [(y - \bar{y})/2, 2(y - \bar{y})] \). On the other hand, it follows from Equation (3.2) that there exist constants \( c_3, c_4, c_5 > 0 \) such that

\[
\Lambda_2 \leq c_3 \int_{\{|z \geq (y - \bar{y})/2\}} (y - \bar{y} + |x - \bar{x} + z| + z) \nu_\alpha(dz) \leq c_4 \int_{\{|z \geq (y - \bar{y})/2\}} (y - \bar{y} + z) \nu_\alpha(dz) \leq c_5(y - \bar{y})^{1-\alpha},
\]

where in the second inequality we used the fact that \( y - \bar{y} \leq |x - \bar{x}| \leq 2(y - \bar{y}) \). Therefore, Equation (3.9) for the case (iii) follows by substituting the two estimates above into Equation (3.15) and taking advantage of Equation (3.10).

\[\square\]

Remark 3.3. The main task of the proof above is to control an upper bound for \((L^*F)(y - \bar{y}, |x - \bar{x}|)\) for all \( y > \bar{y} \geq 0 \) and \( x, \bar{x} \in \mathbb{R} \). In cases (i) and (ii), we can get that

\[
(L^*F)(y - \bar{y}, |x - \bar{x}|) \leq -\lambda |x - \bar{x}| \mathbb{1}_{\{|x - \bar{x}| \geq 2(y - \bar{y})\}} + C_0(y - \bar{y})^{2-\alpha}.
\]

To get Equation (3.6) for those two cases, one can take

\[
V_{c,c_\bar{x}}(s, t) = c \left( s + \int_0^s e^{-c_\lambda u + c_\lambda u} \frac{dF(s, t)}{ds, t} \right)
\]

instead of \( V_{c,\xi}(s, t) \) defined by Equation (3.4) (with possibly choices of \( c, c_\lambda > 0 \)), and apply the coupling by reflection for the first component \((Y_t)_{t \geq 0}\) before two marginal processes meet (i.e.,

\[
-\sqrt{2yy\xi_{12}f(y, \bar{y}, x, \bar{x})} \mathbb{1}_{\{0<\bar{y} - y < 1\}} + \sqrt{2yy\xi_{12}f(y, \bar{y}, x, \bar{x})} \mathbb{1}_{\{y \geq \bar{y}\}}
\]

is replaced by

\[
-\sqrt{2yy\xi_{12}f(y, \bar{y}, x, \bar{x})} \mathbb{1}_{\{y \neq \bar{y}\}} + \sqrt{2yy\xi_{12}f(y, \bar{y}, x, \bar{x})} \mathbb{1}_{\{y = \bar{y}\}}
\]

in the coupling operator \((L^*f)(y, \bar{y}, x, \bar{x})\) given by Equation (2.4)). Note that the nice property of the function \( V_{c,c_\bar{x}}(s, t) \) above is that

\[
V_{c,c_\bar{x}}(s, t) \approx s + t,
\]

which is comparable to the “cost function” in the standard \(L^1\)-Wasserstein distance. However, for the case (iii), to eliminate the last two positive terms in the right-hand side of Equation (3.15), we need not only to apply the test function \( V_{c,\xi}(s, t) \) defined by Equation (3.4), but also to modify the coupling of the Brownian motion \((B_t)_{t \geq 0}\) in the first component \((Y_t)_{t \geq 0}\).

Now, we are in a position to present the
**Proof of Theorem 1.1.** Note that, for all $\theta \in (0,1)$, $(y,\tilde{y}) \mapsto |y-\tilde{y}| + |y-\tilde{y}|^\theta$ is a metric on $\mathbb{R}_+$, and so $(\mu_1,\mu_2) \mapsto W_{\psi_\theta}(\mu_1,\mu_2)$ is a metric on the space of probability measures on $\mathbb{R}_+ \times \mathbb{R}$, where

$$\psi_\theta(u,v) := u + u^\theta + v, \quad u,v \geq 0.$$  

Note that $W_1(\mu_1,\mu_2) \leq W_{\psi_\theta}(\mu_1,\mu_2)$ for any $\theta \in (0,1)$. To prove Theorem 1.1, we actually verify the exponential ergodicity of the process $(Y_t,X_t)_{t \geq 0}$ in terms of $W_{\psi_\theta}$. Furthermore, as mentioned in Section 2.2, the coupling of the first component process preserves the order property, that is, $Y_t \geq \tilde{Y}_t$ for all $t \geq 0$ if $Y_0 = y > \tilde{Y}_0 = \tilde{y}$. By carrying out more or less standard arguments (see, e.g., the proof of [32, Corollary 1.8]), to achieve the desired assertion (1.3) on the exponential ergodicity of the process $(Y_t,X_t)_{t \geq 0}$ in terms of $W_{\psi_\theta}$, it is sufficient to show that for all $y \geq \tilde{y} \geq 0$, $x,\tilde{x} \in \mathbb{R}$ and $t > 0$,

$$W_{\psi_\theta}(\mathbb{P}(t,(y,x),\cdot),\mathbb{P}(t,(\tilde{y},\tilde{x}),\cdot)) \leq C e^{-\eta t} \psi_\theta(y-\tilde{y},|x-\tilde{x}|) \quad (3.16)$$

holds with some constants $C, \eta > 0$ (which are independent of $y,\tilde{y},x,\tilde{x}$ and $t$). This obviously holds true provided that

$$\mathbb{E}^{(y,x)} \psi_\theta(Y_t,|X_t|) < \infty \quad (3.17)$$

and that

$$\mathbb{E}^{(y,x),(\tilde{y},\tilde{x})} \psi_\theta(Y_t - \tilde{Y}_t,|X_t - \tilde{X}_t|) \leq C e^{-\eta t} \psi_\theta(y-\tilde{y},|x-\tilde{x}|). \quad (3.18)$$

Since $\psi_\theta(u,v) = u + u^\theta + v \leq 2(1 + u + v)$, Equation (3.17) is true due to Equation (2.2). Let $V_{c,\theta}(s,t)$ be the function and $\zeta$ the positive constant in Proposition 3.2. By using Equation (3.5), Equation (3.18) is valid once we claim that

$$\mathbb{E}^{(y,x),(\tilde{y},\tilde{x})} V_{c,\theta}(Y_t - \tilde{Y}_t,|X_t - \tilde{X}_t|) \leq e^{-(\lambda + \zeta) t} V_{c,\theta}(y-\tilde{y},|x-\tilde{x}|) \quad (3.19)$$

for all $y \geq \tilde{y} \geq 0, x,\tilde{x} \in \mathbb{R}$ and $t > 0$, where $\lambda > 0$ is the constant in the SDE (1.2). So, in what follows, it remains to show that the assertion (3.19) holds.

Define

$$T_Y = \inf\{t > 0 : Y_t = \tilde{Y}_t\}, \quad T_{Y,n} = \inf\{t > 0 : Y_t - \tilde{Y}_t \leq 1/n\}, \quad n \geq 1.$$  

Let $y \geq \tilde{y} \geq 0$. Note that for all $r \geq 0$,

$$V_{c,\theta}(0,r) = r = r \partial_2 V_{c,\theta}(0,r). \quad (3.20)$$

If $y = \tilde{y} \geq 0$, then $T_Y = 0$ so that for any $t \geq s \geq 0$,

$$e^{(\lambda + \zeta) t} V_{c,\theta}(Y_t - \tilde{Y}_t,|X_t - \tilde{X}_t|) = e^{(\lambda + \zeta) s} V_{c,\theta}(0,|X_t - \tilde{X}_t|)$$

$$= e^{(\lambda + \zeta) s} V_{c,\theta}(0,|X_s - \tilde{X}_s|)$$

$$+ \int_s^t e^{(\lambda + \zeta) r} \{\lambda \zeta V_{c,\theta}(0,|X_r - \tilde{X}_r|) - \lambda |X_r - \tilde{X}_r| \partial_2 V_{c,\theta}(0,|X_r - \tilde{X}_r|)\} dr$$

$$\leq e^{(\lambda + \zeta) s} V_{c,\theta}(0,|X_s - \tilde{X}_s|),$$

where the second identity holds true from the structure of the second equation in Equation (1.2), that is, for any fixed $s \geq 0$ with $Y_s = \tilde{Y}_s$, we have $Y_r = \tilde{Y}_r$ for all $r \geq s$ and so

$$d(X_r - \tilde{X}_r) = -\lambda (X_r - \tilde{X}_r) \, dr, \quad r \geq s, \quad (3.22)$$
and the inequality is owing to Equation (3.20). With Equation (3.21) for $s = 0$ at hand, it is easy to see that Equation (3.19) holds for $y = \bar{y} \geq 0$.

In the following, we only need to verify Equation (3.19) for $y > \bar{y} \geq 0$. Choose $n_0 \geq 1$ sufficiently large such that $y - \bar{y} > 1/n_0$. Note again that $Y_t \geq \bar{Y}_t$ for all $t \geq 0$ (in particular, for all $0 \leq t \leq T_{Y,n}$ and $n \geq 1$) whenever $Y_0 = y > \bar{Y}_0 = \bar{y}$. Then, for any $t > 0$ and $n \geq n_0$, by Itô’s formula, it follows that

$$
\mathbb{E}((y, x),(\bar{y}, \bar{x}))((e^{s \zeta(T_{Y,n})}V_{c,\theta}(Y_{t \wedge T_{Y,n}} - \bar{Y}_{t \wedge T_{Y,n}}, |X_{t \wedge T_{Y,n}} - \bar{X}_{t \wedge T_{Y,n}}|))
\leq V_{c,\theta}(y - \bar{y}, |x - \bar{x}|),
$$

where we utilized Equation (3.6) in the last inequality. Approaching $n \to \infty$ yields that for all $y > \bar{y} \geq 0$, $x, \bar{x} \in \mathbb{R}$ and $t > 0$,

$$
\mathbb{E}((y, x),(\bar{y}, \bar{x}))((e^{s \zeta(T_{Y})}V_{c,\theta}(Y_{t \wedge T_{Y}} - \bar{Y}_{t \wedge T_{Y}}, |X_{t \wedge T_{Y}} - \bar{X}_{t \wedge T_{Y}}|)) \leq V_{c,\theta}(y - \bar{y}, |x - \bar{x}|).
$$

Consequently, Equation (3.19) is available for $y > \bar{y} \geq 0$ by taking advantage of

$$
e^{\lambda \zeta(T_{Y})}V_{c,\theta}(Y_{t \wedge T_{Y}} - \bar{Y}_{t \wedge T_{Y}}, |X_{t \wedge T_{Y}} - \bar{X}_{t \wedge T_{Y}}|)
\leq e^{\lambda \zeta(T_{Y})}V_{c,\theta}(y - \bar{y}, |x - \bar{x}|),
$$

where in the first inequality above we used Equation (3.21) with $s = T_{Y}$.

Remark 3.4. Here, we make some comments on the proof of Theorem 1.1.

(i) Recall that $T_Y = \inf\{t > 0 : Y_t = \bar{Y}_t\}$. Define

$$
T_X = \inf\{t \geq 0 : X_t = \bar{X}_t\}, \quad T = \inf\{t > 0 : Y_t = \bar{Y}_t, X_t = \bar{X}_t\}.
$$

It is clear that $T \geq T_Y \vee T_X$ and $Y_t = \bar{Y}_t$ for $t \geq T_Y$. However, by the structure of the two-factor process defined by Equation (1.2), it can take place that $T > T_Y \vee T_X$, since $X_t = \bar{X}_t$ is not true for all $t \geq T_X$ unless $t \geq T_Y$. The idea for the proof of Theorem 1.1 is to make full use of the coupling time $T_Y$ for the first component, rather than $T$. To consider the coupling process until the coupling time $T_Y$, we need the crucial estimate (3.6). Since before time $T_Y$, it may occur that $X_t = \bar{X}_t$ for some $t \leq T_Y$. Hence, if we will apply the Itô formula for the test function $V_{c,\theta}(y - \bar{y}, |x - \bar{x}|)$ as given in Proposition 3.2, then this function is required to be differentiable on $\{(x, x) : x \in \mathbb{R}\}$ for any fixed $y, \bar{y}$. Furthermore, from $T_Y$ to $T$ the coupling of the first component always stays together, so we only need to couple the second component. For this, we make use of Equation (3.22), which is due to the special characterization of the two-factor process.

(ii) As mentioned above, Proposition 3.2 is crucial for the proof of Theorem 1.1. Instead of Equation (3.7), one natural way to obtain the exponential ergodicity in the $L^1$-Wasserstein distance is to prove that there exists a constant $c_0 > 0$ such that for all $y > \bar{y} \geq 0$ and $x, \bar{x} \in \mathbb{R}$,

$$
(L^*G)(y - \bar{y}, |x - \bar{x}|) \leq -c_0 G(y - \bar{y}, |x - \bar{x}|),
$$

(3.23)
where
\[ G(s, t) \triangleq s + t, \quad s, t \geq 0; \quad (3.24) \]
see, for example, [29, 31, 32]. Since in the setting of Theorem 1.1 the drift term satisfies the so-called monotone condition due to \( b > 0 \) and \( \lambda > 0 \), one may just take
\[ G(s, t) = G_0(s, t) : = s + t, \quad s, t \geq 0. \]
By some calculations, we find that there is a constant \( c_1 > 1 \) so that for all \( y > \tilde{y} \geq 0 \) and \( \tilde{x} \geq x \),
\begin{align*}
- b(y - \tilde{y}) - \lambda(\tilde{x} - x) + c_1^{-1} (y - \tilde{y})(\tilde{x} - x)^{1-\alpha} & \leq (L^*G_0)(y - \tilde{y}, \tilde{x} - x) \\
& \leq - b(y - \tilde{y}) - \lambda(\tilde{x} - x) + c_1 (y - \tilde{y})(\tilde{x} - x)^{1-\alpha}.
\end{align*}
That is, with this choice, Equation (3.23) cannot be true. Hence, some modification of \( G_0(s, t) \) is required. The function \( F(s, t) \) defined by Equation (3.1) and satisfying Equation (3.24) is one possible candidate. Roughly speaking, \( F(s, t) \) is a \( C^2 \)-mollifier for two coordinate functions \( (s, t) \mapsto s \) and \( (s, t) \mapsto t \). Whereas, the function \( F(s, t) \) above is not enough for our aim, we need further to refine it into \( V_{c, \beta}(s, t) \). In particular, the factor \( \beta^2 \) is added to balance some bad estimates from \( (L^*F)(y - \tilde{y}, |x - \tilde{x}|) \); see Remark 3.3 above for some details. This partly explains the intuition for the construction of the function \( V_{c, \beta} \) used in the proof of Theorem 1.1.

4 | EXPONENTIAL ERGODICITY FOR OTHER TWO-FACTOR AFFINE PROCESSES AND BEYOND

4.1 | The framework (1.2) with two spectrally positive stable noises

In this subsection, we are still interested in the two-factor model (1.2) but with the Brownian motion \((B_t)_{t \geq 0}\) in the first equation replaced by a spectrally positive \( \beta \)-stable process \((L_t)_{t \geq 0}\) for some \( \beta \in (1, 2) \). More precisely, we shall work on the SDE on \( \mathbb{R}_+ \times \mathbb{R} \):
\[ \begin{aligned}
\{ & dY_t = (a - bY_t) dt + Y_t^{1/\beta} dL_t, \quad t \geq 0, \ Y_0 \geq 0, \\
& dX_t = (\lambda - \lambda X_t) dt + Y_t^{1/\alpha} dZ_t, \quad t \geq 0, \ X_0 \in \mathbb{R},
\end{aligned} \quad (4.1) \]
where \( a \geq 0, \ b, \lambda \in \mathbb{R}, (L_t)_{t \geq 0} \) (resp. \( (Z_t)_{t \geq 0} \)) is a spectrally positive \( \beta \)-stable (resp. \( \alpha \)-stable) process with the Lévy measure \( \nu_\beta(dz) \) (resp. \( \nu_\alpha(dz) \)). We further assume that \((L_t)_{t \geq 0}\) and \((Z_t)_{t \geq 0}\) are mutually independent. Again, by means of [4, Theorem 2.1], Equation (4.1) has a unique strong solution \((Y_t, X_t)_{t \geq 0}\). Then, we have the following statement for the affine process associated with the SDE (4.1).

**Theorem 4.1.** Let \((Y_t, X_t)_{t \geq 0}\) be the unique strong solution to Equation (4.1). If \( b > 0 \) and \( \lambda > 0 \), then the process \((Y_t, X_t)_{t \geq 0}\) is exponentially ergodic with respect to the \( L^1 \)-Wasserstein distance \( W_1 \).

For the SDE (4.1), we shall apply the synchronous coupling for both components \((Y_t)_{t \geq 0}\) and \((X_t)_{t \geq 0}\). For this, we consider the SDE on \( \mathbb{R}_+ \times \mathbb{R} \):
\[ \begin{aligned}
\{ & d\tilde{Y}_t = (a - b\tilde{Y}_t) dt + \tilde{Y}_t^{1/\beta} dL_t, \quad t \geq 0, \ \tilde{Y}_0 > 0, \\
& d\tilde{X}_t = (\lambda - \lambda \tilde{X}_t) dt + \tilde{Y}_t^{1/\alpha} dZ_t, \quad t \geq 0, \ \tilde{X}_0 \in \mathbb{R},
\end{aligned} \]
It is clear that \((Y_t, X_t, (\tilde{Y}_t, \tilde{X}_t))_{t \geq 0}\) is a coupling of the process \((Y_t, X_t)_{t \geq 0}\). According to [29, Corollary 2.3], the coupling of the first component process also preserves the order property; that is, \( Y_t \geq \tilde{Y}_t \) for all \( t \geq 0 \), in case of \( Y_0 \geq \tilde{Y}_0 \).
Furthermore, the infinitesimal generator (i.e., the coupling operator) of the coupling process \((Y_t, X_t, \tilde{Y}_t, \tilde{X}_t)_{t \geq 0}\) is given by

\[
\begin{align*}
(L^*G)(y - \tilde{y}, |x - \tilde{x}|) &= -b\partial_1 G(y - \tilde{y}, |x - \tilde{x}|)(y - \tilde{y}) - \lambda \partial_2 G(y - \tilde{y}, |x - \tilde{x}|)z \nu_\beta(dz) \\
&\quad + (y - \tilde{y}) \int_0^\infty \left( G(y - \tilde{y} + z, |x - \tilde{x}|) - G(y - \tilde{y}, |x - \tilde{x}|) \right) \nu_\beta(dz) \\
&\quad + (y - \tilde{y}) \int_0^\infty \left( G(y - \tilde{y}, |x + z - \tilde{x}|) - G(y - \tilde{y}, |x - \tilde{x}|) \right) \partial_2 G(y - \tilde{y}, |x - \tilde{x}|) \frac{z}{|x - \tilde{x}|} \nu_\alpha(dz)
\end{align*}
\]

(4.2)

for \(G \in C^2(\mathbb{R}^2_+)\) and \((y, \tilde{y}, x, \tilde{x}) \in \mathbb{R}^2_+ \times \mathbb{R}^2\) with \(0 \leq \tilde{y} \leq y\). Similarly, we can write the expression of \((L^*G)(y - \tilde{y}, |x - \tilde{x}|)\) for \(G \in C^2(\mathbb{R}^2_+)\) and \((y, \tilde{y}, x, \tilde{x}) \in \mathbb{R}^2_+ \times \mathbb{R}^2\) with \(0 \leq y \leq \tilde{y}\).

For any \(c > 0\) and \(\theta \in (0, 1)\), define

\[
V_{c, \theta}(s, t) = c(s + s^\theta) + F(s, t),
\]

where \(F\) was introduced in Equation (3.1). It is obvious that for any \(c > 0\),

\[
V_{c, \theta}(s, t) \asymp (s + s^\theta) + t.
\]

The proof of Theorem 4.1 is based on the following proposition.

**Proposition 4.2.** For any \(\theta \in (0, 2 - (\alpha \vee \beta)]\), there exist constants \(c, \eta > 0\) such that

\[
(L^*V_{c, \theta})(y - \tilde{y}, |x - \tilde{x}|) \leq -\eta V_{c, \theta}(y - \tilde{y}, |x - \tilde{x}|), \quad \tilde{y} \in [0, y), x, \tilde{x} \in \mathbb{R}.
\]

(4.3)

**Proof.** Some non-trivial modifications of the proof for Proposition 4.2 are required although the corresponding idea is similar to that of Proposition 3.2. In the following, we highlight some key differences. For \(c > 0\) and \(\theta \in (0, 2 - (\alpha \vee \beta)]\), set

\[
U_{c, \theta}(s, t) := c(s + s^\theta), \quad s, t \geq 0.
\]

Below, we set \(0 \leq \tilde{y} < y\) and let \(x, \tilde{x} \in \mathbb{R}\). According to Equation (4.2), we deduce

\[
(L^*U_{c, \theta})(y - \tilde{y}, |x - \tilde{x}|) = -b c \{y - \tilde{y} + \theta(y - \tilde{y})^\theta\} + c(y - \tilde{y}) \int_0^\infty \frac{(y - \tilde{y} + z)^{\theta} - (y - \tilde{y})^\theta - \theta(y - \tilde{y})^{\theta - 1} z}{|x - \tilde{x}|} \nu_\beta(dz)
\]

\[
\leq -b c \{y - \tilde{y} + \theta(y - \tilde{y})^\theta\},
\]

(4.4)

where in the inequality we used the fact that for all \(0 \leq \tilde{y} < y\) and \(z \geq 0\),

\[
(y - \tilde{y} + z)^{\theta} - (y - \tilde{y})^\theta - \theta(y - \tilde{y})^{\theta - 1} z \leq 0, \quad \theta \in (0, 1)
\]

thanks to the mean value theorem. Thus, Equation (4.3) follows provided that there exists a constant \(C_1 > 0\) such that

\[
(L^*F)(y - \tilde{y}, |x - \tilde{x}|) \leq -\lambda |x - \tilde{x}| + C_1 \{y - \tilde{y} + (y - \tilde{y})^\theta\}
\]

(4.5)

by taking Equation (4.4) into account and choosing \(c > 0\) such that \(cb \theta > C_1\). Nevertheless, to derive Equation (4.5), it suffices to show that there exists a constant \(C_2 > 0\) such that

\[
(L^*F)(y - \tilde{y}, |x - \tilde{x}|) \leq -\lambda |x - \tilde{x}| 1_{\{|x - \tilde{x}| \geq 2(y - \tilde{y})\}} + C_2 \{y - \tilde{y} + (y - \tilde{y})^{2-\alpha} + (y - \tilde{y})^{2-\beta}\}
\]

(4.6)
In fact, Equation (4.6) yields Equation (4.5) by noting that there exists a constant $C_3 > 0$ such that

$$-\lambda|x - \bar{x}|1_{|x-\bar{x}|\geq 2|y-\bar{y}|} + C_4(y - \bar{y} + (y - \bar{y})^{2-\alpha} + (y - \bar{y})^{2-\beta})$$

$$\leq -\lambda|x - \bar{x}| + 2\lambda(y - \bar{y}) + C_3(y - \bar{y} + (y - \bar{y})^{2-(\alpha \lor \beta)})$$

$$\leq -\lambda|x - \bar{x}| + C_3(y - \bar{y} + (y - \bar{y})^{2-(\alpha \lor \beta)})$$

$$\leq -\lambda|x - \bar{x}| + C_4(y - \bar{y} + (y - \bar{y})^{2-\alpha} + (y - \bar{y})^{2-\beta})$$

$$\leq -\lambda|x - \bar{x}| + 2\lambda(y - \bar{y}) + C_2(y - \bar{y} + (y - \bar{y})^2 - \alpha + (y - \bar{y})^2 - \beta)$$

where in the second inequality we used, due to $\alpha, \beta > 1$, $(y - \bar{y})^{1-\alpha} \wedge (y - \bar{y})^{1-\beta} \leq 1$ for $y - \bar{y} \geq 1$ and $1 \leq (y - \bar{y})^{-\alpha} \lor (y - \bar{y})^{-\beta} \leq (y - \bar{y})^{-(\alpha \lor \beta)}$ for $0 < y - \bar{y} \leq 1$, and in the last two inequality we utilized $\beta \in (0, 2 - (\alpha \lor \beta))$. Therefore, to obtain the desired assertion (4.3), we need only to verify Equation (4.6).

From Equation (4.2), we have

$$(L^* F)(y - \bar{y}, |x - \bar{x}|) = -b\delta_1 F(y - \bar{y}, |x - \bar{x}|)(y - \bar{y}) - \lambda \delta_2 F(y - \bar{y}, |x - \bar{x}|) |x - \bar{x}|$$

$$+ (y - \bar{y}) \int_0^\infty \left( F(y - \bar{y} + z, |x - \bar{x}|) - F(y - \bar{y}, |x - \bar{x}|) - \delta_1 F(y - \bar{y}, |x - \bar{x}|)z \right) \nu_\beta(dz)$$

$$+ (y - \bar{y}) \int_0^\infty \left( F(y - \bar{y}, |x + z - \bar{x}|) - F(y - \bar{y}, |x - \bar{x}|) - \delta_2 F(y - \bar{y}, |x - \bar{x}|)(x - \bar{x})z \right) \nu_\alpha(dz)$$

$$= : Y_1 + Y_2 + (y - \bar{y})Y_3 + (y - \bar{y})Y_4.$$  (4.7)

Thus, Equation (4.6) is available once we prove that there is a constant $C_4 > 0$ so that

$$Y_1 + Y_2 + (y - \bar{y})Y_4 \leq -\lambda|x - \bar{x}|1_{|x-\bar{x}|\geq 2|y-\bar{y}|} + C_4(y - \bar{y} + (y - \bar{y})^{2-\alpha}),$$  (4.8)

and

$$Y_3 \leq C_4(y - \bar{y})^{1-\beta}.$$  (4.9)

In what follows, we aim to prove Equations (4.8) and (4.9), respectively. Let (i)–(iii) be the three cases listed in the proof of Proposition 3.2. By a close inspection of argument for Proposition 3.2, there is a constant $c_1 > 0$ such that

$$Y_2 + (y - \bar{y})Y_4 \leq -\lambda|x - \bar{x}|1_{|x-\bar{x}|\geq 2|y-\bar{y}|} + c_1(y - \bar{y} + (y - \bar{y})^{2-\alpha}).$$  (4.10)

From Equation (3.2), it is easy to see that $Y_1 \leq b c_0(y - \bar{y})$. As a result, Equation (4.8) follows immediately. Next, we turn to the proof of Equation (4.9).

**Proof of Equation (4.9) for the case (i).** For this case, it follows that

$$Y_3 = \int_{|x-\bar{x}|\leq y-\bar{y}+z} (y - \bar{y} + z - |x - \bar{x}|) \nu_\beta(dz)$$

$$+ \int_{y-\bar{y}+z<|x-\bar{x}|<2(y-\bar{y}+z)} \left(1 - g\left(\frac{|x - \bar{x}|}{y - \bar{y} + z}\right)\right)(y - \bar{y} + z - |x - \bar{x}|) \nu_\beta(dz)$$

$$= : Y_{31} + Y_{32}.$$
On one hand, since \( y - \bar{y} + z < |x - \bar{x}| \) and \( g \in [0, 1] \), \( Y_{32} \leq 0 \). On the other hand, we have

\[
Y_{31} \leq \int_{z \geq y - \bar{y}} (y - \bar{y} + z) \nu_\beta(dz) \leq c_2(y - \bar{y})^{1 - \beta}
\]

for some constant \( c_2 > 0 \), where we used \( |x - \bar{x}| \geq 2(y - \bar{y}) \) in the first inequality. Hence, we arrive at

\[
Y_3 \leq c_2(y - \bar{y})^{1 - \beta}.
\]

Whence, we infer that Equation (4.9) holds true for the case (i).

**Proof of Equation (4.9) for the case (ii).** Indeed, with regard to this case, Equation (4.9) is available by using

\[
Y_3 = \int_{[|x - \bar{x}| < y - \bar{y} + z]} \left( \left( 1 - g\left( \frac{|x - \bar{x}|}{y - \bar{y} + z} \right) \right)(y - \bar{y} + z) - (y - \bar{y}) - z \right) \nu_\beta(dz)
\]

\[
\leq \int_{[|x - \bar{x}| < y - \bar{y} + z]} (y - \bar{y} + z - (y - \bar{y}) - z) \nu_\beta(dz) = 0,
\]

where the inequality is due to \( g \in [0, 1] \).

**Proof of Equation (4.9) for the case (iii).** As for this case, \( Y_3 \) can be rewritten as

\[
Y_3 = \int_{z \leq (y - \bar{y})/2} (F(y - \bar{y} + z, |x - \bar{x}|) - F(y - \bar{y}, |x - \bar{x}|) - \partial_1 F(y - \bar{y}, |x - \bar{x}|) z) \nu_\beta(dz)
\]

\[
+ \int_{z \geq (y - \bar{y})/2} (F(y - \bar{y} + z, |x - \bar{x}|) - F(y - \bar{y}, |x - \bar{x}|) - \partial_1 F(y - \bar{y}, |x - \bar{x}|) z) \nu_\beta(dz)
\]

\[= : Y_{31} + Y_{32}.
\]

By Taylor’s expansion, there exists a constant \( \xi \in [y - \bar{y}, 3(y - \bar{y})/2] \) such that

\[
Y_{31} = \frac{1}{2} \partial_{11} F(\xi, |x - \bar{x}|) \int_{z \leq (y - \bar{y})/2} z^2 \nu_\beta(dz) \leq c_3(y - \bar{y})^{1 - \beta}
\]

(4.11)

for some constant \( c_3 > 0 \), where the inequality above follows from Equation (3.2). Furthermore, in terms of the definition of \( F \) and Equation (3.2), there is a constant \( c_4 > 0 \) such that

\[
Y_{32} \leq (1 + c_0) \int_{z \geq (y - \bar{y})/2} (y - \bar{y} + |x - \bar{x}| + 2z) \nu_\beta(dz)
\]

\[
\leq 3(1 + c_0) \int_{z \geq (y - \bar{y})/2} (y - \bar{y} + 2z) \nu_\beta(dz) \leq c_4(y - \bar{y})^{1 - \beta},
\]

(4.12)

where the second inequality is owing to \( y - \bar{y} \leq |x - \bar{x}| \leq 2(y - \bar{y}) \). Consequently, for the case (iii), Equation (4.9) follows by combining Equation (4.11) with Equation (4.12).

\[\square\]

**Proof of Theorem 4.1.** Similar to Theorem 1.1, we indeed can claim that, for any \( \vartheta \in (0, 2 - (\alpha \vee \beta)] \), there exist a unique probability measure \( \mu \) on \( \mathbb{R}_+ \times \mathbb{R} \) and a constant \( \eta := \eta(\vartheta) > 0 \) such that for all \( y \in \mathbb{R}_+, x \in \mathbb{R} \) and \( t > 0 \),

\[
W_{\vartheta}(P(t, (y, x), \cdot), \mu) \leq C(y, x, \vartheta)e^{-\eta t},
\]

where

\[
\psi_\vartheta(u, v) := u + u^\vartheta + v, \quad u, v \geq 0
\]
and \( C(y, x, \theta) > 0 \) is independent of \( t \). Note that, as far as Equation (4.1) is concerned, one can check that Equation (2.2) still holds. With Equation (2.2) and Proposition 4.2 at hand, the proof of the assertion above can be completed by implementing the same argument of Theorem 1.1, and so we herein omit the corresponding details.

Remark 4.3. In order to prove the exponential ergodicity for the two-factor model (4.1) driven by two independent spectrally positive stable noises, respectively, we just apply the synchronous coupling. Similarly, one can apply the synchronous coupling (instead of the coupling by reflection for a small distance and the synchronous coupling for a big distance to the first component and the synchronous coupling to the second component) to the two-factor model (1.2). However, with the aid of the synchronous coupling the approach can only yield the exponential ergodicity under a weaker Wasserstein distance with the cost function

\[
(u, v) \mapsto \psi^*(u, v) := u + u^{1/2} + v, \quad u, v \geq 0;
\]

see [17, Theorem 1.5(a)] or Remark (i) in the last but one paragraph in Section 1.

### 4.2 General two-factor affine processes

In this part, we emphasize that the approaches applied to Theorems 1.1 and 4.1 still work for two other general two-factor affine processes. In particular, we can show the exponential ergodicity for the following two kinds of two-factor affine processes with respect to the \( L^1 \)-Wasserstein distance \( W_1 \).

Two-factor affine process (I):

\[
\begin{align*}
\frac{dY_t}{dt} &= (a - b Y_t) dt + Y_t^{1/2} dB_t^{(1)} + Y_t^{1/\beta} dZ_t^{(\beta)}, \quad t \geq 0, Y_0 \geq 0, \\
\frac{dX_t}{dt} &= (\nu - \lambda X_t) dt + Y_t^{1/2} (\rho dB_t^{(1)} + \sqrt{1 - \rho^2} dB_t^{(2)}) + Y_t^{1/\alpha} dZ_t^{(\alpha)}, \quad t \geq 0, X_0 \in \mathbb{R},
\end{align*}
\]

where \( a \geq 0, \nu \in \mathbb{R}, b, \lambda > 0, \rho \in [-1, 1], \beta, \alpha \in (1, 2), (B_t^{(1)}, B_t^{(2)})_{t \geq 0} \) is a two-dimensional standard Brownian motion, \((Z_t^{(\beta)})_{t \geq 0}\) (resp. \((Z_t^{(\alpha)})_{t \geq 0}\)) is a spectrally positive \( \beta \)-stable (resp. \( \alpha \)-stable) Lévy process. Moreover, the processes \((B_t^{(1)}, B_t^{(2)})_{t \geq 0}, (Z_t^{(\beta)})_{t \geq 0}\) and \((Z_t^{(\alpha)})_{t \geq 0}\) are mutually independent.

Two-factor affine process (II):

\[
\begin{align*}
\frac{dY_t}{dt} &= (a - b Y_t) dt + Y_t^{1/2} dB_t, \quad t \geq 0, Y_0 \geq 0, \\
\frac{dX_t}{dt} &= (\nu - \lambda X_t - \gamma Y_t) dt + Y_t^{1/\alpha} dZ_t, \quad t \geq 0, X_0 \in \mathbb{R},
\end{align*}
\]

where \( a \geq 0, b, \lambda > 0, \gamma \in \mathbb{R}, (B_t)_{t \geq 0} \) is a standard Brownian motion, and \((Z_t)_{t \geq 0}\) is an independently spectrally positive \( \alpha \)-stable process with \( \alpha \in (1, 2) \).

We have proven the exponential ergodicity for the SDEs (1.2) and (4.1) (with respect to the \( L^1 \)-Wasserstein distance \( W_1 \)). Combining both arguments together, we can show that the following two-factor affine process

\[
\begin{align*}
\frac{dY_t}{dt} &= (a - b Y_t) dt + Y_t^{1/2} dB_t + Y_t^{1/\beta} dZ_t^{(\beta)}, \quad t \geq 0, Y_0 \geq 0, \\
\frac{dX_t}{dt} &= (\nu - \lambda X_t) dt + Y_t^{1/\alpha} dZ_t^{(\alpha)}, \quad t \geq 0, X_0 \in \mathbb{R},
\end{align*}
\]

is exponentially ergodic. The difference between the SDE (4.13) and type (I) above is that in (I) there is an additional diffusion term driven by Brownian motions for the second equation. Then, one can apply the synchronous coupling to this additional term and follow the argument showing that Equation (4.13) is exponentially ergodic to derive the exponential ergodicity of the type (I).

The difference between the SDE (1.2) and type (II) above is due to the drift term in the second equation for the process \((Y_t)_{t \geq 0}\) of the associated affine process. Concerning the two-factor affine process (II), we can also show that \((Y_t, X_t)\) is exponentially ergodic with respect to the \( L^1 \)-Wasserstein distance \( W_1 \) by following the argument of Theorem 1.1 if the counterpart of Proposition 3.2 is still valid, see Proposition 4.4 for more details.

To proceed, we introduce some additional notation. Below, let \( F \) be defined as in Equation (3.1) but with \( g \in C^2_b(\mathbb{R}_+) \) such that \( g' \geq 0 \) and
for some constant $\delta > 0$. With $F$ above at hand, the function $V_{c,\beta}$ is defined exactly as in Equation (3.4).

**Proposition 4.4.** There exist constants $c, \eta > 0$ such that for any $y > \bar{y} \geq 0$ and $x, \bar{x} \in \mathbb{R}$,

$$(L^* V_{c,\beta})(y - \bar{y}, |x - \bar{x}|) \leq -\eta V_{c,\beta}(y - \bar{y}, |x - \bar{x}|),$$

where $L^*$ is the coupling operator of the two-factor affine process (II) given in Equation (2.4) with $\kappa - \lambda x$ replaced by $\kappa - \lambda x - \gamma y$.

**Proof.** Below, we shall fix $y > \bar{y} \geq 0$ and $x, \bar{x} \in \mathbb{R}$. By following the argument of Proposition 3.2, to end the proof of Proposition 4.4, it is sufficient to show that

$$
(L^* F)(y - \bar{y}, |x - \bar{x}|) \leq \frac{\lambda}{2} |x - \bar{x}| 1_{|x - \bar{x}| \geq \kappa_0(y - \bar{y})} + C(y - \bar{y}) + C \left( \sqrt{y} + \sqrt{\bar{y}} \right)^2 (y - \bar{y})^{-1} 1_{0 < y - \bar{y} < 1} \tag{4.14}
$$

for some constant $C > 0$. In what follows, we are going to verify Equation (4.14) in terms of the three cases below:

(i) $|x - \bar{x}| > \kappa_0(y - \bar{y})$;
(ii) $|x - \bar{x}| \leq y - \bar{y}$;
(iii) $y - \bar{y} < |x - \bar{x}| \leq \kappa_0(y - \bar{y})$.

For the setting (i), we have $F(y - \bar{y}, |x - \bar{x}|) = |x - \bar{x}|$. This, in addition to Equation (3.7) with $-\lambda|x - \bar{x}| - \gamma(x - \bar{x})(y - \bar{y})/|x - \bar{x}|$ instead of $-\lambda|x - \bar{x}|$ therein, yields

$$(L^* F)(y - \bar{y}, |x - \bar{x}|) = -\lambda|x - \bar{x}| - \frac{\gamma(x - \bar{x})(y - \bar{y})}{|x - \bar{x}|} + (y - \bar{y})(I_1 + I_2 + I_3)$$

$$
\leq -\lambda|x - \bar{x}| + |\gamma|(y - \bar{y}) + (y - \bar{y})(I_1 + I_2 + I_3)$$

$$\leq -\frac{\lambda}{2} |x - \bar{x}| + (y - \bar{y})(I_1 + I_2 + I_3),$$

in which $I_1, I_2, I_3$ were introduced in Equation (3.11) with the factor 2 in the splitting intervals replaced by the number $\kappa_0$, and in the last display we used

$$|\gamma|(y - \bar{y}) \leq \frac{\lambda}{2} |x - \bar{x}|.$$

Thus, combining the estimates on $I_1, I_2, I_3$, we infer that (4.14) holds true for the case (i). Observe that, as for the case (ii), $L^* F$ shares the same expression with the counterpart for the setup (ii) in the proof of Proposition 3.2. Subsequently, with some mild modifications of the associated details (where, in particular, replace the factor 2 in the splitting intervals by $\kappa_0$), we conclude that Equation (4.14) is still available for the case (ii). Note that Lemma 3.1 is still valid with the number 2 in Equation (3.2) replaced by $\kappa_0$ so that Equation (3.15) remains true for the new expression $F$. As a result, Equation (4.14) follows by carrying out the same argument to derive Equation (3.9) for the case (iii). $\square$

We mention that, applying the similar idea, one can also prove the exponential ergodicity with respect to the $L^1$-Wasserstein distance $W_1$ for type (II) if the first equation is replaced by

$$
dY_t = (a - bY_t) dt + Y_t^{1/\beta} dZ_t^{(\beta)}, \quad t \geq 0, \ Y_0 \geq 0
$$
or

\[ dY_t = (a - bY_t) \, dt + Y_t^{1/2} \, dB_t + Y_t^{1/\beta} \, dZ_t^{(\beta)}, \quad t \geq 0, \quad Y_0 \geq 0, \]

where \( a \geq 0, b > 0 \), \((B_t)_{t \geq 0}\) is a standard Brownian motion, and \((Z_t)_{t \geq 0}\) is an independently spectrally positive \(\alpha\)-stable process with \(\alpha \in (1, 2]\).

### 4.3 Beyond two-factor affine processes

In this part, we will briefly mention that our method also works for the following models beyond two-factor affine processes. Let \((Y_t, X_t)_{t \geq 0}\) be a time-homogeneous Markov process on \(\mathbb{R}^+ \times \mathbb{R}\) such that

\[
\begin{cases}
    dY_t = b_1(Y_t) \, dt + Y_t^{1/\beta} \, dL_t, & t \geq 0, \quad Y_0 \geq 0, \\
    dX_t = b_2(X_t) \, dt + Y_t^{1/\alpha} \, dZ_t, & t \geq 0, \quad X_0 \in \mathbb{R},
\end{cases}
\tag{4.15}
\]

where \((L_t)_{t \geq 0}\) is a spectrally positive \(\beta\)-stable process with \(\beta \in (1, 2]\), \((Z_t)_{t \geq 0}\) is an independently spectrally positive \(\alpha\)-stable process with \(\alpha \in (1, 2]\), and \(b_1\) (resp. \(b_2\)) is continuous on \(\mathbb{R}^+\) (resp. \(\mathbb{R}\)) so that there are constants \(\lambda_i > 0\) (\(i = 1, 2\)) such that for \(i = 1, 2\) and any \(x > y\),

\[ b_i(x) - b_i(y) \leq -\lambda_i(x - y). \tag{4.16} \]

Note that, according to [18, Theorem 5.6], there exists a unique strong solution to the first component \((Y_t)_{t \geq 0}\) of the SDE (4.15). Once \((Y_t)_{t \geq 0}\) is fixed, the unique strong solution to the second component \((X_t)_{t \geq 0}\) is guaranteed by the monotone condition (4.16) for the drift term \(b_2(x)\). Therefore, the SDE (4.15) has a unique strong solution \((Y_t, X_t)_{t \geq 0}\). Furthermore, according to [29, Corollary 2.3 and Remark 2.4], the coupling of the first component \((Y_t)_{t \geq 0}\) can be chosen to preserve the order property. With these facts and Equation (4.16) again at hand, we can use the ideas of proofs for Theorems 1.1 and 4.1 to conclude that the process \((Y_t, X_t)_{t \geq 0}\) is exponentially ergodic with respect to the \(L^1\)-Wasserstein distance \(W_1\). Note that the contractive property like Equation (3.22) in the proof of Theorem 1.1 is a consequence of Equation (4.16) for the drift term \(b_2(x)\). As we mentioned before, since all known approaches dealing with the ergodicity of affine processes depend on their especially structural characterizations, they seem to be invalid in establishing the exponential ergodicity of the SDE (4.15).

**REFERENCES**

[1] E. AbiJaber, M. Larsson, and S. Pulido, *Affine Volterra processes*, Ann. Appl. Probab. 29 (2019), 3155–3200.

[2] A. Alfonsi, *Affine diffusions and related processes: simulation, theory and applications*, Springer, Cham, 2015.

[3] M. Barczy, L. Döring, Z. Li, and G. Pap, *Parameter estimation for a subcritical affine two factor model*, J. Statist. Plann. Inference 151/152 (2014), 37–59.

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

[5] B. Bolyog and G. Pap, *Conditions for stationarity and ergodicity of two-factor affine diffusions*, Commun. Stoch. Anal. 10 (2016), 587–610.

[6] B. Bolyog and G. Pap, *On conditional least square estimation for affine diffusion based on continuous observations*, Stat. Inference Stoch. Process. 22 (2019), 41–75.

[7] B. Böttcher, R. L. Schilling, and J. Wang, *Constructions of coupling processes for Lévy processes*, Stoch. Proc. Appl. 121 (2011), 1201–1216.

[8] B. Böttcher, R.L. Schilling, and J. Wang, *Lévy-type processes: construction, approximation and sample path properties*, Lecture Notes in Mathematics, vol. 2099, Lévy Matters III, Springer, Berlin, 2014.

[9] M. E. Caballero, J. L. P. Garmendia, and G. Uribe Bravo, *Affine processes on \(\mathbb{R}^m \times \mathbb{R}^m\) and multiparameter time changes*, Ann. Inst. Henri Poincaré Probab. Stat. 13 (2017), 1280–1304.

[10] H. Chen and S. Joslin, *Generalized transform analysis of affine processes and applications in finance*, Rev. Financial Stud. 25 (2012), 2225–2256.

[11] J. C. Cox, J. E. Ingersoll, and S. A. Ross, *A theory of the term structure of interest rates*, Econometrica 53 (1985), 385–407.

[12] D. Duffie, D. Filipović, and W. Schachermayer, *Affine processes and applications in finance*, Ann. Appl. Probab. 13 (2003), 984–1053.

[13] D. Duffie, J. Pan, and K. Singleton, *Transform analysis and asset pricing for affine jump-diffusions*, Econometrica 68 (2000), 1343–1376.

[14] D. Filipović, E. Mayerhofer, and P. Schneider, *Density approximations for multivariate affine jump-diffusion processes*, J. Econometrics 176 (2013), 93–111.
[15] M. Friesen, P. Jin, J. Kremer, and B. Rüdiger, *Exponential ergodicity for stochastic equations of nonnegative processes with jumps*, arXiv:1902.02833, https://arxiv.org/archive/math.

[16] M. Friesen, P. Jin, J. Kremer, and B. Rüdiger, *Ergodicity of affine processes on the cone of symmetric positive semidefinite matrices*, Adv. Appl. Probab. 52 (2020), 825–854.

[17] M. Friesen, P. Jin, and B. Rüdiger, *Stochastic equation and exponential ergodicity in Wasserstein distances for affine processes*, Ann. Appl. Probab. 30 (2020), 2165–2195.

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

[19] G. Grafendorfer, *Infinite-dimensional affine processes*, PhD thesis. ETH, Zurich, 2016.

[20] Y. Jiao, C. Ma, and S. Scotti, *Alpha-CIR model with branching processes in sovereign interest rate modeling*, Finance Stoch. 21 (2017), 789–813.

[21] P. Jin, J. Kremer, and B. Rüdiger, *Exponential ergodicity of an affine two-factor model based on the \( \alpha \)-root process*, Adv. Appl. Probab. 49 (2017), 1144–1169.

[22] P. Jin, J. Kremer, and B. Rüdiger, *Existence of limiting distribution for affine processes*, J. Math. Anal. Appl. 486 (2020), 123912.

[23] P. Jin, B. Rüdiger, and C. Trabelsi, *Positive Harris recurrence and exponential ergodicity of the basic affine jump-diffusion*, Stoch. Anal. Appl. 34 (2016), 75–95.

[24] M. Keller-Ressel and E. Mayerhofer, *Exponential moments of affine processes*, Ann. Appl. Probab. 25 (2015), 714–752.

[25] M. Keller-Ressel and W. Schachermayer, *Affine processes are regular*, Probab. Theory Relat. Fields 151 (2011), 591–611.

[26] M. Keller-Ressel, W. Schachermayer, and J. Teichmann, *Regularity of affine processes on general state spaces*, Electron. J. Probab. 18 (2013), no. 43, 17.

[27] M. Keller-Ressel, T. Schmidt, and R. Wardenga, *Affine processes beyond stochastic continuity*, Ann. Appl. Probab. 29 (2019), 3387–3437.

[28] Z. Li and C. Ma, *Asymptotic properties of estimators in a stable Cox-Intersoll–Ross model*, Stoch. Proc. Appl. 125 (2015), 3196–3233.

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

[30] M. B. Majka, *Coupling and exponential ergodicity for stochastic differential equations driven by Lévy processes*, Stoch. Proc. Appl. 127 (2017), 4083–4125.

[31] H. Masuda, *On multidimensional Ornstein–Uhlenbeck processes driven by a general Lévy process*, Bernoulli 10 (2004), 97–120.

[32] E. Mayerhofer, R. Stelzer, and J. Vestweber, *Geometric ergodicity of affine processes on cones*, Stoch. Proc. Appl. 130 (2020), 4141–4173.

[33] P. S. Li and J. Wang, *Stability of Markovian processes III: Foster–Lyapunov criteria for continuous time processes*, Adv. Appl. Probab. 25 (1993), 518–548.

[34] R. L. Schilling, P. Sztonyk, and J. Wang, *Coupling property and gradient estimates of Lévy processes via symbol*, Bernoulli 18 (2012), 1128–1149.

[35] T. Schmidt, S. Tappe, and W. Yu, *Infinite-dimensional affine processes*, Stoch. Proc. Appl. 130 (2020), 7131–7169.

[36] J. Wang, *On the exponential ergodicity of Lévy-driven Ornstein–Uhlenbeck processes*, J. Appl. Probab. 49 (2012), 990–1004.