Bernoulli **22**(3), 2016, 1598–1616
DOI: 10.3150/15-BEJ705

L^p-Wasserstein distance for stochastic differential equations driven by Lévy processes

JIAN WANG

School of Mathematics and Computer Science, Fujian Normal University, Fuzhou 350007, China.
E-mail: jianwang@fjnu.edu.cn

Coupling by reflection mixed with synchronous coupling is constructed for a class of stochastic differential equations (SDEs) driven by Lévy noises. As an application, we establish the exponential contractivity of the associated semigroups $(P_t)_{t\ge 0}$ with respect to the standard L^p-Wasserstein distance for all $p \in [1, \infty)$. In particular, consider the following SDE:

$$\mathrm{d}X_t = \mathrm{d}Z_t + b(X_t)\,\mathrm{d}t,$$

where $(Z_t)_{t\ge 0}$ is a symmetric α-stable process on $\mathbb{R}^d$ with $\alpha \in (1, 2)$. We show that if the drift term b satisfies that for any $x, y \in \mathbb{R}^d$,

$$\langle b(x) - b(y), x - y\rangle \le \begin{cases} K_1|x-y|^2, & |x-y| \le L_0; \\ -K_2|x-y|^\theta, & |x-y| > L_0 \end{cases}$$

holds with some positive constants K_1, K_2, $L_0 > 0$ and $\theta \ge 2$, then there is a constant $\lambda := \lambda(\theta, K_1, K_2, L_0) > 0$ such that for all $p \in [1, \infty)$, $t > 0$ and $x, y \in \mathbb{R}^d$,

$$W_p(\delta_x P_t, \delta_y P_t) \le C(p, \theta, K_1, K_2, L_0)\mathrm{e}^{-\lambda t/p}\left[\frac{|x-y|^{1/p} \vee |x-y|}{1 + |x-y|\mathbf{1}_{(1,\infty)\times(2,\infty)}(t, \theta)}\right].$$

Keywords: coupling by reflection; exponential contractivity; L^p-Wasserstein distance; stochastic differential equation driven by Lévy noise; symmetric stable process

1. Introduction

In this paper, we consider the following stochastic differential equation (SDE) driven by Lévy noises:

$$\mathrm{d}X_t = \mathrm{d}Z_t + b(X_t)\,\mathrm{d}t, \tag{1.1}$$

where $(Z_t)_{t\ge 0}$ is a d-dimensional Lévy process, and $b : \mathbb{R}^d \to \mathbb{R}^d$ is a continuous vector field such that for any $x, y \in \mathbb{R}^d$,

$$\langle b(x) - b(y), x - y\rangle \le C|x-y|^2$$

holds for some constant $C > 0$. It is a standard result that in this case the SDE (1.1) enjoys the unique strong solution.

Denote by $(P_t)_{t\geq 0}$ the semigroup associated to (1.1). If the initial value X_0 is distributed as μ, then for any $t > 0$, the distribution of X_t is μP_t. We are concerned with the exponential contractivity of the map $\mu \mapsto \mu P_t$ with respect to the standard L^p-Wasserstein distance W_p for all $p \geq 1$. Given two probability measures μ and ν on $\mathbb{R}^d$, the standard L^p-Wasserstein distance W_p for all $p \in [1, \infty)$ (with respect to the Euclidean norm $|\cdot|$) is given by

$$W_p(\mu,\nu) = \inf_{\Pi \in \mathcal{C}(\mu,\nu)} \left(\int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^p \, \mathrm{d}\Pi(x,y) \right)^{1/p}.$$

Equipped with W_p, the totality $\mathcal{P}_p(\mathbb{R}^d)$ of probability measures having finite moment of order p becomes a complete metric space.

The following result is well known.

Theorem 1.1. *Suppose that there exists a constant $K > 0$ such that*

$$\langle b(x) - b(y), x - y \rangle \leq -K|x-y|^2 \qquad \textit{for all } x, y \in \mathbb{R}^d. \tag{1.2}$$

Then, for any $p \geq 1$ and $t > 0$,

$$W_p(\mu P_t, \nu P_t) \leq \mathrm{e}^{-Kt} W_p(\mu,\nu) \qquad \textit{for all } \mu, \nu \in \mathcal{P}_p(\mathbb{R}^d). \tag{1.3}$$

The proof of this result is quite straightforward, by simply using the synchronous coupling, which is also called the basic coupling or the coupling of marching soldiers in the literature (see, e.g., [8], Definition 2.4 and [7], Example 2.16). The reader can refer to [3], page 2432, the proof of Theorem 1.1 for the case of diffusion processes. (1.2) is the so-called uniformly dissipative condition, which seems to be a limit in applications. For diffusion processes, it follows from [21], Theorem 1, or [3], Remark 3.6 (also see [5], Theorem 3.6) that (1.3) holds for any probability measures μ and ν if and only if (1.2) holds for all $x, y \in \mathbb{R}^d$. The first breakthrough to get rid of such restrictive condition in this direction for L^1-Wasserstein distance W_1 was done recently by Eberle in [10,11], at the price of multiplying a constant $C \geq 1$ on the right-hand side of (1.3). See [10], Corollary 2.3, for more details, and [15], Theorem 1.3, for related developments on L^p-Wasserstein distance W_p with all $p \in [1, \infty)$ on this topic. However, the corresponding result for SDEs driven by Lévy noises is not available yet now. Indeed, we will see later that in this case we need a completely different idea for the construction of the coupling processes, and a new approach by using the coupling argument, in particular the more delicate choice of auxiliary functions.

Throughout this paper, we assume that the driving Lévy process has a symmetric α-stable process as a component. That is, let ν be the Lévy measure of the process $(Z_t)_{t\geq 0}$, then

$$\nu(\mathrm{d}z) \geq \frac{C_{d,\alpha}}{|z|^{d+\alpha}} \, \mathrm{d}z,$$

where $C_{d,\alpha}=2^{\alpha}\Gamma((d+\alpha)/2)\pi^{-d/2}|\Gamma(-\alpha/2)|^{-1}$ is a constant associated with the Lévy measure of a symmetric α-stable process or fractional Laplacian, that is,

$$-(-\Delta)^{\alpha/2}f(x)=\int\big(f(x+z)-f(x)-\langle\nabla f(x),z\rangle\mathbf{1}_{\{|z|\le 1\}}\big)\frac{C_{d,\alpha}}{|z|^{d+\alpha}}\,\mathrm{d}z.$$

Denote by $\omega_d=2\pi^{d/2}/\Gamma(d/2)$ the surface measure of the unit sphere in $\mathbb{R}^d$. Our main contribution of this paper is as follows.

Theorem 1.2. *Assume that for any $x,y\in\mathbb{R}^d$,*

$$\langle b(x)-b(y),x-y\rangle\le\begin{cases}K_1|x-y|^2, & |x-y|\le L_0;\\ -K_2|x-y|^{\theta}, & |x-y|>L_0\end{cases}\tag{1.4}$$

holds with some positive constants K_1, K_2, $L_0>0$ and $\theta\ge 2$. Then for all $\alpha\in(1,2)$ or for all $\alpha\in(0,1]$ with

$$\frac{\alpha C_{d,\alpha}\omega_d 3^{\alpha-1}}{8(2-\alpha)d}>K_1L_0^{\alpha},\tag{1.5}$$

there exists a positive constant $\lambda:=\lambda(\theta,K_1,K_2,L_0)>0$, such that for any $p\ge 1$ the following two statements hold:

(i) *if $\theta=2$, then for all $x,y\in\mathbb{R}^d$ and any $t>0$,*

$$W_p(\delta_x P_t,\delta_y P_t)\le C\mathrm{e}^{-\lambda t/p}\big(|x-y|^{1/p}\vee|x-y|\big);\tag{1.6}$$

(ii) *if $\theta>2$, then for all $x,y\in\mathbb{R}^d$ and any $t>0$,*

$$W_p(\delta_x P_t,\delta_y P_t)\le C\mathrm{e}^{-\lambda t/p}\left[\frac{|x-y|^{1/p}\vee|x-y|}{1+|x-y|\mathbf{1}_{(1,\infty)}(t)}\right],\tag{1.7}$$

where $C>0$ is a positive constant depending on θ, K_1, K_2, L_0 and p.

Theorem 1.2 above does provide new conditions on the drift term b such that the associated semigroup $(P_t)_{t\ge 0}$ is exponentially contractive with respect to the L^p-Wasserstein distance W_p for all $p\ge 1$. In particular, when $\alpha\in(1,2)$, the conclusion of Theorem 1.2 is the same as that of [15], Theorem 1.3, for diffusion processes; while for $\alpha\in(0,1]$ we need the restrictive condition (1.5); see Remark 3.3 for a further comment. Indeed, (1.5) is natural in the sense that, when $\alpha\in(0,1]$ the drift term plays the dominant role or the same role (just in case that $\alpha=1$) for the behavior of SDEs driven by symmetric α-stable processes, see, for example, [2,9] for (Dirichlet) heat kernel estimates and [24] for dimensional free Harnack inequalities on this topic. Similarly, in considering the exponential contractivity of SDE (1.1), we need (1.5) to control the locally non-dissipative part of the drift term. Note that (1.5) holds true when K_1, L_0 are small enough.

To show the power of Theorem 1.2, we consider the following example about the SDE driven by symmetric α-stable processes with $\alpha\in(0,2)$, which yields the exponential contractivity of the semigroup $(P_t)_{t\ge 0}$ with respect to the L^p-Wasserstein distance W_p $(p\ge 1)$ for super-convex potentials.

Example 1.3. Let $(Z_t)_{t\ge 0}$ be a symmetric α-stable process in $\mathbb{R}^d$ with $\alpha\in(0,2)$, and $b(x)=\nabla V(x)$ with $V(x)=-|x|^{2\beta}$ and $\beta>1$. Then there exists a constant $\lambda:=\lambda(\alpha,\beta)>0$ such that for all $p\ge 1$, $x,y\in\mathbb{R}^d$ and $t>0$,

$$W_p(\delta_x P_t,\delta_y P_t)\le C(\alpha,\beta,p)\mathrm{e}^{-\lambda t/p}\left[\frac{|x-y|^{1/p}\vee|x-y|}{1+|x-y|\mathbf{1}_{(1,\infty)}(t)}\right].$$

Note that the uniformly dissipative condition (1.2) fails for Example 1.3; see, for example, (3.8) below. That is, one cannot deduce directly from Theorem 1.1 the exponential contractivity with respect to the L^p-Wasserstein distance W_p for all $p\in[1,\infty)$.

The remainder of this paper is arranged as follows. In the next section, we will present the coupling by reflection mixed with the synchronous coupling for the SDE (1.1) driven by Lévy noise, and also prove the existence of coupling process associated with this coupling (operator). Section 3 is mainly devoted to the proof of Theorem 1.2. For this, we need more delicate choice of auxiliary functions and some key estimates for them, which are different between $\alpha\in(1,2)$ and $\alpha\in(0,1]$. The sketch of the proof of Example 1.3 is also given here.

2. Coupling operator and coupling process for SDEs with jumps

2.1. Coupling by reflection and synchronous coupling

It is easy to see that the generator of the process $(X_t)_{t\ge 0}$ acting on $C_b^2(\mathbb{R}^d)$ is

$$Lf(x)=\int\big(f(x+z)-f(x)-\langle\nabla f(x),z\rangle\mathbf{1}_{\{|z|\le 1\}}\big)\nu(\mathrm{d}z)+\langle b(x),\nabla f(x)\rangle. \tag{2.1}$$

In this part, we construct a coupling operator for the generator L above. For any x, y and $z\in\mathbb{R}^d$, we set

$$\varphi_{x,y}(z):=\begin{cases} z-\dfrac{2\langle x-y,z\rangle}{|x-y|^2}(x-y), & x\ne y;\\ -z, & x=y.\end{cases}$$

It is clear that $\varphi_{x,y}:\mathbb{R}^d\to\mathbb{R}^d$ has the following three properties:

(A1) $\varphi_{x,y}(z)=\varphi_{y,x}(z)$ and $\varphi^2_{x,y}(z)=z$, that is, $\varphi^{-1}_{x,y}(z)=\varphi_{x,y}(z)$;
(A2) $|\varphi_{x,y}(z)|=|z|$;
(A3) $(z-\varphi_{x,y}(z))\,/\!/\,(x-y)$ and $(z+\varphi_{x,y}(z))\perp(x-y)$.

Next, for any $f\in C_b^2(\mathbb{R}^{2d})$, let

$$\nabla_x f(x,y):=\left(\frac{\partial f(x,y)}{\partial x_i}\right)_{1\le i\le d},\qquad \nabla_y f(x,y):=\left(\frac{\partial f(x,y)}{\partial y_i}\right)_{1\le i\le d}.$$

Now, let L_0 be the constant appearing in (1.4). We will split the construction of the coupling operator into two parts, according to $x,y\in\mathbb{R}^d$ with $|x-y|\le L_0$ or with $|x-y|>L_0$. First, for any $f \in C_b^2(\mathbb{R}^{2d})$ and $x, y \in \mathbb{R}^d$ with $|x-y| \le L_0$, we define

$$\begin{aligned}
\widetilde{L} f(x,y) := \frac{1}{2}\bigg[\int_{\{|z|\le a|x-y|\}} &\big(f\big(x+z, y+\varphi_{x,y}(z)\big) - f(x,y) - \big\langle \nabla_x f(x,y), z\big\rangle \mathbf{1}_{\{|z|\le 1\}} \\
&- \big\langle \nabla_y f(x,y), \varphi_{x,y}(z)\big\rangle \mathbf{1}_{\{|z|\le 1\}}\big) \frac{C_{d,\alpha}}{|z|^{d+\alpha}}\,\mathrm{d}z \\
&+ \int_{\{|z|\le a|x-y|\}} \big(f\big(x+\varphi_{x,y}(z), y+z\big) - f(x,y) - \big\langle \nabla_y f(x,y), z\big\rangle \mathbf{1}_{\{|z|\le 1\}} \\
&- \big\langle \nabla_x f(x,y), \varphi_{x,y}(z)\big\rangle \mathbf{1}_{\{|z|\le 1\}}\big) \frac{C_{d,\alpha}}{|z|^{d+\alpha}}\,\mathrm{d}z\bigg] \\
&+ \int_{\{|z|\le a|x-y|\}} \big(f(x+z, y+z) - f(x,y) - \big\langle \nabla_x f(x,y), z\big\rangle \mathbf{1}_{\{|z|\le 1\}} \\
&- \big\langle \nabla_y f(x,y), z\big\rangle \mathbf{1}_{\{|z|\le 1\}}\big)\bigg(\nu(\mathrm{d}z) - \frac{C_{d,\alpha}}{|z|^{d+\alpha}}\,\mathrm{d}z\bigg) \\
&+ \int_{\{|z|> a|x-y|\}} \big(f(x+z, y+z) - f(x,y) - \big\langle \nabla_x f(x,y), z\big\rangle \mathbf{1}_{\{|z|\le 1\}} \\
&- \big\langle \nabla_y f(x,y), z\big\rangle \mathbf{1}_{\{|z|\le 1\}}\big)\nu(\mathrm{d}z) \\
&+ \big\langle \nabla_x f(x,y), b(x)\big\rangle + \big\langle \nabla_y f(x,y), b(y)\big\rangle,
\end{aligned}$$

where $a \in (0, 1/2)$ is a constant determined by later.

On the other hand, for any $f \in C_b^2(\mathbb{R}^{2d})$ and $x, y \in \mathbb{R}^d$ with $|x-y| > L_0$, we define

$$\begin{aligned}
\widetilde{L} f(x,y) := \int \big(&f(x+z, y+z) - f(x,y) - \big\langle \nabla_x f(x,y), z\big\rangle \mathbf{1}_{\{|z|\le 1\}} \\
&- \big\langle \nabla_y f(x,y), z\big\rangle \mathbf{1}_{\{|z|\le 1\}}\big)\nu(\mathrm{d}z) \\
&+ \big\langle \nabla_x f(x,y), b(x)\big\rangle + \big\langle \nabla_y f(x,y), b(y)\big\rangle.
\end{aligned}$$

We can conclude the following.

Proposition 2.1. *The operator $\widetilde{L}$ defined by above is the coupling operator of the operator L given by* (2.1).

Proof. Since $\widetilde{L}$ is a linear operator, it suffices to verify that

$$\widetilde{L} f(x) = L f(x), \qquad f \in C_b^2\big(\mathbb{R}^d\big), \tag{2.2}$$

where, on the left-hand side, f is regarded as a bivariate function on $\mathbb{R}^{2d}$, that is, $f(x) = f(x,y)$ for all $x, y \in \mathbb{R}^d$.

For any $x, y \in \mathbb{R}^d$ with $|x-y| > L_0$, it is trivial to see that (2.2) holds true, and so we only need to verify that for $x, y \in \mathbb{R}^d$ with $|x-y| \le L_0$. First, we have

$$\begin{aligned}
\widetilde{L} f(x) = {} & \frac{1}{2}\bigg[\int_{\{|z| \le a|x-y|\}} \big(f(x+z) - f(x) - \langle \nabla f(x), z\rangle \mathbf{1}_{\{|z|\le 1\}}\big) \frac{C_{d,\alpha}}{|z|^{d+\alpha}}\,\mathrm{d}z \\
& + \int_{\{|z| \le a|x-y|\}} \big(f\big(x+\varphi_{x,y}(z)\big) - f(x) - \langle \nabla f(x), \varphi_{x,y}(z)\rangle \mathbf{1}_{\{|z|\le 1\}}\big) \frac{C_{d,\alpha}}{|z|^{d+\alpha}}\,\mathrm{d}z\bigg] \\
& + \int_{\{|z| \le a|x-y|\}} \big(f(x+z) - f(x) - \langle \nabla f(x), z\rangle \mathbf{1}_{\{|z|\le 1\}}\big) \left(\nu(\mathrm{d}z) - \frac{C_{d,\alpha}}{|z|^{d+\alpha}}\,\mathrm{d}z\right) \\
& + \int_{\{|z| > a|x-y|\}} \big(f(x+z) - f(x) - \langle \nabla f(x), z\rangle \mathbf{1}_{\{|z|\le 1\}}\big) \nu(\mathrm{d}z) \\
& + \langle b(x), \nabla f(x)\rangle.
\end{aligned}$$

By (A1) and (A2), we know that the measure $\frac{C_{d,\alpha}}{|z|^{d+\alpha}}\,\mathrm{d}z$ is invariant under the transformation $z \mapsto \varphi_{x,y}(z)$. This, along with (A2) and the equality above, leads to

$$\begin{aligned}
\widetilde{L} f(x) = {} & \int_{\{|z| \le a|x-y|\}} \big(f(x+z) - f(x) - \langle \nabla f(x), z\rangle \mathbf{1}_{\{|z|\le 1\}}\big) \frac{C_{d,\alpha}}{|z|^{d+\alpha}}\,\mathrm{d}z \\
& + \int_{\{|z| \le a|x-y|\}} \big(f(x+z) - f(x) - \langle \nabla f(x), z\rangle \mathbf{1}_{\{|z|\le 1\}}\big) \left(\nu(\mathrm{d}z) - \frac{C_{d,\alpha}}{|z|^{d+\alpha}}\,\mathrm{d}z\right) \\
& + \int_{\{|z| > a|x-y|\}} \big(f(x+z) - f(x) - \langle \nabla f(x), z\rangle \mathbf{1}_{\{|z|\le 1\}}\big) \nu(\mathrm{d}z) \\
& + \langle b(x), \nabla f(x)\rangle \\
= {} & Lf(x).
\end{aligned}$$

This completes the proof. □

Remark 2.2. (1) Here, we give an interpretation of the construction of the coupling operator $\widetilde{L}$ above. If $|x-y| > L_0$, we use the synchronous coupling. If $|x-y| \le L_0$, then the coupling operator $\widetilde{L}$ constructed above consists of two parts. Fix any $x, y \in \mathbb{R}^d$. If $|z| \le a|x-y|$, then we adopt the coupling by reflection by making full use of the rotationally invariant measure $\frac{C_{d,\alpha}}{|z|^{d+\alpha}}\,\mathrm{d}z$; while for the remainder term, we use the synchronous coupling again, where the components maintain at each step the same length of jumps (i.e., from (x, y) to $(x+z, y+z)$) with the biggest rate $\nu(\mathrm{d}z)$ when $|z| > a|x-y|$, and with the rate $\nu(\mathrm{d}z) - \frac{C_{d,\alpha}}{|z|^{d+\alpha}}\,\mathrm{d}z$ when $|z| \le a|x-y|$. For the coupling by reflection for Brownian motion and diffusion processes, we refer to [6,14,22].

(2) Recently, the coupling property of Lévy processes has been developed in [4,17,18]. The corresponding property for Ornstein–Uhlenbeck processes with jumps also has been successfully studied in [19,23]. Unlike Lévy processes and Ornstein–Uhlenbeck processes with jumps, it is impossible to write out an explicit expression for transition functions of the solution to the SDE

(1.1) with general drift term $b(x)$. This observation indicates that all the approaches in [4,17–19,23] are not efficient in the present setting. This difficulty will be overcome by constructing proper coupling operators for the Markov generator corresponding to the solution of the SDE (1.1), as done in [25]. However, different from [25] which deals with the corresponding coupling property by making full use of large jumps part of Lévy processes, here to consider the exponential contractivity of the associated semigroups $(P_t)_{t\ge 0}$ with respect to Wasserstein distances we need a new construction of the coupling operator. As seen from Propositions 3.1 and 3.2 below, the coupling for small jumps part of Lévy processes [i.e., the coupling by reflection as mentioned in (1)] is key for our purpose.

2.2. Coupling process

In this part, we will construct a coupling process associated with the coupling operator $\widetilde{L}$. For this, we will frequently talk about the martingale problem for the operator L given by (2.1) and the coupling operator $\widetilde{L}$. Let $\mathcal{D}([0,\infty);\mathbb{R}^d)$ be the space of right continuous $\mathbb{R}^d$-valued functions having left limits on $[0,\infty)$, equipped with the Skorokhod topology. For $t\ge 0$, denote by X_t the projection coordinate map on $\mathcal{D}([0,\infty);\mathbb{R}^d)$. A probability measure $\mathbb{P}^x$ on the Skorokhod space $\mathcal{D}([0,\infty);\mathbb{R}^d)$ is said to be a solution to the martingale problem for $(L, C_c^2(\mathbb{R}^d))$ with initial value $x\in\mathbb{R}^d$ if $\mathbb{P}^x(X_0=x)=1$ and for every $f\in C_c^2(\mathbb{R}^d)$

$$\left\{f(X_t)-f(x)-\int_0^t Lf(X_s)\,\mathrm{d}s, t\ge 0\right\}$$

is a $\mathbb{P}^x$-martingale. The martingale problem for $(L, C_c^2(\mathbb{R}^d))$ is said to be well-posed if it has a unique solution for every initial value $x\in\mathbb{R}^d$. Similarly, we can define a solution to the martingale problem for the coupling operator $\widetilde{L}$ on $C_c^2(\mathbb{R}^{2d})$. Note that, in [13] an equivalence is proved between the existence of weak solutions to SDEs with jumps and the existence of solutions to the corresponding martingale problem, by using a martingale representation theorem. Recently, Kurtz [12] studied equivalence between the uniqueness (in sense of distribution) of weak solutions to a class of SDEs driven by Poisson random measures and the well-posed solution to martingale problems for a class of non-local operators using a non-constructive approach. Note that in our setting the SDE (1.1) has the pathwise unique strong solution. According to [1], Theorem 1, page 2, the weak solution to the SDE (1.1) enjoys the unique (in sense of distribution) weak solution. This, along with [12], Corollary 2.5, yields that the martingale problem for $(L, C_c^2(\mathbb{R}^d))$ is well posed.

Let L_0, a be the constants in the definition of the coupling operator $\widetilde{L}$. For any $x, y\in\mathbb{R}^d$ and $A\in\mathcal{B}(\mathbb{R}^{2d})$, set

$$\begin{aligned}\mu(x,y,A) := {} & \frac{1}{2}\int_{\{(z,\varphi_{x,y}(z))\in A, |z|\le a|x-y|, |x-y|\le L_0\}}\frac{C_{d,\alpha}}{|z|^{d+\alpha}}\,\mathrm{d}z\\ & + \frac{1}{2}\int_{\{(\varphi_{x,y}(z),z)\in A, |z|\le a|x-y|, |x-y|\le L_0\}}\frac{C_{d,\alpha}}{|z|^{d+\alpha}}\,\mathrm{d}z\end{aligned}$$

$$+\int_{\{(z,z)\in A,|z|\le a|x-y|,|x-y|\le L_0\}}\left(\nu(\mathrm{d}z)-\frac{C_{d,\alpha}}{|z|^{d+\alpha}}\,\mathrm{d}z\right)$$
$$+\int_{\{(z,z)\in A,|z|>a|x-y|,|x-y|\le L_0\}\cup\{(z,z)\in A,|x-y|>L_0\}}\nu(\mathrm{d}z).$$

Then, for any $x, y\in\mathbb{R}^d$ and $f\in C_b^2(\mathbb{R}^{2d})$, we have

$$\begin{aligned}
&\widetilde{L}f(x,y)\\
&\quad=\int_{\mathbb{R}^{2d}}\Big[f\big((x,y)+(u_1,u_2)\big)-f(x,y)\\
&\qquad\qquad-\big\langle\big(\nabla_x f(x,y),\nabla_y f(x,y)\big),(u_1,u_2)\big\rangle\mathbf{1}_{\{|u_1|\le1,|u_2|\le1\}}\Big]\mu(x,y,\mathrm{d}u_1,\mathrm{d}u_2)\\
&\qquad+\big\langle\nabla_x f(x,y),b(x)\big\rangle+\big\langle\nabla_y f(x,y),b(y)\big\rangle.
\end{aligned}$$

Furthermore, for any $h\in C_b(\mathbb{R}^{2d})$, by (A2),

$$\begin{aligned}
&\int_{\mathbb{R}^{2d}}h(u)\frac{|u|^2}{1+|u|^2}\mu(x,y,\mathrm{d}u)\\
&\quad=\int_{\{|z|\le a|x-y|,|x-y|\le L_0\}}h\big(z,\varphi_{x,y}(z)\big)\frac{|z|^2}{1+2|z|^2}\frac{C_{d,\alpha}}{|z|^{d+\alpha}}\,\mathrm{d}z\\
&\qquad+\int_{\{|z|\le a|x-y|,|x-y|\le L_0\}}h\big(\varphi_{x,y}(z),z\big)\frac{|z|^2}{1+2|z|^2}\frac{C_{d,\alpha}}{|z|^{d+\alpha}}\,\mathrm{d}z\\
&\qquad+2\int_{\{|z|\le a|x-y|,|x-y|\le L_0\}}h(z,z)\frac{|z|^2}{1+2|z|^2}\left(\nu(\mathrm{d}z)-\frac{C_{d,\alpha}}{|z|^{d+\alpha}}\,\mathrm{d}z\right)\\
&\qquad+2\int_{\{|z|>a|x-y|,|x-y|\le L_0\}\cup\{|x-y|>L_0\}}h(z,z)\frac{|z|^2}{1+2|z|^2}\nu(\mathrm{d}z),
\end{aligned}$$

which implies that $(x,y)\mapsto\int h(u)\frac{|u|^2}{1+|u|^2}\mu(x,y,\mathrm{d}u)$ is a continuous function on $\mathbb{R}^{2d}$. Note that $b(x)$ is a continuous function on $\mathbb{R}^d$. According to [20], Theorem 2.2, there is a solution to the martingale problem for $\widetilde{L}$, that is, there exist a probability space $(\widetilde{\Omega},\widetilde{\mathcal{F}},(\widetilde{\mathcal{F}}_t)_{t\ge0},\widetilde{\mathbb{P}})$ and an $\bar{\mathbb{R}}^{2d}$-valued process $(\widetilde{X}_t)_{t\ge0}$ such that $(\widetilde{X}_t)_{t\ge0}$ is $(\widetilde{\mathcal{F}}_t)_{t\ge0}$-progressively measurable, and for every $f\in C_b^2(\mathbb{R}^{2d})$,

$$\left\{f(\widetilde{X}_t)-\int_0^{t\wedge e}\widetilde{L}f(\widetilde{X}_u)\,\mathrm{d}u,\ t\ge0\right\}$$

is an $(\widetilde{\mathcal{F}}_t)_{t\ge0}$-local martingale, where e is the explosion time of $(\widetilde{X}_t)_{t\ge0}$, that is,

$$e=\lim_{n\to\infty}\inf\big\{t\ge0\colon|\widetilde{X}_t|\ge n\big\}.$$

Let $(\widetilde{X}_t)_{t\ge 0} := (X'_t, X''_t)_{t\ge 0}$. Then $(X'_t)_{t\ge 0}$ and $(X''_t)_{t\ge 0}$ are two stochastic processes on $\mathbb{R}^d$. Since $\widetilde{L}$ is the coupling operator of L, the generator of each marginal process $(X'_t)_{t\ge 0}$ and $(X''_t)_{t\ge 0}$ is just the operator L, and hence both distributions of the processes $(X'_t)_{t\ge 0}$ and $(X''_t)_{t\ge 0}$ are solutions to the martingale problem of L. In particular, by our assumption and the remark in the beginning of this subsection, the processes $(X'_t)_{t\ge 0}$ and $(X''_t)_{t\ge 0}$ are non-explosive, hence one has $e=\infty$ a.s. Therefore, the coupling operator $\widetilde{L}$ generates a non-explosive process $(\widetilde{X}_t)_{t\ge 0}$.

Let T be the coupling time of $(X'_t)_{t\ge 0}$ and $(X''_t)_{t\ge 0}$, that is,

$$T = \inf\{t \ge 0\colon\ X'_t = X''_t\}.$$

Then T is an $(\widetilde{\mathcal{F}}_t)_{t\ge 0}$-stopping time. Define a new process $(Y'_t)_{t\ge 0}$ as follows:

$$Y'_t = \begin{cases} X''_t, & t < T; \\ X'_t, & t \ge T. \end{cases}$$

For any $f \in C_b^2(\mathbb{R}^d)$ and $t > 0$,

$$\begin{aligned} f(Y'_t) - \int_0^t Lf(Y'_s)\,\mathrm{d}s &= f(Y'_{t\wedge T}) - \int_0^{t\wedge T} Lf(Y'_s)\,\mathrm{d}s \\ &\quad + f(Y'_t) - f(Y'_{t\wedge T}) - \int_{t\wedge T}^t Lf(Y'_s)\,\mathrm{d}s \\ &= f(X''_{t\wedge T}) - \int_0^{t\wedge T} Lf(X''_s)\,\mathrm{d}s \\ &\quad + f(X'_t) - f(X'_{t\wedge T}) - \int_{t\wedge T}^t Lf(X'_s)\,\mathrm{d}s \\ &=: M_t^1 + M_t^2. \end{aligned}$$

By the optimal stopping theorem and the facts that both $(X'_t)_{t\ge 0}$ and $(X''_t)_{t\ge 0}$ are solutions to the martingale problem of L, $(M_t^1)_{t\ge 0}$ and $(M_t^2)_{t\ge 0}$ are martingales and so is $(Y'_t)_{t\ge 0}$ (see, e.g., [16], Section 3.1, page 251). Since the martingale problem for the operator L is well-posed, $(Y'_t)_{t\ge 0}$ and $(X''_t)_{t\ge 0}$ are equal in the distribution. Therefore, we conclude that $(X'_t, Y'_t)_{t\ge 0}$ is also a non-explosive coupling process of $(X_t)_{t\ge 0}$ such that $X'_t = Y'_t$ for any $t \ge T$ and the generator of $(X'_t, Y'_t)_{t\ge 0}$ before the coupling time T is just the coupling operator $\widetilde{L}$. In particular, according to [8], Lemma 2.1, we know that for any $x, y \in \mathbb{R}^d$ and $f \in B_b(\mathbb{R}^d)$,

$$P_t f(x) = \mathbb{E}^x f(X'_t) = \widetilde{\mathbb{E}}^{(x,y)} f(X'_t)$$

and

$$P_t f(y) = \mathbb{E}^y f(Y'_t) = \widetilde{\mathbb{E}}^{(x,y)} f(Y'_t),$$

where $\widetilde{\mathbb{E}}^{(x,y)}$ is the expectation of the process $(X'_t, Y'_t)_{t\ge 0}$ with starting point (x, y).

3. Proofs

3.1. Key estimates

We first assume that $\alpha \in (1,2)$. For any $r > 0$, define

$$\psi(r) := \begin{cases} 1 - \mathrm{e}^{-c_1 r}, & r \in [0, 2L_0]; \\ A\mathrm{e}^{c_2(r-2L_0)} + B(r-2L_0)^2 + \big(1 - \mathrm{e}^{-2c_1 L_0} - A\big), & r \in [2L_0, \infty), \end{cases}$$

where

$$A = \frac{c_1}{c_2}\mathrm{e}^{-2L_0 c_1}, \qquad B = -\frac{(c_1+c_2)c_1}{2}\mathrm{e}^{-2L_0 c_1},$$

c_2 is a positive constant such that $c_2 \geq 20c_1$, that is,

$$\log\frac{2(c_1+c_2)}{c_2} \leq \log 2.1,$$

and c_1 is a positive constant determined by later. With the choice of the constants A and B above, it is easy to see that $\psi \in C^2([0,\infty))$. Then we have:

Proposition 3.1. *Assume that* $\alpha \in (1,2)$. *Then there exists a constant* $\lambda > 0$ *such that for any* $x, y \in \mathbb{R}^d$,

$$\widetilde{L}\psi\big(|x-y|\big) \leq -\lambda\psi\big(|x-y|\big).$$

Proof. (1) In this part, we treat the case that $x, y \in \mathbb{R}^d$ with $|x-y| \leq L_0$. First, for any $x, y, z \in \mathbb{R}^d$, by (A3),

$$\big\langle x-y, z+\varphi_{x,y}(z)\big\rangle = 0$$

and so

$$\big\langle \nabla_x\psi\big(|x-y|\big), z+\varphi_{x,y}(z)\big\rangle = 0 \quad \text{and} \quad \big\langle \nabla_y\psi\big(|x-y|\big), z+\varphi_{x,y}(z)\big\rangle = 0.$$

Therefore,

$$\begin{aligned} &\widetilde{L}\psi\big(|x-y|\big) \\ &\quad = \frac{1}{2}\bigg[\int_{\{|z|\leq a|x-y|\}}\Big(\psi\big(\big|x-y+\big(z-\varphi_{x,y}(z)\big)\big|\big) + \psi\big(\big|x-y-\big(z-\varphi_{x,y}(z)\big)\big|\big) \\ &\qquad - 2\psi\big(|x-y|\big)\Big)\frac{C_{d,\alpha}}{|z|^{d+\alpha}}\,\mathrm{d}z\bigg] \\ &\qquad + \psi'\big(|x-y|\big)\frac{\langle b(x)-b(y), x-y\rangle}{|x-y|}. \end{aligned}$$

It is easy to see that $\psi \in C^3([0,2L_0))$ such that $\psi' > 0$, $\psi'' < 0$ and $\psi''' > 0$ on $[0,2L_0)$. Then, for any $0 \le \delta < r \le L_0$,

$$\psi(r+\delta)+\psi(r-\delta)-2\psi(r)=\int_r^{r+\delta}\mathrm{d}s\int_{s-\delta}^{s}\psi''(u)\,\mathrm{d}u\le\psi''(r+\delta)\delta^2,$$

where in the inequality we have used the fact that $\psi''' > 0$ on $[0,2L_0)$. Hence, according to the definition of $\varphi_{x,y}(z)$ and the inequality above, for all $x,y,z\in\mathbb{R}^d$ with $|x-y|\le L_0$ and $|z|\le a|x-y|$ with $a\in(0,1/2)$, we have

$$\begin{aligned}
&\psi\big(\big|x-y+\big(z-\varphi_{x,y}(z)\big)\big|\big)+\psi\big(\big|x-y-\big(z-\varphi_{x,y}(z)\big)\big|\big)-2\psi\big(|x-y|\big)\\
&\quad=\psi\bigg(|x-y|+\frac{2\langle x-y,z\rangle}{|x-y|}\bigg)+\psi\bigg(|x-y|-\frac{2\langle x-y,z\rangle}{|x-y|}\bigg)-2\psi\big(|x-y|\big)\\
&\quad\le 4\psi''\big((1+2a)|x-y|\big)\frac{\langle x-y,z\rangle^2}{|x-y|^2}.
\end{aligned}\tag{3.1}$$

Then we deduce that for any $x,y\in\mathbb{R}^d$ with $|x-y|\le L_0$,

$$\begin{aligned}
\widetilde{L}\psi\big(|x-y|\big)&\le 2\psi''\big((1+2a)|x-y|\big)\int_{\{|z|\le a|x-y|\}}\frac{|\langle x-y,z\rangle|^2}{|x-y|^2}\frac{C_{d,\alpha}}{|z|^{d+\alpha}}\,\mathrm{d}z\\
&\quad+\psi'\big(|x-y|\big)\frac{\langle b(x)-b(y),x-y\rangle}{|x-y|}\\
&=2\psi''\big((1+2a)|x-y|\big)\int_{\{|z|\le a|x-y|\}}|z_1|^2\frac{C_{d,\alpha}}{|z|^{d+\alpha}}\,\mathrm{d}z\\
&\quad+\psi'\big(|x-y|\big)\frac{\langle b(x)-b(y),x-y\rangle}{|x-y|}\\
&=\frac{2C_{d,\alpha}}{d}\psi''\big((1+2a)|x-y|\big)\int_{\{|z|\le a|x-y|\}}|z|^2\frac{1}{|z|^{d+\alpha}}\,\mathrm{d}z\\
&\quad+\psi'\big(|x-y|\big)\frac{\langle b(x)-b(y),x-y\rangle}{|x-y|}\\
&\le\bigg[-\frac{2C_{d,\alpha}\omega_d L_0^{1-\alpha}}{d(2-\alpha)}c_1a^{2-\alpha}\mathrm{e}^{-2c_1aL_0}+K_1\bigg]c_1\mathrm{e}^{-c_1|x-y|}|x-y|,
\end{aligned}\tag{3.2}$$

where in the inequality z_1 denotes the first coordinate of z, that is, $z=(z_1,z_2,\dots,z_d)$, both equalities above follow from the rotationally invariant property of the measure $\frac{C_{d,\alpha}}{|z|^{d+\alpha}}\,\mathrm{d}z$, and in the last inequality we have used (1.4) and the fact that $\alpha>1$.

Now, taking

$$C=\frac{2C_{d,\alpha}\omega_d L_0^{1-\alpha}}{d(2-\alpha)},\qquad c_1=(2K_1/C)^{1/(\alpha-1)}\mathrm{e}^{2L_0/(\alpha-1)}+2,\qquad a=1/c_1,$$

we find that for any $x, y \in \mathbb{R}^d$ with $|x-y| \le L_0$,

$$\widetilde{L}\psi\big(|x-y|\big) \le -\frac{C}{2}c_1^\alpha e^{-2L_0}e^{-c_1|x-y|}|x-y|.$$

Since $\psi(0)=0$ and $\psi'' \le 0$ on $[0, 2L_0)$,

$$\psi(r) \le \psi'(r)r = c_1 e^{-c_1 r} r, \qquad r \in [0, L_0],$$

which along with the estimate above yields that for any $x, y \in \mathbb{R}^d$ with $|x-y| \le L_0$,

$$\widetilde{L}\psi\big(|x-y|\big) \le -\lambda_1\psi\big(|x-y|\big),$$

where $\lambda_1 = Cc_1^{\alpha-1}e^{-2L_0}/2$.

(2) Second, we consider the case that $x, y \in \mathbb{R}^d$ with $|x-y| > L_0$. For any $x, y \in \mathbb{R}^d$ with $L_0 < |x-y| \le 2L_0$, by (1.4) and the fact that $\psi' > 0$,

$$\widetilde{L}\psi\big(|x-y|\big) \le -c_1 K_2 e^{-c_1|x-y|}|x-y|^{\theta-1} \le -c_1 K_2 L_0^{\theta-2} e^{-c_1|x-y|}|x-y|.$$

On the other hand, also by (1.4) and the fact that $\psi' > 0$, for any $x, y \in \mathbb{R}^d$ with $|x-y| \ge 2L_0$,

$$\widetilde{L}\psi\big(|x-y|\big) \le -K_2\big[Ac_2 e^{c_2(|x-y|-2L_0)} + 2B\big(|x-y|-2L_0\big)\big]|x-y|^{\theta-1}.$$

Next, we consider the function

$$g(r) = \tfrac{1}{2}Ac_2 e^{c_2(r-2L_0)} + 2B(r-2L_0)$$

on $[2L_0, \infty)$. It is easy to see that due to the definitions of the constants A and B, there is a unique $r_1 \in [2L_0, \infty)$ such that $g'(r_1) = 0$ and

$$g(r_1) = \frac{-2B}{c_2}\left[1 - \log\frac{-4B}{Ac_2^2}\right] = \frac{-2B}{c_2}\left[1 - \log\frac{2(c_1+c_2)}{c_2}\right].$$

Since $c_2 > 0$ is large enough such that

$$\log\frac{2(c_1+c_2)}{c_2} \le \log 2.1,$$

we have $g(r_1) > 0$, which implies that $g(r) > 0$ for all $r \in [2L_0, \infty)$. In particular,

$$\tfrac{1}{2}Ac_2 e^{c_2(|x-y|-2L_0)} + 2B\big(|x-y|-2L_0\big) \ge 0$$

for any $x, y \in \mathbb{R}^d$ with $|x-y| \ge 2L_0$. That is, for any $x, y \in \mathbb{R}^d$ with $|x-y| \ge 2L_0$,

$$\widetilde{L}\psi\big(|x-y|\big) \le -\tfrac{1}{2}K_2 Ac_2 e^{c_2(|x-y|-2L_0)}|x-y|^{\theta-1} \le 2^{\theta-3}K_2 Ac_2 L_0^{\theta-2} e^{c_2(|x-y|-2L_0)}|x-y|.$$

Combining both estimates above with the definition of ψ, we finally conclude that there is a constant $\lambda_2 > 0$ such that for any $x, y \in \mathbb{R}^d$ with $|x-y| > L_0$,

$$\widetilde{L}\psi\big(|x-y|\big) \le -\lambda_2\psi\big(|x-y|\big).$$

This along with the conclusion of part (1) yields the desired assertion. □

Next, we turn to the case of $\alpha \in (0,1]$. For this, we first take the constant $a = \frac{1}{4}$ in the definition of the coupling operator $\widetilde{L}$, and then change the test function ψ as follows, which is different from that in the case $\alpha \in (1,2)$. For any $r > 0$, we define

$$\psi(r) := \begin{cases} r - cr^{1+\alpha}, & r \in [0, 2L_0]; \\ A\mathrm{e}^{c_0(r-2L_0)} + B(r-2L_0)^2 + \big(2L_0 - c(2L_0)^{1+\alpha} - A\big), & r \in [2L_0, \infty), \end{cases}$$

where

$$c = \frac{1}{2^{1+\alpha}(1+\alpha)L_0^\alpha}, \qquad A = \frac{1}{2c_0}, \qquad B = -\frac{1}{2}\left[\frac{\alpha}{4L_0} + \frac{c_0}{2}\right], \qquad c_0 = \frac{10\alpha}{L_0}.$$

Due to the choice of the constants above, $\psi \in C^2([0,\infty))$ and $\psi'(r) > 0$ for all $r > 0$.

Proposition 3.2. *Assume that $\alpha \in (0,1]$. If*

$$\frac{\alpha C_{d,\alpha}\omega_d 3^{\alpha-1}}{8(2-\alpha)d} > K_1 L_0^\alpha, \tag{3.3}$$

then there exists a constant $\lambda > 0$ such that for any $x, y \in \mathbb{R}^d$ with $x \ne y$,

$$\widetilde{L}\psi\big(|x-y|\big) \le -\lambda\psi\big(|x-y|\big).$$

Proof. We mainly follow the proof of Proposition 3.1, and here we only present the main different steps. For $x, y \in \mathbb{R}^d$ with $|x-y| \le L_0$, we have

$$\begin{aligned} &\widetilde{L}\psi\big(|x-y|\big) \\ &\quad= \frac{1}{2}\bigg[\int_{\{|z| \le (1/4)|x-y|\}} \Big(\psi\big(\big|x-y+\big(z-\varphi_{x,y}(z)\big)\big|\big) + \psi\big(\big|x-y-\big(z-\varphi_{x,y}(z)\big)\big|\big) \\ &\qquad - 2\psi\big(|x-y|\big)\Big)\frac{C_{d,\alpha}}{|z|^{d+\alpha}}\,\mathrm{d}z\bigg] \\ &\qquad + \psi'\big(|x-y|\big)\frac{\langle b(x)-b(y), x-y\rangle}{|x-y|}. \end{aligned}$$

Since $\psi \in C^3((0,2L_0))$ such that $\psi' > 0$, $\psi'' < 0$ and $\psi''' > 0$ on $(0,2L_0)$, one can follow the proof of (3.1), and get that for all $x, y, z \in \mathbb{R}^d$ with $0 < |x-y| \le L_0$ and $|z| \le \frac{1}{4}|x-y|$,

$$\begin{aligned} &\psi\big(\big|x-y+\big(z-\varphi_{x,y}(z)\big)\big|\big) + \psi\big(\big|x-y-\big(z-\varphi_{x,y}(z)\big)\big|\big) - 2\psi\big(|x-y|\big) \\ &\quad \le 4\psi''\bigg(\frac{3}{2}|x-y|\bigg)\frac{\langle x-y, z\rangle^2}{|x-y|^2}. \end{aligned}$$

Then we follow the argument of (3.2) and deduce that for any $x, y \in \mathbb{R}^d$ with $0 < |x-y| \le L_0$,

$$\widetilde{L}\psi\big(|x-y|\big) \le \left[-\frac{\alpha C_{d,\alpha}\omega_d 3^{\alpha-1}}{8(2-\alpha)dL_0^\alpha} + K_1\right]|x-y|.$$

By assumption (3.3), we know that for all $x, y \in \mathbb{R}^d$ with $0 < |x-y| \le L_0$,

$$\begin{aligned}\widetilde{L}\psi\big(|x-y|\big) &\le -\left(\frac{\alpha C_{d,\alpha}\omega_d 3^{\alpha-1}}{8(2-\alpha)dL_0^\alpha} - K_1\right)|x-y| \le -\left(\frac{\alpha C_{d,\alpha}\omega_d 3^{\alpha-1}}{8(2-\alpha)dL_0^\alpha} - K_1\right)\psi\big(|x-y|\big)\\ &=: -\lambda_1\psi\big(|x-y|\big).\end{aligned}$$

Next, we turn to the case that $x, y \in \mathbb{R}^d$ with $|x-y| > L_0$. For any $x, y \in \mathbb{R}^d$ with $L_0 < |x-y| \le 2L_0$, by (1.4) and $\psi' > 0$,

$$\begin{aligned}\widetilde{L}\psi\big(|x-y|\big) &\le -K_2\big(1 - c(1+\alpha)|x-y|^\alpha\big)|x-y|^{\theta-1}\\ &\le -K_2L_0^{\theta-2}\left(1 - \frac{1}{2^{1+\alpha}L_0^\alpha}|x-y|^\alpha\right)|x-y|.\end{aligned}$$

On the other hand, also by (1.4) and the fact that $\psi' > 0$, for any $x, y \in \mathbb{R}^d$ with $|x-y| \ge 2L_0$,

$$\widetilde{L}\psi\big(|x-y|\big) \le -K_2\big[Ac_0\mathrm{e}^{c_0(|x-y|-2L_0)} + 2B\big(|x-y| - 2L_0\big)\big]|x-y|^{\theta-1}.$$

Now, we consider again the function

$$g(r) = \tfrac{1}{2}Ac_0\mathrm{e}^{c_0(r-2L_0)} + 2B(r-2L_0)$$

on $[2L_0, \infty)$. It is easy to see that due to the definitions of the constants A and B, there is a unique $r_1 \in [2L_0, \infty)$ such that $g'(r_1) = 0$ and

$$g(r_1) = \frac{-2B}{c_0}\left[1 - \log\frac{-4B}{Ac_0^2}\right] = \frac{-2B}{c_0}\left[1 - \log\left(2 + \frac{\alpha}{L_0c_0}\right)\right].$$

Noticing that $c_0 = 10\alpha {L_0}^{-1}$, we get

$$\log\left(2 + \frac{\alpha}{L_0c_0}\right) = \log 2.1,$$

and so $g(r_1) > 0$, which implies that $g(r) > 0$ for all $r \in [2L_0, \infty)$. In particular,

$$\tfrac{1}{2}Ac_0\mathrm{e}^{c_0(|x-y|-2L_0)} + 2B\big(|x-y| - 2L_0\big) \ge 0$$

for any $x, y \in \mathbb{R}^d$ with $|x-y| \ge 2L_0$. That is, for any $x, y \in \mathbb{R}^d$ with $|x-y| \ge 2L_0$,

$$\widetilde{L}\psi\big(|x-y|\big) \le -\tfrac{1}{2}K_2Ac_0\mathrm{e}^{c_0(|x-y|-2L_0)}|x-y|^{\theta-1}.$$

According to both estimates above and the definition of ψ, we finally conclude that there is a constant $\lambda_2 > 0$ such that for any $x, y \in \mathbb{R}^d$ with $|x-y| > L_0$,

$$\widetilde{L}\psi\big(|x-y|\big) \le -\lambda_2 \psi\big(|x-y|\big).$$

This along with the conclusion above yields the desired assertion. □

Remark 3.3. According to the argument above, we can easily improve (3.3), for example, by taking $\psi(r) = c_1 r - c_2 r^{1+\alpha'}$ for $r \in [0, 2L_0]$ and changing the integral domain $\{z\colon |z| \le \frac{1}{4}|x-y|\}$ in the definition of the coupling operator $\widetilde{L}$ into $\{z\colon |z| \le a|x-y|\}$ with some proper choices of $c_1, c_2 > 0$, $\alpha' \in (0, \alpha]$ and $a \in (0, 1/2)$. For simplicity, here we just set $c_1 = 1$, $c_2 = c$, $\alpha' = \alpha$ and $a = 1/4$.

3.2. Proofs of Theorem 1.2 and Example 1.3

We divide the proof of Theorem 1.2 into two parts.

Proof of Theorem 1.2 for $|x-y| \le L_0$ or $\theta = 2$. We will make full use of the coupling process $(X'_t, Y'_t)_{t\ge 0}$ constructed in Section 2.2. Denote by $\widetilde{\mathbb{P}}^{(x,y)}$ and $\widetilde{\mathbb{E}}^{(x,y)}$ the distribution and the expectation of $(X'_t, Y'_t)_{t\ge 0}$ starting from (x, y), respectively. For any $t > 0$ set $r_t = |X'_t - Y'_t|$, and for $n \ge 1$ define the stopping time

$$T_n = \inf\big\{t > 0\colon r_t \notin [1/n, n]\big\}.$$

For any $x, y \in \mathbb{R}^d$ with $|x-y| > 0$, we take n large enough such that $1/n < |x-y| < n$. Let ψ be the function given in Proposition 3.1 if $\alpha \in (1, 2)$ or the function given in Proposition 3.2 if $\alpha \in (0, 1]$. Then

$$\begin{aligned}
&\widetilde{\mathbb{E}}^{(x,y)}\psi\big(\big|X'_{t\wedge T_n} - Y'_{t\wedge T_n}\big|\big)\\
&\quad = \psi\big(|x-y|\big) + \widetilde{\mathbb{E}}^{(x,y)}\bigg(\int_0^{t\wedge T_n} \widetilde{L}\psi\big(\big|X'_s - Y'_s\big|\big)\,\mathrm{d}s\bigg)\\
&\quad \le \psi\big(|x-y|\big) - \lambda\widetilde{\mathbb{E}}^{(x,y)}\bigg(\int_0^{t} \psi\big(\big|X'_{s\wedge T_n} - Y'_{s\wedge T_n}\big|\big)\,\mathrm{d}s\bigg).
\end{aligned}$$

Therefore,

$$\mathbb{E}\big[\psi(r_{t\wedge T_n})\big] \le \psi(r_0)\mathrm{e}^{-\lambda t}.$$

Since the coupling process $(X'_t, Y'_t)_{t\ge 0}$ is non-explosive, we have $T_n \uparrow T$ a.s. as $n \to \infty$, where T is the coupling time of the process (X'_t, Y'_t). Thus, by Fatou's lemma, letting $n \to \infty$ in the above inequality gives us

$$\mathbb{E}\big[\psi(r_{t\wedge T})\big] \le \psi(r_0)\mathrm{e}^{-\lambda t}.$$

Thanks to our convention that $Y'_t = X'_t$ for $t \ge T$, we have $r_t = 0$ for all $t \ge T$, and so

$$\mathbb{E}\psi(r_t) \le \psi(r_0)\mathrm{e}^{-\lambda t}.$$

That is,

$$\mathbb{E}\psi\big(|X_t - Y_t|\big) \le \psi\big(|x-y|\big)\mathrm{e}^{-\lambda t}.$$

As a result, if $|x-y| \le L_0$, then for any $p \ge 1$ and $t > 0$,

$$\mathbb{E}|X_t - Y_t|^p \le C(p)\mathbb{E}\psi\big(|X_t - Y_t|\big) \le C_1\mathrm{e}^{-\lambda t}|x-y|, \tag{3.4}$$

where the first inequality follows from the definitions of the test function ψ in Propositions 3.1 and 3.2.

Now for any $x, y \in \mathbb{R}^d$ with $|x-y| > L_0$, take $n := [|x-y|/L_0] + 1 \ge 2$. We have

$$\frac{n}{2} \le n-1 \le \frac{|x-y|}{L_0} \le n. \tag{3.5}$$

Set $x_i = x + i(y-x)/n$ for $i = 0, 1, \dots, n$. Then $x_0 = x$ and $x_n = y$; moreover, (3.5) implies $|x_{i-1} - x_i| = |x-y|/n \le L_0$ for all $i = 1, 2, \dots, n$. Therefore, by (3.4) and (3.5),

$$\begin{aligned}
W_p(\delta_x P_t, \delta_y P_t) &\le \sum_{i=1}^n W_p(\delta_{x_{i-1}} P_t, \delta_{x_i} P_t) \\
&\le C_1^{1/p}\mathrm{e}^{-\lambda t/p} \sum_{i=1}^n |x_{i-1} - x_i|^{1/p} \\
&\le C_1^{1/p}\mathrm{e}^{-\lambda t/p} n L_0^{1/p} \\
&\le 2C_1^{1/p} L_0^{1/p-1}\mathrm{e}^{-\lambda t/p}|x-y| \\
&=: C_2\mathrm{e}^{-\lambda t/p}|x-y|.
\end{aligned}$$

In particular, the proof of the first assertion for $\theta = 2$ in Theorem 1.2 is completed. On the other hand, from (3.4) and the conclusion above, we also get the second assertion for $\theta > 2$ with $|x-y| \le 1$ and all $t > 0$, or with $|x-y| > 1$ and $0 < t \le 1$. □

Next, we turn to:

Proof of Theorem 1.2 for $|x-y| > L_0$ and $\theta > 2$. For $|x-y| > L_0$, we use the synchronous coupling and the assertion of Theorem 1.2 for $|x-y| \le L_0$. In detail, with (1.1), let $(X_t, Y_t^{(2)})_{t \ge 0}$ be the coupling process on $\mathbb{R}^{2d}$ such that its distribution is the same as that of $(X'_t, Y'_t)_{t \ge 0}$ constructed in Section 2.2. We now consider

$$\mathrm{d}Y_t = \begin{cases} \mathrm{d}Z_t + b(Y_t)\,\mathrm{d}t, & 0 \le t < T_{L_0}, \\ \mathrm{d}Y_t^{(2)}, & T_{L_0} \le t < T, \end{cases} \tag{3.6}$$

where

$$T_{L_0} = \inf\{t > 0: |X_t - Y_t| \le L_0\}$$

and $T = \inf\{t > 0: X_t = Y_t\}$ is the coupling time. For $t \ge T$, we still set $Y_t = X_t$. Therefore, the difference process $(D_t)_{t\ge 0} := (X_t - Y_t)_{t\ge 0}$ satisfies

$$\mathrm{d}D_t = \big(b(X_t) - b(Y_t)\big)\,\mathrm{d}t, \qquad t < T_{L_0}.$$

Note that the equality above implies that $t \mapsto D_t$ is a continuous function on $[0, T_{L_0})$ such that $\lim_{t\to T_{L_0}-} |D_t| = L_0$. As a result,

$$\mathrm{d}|D_t|^2 = 2\big\langle D_t, b(X_t) - b(Y_t)\big\rangle\,\mathrm{d}t, \qquad t < T_{L_0}.$$

Still denoting by $r_t = |D_t|$, we get from (1.4) that

$$\mathrm{d}r_t \le -K_2 r_t^{\theta-1}\,\mathrm{d}t, \qquad t < T_{L_0},$$

which implies that

$$T_{L_0} \le \frac{1}{K_2(2-\theta)}\big(|x-y|^{2-\theta} - L_0^{2-\theta}\big) \le \frac{L_0^{2-\theta}}{K_2(\theta-2)} =: t_0 \tag{3.7}$$

since $\theta > 2$ and the continuity of $t \mapsto r_t$ on $[0, T_{L_0})$.

Therefore, for any $x, y \in \mathbb{R}^d$ with $|x-y| > L_0$, $p \ge 1$ and $t > t_0$, we have

$$\begin{aligned}
\mathbb{E}|X_t - Y_t|^p &= \mathbb{E}\big[\mathbb{E}^{(X_{T_{L_0}}, Y_{T_{L_0}})}|X_{t-T_{L_0}} - Y_{t-T_{L_0}}|^p\big] \\
&\le C_1\mathbb{E}\big[|X_{T_{L_0}} - Y_{T_{L_0}}|\mathrm{e}^{-\lambda(t-T_{L_0})}\big] \\
&\le C_1 L_0 \exp(\lambda t_0)\mathrm{e}^{-\lambda t},
\end{aligned}$$

where in the first inequality we have used (3.4), and the last inequality follows from (3.7) and the fact that $|X_{T_{L_0}} - Y_{T_{L_0}}| \le L_0$. In particular, we have for all $|x-y| > L_0$ and $t > t_0$,

$$\mathbb{E}|X_t - Y_t|^p \le C_3\mathrm{e}^{-\lambda t}.$$

Combining with all conclusions above, we complete the proof of the second assertion in Theorem 1.2. □

We finally present the following.

Proof of Example 1.3. In this example,

$$b(x) = \nabla V(x) = 2\beta|x|^{2\beta-2}x.$$

It follows from the proof of [5], Example 5.3, that for any $x, y \in \mathbb{R}^d$,

$$\big\langle b(x) - b(y), x - y\big\rangle \le -\beta 2^{4-3\beta}|x-y|^{2\beta}. \tag{3.8}$$

Then, (1.4) holds with $K_2 = \beta 2^{4-3\beta}$, $\theta = 2\beta$ and any positive constants K_1, L_0. In particular, (1.5) holds for all $\alpha \in (0, 1]$ and $K_1, L_0 > 0$ small enough. Then the required assertion is a direct consequence of Theorem 1.2. □

Acknowledgements

The author would like to thank Professor Feng-Yu Wang and the referee for helpful comments and careful corrections. Financial support through NSFC (No. 11201073), JSPS (No. 26·04021), NSF-Fujian (No. 2015J01003) and the Program for Nonlinear Analysis and Its Applications (No. IRTL1206) are gratefully acknowledged.

References

[1] Barczy, M., Li, Z. and Pap, G. (2015). Yamada–Watanabe results for stochastic differential equations with jumps. *Int. J. Stoch. Anal.* Art. ID 460472, 23. MR3298537

[2] Bogdan, K. and Jakubowski, T. (2007). Estimates of heat kernel of fractional Laplacian perturbed by gradient operators. *Comm. Math. Phys.* **271** 179–198. MR2283957

[3] Bolley, F., Gentil, I. and Guillin, A. (2012). Convergence to equilibrium in Wasserstein distance for Fokker–Planck equations. *J. Funct. Anal.* **263** 2430–2457. MR2964689

[4] Böttcher, B., Schilling, R.L. and Wang, J. (2011). Constructions of coupling processes for Lévy processes. *Stochastic Process. Appl.* **121** 1201–1216. MR2794973

[5] Cattiaux, P. and Guillin, A. (2014). Semi log-concave Markov diffusions. In *Séminaire de Probabilités XLVI. Lecture Notes in Math.* **2123** 231–292. Berlin: Springer.

[6] Chen, M.-F. (2004). *From Markov Chains to Non-equilibrium Particle Systems*, 2nd ed. River Edge, NJ: World Scientific. MR2091955

[7] Chen, M.-F. (2005). *Eigenvalues, Inequalities, and Ergodic Theory. Probability and Its Applications (New York)*. London: Springer. MR2105651

[8] Chen, M.F. and Li, S.F. (1989). Coupling methods for multidimensional diffusion processes. *Ann. Probab.* **17** 151–177. MR0972776

[9] Chen, Z.-Q., Kim, P. and Song, R. (2012). Dirichlet heat kernel estimates for fractional Laplacian with gradient perturbation. *Ann. Probab.* **40** 2483–2538. MR3050510

[10] Eberle, A. Reflection couplings and contraction rates for diffusions. Available at arXiv:1305.1233v3.

[11] Eberle, A. (2011). Reflection coupling and Wasserstein contractivity without convexity. *C. R. Math. Acad. Sci. Paris* **349** 1101–1104. MR2843007

[12] Kurtz, T.G. (2011). Equivalence of stochastic equations and martingale problems. In *Stochastic Analysis* 2010 113–130. Heidelberg: Springer. MR2789081

[13] Lepeltier, J.-P. and Marchal, B. (1976). Problème des martingales et équations différentielles stochastiques associées à un opérateur intégro-différentiel. *Ann. Inst. H. Poincaré Sect. B (N.S.)* **12** 43–103. MR0413288

[14] Lindvall, T. and Rogers, L.C.G. (1986). Coupling of multidimensional diffusions by reflection. *Ann. Probab.* **14** 860–872. MR0841588

[15] Luo, D. and Wang, J. Exponential contractivity in the L^p-Wasserstein distance for diffusion processes. Available at arXiv:1407.1986.

[16] Priola, E. and Wang, F.-Y. (2006). Gradient estimates for diffusion semigroups with singular coefficients. *J. Funct. Anal.* **236** 244–264. MR2227134

[17] Schilling, R.L., Sztonyk, P. and Wang, J. (2012). Coupling property and gradient estimates of Lévy processes via the symbol. *Bernoulli* **18** 1128–1149. MR2995789

[18] Schilling, R.L. and Wang, J. (2011). On the coupling property of Lévy processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **47** 1147–1159. MR2884228

[19] Schilling, R.L. and Wang, J. (2012). On the coupling property and the Liouville theorem for Ornstein–Uhlenbeck processes. *J. Evol. Equ.* **12** 119–140. MR2891204

[20] Stroock, D.W. (1975). Diffusion processes associated with Lévy generators. *Z. Wahrsch. Verw. Gebiete* **32** 209–244. MR0433614

[21] von Renesse, M.-K. and Sturm, K.-T. (2005). Transport inequalities, gradient estimates, entropy, and Ricci curvature. *Comm. Pure Appl. Math.* **58** 923–940. MR2142879

[22] Wang, F.-Y. (2005). *Functional Inequalities, Markov Semigroups and Spectral Theory*. Beijing: Science Press.

[23] Wang, F.-Y. (2011). Coupling for Ornstein–Uhlenbeck processes with jumps. *Bernoulli* **17** 1136–1158. MR2854767

[24] Wang, F.-Y. and Wang, J. (2014). Harnack inequalities for stochastic equations driven by Lévy noise. *J. Math. Anal. Appl.* **410** 513–523. MR3109860

[25] Wang, J. (2014). On the existence and explicit estimates for the coupling property of Lévy processes with drift. *J. Theoret. Probab.* **27** 1021–1044. MR3245997

Received August 2014 and revised January 2015