Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2020, Vol. 56, No. 1, 175–229
https://doi.org/10.1214/19-AIHP959

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Ergodicity of stochastic differential equations with jumps and singular coefficients

Longjie Xie[a] and Xicheng Zhang[b]

[a]*School of Mathematics and Statistics, Jiangsu Normal University, Xuzhou, Jiangsu 221000, P.R. China. E-mail:* xlj.98@whu.edu.cn
[b]*School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei 430072, P.R. China. E-mail:* XichengZhang@gmail.com

Received 11 August 2017; revised 4 December 2018; accepted 15 January 2019

Abstract. We show the strong well-posedness of SDEs driven by general multiplicative Lévy noises with Sobolev diffusion and jump coefficients and integrable drifts. Moreover, we also study the strong Feller property, irreducibility as well as the exponential ergodicity of the corresponding semigroup when the coefficients are time-independent and singular dissipative. In particular, the large jump is allowed in the equation. To achieve our main results, we present a general approach for treating the SDEs with jumps and singular coefficients so that one just needs to focus on Krylov's a priori estimates for SDEs.

Résumé. Nous montrons que les EDS dirigées par un bruit de Lévy multiplicatif général avec des coefficients de diffusion et de saut Sobolev, et une dérive intégrable, sont fortement bien posées. De plus, nous étudions la propriété forte de Feller, l'irréductibilité ainsi que l'ergodicité exponentielle des semi-groupes correspondants quand les coefficients sont indépendants du temps et singulièrement dissipatifs. En particulier, les grands sauts sont autorisés dans l'équation. Pour aboutir au résultat principal, nous présentons une approche générale pour traiter les EDS avec sauts et coefficients singuliers, de telle sorte que nous devons seulement nous intéresser aux estimées a priori de Krylov pour les EDS.

MSC: 60H10; 60J60

Keywords: Pathwise uniqueness; Krylov's estimate; Zvonkin's transformation; Ergodicity; Heat kernel

Contents

Research of L. Xie is partially supported by NNSF of China (No. 11701233), NSF of Jiangsu (No. BK20170226) and the Project Funded by the PAPD of Jiangsu Higher Education Institutions. X. Zhang is partially supported by NNSF of China (No. 11731009) and the DFG through the CRC 1283 "Taming uncertainty and profiting from randomness and low regularity in analysis, stochastics and their applications".

1. Introduction

Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space, which satisfies the usual conditions. On this probability space, let $(W_t)_{t\geq 0}$ be a d-dimensional standard $\mathscr{F}_t$-Brownian motion and N an $\mathscr{F}_t$-Poisson random measure with intensity measure $\mathrm{d}t\,\nu(\mathrm{d}z)$, where ν is a Lévy measure on $\mathbb{R}^d$, that is,

$$\int_{\mathbb{R}^d} \big(|z|^2 \wedge 1\big)\nu(\mathrm{d}z) < +\infty, \quad \nu\big(\{0\}\big) = 0. \tag{1.1}$$

The compensated Poisson random measure $\tilde{N}$ is defined as

$$\tilde{N}(\mathrm{d}t, \mathrm{d}z) := N(\mathrm{d}t, \mathrm{d}z) - \mathrm{d}t\,\nu(\mathrm{d}z).$$

Consider the following stochastic differential equation (SDE) with jumps in $\mathbb{R}^d$:

$$\mathrm{d}X_t = \sigma_t(X_t)\,\mathrm{d}W_t + b_t(X_t)\,\mathrm{d}t + \int_{|z|<R} g_t(X_{t-}, z)\tilde{N}(\mathrm{d}t, \mathrm{d}z) + \int_{|z|\geq R} g_t(X_{t-}, z)N(\mathrm{d}t, \mathrm{d}z), \tag{1.2}$$

where $R > 0$ is a fixed constant, and $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$, $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $g : \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ are Borel measurable functions, which are called diffusion, drift and jump coefficients, respectively. Recall that an $\mathscr{F}_t$-adapted càdlàg (right continuous with left limit) process X is called a (strong) solution of SDE (1.2) if for each $t > 0$, the following random variables are finite $\mathbb{P}$-almost surely,

$$\int_0^t \big\|\sigma_s(X_s)\big\|^2\,\mathrm{d}s, \qquad \int_0^t \big|b_s(X_s)\big|\,\mathrm{d}s, \qquad \int_0^t\int_{|z|<R} \big|g_s(X_s, z)\big|^2\nu(\mathrm{d}z)\,\mathrm{d}s, \qquad \int_0^t\int_{|z|\geq R} \big|g_s(X_s, z)\big|\nu(\mathrm{d}z)\,\mathrm{d}s,$$

and

$$X_t = X_0 + \int_0^t \sigma_s(X_s)\,\mathrm{d}W_s + \int_0^t b_s(X_s)\,\mathrm{d}s + \int_0^t\int_{|z|<R} g_s(X_{s-}, z)\tilde{N}(\mathrm{d}s, \mathrm{d}z)$$
$$+ \int_0^t\int_{|z|\geq R} g_s(X_{s-}, z)N(\mathrm{d}s, \mathrm{d}z), \quad \mathbb{P}\text{-a.s.}$$

In this paper, we shall study the existence and uniqueness of strong solutions to the above SDE under some mild assumptions on the coefficients, both in the non-degenerate diffusion case and in the multiplicative pure jump case. Moreover, we also study the strong Feller property, irreducibility as well as the ergodicity of the semigroup associated with the above SDE when the coefficients are time-independent and singular dissipative.

1.1. *Well-posedness*

In the past decades, SDEs with singular drifts and driven by Brownian motions have been extensively studied. In the case that $g \equiv 0$ and $\sigma \equiv \mathbb{I}_{d\times d}$ (the identity matrix), a remarkable result due to Krylov and Röckner [31] says that SDE (1.2) has a unique strong solution provided that

$$b \in L^q_{\mathrm{loc}}\big(\mathbb{R}_+; L^p\big(\mathbb{R}^d\big)\big) \quad \text{with } \frac{d}{p} + \frac{2}{q} < 1.$$

Latter, the second named author [54,55] extended their results to the multiplicative noise under some non-degenerate and Sobolev conditions on the diffusion coefficient. On the other hand, by studying the stochastic homeomorphism flow property of the SDEs with irregular drifts, Flandoli, Gubinelli and Priola [20] obtained a well-posedness result for a class of stochastic transport equations with irregular coefficients. After that, there are many works devoted to the study of the regularities of the unique strong solution to SDEs with rough coefficients, such as the Sobolev differentiability with respect to the initial value, stochastic homeomorphism flow and the Malliavin differentiability with respect to the sample path. The interested readers are referred to [18,19,35,38,49,51,57] and references therein.

In recent years, SDEs driven by pure jump Lévy processes (i.e., $\sigma \equiv 0$) and with irregular drifts have also attracted great interests since it behaves quite differently. In fact, when $d = 1$ and $(L_t)_{t\geq 0}$ is a symmetric α-stable process with $\alpha \in (0, 1)$, Tanaka, Tsuchiya and Watanabe [47] showed that even if b is time-independent, bounded and β-Hölder continuous with $\beta < 1 - \alpha$, the following SDE

$$\mathrm{d}X_t = \mathrm{d}L_t + b(X_t)\,\mathrm{d}t, \quad X_0 = x \in \mathbb{R}^d \tag{1.3}$$

may not have a pathwise uniqueness strong solution, see also [3] for related results. On the other hand, when $\alpha \in [1, 2)$ and

$$b \in C_b^\beta\big(\mathbb{R}^d\big) \quad \text{with } \beta > 1 - \frac{\alpha}{2},$$

it was shown by Priola [40] that there exists a unique strong solution $X_t(x)$ to SDE (1.3) for each $x \in \mathbb{R}^d$, which forms a stochastic C^1-diffeomorphism flow. Under the same condition, Haadem and Proske [23] obtained the unique strong solution by using the Malliavin calculus. Recently, Zhang [56] obtained the pathwise uniqueness to SDE (1.3) when $\alpha \in (1, 2)$, b is bounded and in some fractional Sobolev spaces. See also [9,12,41,42] for related results. It is noticed that all the works mentioned above for SDE (1.2) with $\sigma \equiv 0$ are restricted to the additive noise case. We also mention that Bogachev and Pilipenko [10] treated the SDE with general Lévy noise and discontinuous drifts based on heat kernel estimates.

The first aim of this paper is to study the well-posedness of the SDE (1.2) with Sobolev diffusion and jump coefficients and integrable drifts. In the mixing and non-degenerate diffusion case, we shall not make any assumptions on the pure jump Lévy noise (or the Lévy measure ν in (1.1)), see Theorem 2.1. Our result extends the existing results concerning singular SDEs driven by Brownian motion (see [10,26,31,55]). In the pure jump case, we shall assume that the Lévy measure ν is symmetric and rotationally invariant α-stable type in order to use the heat kernel estimates established in [13,14], see Theorem 2.4. Compared with [23,40–42,56], we are considering the multiplicative noise and drop the boundedness assumption on drift b.

Let us now introduce the main argument adopted in the present paper: Zvonkin's transformation. Let $\mathcal{L}_2^\sigma$ be the second order differential operator associated with the diffusion coefficient σ, that is,

$$\mathcal{L}_2^\sigma u(x) := \frac{1}{2}\big(\sigma_t^{ik}\sigma_t^{jk}\partial_i\partial_j u\big)(x). \tag{1.4}$$

Here and below, we use Einstein's convention for summation that the repeated indices in a product will be summed automatically. Let $\mathcal{L}_1^b$ be the first order differential operator associated with the drift coefficient b, that is,

$$\mathcal{L}_1^b u(x) := \big(b_t^i\partial_i u\big)(x),$$

and $\mathcal{L}_\nu^g$ the nonlocal operator associated with the jump coefficient g, that is,

$$\begin{aligned}\mathcal{L}_\nu^g u(x) :=& \int_{|z|<R}\big[u\big(x + g_t(x,z)\big) - u(x) - g_t(x,z)\cdot\nabla u(x)\big]\nu(\mathrm{d}z)\\ &+ \int_{|z|\geq R}\big[u\big(x + g_t(x,z)\big) - u(x)\big]\nu(\mathrm{d}z) =: \mathcal{L}_{\nu,R}^g u(x) + \bar{\mathcal{L}}_{\nu,R}^g u(x).\end{aligned} \tag{1.5}$$

For a fixed time $T > 0$, we consider the following Kolmogorov's backward equation:

$$\partial_t\Phi + \big(\mathcal{L}_2^\sigma + \mathcal{L}_1^b + \mathcal{L}_{\nu,R}^g\big)\Phi = 0, \quad \Phi_T(x) = x \in \mathbb{R}^d. \tag{1.6}$$

Suppose that this equation has a regular enough solution Φ so that for each $t \in [0,T]$, the map $x \mapsto \Phi_t(x)$ forms a C^2-diffeomorphism on $\mathbb{R}^d$. Then, by Itô's formula, one gets that

$$\begin{aligned}\Phi_t(X_t) &= \Phi_0(X_0) + \int_0^t \nabla\Phi_s(X_s)\sigma_s(X_s)\,\mathrm{d}W_s \\ &\quad + \int_0^t\int_{|z|<R}\big(\Phi_s\big(X_{s-}+g_s(X_{s-},z)\big)-\Phi_s(X_{s-})\big)\tilde{N}(\mathrm{d}s,\mathrm{d}z) \\ &\quad + \int_0^t\int_{|z|\ge R}\big(\Phi_s\big(X_{s-}+g_s(X_{s-},z)\big)-\Phi_s(X_{s-})\big)N(\mathrm{d}s,\mathrm{d}z).\end{aligned}$$

Thus, if we let $Y_t := \Phi_t(X_t)$ and

$$\tilde{\sigma}_t(y) := (\nabla\Phi_s\cdot\sigma_s)\circ\Phi_t^{-1}(y), \qquad \tilde{g}_t(y,z) := \Phi_t\big(\Phi_t^{-1}(y)+g_t\big(\Phi_t^{-1}(y),z\big)\big)-y,$$

then Y_t satisfies the following *new* SDE with *disappeared* drift:

$$\mathrm{d}Y_t = \tilde{\sigma}_t(Y_t)\,\mathrm{d}W_t + \int_{|z|<R}\tilde{g}_t(Y_{t-},z)\tilde{N}(\mathrm{d}s,\mathrm{d}z) + \int_{|z|\ge R}\tilde{g}_t(Y_{t-},z)N(\mathrm{d}t,\mathrm{d}z), \tag{1.7}$$

and vice versa, that is, if Y_t solves (1.7), then $X_t := \Phi_t^{-1}(Y_t)$ solves SDE (1.2). Notice that in the case $g \equiv 0$, if σ is uniformly elliptic and Lipschitz continuous, and b is only Hölder continuous, then $\tilde{\sigma}$ could be also Lipschitz continuous due to the second order regularization effect of equation (1.6), see [55,57]. Consequently, the well-posedness for SDE (1.2) with Hölder drifts follows by the well-posedness of SDE (1.7). Thus, the main task is to solve equation (1.6) so that Φ has the desired properties.

However, as we shall see below in Theorems 4.3 and 4.5 that for $b \in L^q_{\mathrm{loc}}(\mathbb{R}_+; L^p(\mathbb{R}^d))$, it is in general not possible to construct a C^2-solution Φ to equation (1.6), and thus the transformed coefficients $\tilde{\sigma}$ and $\tilde{g}$ in (1.7) are not expected to be Lipschitz continuous, but at most in the first order Sobolev space $\mathbb{W}^{1,p}_{\mathrm{loc}}(\mathbb{R}^d)$. In other words, we need to first study SDE (1.7) with coefficients being in $\mathbb{W}^{1,p}_{\mathrm{loc}}(\mathbb{R}^d)$. To this end, a key ingredient that needed is the following *a priori* Krylov's estimate: for any solution Y of (1.7), any $T>0$ and $f \in L^q_{\mathrm{loc}}(\mathbb{R}_+; L^p(\mathbb{R}^d))$ with certain $p,q \ge 1$,

$$\mathbb{E}\bigg(\int_0^T f(t,Y_t)\,\mathrm{d}t\bigg) \le C\bigg(\int_0^T\bigg(\int_{\mathbb{R}^d}\big|f(t,x)\big|^p\,\mathrm{d}x\bigg)^{q/p}\mathrm{d}t\bigg)^{1/q}. \tag{1.8}$$

For general continuous Itô's process, such an estimate was established in [29] for $p=q\ge d+1$. For SDE (1.2) with $g\equiv 0$ and general p,q satisfying $\frac{d}{p}+\frac{2}{q}<2$, we refer to [55,57,58]. However, for discontinuous semimartingales, there are few results. We mention that the authors in [33] and [39,45] obtained some rough Krylov's estimates, which are not enough for our purpose. In Section 5, we shall devote to a detailed study about the above Krylov estimate for any solution of SDE (1.2) under certain optimal conditions on p, q. In the non-degenerate diffusion case, we first use a Krylov's lemma to show that estimate (1.8) holds for any solution of SDE (1.2) and for any $p=q\ge d+1$, see Theorem 5.2. Then we combine the results proved in Theorem 4.3 for non-homogeneous Kolmogorov's backward equation with Girsanov's theorem to generalize the Krylov estimate to general p,q with $\frac{d}{p}+\frac{2}{q}<2$, see Theorem 5.7. While in the purely nonlocal α-stable-like noise case, we shall use the result proved in Theorem 4.5 for non-homogeneous nonlocal Kolmogorov's backward equation and smoothing coefficients technique to show that (1.8) holds for any p, q with $\frac{d}{p}+\frac{\alpha}{q}<\alpha$, see Theorem 5.10.

1.2. *Ergodicity*

Although the well-posedness and regularity properties of strong solutions for SDEs with singular coefficients have been intensively studied, it seems that there are few works devoted to studying the existence and uniqueness of invariant probability measures for time-independent SDEs with singular coefficients. As we know, a general approach for proving the existence of invariant probability measures is to verify the Lyapunov condition. More precisely, if there exists a positive function $\Phi_1 \in C^2(\mathbb{R}^d)$ and a positive compact function Φ_2 such that

$$\big(\mathscr{L}_2^\sigma + \mathscr{L}_1^b + \mathscr{L}_\nu^g\big)\Phi_1 \le C - \Phi_2 \tag{1.9}$$

holds for some constant $C > 0$, then the associated semigroup of SDE (1.2) has an invariant probability measure μ with $\mu(\Phi_2) < \infty$, see for instance [25]. Obviously, if $b \in L^p(\mathbb{R}^d)$ for some $p > d$, then compact function Φ_2 would not exist since b can be singular at infinity. In this direction, to the authors' best knowledge, Wang [50] obtained a first result about the ergodicity for SDEs with singular drifts by using perturbation argument and his local dimension-free Harnack inequality. In particular, the main result in [50] is applied to the following singular SDE so that it admits a unique invariant probability measure:

$$\mathrm{d}X_t = \big(b(X_t) - \lambda_0 X_t\big)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t, \quad X_0 = x \in \mathbb{R}^d,$$

where $\lambda_0 > 0$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ satisfies

$$\int_{\mathbb{R}^d} \mathrm{e}^{\lambda|b(x)|^2 - \lambda_0|x|^2/2}\,\mathrm{d}x < \infty \quad \text{for some } \lambda > \frac{1}{2\lambda_0}. \tag{1.10}$$

To prove the uniqueness of invariant measures, a usual way is to show the strong Feller property and irreducibility of the associated semigroup. In [53], we have studied these two properties for SDEs driven by Brownian motions under some local conditions on the coefficients.

In the pure jump case, when the coefficients are locally Lipschitz continuous and satisfy a Lyapunov-type dissipative condition such as (1.9), it has been shown in [32,34] that there is a unique invariant probability measure associated to the SDE. The exponential ergodicity is also studied therein under some abstract conditions. In a recent work [2], the authors also introduced some Lyapunov stability conditions for the existence of invariant probability measures for general nonlocal operators. Clearly, the singular drift does not satisfy the Lyapunov conditions as required in all the works mentioned above.

The second aim of this paper is to show the existence and uniqueness of invariant probability measures associated to SDE (1.2) both in non-degenerate diffusion case and in pure jump case under suitable singular and dissipative assumptions. Our basic idea is as follows: suppose that b can be decomposed into two parts:

$$b = b_1 + b_2,$$

where b_1 is the singular part and b_2 is the dissipative part, see $(\mathbf{H}^b)$ and $(\widetilde{\mathbf{H}}^b)$ below. We shall use Zvonkin's transformation to kill only the singular part b_1, and obtain a new SDE. Of course, the dissipative part b_2 will roll together with the transforming function. The key observation is that the dissipativity will be preserved in the transformed SDE. Since the transform is one-to-one, we can get the ergodicity of the original equation from the new one. Here, to perform Zvonkin's transformation, we need to solve a nonlocal elliptic equation rather than the parabolic equation (1.6), see Theorems 7.6 and 7.10. Moreover, Krylov's estimates obtained in Section 5 are not applicable any more since the coefficients may have polynomial growth. Instead, we shall show the non-explosion and a priori Krylov's estimates for any solutions of SDE (1.2) with singular and dissipative drift, see Lemmas 7.5 and 7.9.

Table 1 figures out the method of showing the existence and uniqueness of invariant probability measures for time-independent SDE (1.2) with dissipative drifts. Here, IPM, GT, HK and DHK stands for invariant probability measure, Girsanov's transform, heat kernel and Dirichlet heat kernel, respectively.

1.3. *Examples*

Below we provide two simple examples to illustrate the main results obtained in this paper.

Table 1
Method of showing ergodicity

Existence of IPMs	Strong Feller	Irreducibility
	Diffusion with jump	
Lyapunov condition	Derivative formula	Coupling + GT
	Pure jump SDE	
Lyapunov condition	Continuity of HK	Positivity of DHK

Example 1.1. Consider the following SDE of OU-type:

$$\mathrm{d}X_t = \mathrm{d}L_t - \lambda_0 X_t \,\mathrm{d}t + b(X_t)\,\mathrm{d}t, \quad X_0 = x \in \mathbb{R}^d.$$

When L_t is a d-dimensional standard Brownian motion, we assume $b \in L^p(\mathbb{R}^d)$ for some $p > d$. When L_t is a rotationally invariant symmetric α-stable process with $\alpha \in (1,2)$, we assume $b \in H_p^\theta(\mathbb{R}^d)$ for some $\theta > 1-\alpha/2$ and $p > 2d/\alpha$, where $H_p^\theta(\mathbb{R}^d)$ is the Bessel potential space. By Theorems 2.9 and 2.12 below, the above SDE admits a unique strong solution and there exists a unique invariant probability measure associated with it. Note that in both cases, the classical Lyapunov condition (1.9) can not be verified, our result is new even in the existence of invariant probability measures. Moreover, compared with Wang's global condition (1.10), our global assumption $b \in L^p(\mathbb{R}^d)$ is weaker locally and not comparable at infinity.

Example 1.2. Consider the following mixing SDE with jumps:

$$\mathrm{d}X_t = \mathrm{d}W_t + \lambda_1 |X_{t-}|^\beta \,\mathrm{d}L_t - \lambda_0 X_t |X_t|^{\gamma-1}\,\mathrm{d}t, \quad X_0 = x \in \mathbb{R}^d,$$

where $\beta \in (0,1)$, $\gamma \in (0,\infty)$ and $\lambda_0 > 0$, $\lambda_1 \in \mathbb{R}$, L_t is a d-dimensional pure jump Lévy process. The main features of this SDE are that the jump coefficient $x \mapsto |x|^\beta$ is Hölder continuous and the drift term may have polynomial growth. By Theorem 7.4 below, the above SDE has a unique strong solution. Moreover, there exists a unique invariant probability measure and the SDE is V-ergodic (see Definition 2.7) in the case $\gamma \in (0,1]$ and exponential ergodic in the case $\gamma > 1$.

Finally, we recall that a probability measure μ on $\mathbb{R}^d$ is called an invariant probability measure of the operator $\mathscr{L} := \mathscr{L}_2^\sigma + \mathscr{L}_1^b + \mathscr{L}_\nu^g$ if it satisfies the following Fokker–Planck–Kolmogorov equation:

$$\mathscr{L}^*\mu = 0 \quad \Leftrightarrow \quad \mu(\mathscr{L}\varphi) = 0, \quad \varphi \in C_0^\infty(\mathbb{R}^d), \tag{1.11}$$

where the asterisk stands for the formal adjoint operator. Obviously, any invariant probability measure of the semigroup associated with SDE (1.2) satisfies (1.11). When $g \equiv 0$, the existence of solutions to (1.11) was obtained in [7] by analytic methods under a Lyapunov-type condition, which is much weaker than those needed for the existence of a solution to SDE (1.2). Moreover, under some quite weak conditions, the uniqueness and regularities of the solutions for (1.11) are also studied in [5,6,8], see also [50]. To our knowledge, these results cannot cover our results stated above.

1.4. *Layout*

The plan of this paper is as follows: In Section 2, we state the main results including the existence-uniqueness and ergodicity for SDE (1.2). Since the proofs of well-posedness and ergodicity rely on approximations and Zvonkin's argument, in Section 3 we present two general results: Stability and Zvonkin's transformation for SDE (1.2). Moreover, we also prove a useful stochastic Gronwall's inequality, which extends Scheutzow's result [44] to the discontinuous martingales. In Section 4, we study the solvability and regularity of parabolic integro-differential equations. In Section 5, applying the results obtained in the previous section, we show various Krylov's estimates for the solution of SDE (1.2). By the general results in Section 3, the strong well-posedness results are proved in Section 6. The strong Feller property and irreducibility as well as the ergodicity for SDE (1.2) are proven in Section 7. Finally, some details and auxiliary materials are given in the Appendix. To make the structure of the paper more transparent for the reader, we provide Figure 1 which describes the relations among the main results.

Throughout this paper, we use the following conventions: c with or without subscripts will denote a positive constant, whose value may change in different places. Moreover, we use $A \lesssim B$ to denote $A \leq cB$ for some unimportant constant $c > 0$.

2. Statement of main results

2.1. *Strong well-posedness of singular SDEs with jumps*

To state our main results, we first introduce some spaces and notations. For $p, q \in [1,\infty]$ and $0 \leq S < T < \infty$, let $\mathbb{L}_p^q(S,T)$ be the space of all Borel functions on $[S,T] \times \mathbb{R}^d$ with norm

$$\|f\|_{\mathbb{L}_p^q(S,T)} := \left(\int_S^T \left(\int_{\mathbb{R}^d} |f(t,x)|^p \,\mathrm{d}x\right)^{q/p} \mathrm{d}t\right)^{1/q} < \infty.$$

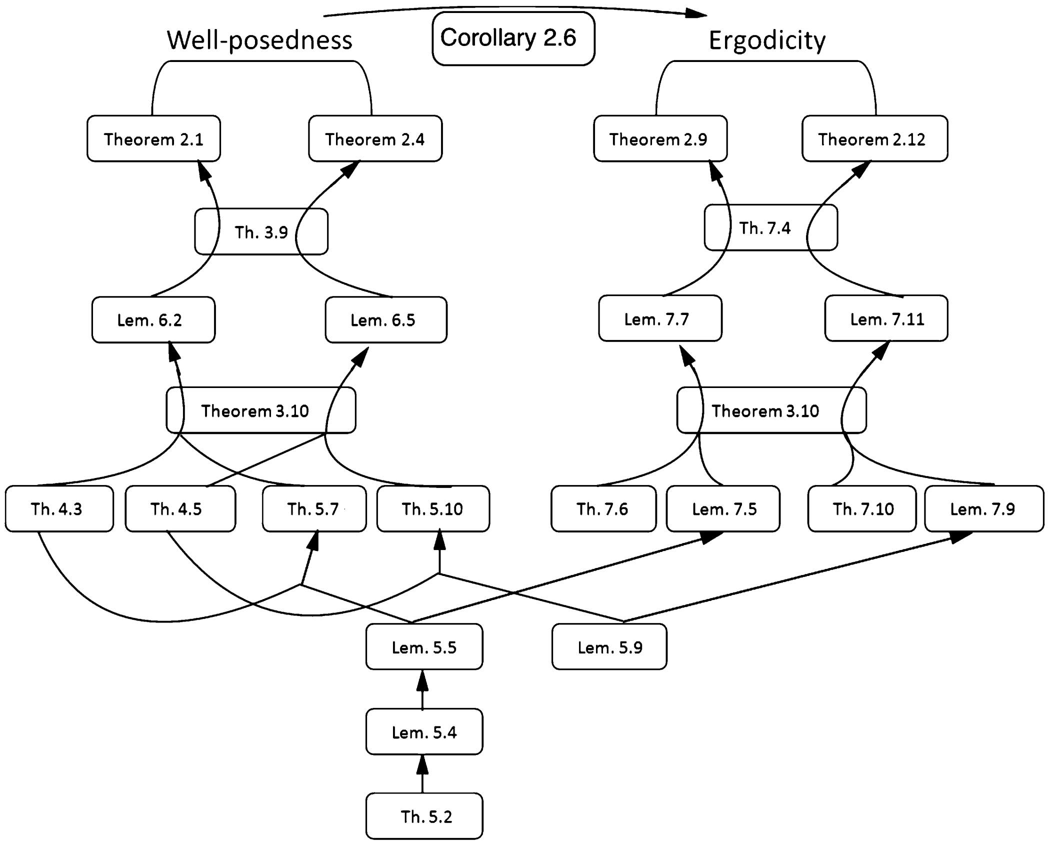

Fig. 1. Relations among theorems and lemmas.

For $p = \infty$ or $q = \infty$, the above norm is understood as the usual L^∞-norm. We shall simply write

$$\mathbb{L}^q_p(T) := \mathbb{L}^q_p(0, T), \qquad \mathbb{L}^p(T) := \mathbb{L}^p_p(T).$$

Given a $R > 0$, we shall write $B_R := \{x \in \mathbb{R}^d : |x| < R\}$. For a measurable function $g_t(x, z) : \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ and $0 \le \varepsilon < R \le \infty$, we introduce the following functions, which will be used frequently below: for $j = 0, 1$ and $\alpha \ge 1$,

$$\Gamma^{j,\alpha}_{\varepsilon,R}(g)(t, x) := \Gamma^{j,\alpha}_{\varepsilon,R}(g_t)(x) := \big\| \nabla^j_x g_t(x, \cdot)\big\|^\alpha_{L^\alpha(B_R \setminus B_\varepsilon; \nu)} := \int_{\varepsilon<|z|<R} \big|\nabla^j_x g_t(x, z)\big|^\alpha \nu(\mathrm{d}z). \tag{2.1}$$

Here and below, ∇_x denotes the generalized gradient with respect to x.

We make the following assumptions on the diffusion coefficient σ:

($\mathbf{H}^\sigma_\beta$) There are constants $c_0 \ge 1$ and $\beta \in (0, 1)$ such that for all $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$,

$$c_0^{-1}|\xi|^2 \le \big|\sigma^*_t(x)\xi\big|^2 \le c_0|\xi|^2, \quad \forall \xi \in \mathbb{R}^d,$$

where the asterisk stands for the transpose of a matrix, and

$$\big\|\sigma_t(x) - \sigma_t\big(x'\big)\big\| \le c_0\big|x - x'\big|^\beta.$$

Here and below, $\|\cdot\|$ denotes the Hilbert–Schmidt norm of a matrix.

Our first main result of this paper is:

Theorem 2.1 (Non-degenerate diffusion with jumps). *Let $\Gamma^{j,2}_{0,R}(g)$ be defined as in (2.1). Suppose that* ($\mathbf{H}^\sigma_\beta$) *holds and for any $T > 0$,*

$$\Gamma^{0,2}_{0,R}(g) \in \mathbb{L}^\infty(T), \quad \lim_{\varepsilon\to 0}\big\|\Gamma^{0,2}_{0,\varepsilon}(g)\big\|_{\mathbb{L}^\infty(T)} = 0,$$

and for some $p, q \in (2, \infty)$ *with* $\frac{d}{p} + \frac{2}{q} < 1$,

$$|\nabla\sigma|, b, \big(\Gamma^{1,2}_{0,R}(g)\big)^{1/2} \in \mathbb{L}^q_p(T).$$

Then for any initial value $X_0 = x \in \mathbb{R}^d$, *SDE* (1.2) *admits a unique strong solution* $X_t(x)$. *Moreover, for any* $T > 0$, *there is a constant* $c_T > 0$ *such that for all* $t \in (0, T]$, $x, y \in \mathbb{R}^d$ *and bounded measurable* φ,

$$\big|\mathbb{E}\varphi\big(X_t(x)\big) - \mathbb{E}\varphi\big(X_t(y)\big)\big| \le \frac{c_T}{\sqrt{t}}\|\varphi\|_\infty|x-y|. \tag{2.2}$$

Let us make some comments on the above result.

Remark 2.2. If $g_t(x, z) = \bar\sigma_t(x)z$ with $\bar\sigma_t(x) \in \mathbb{L}^\infty(T)$ and $\nabla\bar\sigma_t(x) \in \mathbb{L}^q_p(T)$ in the above theorem, then the assumptions on $\Gamma^{j,2}_{0,R}(g)$ automatically hold. In particular, if $\bar\sigma_t(x) = \bar\sigma(x) = |x|^\beta\mathbb{I}$ for some $\beta \in (0, 1)$, then one can check $\nabla\bar\sigma \in L^p_{\mathrm{loc}}(\mathbb{R}^d)$ for any $p < d/(1-\beta)$.

Remark 2.3. It is noticed that in the estimate (2.2), we do not make any assumption about the large jump coefficient since the large jump part is independent from the small jump part and has only finitely many jumps in any finite time interval.

In the above mixing case, the non-degenerate diffusion part plays a dominant role. In the pure jump case, we need to use the regularization effect of the jump noise. For this, we assume $\nu(\mathrm{d}z) = |z|^{-d-\alpha}\,\mathrm{d}z$ for some $\alpha \in (1, 2)$, and g satisfies that

$(\mathbf{H}^g_\beta)$ $g_t(x, 0) = 0$ and there is a constant $c_1 \ge 1$ such that for all $t \ge 0$, $x, x', z, z' \in \mathbb{R}^d$,

$$c_1^{-1}\big|z - z'\big| \le \big|g_t(x, z) - g_t\big(x, z'\big)\big| \le c_1\big|z - z'\big|, \tag{2.3}$$

and for some $\beta \in (0, 1)$ and $j = 0, 1$,

$$\big|\nabla^j_z g_t(x, z) - \nabla^j_z g_t\big(x', z\big)\big| \le c_1\big|x - x'\big|^\beta\big(|z| + |z|^{1-j}\big), \tag{2.4}$$

$$\big|\nabla_z g_t(x, z) - \nabla_z g_t\big(x, z'\big)\big| \le c_1\big|z - z'\big|. \tag{2.5}$$

Our second well-posedness result is:

Theorem 2.4 (Multiplicative pure jump noise). *Suppose that* $\sigma \equiv 0$, $\nu(\mathrm{d}z) = |z|^{-d-\alpha}\,\mathrm{d}z$ *for some* $\alpha \in (1, 2)$, *and* $(\mathbf{H}^g_\beta)$ *holds with* $\beta > 1 - \alpha/2$. *Moreover, we also suppose that for some* $\theta \in (1 - \frac{\alpha}{2}, 1)$, $p \in (\frac{2d}{\alpha} \vee 2, \infty)$ *and* $q \in (\frac{\alpha}{\alpha-1}, \infty)$,

$$\big(\Gamma^{1,2}_{0,R}(g)\big)^{1/2}, \Gamma^{1,1}_{R,\infty}(g), (\mathbb{I} - \Delta)^{\theta/2}b \in \bigcap_{T>0}\mathbb{L}^q_p(T).$$

Then for each initial value $X_0 = x \in \mathbb{R}^d$, *SDE* (1.2) *admits a unique strong solution* $X_t(x)$. *Moreover,* $X_t(x)$ *has a density* $\rho(t, x, y)$, *which enjoys the following estimates*:

(i) (*Two-sided estimate*) *For any* $T > 0$ *and* $\varepsilon \in (0, 1)$, *there are two constants* $c_1, c_2 > 0$ *such that for all* $t \in (0, T)$ *and* $x, y \in \mathbb{R}^d$,

$$c_1 \le t^{-1}\big(t^{1/\alpha} + |x - y|\big)^{d+\alpha} \cdot \rho(t, x, y) \le c_2\big(1 + \big(t^{1/\alpha} + |x - y|\big)^\varepsilon\big). \tag{2.6}$$

(ii) (*Gradient estimate*) *For any* $T > 0$ *and* $\varepsilon \in (0, 1)$, *there is a constant* $c_3 > 0$ *such that for all* $t \in (0, T)$ *and* $x, y \in \mathbb{R}^d$,

$$\big|\nabla_x \log\rho(t, x, y)\big| \le c_3 t^{-1/\alpha}\big(1 + \big(t^{1/\alpha} + |x - y|\big)^\varepsilon\big). \tag{2.7}$$

We would like to make the following comment.

Remark 2.5. If $g_t(x,z)=\bar\sigma_t(x)z$ with $\bar\sigma$ satisfying $(\mathbf{H}^\sigma_\beta)$ and $\nabla\bar\sigma\in\mathbb{L}^\infty_p(T)$ with $p>\frac{2d}{\alpha}$, then the conditions on g in Theorem 2.4 hold. Since in this case, $(\Gamma^{1,2}_{0,R}(g))^{1/2}=c|\nabla\bar\sigma|$ for some $c>0$ and by Sobolev's embedding (see (4.2) below), $\bar\sigma\in L^\infty([0,T];C^\beta_b(\mathbb{R}^d))$ with $\beta=1-d/p>1-\alpha/2$. Compared with the additive noise case considered in [23,40,41,56], we drop the boundness condition on the drift b, which is essentially used in their proof. Moreover, in this case, from the proof below, one sees that the ε in (2.6) and (2.7) can be zero. For the discontinuous drift b, see [56].

Let $\chi:\mathbb{R}^d\to[0,1]$ be a smooth function with $\chi(x)=0$ for $|x|\ge 2$ and $\chi(x)=1$ for $|x|\le 1$. For $m\in\mathbb{N}$, define the cutoff function χ_m by

$$\chi_m(x):=\chi\big(m^{-1}x\big). \tag{2.8}$$

Using suitable localization technique, we have:

Corollary 2.6 (Local well-posedness). *Suppose that for each $m\in\mathbb{N}$,*

$$\sigma^m_t(x):=\sigma_t\big(x\chi_m(x)\big),\qquad b^m_t(x):=b_t(x)\chi_m(x),\qquad g^m_t(x,z):=g_t\big(x\chi_m(x),z\big)$$

satisfy the same assumptions as in Theorem 2.1 or Theorem 2.4. Then SDE (1.2) admits a unique strong solution X_t up to the explosion time ζ, that is, $\lim_{t\uparrow\zeta}X_t=\infty$.

Proof. For each $m\in\mathbb{N}$, by Theorem 2.1 or Theorem 2.4, there exist a unique global strong solution X^m_t to SDE (1.2) with coefficients σ^m, g^m and b^m. For $m\ge k$, define

$$\zeta_{m,k}:=\inf\big\{t\ge 0:\big|X^m_t\big|\ge k\big\}\wedge m.$$

By the uniqueness of the solution, we have

$$\mathbb{P}\big(X^m_t=X^k_t,\forall t\in[0,\zeta_{m,k})\big)=1,$$

which implies that for $m\ge k$,

$$\zeta_{k,k}\le\zeta_{m,k}\le\zeta_{m,m},\quad \text{a.s.}$$

Hence, if we let $\zeta_k:=\zeta_{k,k}$, then $(\zeta_k)_{k\in\mathbb{N}}$ is an increasing sequence of $(\mathscr{F}_t)$-stopping times and for $m\ge k$,

$$\mathbb{P}\big(X^m_t=X^k_t,\forall t\in[0,\zeta_k)\big)=1.$$

Now, for each $k\in\mathbb{N}$, we can define $X_t:=X^k_t$ for $t<\zeta_k$ and $\zeta:=\lim_{k\to\infty}\zeta_k$. It is easy to see that X_t is the unique solution of SDE (1.2) up to the explosion time ζ and $\lim_{t\uparrow\zeta}X_t=\infty$ a.s. □

As for the non-explosion, under some Lyapunov conditions, we may show the existence of global solutions (for instance, see Lemma 7.1 below).

2.2. *Ergodicity of SDEs with singular dissipative coefficients*

Below we turn to the study of the ergodicity of SDE (1.2). We first recall some basic notions about the ergodicity. Let $(P_t)_{t\ge0}$ be a semigroup of bounded linear operators on Banach space $\mathcal{B}_b(\mathbb{R}^d)$, where $\mathcal{B}_b(\mathbb{R}^d)$ denotes the space of all bounded Borel measurable functions. Let μ be a probability measure on Borel space $(\mathbb{R}^d,\mathscr{B}(\mathbb{R}^d))$. We use the following standard notation:

$$\mu(\varphi):=\int_{\mathbb{R}^d}\varphi(x)\mu(\mathrm{d}x).$$

- μ is said to be an invariant probability measure (or stationary distribution) of P_t if

$$\mu(P_t\varphi)=\mu(\varphi),\quad \forall t>0,\forall\varphi\in\mathcal{B}_b\big(\mathbb{R}^d\big).$$

- One says that P_t is ergodic if P_t admits a unique invariant probability measure μ, which amounts to say that

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t P_s f(x)\,\mathrm{d}s=\mu(f),\quad f\in\mathcal{B}_b\big(\mathbb{R}^d\big). \tag{2.9}$$

- One says that P_t has the C_b-strong Feller property if for all $\varphi\in\mathcal{B}_b(\mathbb{R}^d)$, $P_t\varphi\in C_b(\mathbb{R}^d)$.
- P_t is said to be irreducible if for each open ball B and $x\in\mathbb{R}^d$, $P_t 1_B(x)>0$.

About the ergodicity, we have the following classification (cf. [24] and [36]).

Definition 2.7. Let $V:\mathbb{R}^d\to[1,\infty)$ be a measurable function and μ an invariant probability measure of P_t. We say P_t to be V-uniformly exponential ergodic if there exist $c_0,\gamma>0$ such that for all $t\geq 0$ and $x\in\mathbb{R}^d$,

$$\sup_{\|\varphi\|_V\leq 1}\big|P_t\varphi(x)-\mu(\varphi)\big|\leq c_0V(x)\mathrm{e}^{-\gamma t},$$

where $\|\varphi\|_V:=\sup_{x\in\mathbb{R}^d}\frac{|\varphi(x)|}{V(x)}<+\infty$. If $V\equiv 1$, then P_t is said to be uniformly exponential ergodic, which is equivalent to

$$\big\|P_t(x,\cdot)-\mu\big\|_{\mathrm{Var}}\leq c_0\mathrm{e}^{-\gamma t},\quad \forall x\in\mathbb{R}^d,$$

where $\|\cdot\|_{\mathrm{Var}}$ is the total variation of a signed measure, $P_t(x,\cdot)$ is the kernel of bounded linear operator P_t.

It is useful to observe that the above notions are invariant under homeomorphism transformation of the phase space. More precisely, let $\Phi:\mathbb{R}^d\to\mathbb{R}^d$ be a homeomorphism. Define a new semigroup of bounded linear operators on $\mathcal{B}_b(\mathbb{R}^d)$ by

$$P_t^\Phi\varphi(y):=\big[P_t(\varphi\circ\Phi)\big]\big(\Phi^{-1}(y)\big),$$

where Φ^{-1} is the inverse of Φ. We have the following simple observations, which are direct by definition.

Proposition 2.8.

(i) *μ is an invariant probability measure of P_t if and only if $\mu\circ\Phi^{-1}$ is an invariant probability measure of P_t^Φ.*
(ii) *P_t has the C_b-strong Feller property if and only if P_t^Φ has the C_b-strong Feller property.*
(iii) *P_t is irreducible if and only if P_t^Φ is irreducible.*
(iv) *P_t is V-uniformly exponential ergodic if and only if P_t^Φ is $V\circ\Phi^{-1}$-uniformly exponential ergodic.*

To study the ergodicity of SDE (1.2), we shall assume that the coefficients are time-independent, i.e.,

$$\mathrm{d}X_t=\sigma(X_t)\,\mathrm{d}W_t+b(X_t)\,\mathrm{d}t+\int_{|z|<R}g(X_{t-},z)\tilde{N}(\mathrm{d}t,\mathrm{d}z)+\int_{|z|\geq R}g(X_{t-},z)N(\mathrm{d}t,\mathrm{d}z). \tag{2.10}$$

We show two new ergodicity results, which allow the drift to be singular at infinity. We first assume that

($\mathbf{H}^b$) $b=b_1+b_2$, where b_1 is the singular part and for some $p>d$,

$$b_1\in L^p\big(\mathbb{R}^d\big),$$

and b_2 is the dissipative part which satisfies for some $\kappa_1,\kappa_2,\kappa_3>0$ and $r>-1$,

$$\big\langle x,b_2(x)\big\rangle\leq-\kappa_1|x|^{2+r}+\kappa_2\quad\text{and}\quad\big|b_2(x)\big|\leq\kappa_3\big(1+|x|^{1+r}\big). \tag{2.11}$$

We have the following ergodicity result.

Theorem 2.9 (Ergodicity for diffusion with jumps). *Suppose that* ($\mathbf{H}^\sigma_\beta$) *and* ($\mathbf{H}^b$) *hold and for the same p in* ($\mathbf{H}^b$),

$$|\nabla\sigma|,\ \big(\Gamma^{1,2}_{0,R}(g)\big)^{1/2}\in L^p\big(\mathbb{R}^d\big),$$

and for any $\lambda \geq R$,

$$\Gamma^{0,2}_{0,\lambda}(g), \Gamma^{0,1}_{\lambda,\infty}(g) \in L^\infty(\mathbb{R}^d), \quad \lim_{\varepsilon\to 0}\|\Gamma^{0,2}_{0,\varepsilon}(g)\|_\infty = 0. \tag{2.12}$$

Then, for each $X_0 = x \in \mathbb{R}^d$, *SDE* (2.10) *has a unique global strong solution* $X_t(x)$ *which is* C_b*-strong Feller and irreducible. If we let* $P_t\varphi(x) := \mathbb{E}\varphi(X_t(x))$, *then* P_t *admits a unique invariant probability measure* μ, *and* μ *has a density* $\rho \in L^q(\mathbb{R}^d)$ *with* $q < d/(d-1)$. *Moreover, if* $r = 0$, *then* P_t *is* V*-uniformly exponential ergodic with* $V(x) = 1 + |x|$; *if* $r > 0$, *then* P_t *is uniformly exponential ergodic*.

Remark 2.10. If $b \in L^p_{\text{loc}}(\mathbb{R}^d)$ for some $p > d$, and for some $m > 0$, b satisfies (2.11) for all $|x| > m$, then **(H**^{b}**)** holds. In fact, it suffices to take $b_1 = \chi_m b$ and $b_2 = (1 - \chi_m)b$. The typical function satisfying (2.11) is given by $b_2(x) = -x|x|^r c(x)$ with $0 < c_0 \leq c(x) \leq c_1$. Moreover, if $g(x,z) = \bar\sigma(x)z$ with $\bar\sigma(x) \in L^\infty(\mathbb{R}^d)$ and $\nabla\bar\sigma(x) \in L^p(\mathbb{R}^d)$ for some $p > d$ and $\int_{|z|>1}|z|\nu(\mathrm{d}z) < \infty$, then all the assumptions on g in the above theorem hold.

Remark 2.11. Under some minimal assumptions, Bogachev, Röckner and Shaposhnikov [8] have already shown the absolute continuity of μ with respect to the Lebesgue measure. However, it seems that their results can not be used to our singular case since it is not known whether $b \in L^1_{\text{loc}}(\mu)$.

In the pure jump case, we assume $\nu(\mathrm{d}z) = |z|^{-d-\alpha}\,\mathrm{d}z$ for some $\alpha \in (1,2)$ and

($\widetilde{\mathbf{H}}^b$**)** $b = b_1 + b_2$, where b_2 satisfies (2.11) and the local condition for b in Corollary 2.6, and b_1 satisfies that for some $\theta \in (1-\alpha/2, 1)$ and $p > 2d/\alpha$,

$$(\mathbb{I} - \Delta)^{\theta/2} b_1 \in L^p(\mathbb{R}^d).$$

We have

Theorem 2.12 (Ergodicity for pure jump SDE). *Suppose that* **(H**$^g_\beta$**)** *and* **(**$\widetilde{\mathbf{H}}^b$**)** *hold and for the same* p *in* **(**$\widetilde{\mathbf{H}}^b$**)**,

$$(\Gamma^{1,2}_{0,R}(g))^{1/2}, \Gamma^{1,1}_{R,\infty}(g) \in L^p(\mathbb{R}^d).$$

Then the same conclusions of Theorem 2.9 *hold and the invariant probability measure* μ *has a density* $\rho \in L^q(\mathbb{R}^d)$ *with* $q < d/(d - \alpha + 1)$.

3. General stability and Zvonkin's transformation

In this section, we prepare two basic results: Stability and Zvonkin's transformation for SDE (1.2) under general assumptions. First of all, we introduce the following important notion about Krylov's estimate.

Definition 3.1. Let $X = (X_t)_{t\geq 0}$ be an $\mathcal{F}_t$-adapted process and $p, q \in [1,\infty)$. We say that Krylov's estimate holds for X with index p, q, if for all $T > 0$, there is a constant $c_0 > 0$ such that for all $0 \leq t_0 \leq t_1 \leq T$ and $f \in \mathbb{L}^q_p(t_0, t_1)$,

$$\mathbb{E}\left(\int_{t_0}^{t_1} f(s, X_s)\,\mathrm{d}s \,\Big|\, \mathcal{F}_{t_0}\right) \leq c_0\|f\|_{\mathbb{L}^q_p(t_0,t_1)}, \tag{3.1}$$

where c_0 will be called Krylov's constant of X.

Remark 3.2. Krylov's estimate (3.1) implies that for Lebesgue almost all s, the distribution of random variable X_s admits a density $\rho_s(y)$ with respect to the Lebesgue measure so that

$$\|\rho\|_{\mathbb{L}^{q'}_{p'}(T)} \leq c_0, \qquad \frac{1}{p'} + \frac{1}{p} = 1, \qquad \frac{1}{q'} + \frac{1}{q} = 1,$$

where c_0 is the Krylov constant of X. See [58].

Remark 3.3. Suppose that for some $p, q \in [1, \infty)$, Krylov's estimate holds for X with index p, q. Then the Krylov estimate for X also holds for any $p' \in [p, \infty)$ and q' with $p' - \frac{p'}{q'} = p - \frac{p}{q}$. In fact, by Remark 3.2, it automatically holds that

$$\mathbb{E}\left(\int_{t_0}^{t_1} f(s, X_s)\,\mathrm{d}s \,\Big|\, \mathcal{F}_{t_0}\right) \le \|f\|_{\mathbb{L}^1_\infty(t_0,t_1)}.$$

Notice that by the interpolation theorem (see [4, Theorem 5.1.2]), we have

$$\left(\mathbb{L}^1_\infty(T), \mathbb{L}^q_p(T)\right)_{[\theta]} = \mathbb{L}^{q'}_{p'}(T),$$

where $\theta \in (0,1)$, $\frac{1}{q'} = 1 - \theta + \frac{\theta}{q}$ and $\frac{1}{p'} = \frac{\theta}{p}$, $(\cdot,\cdot)_{[\theta]}$ stands for the complex interpolation. The desired Krylov estimate for $p' \in [p, \infty)$ and q' with $p' - \frac{p'}{q'} = p - \frac{p}{q}$ follows by the interpolation theorem (see [58]).

Remark 3.4. Let $\{X^{(n)}, n \in \mathbb{N}\}$ be a sequence of $\mathcal{F}_t$-adapted processes. Suppose that $X^{(n)}$ satisfies Krylov's estimate with the same index $p, q \in [1, \infty)$ and Krylov's constant c_0. If for each t, $X^{(n)}_t$ converges to X_t in probability as $n \to \infty$, then by the dominated convergence theorem, for every $f \in C_0^\infty(\mathbb{R}_+ \times \mathbb{R}^d)$,

$$\mathbb{E}\left(\int_{t_0}^{t_1} f(s, X_s)\,\mathrm{d}s \,\Big|\, \mathcal{F}_{t_0}\right) = \lim_{n\to\infty} \mathbb{E}\left(\int_{t_0}^{t_1} f\big(s, X^{(n)}_s\big)\,\mathrm{d}s \,\Big|\, \mathcal{F}_{t_0}\right) \le c_0 \|f\|_{\mathbb{L}^q_p(t_0,t_1)}.$$

By a standard monotone class argument, the above inequality still holds for all $f \in \mathbb{L}^q_p(t_0, t_1)$. In other words, X still satisfies the Krylov estimate with the same index p, q and Krylov's constant c_0.

The above definition about Krylov's estimate has the following useful consequence.

Lemma 3.5 (Khasminskii's type estimate). *Let $X = (X_t)_{t\ge 0}$ be an $\mathcal{F}_t$-adapted process. Suppose that X satisfies Krylov's estimate for some $p, q \in [1, \infty)$. Then for any $\lambda, T > 0$, $0 \le t_0 \le t_1 \le T$ and $f \in \mathbb{L}^q_p(T)$,*

$$\mathbb{E}^{\mathcal{F}_{t_0}} \exp\left(\lambda \int_{t_0}^{t_1} \big|f(s, X_s)\big|\,\mathrm{d}s\right) \le 2^n,$$

where $\mathbb{E}^{\mathcal{F}_{t_0}}(\cdot) := \mathbb{E}(\cdot|\mathcal{F}_{t_0})$, and n is chosen so that $\|f\|_{\mathbb{L}^q_p((j-1)T/n, jT/n)} \le \frac{1}{2\lambda c_0}$ for all $j = 1, \dots, n$, and c_0 is the Krylov constant of X.

Proof. Without loss of generality, we assume $t_0 = 0$, $t_1 = T$ and f is nonnegative. For $\lambda > 0$, let us choose n large enough so that for $t_j = \frac{jT}{n}$,

$$\lambda c_0 \|f\|_{\mathbb{L}^p(t_j, t_{j+1})} \le 1/2, \quad j = 0, \dots, n-1. \tag{3.2}$$

For $m \in \mathbb{N}$, noticing that

$$\left(\int_{t_j}^{t_{j+1}} g(s)\,\mathrm{d}s\right)^m = m! \int \cdots \int_{\Delta^m} g(s_1)\cdots g(s_m)\,\mathrm{d}s_1 \cdots \mathrm{d}s_m,$$

where

$$\Delta^m := \big\{(s_1, \dots, s_m) : t_j \le s_1 \le s_2 \le \cdots \le s_m \le t_{j+1}\big\},$$

by (3.1), we have

$$\begin{aligned}\mathbb{E}^{\mathcal{F}_{t_j}}\left(\int_{t_j}^{t_{j+1}} f(s, X_s)\,\mathrm{d}s\right)^m &= m!\,\mathbb{E}^{\mathcal{F}_{t_j}}\left(\int \cdots \int_{\Delta^m} f(s_1, X_{s_1})\cdots f(s_m, X_{s_m})\,\mathrm{d}s_1 \cdots \mathrm{d}s_m\right) \\ &= m!\,\mathbb{E}^{\mathcal{F}_{t_j}}\left(\int \cdots \int_{\Delta^{m-1}} f(s_1, X_{s_1})\cdots f(s_{m-1}, X_{s_{m-1}})\right.\end{aligned}$$

$$\times\, \mathbb{E}^{\mathcal{F}_{s_{m-1}}}\left(\int_{s_{m-1}}^{t_{j+1}} f(s_m, X_{s_m})\,\mathrm{d}s_m\right)\mathrm{d}s_1\cdots\mathrm{d}s_{m-1}\Bigg)$$

$$\le m!\mathbb{E}^{\mathcal{F}_{t_j}}\left(\int\cdots\int_{\Delta^{m-1}} f(s_1, X_{s_1})\cdots f(s_{m-1}, X_{s_{m-1}})\right.$$

$$\left.\times\, c_0\|f\|_{\mathbb{L}^q_p(t_j,t_{j+1})}\,\mathrm{d}s_1\cdots\mathrm{d}s_{m-1}\right) \le \cdots \le m!\big(c_0\|f\|_{\mathbb{L}^q_p(t_j,t_{j+1})}\big)^m,$$

which implies by (3.2) that

$$\mathbb{E}^{\mathcal{F}_{t_j}}\exp\left(\lambda\int_{t_j}^{t_{j+1}} f(s, X_s)\,\mathrm{d}s\right) = \sum_m \frac{1}{m!}\mathbb{E}^{\mathcal{F}_{t_j}}\left(\lambda\int_{t_j}^{t_{j+1}} f(s, X_s)\,\mathrm{d}s\right)^m \le 2.$$

Hence,

$$\mathbb{E}^{\mathcal{F}_0}\exp\left(\lambda\int_0^T f(s, X_s)\,\mathrm{d}s\right) = \mathbb{E}^{\mathcal{F}_0}\left(\prod_{j=0}^{n-1}\exp\left(\lambda\int_{t_j}^{t_{j+1}} f(s, X_s)\,\mathrm{d}s\right)\right)$$

$$= \mathbb{E}^{\mathcal{F}_0}\left(\prod_{j=0}^{n-2}\exp\left(\lambda\int_{t_j}^{t_{j+1}} f(s, X_s)\,\mathrm{d}s\right)\mathbb{E}^{\mathcal{F}_{t_{n-1}}}\exp\left(\lambda\int_{t_{n-1}}^{t_n} f(s, X_s)\,\mathrm{d}s\right)\right)$$

$$\le 2\mathbb{E}^{\mathcal{F}_0}\left(\prod_{j=0}^{n-2}\exp\left(\lambda\int_{t_j}^{t_{j+1}} f(s, X_s)\,\mathrm{d}s\right)\right) \le \cdots \le 2^n.$$

The proof is complete. □

As a result of Lemma 3.5 and using some basic inequalities stated in the Appendix, we have the following result, which will be used below to show the stability of SDEs.

Lemma 3.6. *Let X, Y be two $\mathcal{F}_t$-adapted processes, which satisfy Krylov's estimate with the same index $p, q \in (1, \infty)$ and Krylov's constant c_0. Let $f_t(x) : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ and $g_t(x, z) : \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be two Borel functions. Let $T, R > 0$. Suppose that for some $r \ge 1$,*

$$h(t, x) := \big|\nabla f_t(x)\big|^r + \Gamma^{1,r}_{0,R}(g_t)(x) \in \mathbb{L}^q_p(T).$$

Then there exists an $\mathcal{F}_t$-adapted process ℓ_t with the property

$$\mathbb{E}\mathrm{e}^{\lambda\int_0^T \ell_s\,\mathrm{d}s} \le c\big(\lambda, d, r, c_0, \|h\|_{\mathbb{L}^q_p(T)}\big) < \infty, \quad \forall\lambda > 0,$$

such that for Lebesgue almost all $t \in [0, T]$,

$$\big|f_t(X_t) - f_t(Y_t)\big|^r + \int_{|z|<R}\big|g_t(X_t, z) - g_t(Y_t, z)\big|^r\nu(\mathrm{d}z) \le \ell_t|X_t - Y_t|^r \quad a.s. \tag{3.3}$$

Proof. First of all, using (A.1) in the Appendix with $\mathbb{B} = \mathbb{R}$, we have

$$\big|f_t(x) - f_t(y)\big|^r \le 2^{dr}|x - y|^r\big(\mathcal{M}|\nabla f_t|(x) + \mathcal{M}|\nabla f_t|(y)\big)^r, \tag{3.4}$$

and by (A.1) with $\mathbb{B} = L^r(B_R; \nu)$,

$$\int_{|z|<R}\big|g_t(x, z) - g_t(y, z)\big|^r\nu(\mathrm{d}z) \le 2^{dr}|x - y|^r\big(\mathcal{M}\|\nabla_x g_t\|_{\mathbb{B}}(x) + \mathcal{M}\|\nabla_x g_t\|_{\mathbb{B}}(y)\big)^r$$

$$\le 2^{dr+r}|x - y|^r\big(\mathcal{M}\big(\Gamma^{1,r}_{0,R}(g_t)\big)(x) + \mathcal{M}\big(\Gamma^{1,r}_{0,R}(g_t)\big)(y)\big). \tag{3.5}$$

Now let us define

$$\ell_t = 2^{dr+r}\big[\mathcal{M}h(t,\cdot)(X_t) + \mathcal{M}h(t,\cdot)(Y_t)\big].$$

It follows by (A.3) and Lemma 3.5 that ℓ_t has the desired property. The desired estimate (3.3) follows by (3.4), (3.5) and Remark 3.2. □

Next we show a stochastic Gronwall's inequality, which has independent interest.

Lemma 3.7 (Stochastic Gronwall's inequality). *Let $\xi(t)$ and $\eta(t)$ be two nonnegative càdlàg $\mathcal{F}_t$-adapted processes, A_t a continuous nondecreasing $\mathcal{F}_t$-adapted process with $A_0 = 0$, M_t a local martingale with $M_0 = 0$. Suppose that*

$$\xi(t) \le \eta(t) + \int_0^t \xi(s)\,\mathrm{d}A_s + M_t, \quad \forall t \ge 0. \tag{3.6}$$

Then for any $0 < q < p < 1$ and any stopping time τ, we have

$$\big[\mathbb{E}\big(\xi(\tau)^*\big)^q\big]^{1/q} \le \left(\frac{p}{p-q}\right)^{1/q} \big(\mathbb{E}\mathrm{e}^{pA_\tau/(1-p)}\big)^{(1-p)/p}\mathbb{E}\big(\eta(\tau)^*\big), \tag{3.7}$$

where $\xi(t)^* := \sup_{s\in[0,t]} \xi(s)$.

Proof. We fix a stopping time τ. Without loss of generality, we may assume that the right hand side of (3.7) is finite and $\eta(t)$ is nondecreasing. Otherwise, we may replace $\eta(t)$ with $\eta(t)^*$. Let $\bar\xi(t)$ be the right hand side of (3.6) and $\bar A_t := \int_0^t \xi(s)/\bar\xi(s)\,\mathrm{d}A_s$. Then

$$\xi(t) \le \bar\xi(t) = \eta(t) + \int_0^t \bar\xi(s)\,\mathrm{d}\bar A_s + M_t.$$

By Itô's formula, one has

$$\mathrm{e}^{-\bar A_t}\bar\xi(t) = \eta(0) + \int_0^t \mathrm{e}^{-\bar A_s}\,\mathrm{d}\eta(s) + \int_0^t \mathrm{e}^{-\bar A_s}\,\mathrm{d}M_s.$$

Let $(\tau_n)_{n\in\mathbb{N}}$ be the localization sequence of stopping times of local martingale M, that is, for each $n \in \mathbb{N}$,

$$t \mapsto M_{t\wedge\tau_n} \text{ is a martingale.}$$

Using $\mathrm{e}^{-\bar A_s} \le 1$, we have

$$\mathbb{E}\big(\mathrm{e}^{-\bar A_{t\wedge\tau\wedge\tau_n}}\bar\xi(t\wedge\tau\wedge\tau_n)\big) \le \mathbb{E}\big(\eta(t\wedge\tau\wedge\tau_n)\big) \le \mathbb{E}\big(\eta(t\wedge\tau)\big).$$

Since $\lim_{n\to\infty}\tau_n = \infty$ a.s., by Fatou's lemma, we get

$$\mathbb{E}\big(\mathrm{e}^{-\bar A_\tau}\bar\xi(\tau)\big) \le \mathbb{E}\big(\eta(\tau)\big),$$

which yields by Hölder's inequality, $\xi(t) \le \bar\xi(t)$ and $\bar A_t \le A_t$ that for any $p \in (0,1)$,

$$\mathbb{E}\xi(\tau)^p \le \mathbb{E}\bar\xi(\tau)^p \le \big(\mathbb{E}\mathrm{e}^{pA_\tau/(1-p)}\big)^{1-p}\big[\mathbb{E}\big(\eta(\tau)\big)\big]^p.$$

Now, for any $\lambda > 0$, define a stopping time

$$\tau_\lambda := \inf\big\{s \ge 0 : \xi(s) \ge \lambda\big\}.$$

Since ξ is càdlàg, we have $\xi_{\tau_\lambda} \ge \lambda$ and

$$\lambda^p\mathbb{P}\big(\xi(\tau)^* > \lambda\big) \le \lambda^p\mathbb{P}(\tau_\lambda \le \tau) \le \mathbb{E}\xi(\tau\wedge\tau_\lambda)^p \le \big(\mathbb{E}\mathrm{e}^{pA_\tau/(1-p)}\big)^{1-p}\big[\mathbb{E}\big(\eta(\tau)\big)\big]^p =: \delta,$$

and for any $q \in (0,p)$,

$$\mathbb{E}\big|\xi(\tau)^*\big|^q = q\int_0^\infty \lambda^{q-1}\mathbb{P}\big(\xi(\tau)^* > \lambda\big)\,\mathrm{d}\lambda \le q\int_0^\infty \lambda^{q-1}\big(\big(\lambda^{-p}\delta\big)\wedge 1\big)\,\mathrm{d}\lambda = p\delta^{q/p}/(p-q).$$

The proof is complete. □

Remark 3.8. In [44], Scheutzow proved (3.7) for continuous martingales. His proof depends on a martingale inequality of Burkholder, which does not hold for discontinuous martingale as pointed out by him. Compared with the proof provided in [44], our proof is more elementary. Recently, a discrete version of stochastic Gronwall's inequality is also established by Kruse and Scheutzow [28].

The following general stability result and Zvonkin's transformation will be our cornerstone, which will be used several times in Section 6 and Section 7.

Theorem 3.9 (Stability). *For $i=1,2$, let $X^{(i)}_t$ satisfy the following SDE*

$$X^{(i)}_t = X^{(i)}_0 + \int_0^t \sigma^{(i)}_s\big(X^{(i)}_s\big)\,\mathrm{d}W_s + \int_0^t b^{(i)}_s\big(X^{(i)}_s\big)\,\mathrm{d}s + \int_0^t\int_{|z|<R} g^{(i)}_s\big(X^{(i)}_{s-},z\big)\tilde N(\mathrm{d}s,\mathrm{d}z),$$

where $(\sigma^{(i)}, b^{(i)}, g^{(i)})$ are two families of measurable coefficients. Let $r\geq 1$. Suppose that $X^{(i)}$ satisfies Krylov's estimate with index $p,q\in(1,\infty)$, and for all $T>0$, there are $p_i\in[p,\infty]$ and $q_i = 1/(1-(p-p/q)/p_i)$, $i=1,2,3,4$ such that

$$\hbar := \big\|\nabla\sigma^{(1)}\big\|^2_{\mathbb{L}^{2q_1}_{2p_1}(T)} + \big\|\nabla b^{(1)}\big\|_{\mathbb{L}^{q_2}_{p_2}(T)} + \big\|\Gamma^{1,2}_{0,R}\big(g^{(1)}\big)\big\|_{\mathbb{L}^{q_3}_{p_3}(T)} + \big\|\Gamma^{1,2r}_{0,R}\big(g^{(1)}\big)\big\|_{\mathbb{L}^{q_4}_{p_4}(T)} < \infty,$$

where $\Gamma^{j,\alpha}_{0,R}(g^{(1)})$ is defined by (2.1). Then for any $\theta\in(0,1)$ and $T>0$,

$$\Big[\mathbb{E}\Big(\sup_{t\in[0,T]}\big|X^{(1)}_t - X^{(2)}_t\big|^{2r\theta}\Big)\Big]^{1/\theta} \leq c_1\Big[\mathbb{E}\big|X^{(1)}_0 - X^{(2)}_0\big|^{2r} + \mathbb{E}\Big(\int_0^T \delta_s\big(X^{(2)}_s\big)\,\mathrm{d}s\Big)\Big],$$

where c_1 only depends on $T, r, \theta, p, q, d, \hbar$ and the Krylov constant of $X^{(i)}$, and

$$\begin{aligned}\delta_s(x) :=& \big\|\sigma^{(1)}_s(x) - \sigma^{(2)}_s(x)\big\|^{2r} + \big|b^{(1)}_s(x) - b^{(2)}_s(x)\big|^{2r}\\ &+ \Gamma^{0,2r}_{0,R}\big(g^{(1)}_s - g^{(2)}_s\big)(x) + \big(\Gamma^{0,2}_{0,R}\big(g^{(1)}_s - g^{(2)}_s\big)(x)\big)^r.\end{aligned} \tag{3.8}$$

Proof. For simplicity of notations, we write $Z_t := X^{(1)}_t - X^{(2)}_t$ and

$$\begin{aligned}&\Sigma_t := \sigma^{(1)}_t\big(X^{(1)}_t\big) - \sigma^{(2)}_t\big(X^{(2)}_t\big), \qquad B_t := b^{(1)}_t\big(X^{(1)}_t\big) - b^{(2)}_t\big(X^{(2)}_t\big),\\ &G_t(z) := g^{(1)}_t\big(X^{(1)}_t, z\big) - g^{(2)}_t\big(X^{(2)}_t, z\big).\end{aligned}$$

Since $X^{(i)}$ satisfies Krylov's estimate with index $p,q\in(1,\infty)$, by the assumption, Remark 3.3 and Lemma 3.6, there exist $\mathscr{F}_t$-adapted processes $\ell^{(j)}_t$ with

$$\mathbb{E}\mathrm{e}^{\lambda\int_0^T \ell^{(j)}_s\,\mathrm{d}s} \leq c(\lambda,\hbar) < \infty, \quad \lambda>0,\ j=1,2,3,4, \tag{3.9}$$

such that

$$\begin{aligned}&|\Sigma_t|^2 \leq \ell^{(1)}_t|Z_t|^2 + 2\big\|\sigma^{(1)}_t - \sigma^{(2)}_t\big\|^2\big(X^{(2)}_t\big),\\ &|B_t| \leq \ell^{(2)}_t|Z_t| + \big|b^{(1)}_t - b^{(2)}_t\big|\big(X^{(2)}_t\big),\\ &\int_{|z|<R}\big|G_t(z)\big|^2\nu(\mathrm{d}z) \leq \ell^{(3)}_t|Z_t|^2 + 2\Gamma^{0,2}_{0,R}\big(g^{(1)}_t - g^{(2)}_t\big)\big(X^{(2)}_t\big),\\ &\int_{|z|<R}\big|G_t(z)\big|^{2r}\nu(\mathrm{d}z) \leq \ell^{(3)}_t|Z_t|^{2r} + 2^r\Gamma^{0,2r}_{0,R}\big(g^{(1)}_t - g^{(2)}_t\big)\big(X^{(2)}_t\big),\end{aligned} \tag{3.10}$$

where $r\geq 1$. Now, by Itô's formula, we have

$$\begin{aligned}\mathrm{d}|Z_t|^{2r} =& \big(r\|\Sigma_t\|^2|Z_t|^{2(r-1)} + 2r(r-1)|\Sigma_t Z_t|^2|Z_t|^{2(r-2)} + 2r\langle B_t, Z_t\rangle|Z_t|^{2(r-1)}\big)\,\mathrm{d}t\\ &+ \Big[\int_{|z|<R}\big(\big|Z_t + G_t(z)\big|^{2r} - |Z_t|^{2r} - 2r\big\langle G_t(z), Z_t\big\rangle|Z_t|^{2(r-1)}\big)\nu(\mathrm{d}z)\Big]\,\mathrm{d}t + \mathrm{d}M_t,\end{aligned}$$

where M_t is a local martingale. Noticing that

$$|x+y|^{2r}-|x|^{2r}-2r\langle y,x\rangle|x|^{2(r-1)}\lesssim|y|^{2r}+|y|^2|x|^{2(r-1)},$$

by (3.10) and Young's inequality, we get

$$\mathrm{d}|Z_t|^{2r}\lesssim|Z_t|^{2r}\big(\ell_t^{(1)}+\ell_t^{(2)}+\ell_t^{(3)}+\ell_t^{(4)}+1\big)\,\mathrm{d}t+\delta_t\big(X_t^{(2)}\big)\,\mathrm{d}t+\mathrm{d}M_t,$$

where $\delta_t(x)$ is defined by (3.8). By Lemma 3.7 and (3.9), we obtain the desired estimate. □

The following proposition provides a way of transforming SDE (1.2) into a new SDE, which is called Zvonkin's transformation in the literature.

Theorem 3.10 (Zvonkin's transformation). *For each $t\geq 0$, let $\Phi_t(x)$ be a homeomorphism over $\mathbb{R}^d$. Let $p,q\in(1,\infty)$. Suppose that there exist a sequence of smooth functions Φ^n and a function $\bar{b}\in L^q_{\mathrm{loc}}(\mathbb{R}_+;L^p_{\mathrm{loc}}(\mathbb{R}^d))$ such that for each $T>0$ and $(t,x)\in\mathbb{R}_+\times\mathbb{R}^d$, $m\in\mathbb{N}$,*

$$\sup_{n\in\mathbb{N}}\|\nabla\Phi^n\|_{\mathbb{L}^\infty(T)}<\infty,\qquad\lim_{n\to\infty}\Phi^n_t(x)=\Phi_t(x),\qquad\lim_{n\to\infty}\|\nabla\big(\Phi^n-\Phi\big)\chi_m\|_{\mathbb{L}^q_p(T)}=0,\tag{3.11}$$

and

$$\lim_{n\to\infty}\|\big(\big(\partial_s+\mathscr{L}_2^\sigma+\mathscr{L}_1^b+\mathscr{L}^g_{\nu,R}\big)\Phi^n-\bar{b}\big)\chi_m\|_{\mathbb{L}^q_p(T)}=0,$$

where χ_m is the cutoff function defined by (2.8). If X solves SDE (1.2) and satisfies Krylov's estimate with the above index p,q, then $Y_t:=\Phi_t(X_t)$ solves the following SDE:

$$\mathrm{d}Y_t=\tilde{\sigma}_t(Y_t)\,\mathrm{d}W_t+\tilde{b}_t(Y_t)\,\mathrm{d}t+\int_{|z|<R}\tilde{g}_t(Y_{t-},z)\tilde{N}(\mathrm{d}t,\mathrm{d}z)+\int_{|z|\geq R}\tilde{g}_t(Y_{t-},z)N(\mathrm{d}t,\mathrm{d}z),\tag{3.12}$$

where

$$\tilde{\sigma}_t(y):=(\nabla\Phi_t\cdot\sigma_t)\circ\Phi_t^{-1}(y),\qquad\tilde{b}_t(y):=\bar{b}_t\big(\Phi_t^{-1}(y)\big),$$
$$\tilde{g}_t(y,z):=\Phi_t\big(\Phi_t^{-1}(y)+g_t\big(\Phi_t^{-1}(y),z\big)\big)-y.$$

Proof. By Itô's formula, we have

$$\begin{aligned}\Phi^n_t(X_t)=\Phi^n_0(X_0)&+\int_0^t\big(\nabla\Phi^n_s\cdot\sigma_s\big)(X_s)\,\mathrm{d}W_s\\&+\int_0^t\int_{|z|<R}\big[\Phi^n_s\big(X_{s-}+g_s(X_{s-},z)\big)-\Phi^n_s(X_{s-})\big]\tilde{N}(\mathrm{d}s,\mathrm{d}z)\\&+\int_0^t\int_{|z|\geq R}\big[\Phi^n_s\big(X_{s-}+g_s(X_{s-},z)\big)-\Phi^n_s(X_{s-})\big]N(\mathrm{d}s,\mathrm{d}z)\\&+\int_0^t\big(\big(\partial_s+\mathscr{L}_2^\sigma+\mathscr{L}_1^b+\mathscr{L}^g_{\nu,R}\big)\Phi^n_s\big)(X_s)\,\mathrm{d}s.\end{aligned}$$

Since X satisfies Krylov's estimate with index p,q, by the assumptions and taking limits $n\to\infty$, we obtain SDE (3.12). For example, for each $m\in\mathbb{N}$, define

$$\tau_m:=\inf\left\{t>0:|X_t|+\int_0^t\int_{|z|<R}\big|g_s(X_s,z)\big|^2\nu(\mathrm{d}z)\,\mathrm{d}s>m\right\}.$$

By (3.11) and the dominated convergence theorem, we have

$$\mathbb{E}\bigg|\int_0^{t\wedge\tau_m}\int_{|z|<R}\big[\Phi^n_s\big(X_{s-}+g_s(X_{s-},z)\big)-\Phi^n_s(X_{s-})$$

$$
\begin{aligned}
&- \Phi_s\big(X_{s-}+g_s(X_{s-},z)\big)+\Phi_s(X_{s-})\big]\tilde N(\mathrm{d}s,\mathrm{d}z)\bigg|^2 \\
&= \mathbb{E}\int_0^{t\wedge\tau_m}\int_{|z|<R}\big|\Phi_s^n\big(X_s+g_s(X_s,z)\big)-\Phi_s^n(X_s) \\
&\quad - \Phi_s\big(X_s+g_s(X_s,z)\big)+\Phi_s(X_s)\big|^2\nu(\mathrm{d}z)\,\mathrm{d}s\to 0,\quad n\to\infty.
\end{aligned}
$$

Moreover, by Krylov's estimate for X and the assumption, we also have

$$
\begin{aligned}
&\mathbb{E}\bigg(\int_0^{t\wedge\tau_m}\big|\big(\partial_s+\mathscr{L}_2^\sigma+\mathscr{L}_1^b+\mathscr{L}_{\nu,R}^g\big)\Phi_s^n-\bar b_s\big|(X_s)\,\mathrm{d}s\bigg) \\
&\quad\le \mathbb{E}\bigg(\int_0^t\big|\big(\big(\partial_s+\mathscr{L}_2^\sigma+\mathscr{L}_1^b+\mathscr{L}_{\nu,R}^g\big)\Phi_s^n-\bar b_s\big)\chi_m\big|(X_s)\,\mathrm{d}s\bigg) \\
&\quad\le c\big\|\big(\big(\partial_s+\mathscr{L}_2^\sigma+\mathscr{L}_1^b+\mathscr{L}_{\nu,R}^g\big)\Phi^n-\bar b\big)\chi_m\big\|_{\mathbb{L}_p^q(t)}\to 0,\quad n\to\infty.
\end{aligned}
$$

The proof is complete since $\tau_m\to\infty$ as $m\to\infty$. □

4. A study of parabolic integro-differential equations

This section is devoted to a careful study of the Kolmogorov backward equation associated to SDE (1.2). First of all, we introduce some Sobolev spaces and notations for later use. For $(p,\alpha)\in[1,\infty]\times(0,2]\setminus\{\infty\}\times\{1,2\}$, let $H_p^\alpha:=(\mathbb{I}-\Delta)^{-\alpha/2}(L^p(\mathbb{R}^d))$ be the usual Bessel potential space with norm

$$
\|f\|_{\alpha,p}:=\big\|(\mathbb{I}-\Delta)^{\alpha/2}f\big\|_p\asymp\|f\|_p+\big\|\Delta^{\alpha/2}f\big\|_p,
$$

where $\|\cdot\|_p$ is the usual L^p-norm in $\mathbb{R}^d$, and $(\mathbb{I}-\Delta)^{\alpha/2}f$ and $\Delta^{\alpha/2}f$ are defined through the Fourier transformation

$$
(\mathbb{I}-\Delta)^{\alpha/2}f:=\mathcal{F}^{-1}\big(\big(1+|\cdot|^2\big)^{\alpha/2}\mathcal{F}f\big),\qquad \Delta^{\alpha/2}f:=\mathcal{F}^{-1}\big(|\cdot|^\alpha\mathcal{F}f\big).
$$

For $p=\infty$ and $j=1,2$, we define H_∞^j as the space of functions with finite norm

$$
\|f\|_{j,\infty}:=\|f\|_\infty+\big\|\nabla^j f\big\|_\infty<\infty.
$$

Notice that for $n=1,2$ and $p\in(1,\infty)$, an equivalent norm in H_p^n is given by

$$
\|f\|_{n,p}=\|f\|_p+\big\|\nabla^n f\big\|_p,
$$

and for $\alpha\in(0,2)$, up to a multiple constant, an alternative expression of $\Delta^{\alpha/2}$ is given by

$$
\Delta^{\alpha/2}f(x)=\text{p.v.}\int_{\mathbb{R}^d}\frac{f(x+y)-f(x)}{|y|^{d+\alpha}}\,\mathrm{d}y, \tag{4.1}
$$

where p.v. stands for the Cauchy principal value. We shall need the following Sobolev embedding: for $p\in[1,\infty]$ and $\alpha\in[0,2]$,

$$
\begin{cases}
H_p^\alpha\subset L^q,\quad q\in[p,\frac{dp}{d-\alpha p}],\quad \alpha p<d;\\
H_p^\alpha\subset H_\infty^{\alpha-d/p}\subset C_b^{\alpha-d/p},\quad \alpha p>d,
\end{cases} \tag{4.2}
$$

where C_b^β is the usual Hölder space. Moreover, for $\alpha\in[0,1]$ and $p\in(1,\infty]$, there is a constant $c=c(p,d,\alpha)>0$ such that for all $f\in H_p^\alpha$,

$$
\big\|f(\cdot+y)-f(\cdot)\big\|_p\le c\big(|y|^\alpha\wedge 1\big)\|f\|_{\alpha,p}. \tag{4.3}
$$

The above facts are standard and can be found in [4, Chapter 6] or [48].

The following lemma due to [37, Lemma 5] strengthens the estimate (4.3), which will play an important role in the following.

Lemma 4.1. *For $\alpha \in (0,2]$, write $y^{(\alpha)} := y1_{\alpha\in[1,2]}$. For any $p \in (\frac{d}{\alpha} \vee 1, \infty]$, there is a constant $c = c(p,d,\alpha) > 0$ such that for all $f \in H^\alpha_p$,*

$$\left\| \sup_{y\neq 0} \left| f(x+y) - f(x) - y^{(\alpha)} \cdot \nabla f(x) \right| / |y|^\alpha \right\|_p \leq c\|f\|_{\alpha,p}.$$

In the following, given $0 \leq S \leq T < \infty$, $\alpha \in (0,2]$ and $q, p \in [1,\infty]$, we write

$$\mathbb{H}^{\alpha,q}_p(S,T) := L^q\big([S,T]; H^\alpha_p\big), \qquad \mathbb{H}^{\alpha,q}_p(T) := \mathbb{H}^{\alpha,q}_p(0,T).$$

4.1. *Second order integro-differential equations*

Let $a(t,x) : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{M}^d_{\mathrm{sym}}$ be a Borel measurable function, where $\mathbb{M}^d_{\mathrm{sym}}$ denotes the space of all symmetric $d \times d$-matrices. We introduce the following second order partial differential operator:

$$\mathscr{L}^a_2 u := a^{ij}\partial_i\partial_j u.$$

For $\lambda \geq 0$ and $R, T > 0$, let us consider the following backward second order parabolic integro-differential equation:

$$\partial_t u + \big(\mathscr{L}^a_2 - \lambda\big)u + \mathscr{L}^b_1 u + \mathscr{L}^g_{\nu,R} u = f, \quad u(T,x) = 0, \tag{4.4}$$

where $\mathscr{L}^g_{\nu,R}$ is defined in (1.5). We make the following basic assumptions on a.

($\mathbf{H}^a_\beta$) There are constants $c_0 \geq 1$ and $\beta \in (0,1)$ such that for all $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$,

$$c_0^{-1}|\xi|^2 \leq a^{ij}(t,x)\xi_i\xi_j \leq c_0|\xi|^2, \quad \xi \in \mathbb{R}^d,$$

and

$$\big\|a(t,x) - a(t,y)\big\| \leq c_0|x-y|^\beta.$$

Notice that ($\mathbf{H}^\sigma_\beta$) implies ($\mathbf{H}^a_\beta$) for $a(t,x) = (\sigma\sigma^*)(t,x)/2$.

Under ($\mathbf{H}^a_\beta$), it is well known that $\mathscr{L}^a_2$ admits a fundamental solution (also called heat kernel) $\rho(s,x;t,y)$ so that

$$\partial_s\rho(s,x;t,y) + \mathscr{L}^a_2\rho(s,\cdot;t,y)(x) = 0, \qquad \lim_{s\uparrow t}\rho(s,x;t,y) = \delta_{x-y},$$

and $\rho(s,x;t,y)$ enjoys the following gradient estimates (see [17, Chapter 2] or [11]):

$$\big|\nabla^j_x\rho(s,\cdot;t,y)\big|(x) \leq c_1(t-s)^{-(d+j)/2}\mathrm{e}^{-c_1|x-y|^2/(t-s)}, \quad j = 0,1,2. \tag{4.5}$$

Moreover, we also have the following fractional derivative estimate: for any $\vartheta \in (0,2)$,

$$\big|\Delta^{\vartheta/2}\rho(s,\cdot;t,y)\big|(x) \leq c_2\big(|x-y| + (t-s)^{1/2}\big)^{-d-\vartheta}. \tag{4.6}$$

In the case $b = g \equiv 0$, PDE (4.4) has been well-studied. The following $\mathbb{L}^q_p$-estimate (4.7) of $\nabla^2 u$ was proven by Kim [27] for $1 < p \leq q < \infty$. By duality, one in fact can drop the restriction $p \leq q$.

Lemma 4.2. *Let $\lambda, T \geq 0$ and $p, q \in (1,\infty)$. Under ($\mathbf{H}^a_\beta$), for any $f \in \mathbb{L}^q_p(T)$, there exists a unique solution $u \in \mathbb{H}^{2,q}_p(T)$ to the following backward PDE:*

$$\partial_t u + \big(\mathscr{L}^a_2 - \lambda\big)u = f, \quad u(T,x) = 0,$$

and there exists a constant $c_1 = c_1(d,p,q,T,c_0,\beta) > 0$ such that for all $\lambda \geq 0$,

$$\big\|\nabla^2 u\big\|_{\mathbb{L}^q_p(T)} \leq c_1\|f\|_{\mathbb{L}^q_p(T)}. \tag{4.7}$$

Moreover, for any $\vartheta \in [0,2)$ and $p' \in [p,\infty]$, $q' \in [q,\infty]$ satisfying

$$\frac{d}{p} + \frac{2}{q} < 2 - \vartheta + \frac{d}{p'} + \frac{2}{q'}, \tag{4.8}$$

there exists a constant $c_2 = c_2(d, p, q, \vartheta, p', q', T, c_0) > 0$ *such that for all* $\lambda > 0$ *and* $S \in (0, T)$,

$$\lambda^{\frac{1}{2}(2-\vartheta+\frac{d}{p'}+\frac{2}{q'}-\frac{d}{p}-\frac{2}{q})}\|u\|_{\mathbb{H}^{\vartheta,q'}_{p'}(S,T)} \le c_2\|f\|_{\mathbb{L}^q_p(S,T)}. \tag{4.9}$$

Proof. The existence and uniqueness of solutions and estimate (4.7) can be found in [27]. We only need to show the estimate (4.9). Without loss of generality, we assume $f \in C_c^\infty([0, T] \times \mathbb{R}^d)$. By Duhamel's formula, we can write

$$u(s, x) = \int_s^T \mathrm{e}^{-\lambda(t-s)}\left(\int_{\mathbb{R}^d}\rho(s, x; t, y)f(t, y)\,\mathrm{d}y\right)\mathrm{d}t.$$

Let $r := 1/(1 - 1/p + 1/p')$ and $\varrho_\vartheta(t, x) := (|x| + t^{1/2})^{-d-\vartheta}$. Suppose $(p', \vartheta) \neq (\infty, 1)$. By (4.6) and Young's convolution inequality, we have for $\vartheta \in (0, 2)$,

$$\begin{aligned}
\|\Delta^{\vartheta/2}u(s)\|_{p'} &\le \int_s^T \mathrm{e}^{-\lambda(t-s)}\left\|\int_{\mathbb{R}^d}\Delta_x^{\vartheta/2}\rho(s, \cdot; t, y)f(t, y)\,\mathrm{d}y\right\|_{p'}\mathrm{d}t\\
&\lesssim \int_s^T \mathrm{e}^{-\lambda(t-s)}\left\|\int_{\mathbb{R}^d}\varrho_\vartheta(t-s, \cdot - y)\big|f(t, y)\big|\,\mathrm{d}y\right\|_{p'}\mathrm{d}t\\
&\le \int_s^T \mathrm{e}^{-\lambda(t-s)}\|\varrho_\vartheta(t-s, \cdot)\|_r\|f(t)\|_p\,\mathrm{d}t\\
&\lesssim \int_s^T \mathrm{e}^{-\lambda(t-s)}(t-s)^{(d/r-\vartheta-d)/2}\|f(t)\|_p\,\mathrm{d}t\\
&= \big(h_\lambda * \big(\|f(\cdot)\|_p 1_{[S,T]}\big)\big)(s),
\end{aligned}$$

where $h_\lambda(t) := \mathrm{e}^{-\lambda t}t^{(d/r-\vartheta-d)/2}1_{t>0}$. Hence, by Young's convolution inequality again,

$$\|\Delta^{\vartheta/2}u\|_{\mathbb{L}^{q'}_{p'}(S,T)} \lesssim \|h_\lambda\|_{L^{1/(1+1/q'-1/q)}(0,T-S)}\|f\|_{\mathbb{L}^q_p(S,T)} \lesssim \lambda^{\frac{1}{2}(\vartheta-2-\frac{d}{p'}-\frac{2}{q'}+\frac{d}{p}+\frac{2}{q})}\|f\|_{\mathbb{L}^q_p(S,T)}.$$

For $(p', \vartheta) = (\infty, 1)$, by the gradient estimate (4.5), we still have

$$\|\nabla u\|_{\mathbb{L}^{q'}_\infty(S,T)} \lesssim \lambda^{\frac{1}{2}(-1-\frac{2}{q'}+\frac{d}{p}+\frac{2}{q})}\|f\|_{\mathbb{L}^q_p(S,T)}.$$

Moreover, using the upper bound estimate of the heat kernel, we also have

$$\|u\|_{\mathbb{L}^{q'}_{p'}(S,T)} \lesssim \lambda^{-1-\frac{1}{2}(\frac{d}{p'}+\frac{2}{q'}-\frac{d}{p}-\frac{2}{q})}\|f\|_{\mathbb{L}^q_p(S,T)}.$$

Combining the above calculations, we get (4.9). □

With this result in hand, we now prove the following solvability of the integro-differential equation (4.4).

Theorem 4.3. *Let* $p \in (d/2 \vee 1, \infty)$, $q \in (1, \infty)$ *and* $T > 0$. *Let* $\Gamma^{0,2}_{0,R}(g)$ *be defined by* (2.1). *Assume that* $(\mathbf{H}^a_\beta)$ *holds and*

(i) *for some* $p_1 \in [p, \infty]$ *and* $q_1 \in [q, \infty]$ *with* $\frac{d}{p_1} + \frac{2}{q_1} < 1$, $b \in \mathbb{L}^{q_1}_{p_1}(T)$;
(ii) $\Gamma^{0,2}_{0,R}(g) \in \mathbb{L}^\infty(T)$ *and* $\lim_{\varepsilon\to 0}\|\Gamma^{0,2}_{0,\varepsilon}(g)\|_{\mathbb{L}^\infty(T)} = 0$.

Then for some $\lambda_0 > 0$ *depending on* $\|b\|_{\mathbb{L}^{q_1}_{p_1}(T)}$ *and* $\|\Gamma^{0,2}_{0,R}(g)\|_{\mathbb{L}^\infty(T)}$, *and for all* $\lambda \ge \lambda_0$ *and* $f \in \mathbb{L}^q_p(T)$, *there exists a unique solution* $u \in \mathbb{H}^{2,q}_p(T)$ *to the equation* (4.4). *Moreover, in this case the estimates* (4.7) *and* (4.9) *still hold and* $\partial_t u \in \mathbb{L}^q_p(T)$.

Proof. By standard continuity method, it suffices to show the a priori estimates (4.7) and (4.9) for equation (4.4) under the assumptions in the theorem. First of all, for any $\vartheta \in [0, 2)$ and $p' \in [p, \infty]$, $q' \in [q, \infty]$ satisfying (4.8), by (4.7) and

(4.9) we have

$$\lambda^{\frac{1}{2}(2-\vartheta+\frac{d}{p'}+\frac{2}{q'}-\frac{d}{p}-\frac{2}{q})}\|u\|_{\mathbb{H}^{\vartheta,q'}_{p'}(S,T)} + \|\nabla^2 u\|_{\mathbb{L}^q_p(S,T)} \le c_1 \|f + \mathcal{L}^b_1 u + \mathcal{L}^g_{\nu,R} u\|_{\mathbb{L}^q_p(S,T)}. \tag{4.10}$$

Below, for simplicity of notation, we drop the time variable t. Recalling the definitions of $\mathcal{L}^g_{\nu,R}u$ and $\Gamma^{0,\alpha}_{\varepsilon,R}(g)$ (see (2.1)), we have for any $\varepsilon \in (0, R)$,

$$\begin{aligned}\left|\mathcal{L}^g_{\nu,\varepsilon}u(x)\right| &\le \int_{|z|\le\varepsilon}\left|u\big(x+g(x,z)\big)-u(x)-g(x,z)\cdot\nabla u(x)\right|\nu(\mathrm{d}z)\\ &\le \sup_{y\neq 0}|y|^{-2}\left|u(x+y)-u(y)-y\cdot\nabla u(x)\right|\left|\Gamma^{0,2}_{0,\varepsilon}(g)(x)\right|,\end{aligned}$$

and for $\alpha \in (d/p \vee 1, 2)$,

$$\begin{aligned}\left|\mathcal{L}^g_{\nu,R}u(x)-\mathcal{L}^g_{\nu,\varepsilon}u(x)\right| &\le \int_{\varepsilon<|z|<R}\left|u\big(x+g(x,z)\big)-u(x)-g(x,z)\cdot\nabla u(x)\right|\nu(\mathrm{d}z)\\ &\le \sup_{y\neq 0}|y|^{-\alpha}\left|u(x+y)-u(y)-y\cdot\nabla u(x)\right|\left|\Gamma^{0,\alpha}_{\varepsilon,R}(g)(x)\right|.\end{aligned}$$

Thus, thanks to $p > d/\alpha \vee 1$, by Lemma 4.1, we obtain that for any $\varepsilon \in (0, R)$,

$$\begin{aligned}\|\mathcal{L}^g_{\nu,R}u\|_{\mathbb{L}^q_p(S,T)} &\le \|\mathcal{L}^g_{\nu,\varepsilon}u\|_{\mathbb{L}^q_p(S,T)} + \|\mathcal{L}^g_{\nu,R}u-\mathcal{L}^g_{\nu,\varepsilon}u\|_{\mathbb{L}^q_p(S,T)}\\ &\lesssim \|\Gamma^{0,2}_{0,\varepsilon}(g)\|_{\mathbb{L}^\infty(T)}\|u\|_{\mathbb{H}^{2,q}_p(S,T)} + \|\Gamma^{0,\alpha}_{\varepsilon,R}(g)\|_{\mathbb{L}^\infty(T)}\|u\|_{\mathbb{H}^{\alpha,q}_p(S,T)}.\end{aligned} \tag{4.11}$$

On the other hand, letting $q_2 := qq_1/(q_1-q)$ and $p_2 := pp_1/(p_1-p)$, by Hölder's inequality, we have

$$\|\mathcal{L}^b_1 u\|_{\mathbb{L}^q_p(S,T)} \le \|b\|_{\mathbb{L}^{q_1}_{p_1}(S,T)}\|u\|_{\mathbb{H}^{1,q_2}_{p_2}(S,T)}. \tag{4.12}$$

Now by (4.10), (4.11) and (4.12), there are $c_2, c_3 > 0$ such that for all $\varepsilon \in (0, R)$,

$$\begin{aligned}&\lambda^{\frac{1}{2}(1+\frac{d}{p_2}+\frac{2}{q_2}-\frac{d}{p}-\frac{2}{q})}\|u\|_{\mathbb{H}^{1,q_2}_{p_2}(S,T)} + \lambda^{1-\frac{\alpha}{2}}\|u\|_{\mathbb{H}^{\alpha,q}_p(S,T)} + \|\nabla^2 u\|_{\mathbb{L}^q_p(S,T)}\\ &\quad\le c_2\Big(\|\Gamma^{0,2}_{0,\varepsilon}(g)\|_{\mathbb{L}^\infty(T)}\|u\|_{\mathbb{H}^{2,q}_p(S,T)} + \|\Gamma^{0,\alpha}_{\varepsilon,R}(g)\|_{\mathbb{L}^\infty(T)}\|u\|_{\mathbb{H}^{\alpha,q}_p(S,T)}\Big)\\ &\qquad + c_3\Big(\|b\|_{\mathbb{L}^{q_1}_{p_1}(T)}\|u\|_{\mathbb{H}^{1,q_2}_{p_2}(S,T)} + \|f\|_{\mathbb{L}^q_p(S,T)}\Big),\end{aligned}$$

which implies that for ε small enough and some λ_0 large enough and all $\lambda \ge \lambda_0$,

$$\|u\|_{\mathbb{H}^{1,q_2}_{p_2}(S,T)} + \|u\|_{\mathbb{H}^{\alpha,q}_p(S,T)} + \|\nabla^2 u\|_{\mathbb{L}^q_p(S,T)} \lesssim \|f\|_{\mathbb{L}^q_p(S,T)}.$$

Here we have used that $\lim_{\varepsilon\to 0}\|\Gamma^{0,2}_{0,\varepsilon}(g)\|_{\mathbb{L}^\infty(T)} = 0$ and

$$\|\Gamma^{0,\alpha}_{\varepsilon,R}\|_{\mathbb{L}^\infty(T)} \le \|\Gamma^{0,2}_{\varepsilon,R}\|^{\alpha/2}_{\mathbb{L}^\infty(T)}\nu\big(\{z : \varepsilon < |z| < R\}\big)^{1-\frac{\alpha}{2}}.$$

Substituting this estimate into (4.10), (4.11) and (4.26), we get the estimates (4.7) and (4.9). The proof is finished. □

4.2. *Non-local parabolic equations*

In this subsection we assume $\alpha \in (1, 2)$ and introduce the following nonlocal operator:

$$\mathcal{L}^\kappa_\alpha f(x) := \int_{\mathbb{R}^d}\big[f(x+z)-f(x)-z\cdot\nabla f(x)\big]\frac{\kappa(t,x,z)}{|z|^{d+\alpha}}\,\mathrm{d}z, \tag{4.13}$$

where the kernel function $\kappa(t, x, z) : \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ satisfies

($\mathbf{H}^\kappa_\beta$) There exist constants $\kappa_0 > 1$ and $\beta \in (0, 1]$ such that for all $t \geq 0$ and $x, y, z \in \mathbb{R}^d$,

$$\kappa_0^{-1} \leq \kappa(t, x, z) \leq \kappa_0, \qquad \big|\kappa(t, x, z) - \kappa(t, y, z)\big| \leq \kappa_0 |x - y|^\beta. \tag{4.14}$$

Under ($\mathbf{H}^\kappa_\beta$), it is well known that $\mathcal{L}^\kappa_\alpha$ admits a fundamental solution $\rho_\kappa(s, x; t, y)$ so that (see [13] or [14, Theorem 1.1])

$$\partial_s \rho_\kappa(s, x; t, y) + \mathcal{L}^\kappa_\alpha \rho_\kappa(s, \cdot; t, y)(x) = 0, \qquad \lim_{s \uparrow t} \rho_\kappa(s, x; t, y) = \delta_{x-y},$$

and $\rho_\kappa(s, x; t, y)$ enjoys the following estimates: for $j = 0, 1$ and $T > 0$, there is a constant $c > 0$ such that for all $0 \leq s < t \leq T$ and $x, y \in \mathbb{R}^d$,

$$\big|\nabla^j_x \rho_\kappa(s, \cdot; t, y)\big|(x) \leq c \varrho^{(\alpha)}_{-j}(t - s, x - y), \tag{4.15}$$

and for any $\theta \in (0, (\alpha + \beta) \wedge 2)$,

$$\big|\Delta^{\theta/2}_x \rho_\kappa(s, x; t, y)\big| \leq c \varrho^{(\alpha \wedge \theta)}_{-\theta}(t - s, x - y), \tag{4.16}$$

where for $\eta \geq 0$ and $\gamma \in \mathbb{R}$,

$$\varrho^{(\eta)}_\gamma(t, x) := \frac{t^{(\gamma+\eta)/\alpha}}{(t^{1/\alpha} + |x|)^{d+\eta}} = \frac{t^{(\gamma-d)/\alpha}}{(1 + |x|/t^{1/\alpha})^{d+\eta}}. \tag{4.17}$$

It is easy to see that for any $p \geq 1$, there is a $c > 0$ such that

$$\big\|\varrho^{(\eta)}_\gamma(t)\big\|_p \leq c t^{(\gamma-d)/\alpha + d/(\alpha p)}, \quad t > 0. \tag{4.18}$$

The following lemma is proved in the Appendix, which can be regarded as an extension of Hölder's inequality to H^α_p.

Lemma 4.4. *For any $\alpha, \gamma_1, \gamma_2 \in [0, 1)$ and $p, p_1, p_2 \in (1, \infty]$ with*

$$\frac{1}{p_i} \leq \frac{1}{p} + \frac{\gamma_i}{d}, \qquad \frac{\gamma_i}{d} \leq \frac{1}{p_1} + \frac{1}{p_2} - \frac{1}{p} < \frac{\gamma_1 + \gamma_2 + \alpha}{d}, \quad i = 1, 2,$$

there is a constant $c = c(p_i, \gamma_i, p, \alpha, d) > 0$ such that

$$\|fg\|_{\alpha,p} \leq c\|f\|_{\alpha+\gamma_1,p_1}\|g\|_{\alpha+\gamma_2,p_2}.$$

For $\lambda \geq 0$, we consider the following nonlocal parabolic equation:

$$\partial_t u + \mathcal{L}^\kappa_\alpha u - \lambda u + \mathcal{L}^b_1 u + \mathcal{K}_t u = f, \quad u(T, x) = 0, \tag{4.19}$$

where $\mathcal{K}_t$ is a family of abstract operators. By Duhamel's formula, we shall consider the following mild form:

$$u(s, x) = \int_s^T \mathrm{e}^{-\lambda(t-s)} P^\kappa_{s,t}\big(\mathcal{L}^b_1 u + \mathcal{K}_t u + f\big)(t, x)\, \mathrm{d}t, \tag{4.20}$$

where

$$P^\kappa_{s,t} f(t, x) := \int_{\mathbb{R}^d} \rho_\kappa(s, x; t, y) f(t, y)\, \mathrm{d}y.$$

We show the following main result of this subsection.

Theorem 4.5. *Let $p, q \in (1, \infty)$ and $T > 0$. Suppose that* ($\mathbf{H}^\kappa_\beta$) *holds for some $\beta \in (0, 1)$, and for any $\theta \in [0, \beta)$, there is a constant $c_0 > 0$ such that for all $t \in [0, T]$,*

$$\|\mathcal{K}_t u\|_{\theta,p} \leq c_0\|u\|_{1,p}. \tag{4.21}$$

(i) *Suppose that for some* $p_1 \in [p,\infty]$ *and* $q_1 \in [q,\infty]$ *with* $\frac{d}{p_1}+\frac{\alpha}{q_1}<\alpha-1$,

$$b=b_1+b_2, \quad b_1 \in \mathbb{L}^{q_1}_{p_1}(T), b_2 \in \mathbb{L}^\infty(T).$$

Then for any $f \in \mathbb{L}^q_p(T)$, *there are* $\lambda_0, c_1 \geq 1$ *and unique* u *satisfying* (4.20) *such that for all* $p' \in [p,\infty], q' \in [q,\infty]$ *and* $\vartheta \in [0,\alpha)$ *with*

$$\frac{d}{p}+\frac{\alpha}{q}<\alpha-\vartheta+\frac{d}{p'}+\frac{\alpha}{q'}, \tag{4.22}$$

and all $\lambda \geq \lambda_0$ *and* $S \in [0,T]$,

$$\lambda^{\frac{1}{\alpha}(\alpha-\vartheta+\frac{d}{p'}+\frac{\alpha}{q'}-\frac{d}{p}-\frac{\alpha}{q})}\|u\|_{\mathbb{H}^{\vartheta,q'}_{p'}(S,T)} \leq c_1\|f\|_{\mathbb{L}^q_p(S,T)}. \tag{4.23}$$

(ii) *Let* $\vartheta \in [\alpha,(\alpha+\beta)\wedge 2)$ *and* $\theta \in (\frac{\alpha}{q}+\vartheta-\alpha,(\vartheta-1)\wedge\beta]$. *Suppose that* $b=b_1+b_2$ *with* $b_1 \in \mathbb{H}^{\theta,q}_{p_1}(T)$ *for some* $p_1 \in [p,\infty]\cap(\frac{d}{\vartheta-1},\infty]$ *and* $b_2 \in \mathbb{H}^{\theta,\infty}_\infty(T)$. *Then for any* $f \in \mathbb{H}^{\theta,q}_p(T)$, *there are* $\lambda_0, c_2 \geq 1$ *and unique* u *satisfying* (4.20) *such that for all* $\lambda \geq \lambda_0$,

$$\lambda^{1-\frac{1}{q}-\frac{\vartheta-\theta}{\alpha}}\|u\|_{\mathbb{H}^{\vartheta,\infty}_p(T)} \leq c_2\|f\|_{\mathbb{H}^{\theta,q}_p(T)}. \tag{4.24}$$

Proof. By the fixed point theorem, it suffices to show the a priori estimates (4.23) and (4.24). We divide the proof into four steps.

(*Step 1*) We first show (4.23) for $b=0$ and $\mathcal{K}=0$. For $\vartheta \in (0,\alpha)$, by (4.16) and Young's convolution inequality with $1+\frac{1}{p'}=\frac{1}{r}+\frac{1}{p}$, we have

$$\begin{aligned}\left\|\Delta_x^{\vartheta/2}P^\kappa_{s,t}f\right\|_{p'} &= \left\|\int_{\mathbb{R}^d}\Delta_x^{\vartheta/2}\rho_\kappa(s,\cdot;t,y)f(y)\,\mathrm{d}y\right\|_{p'} \lesssim \left\|\varrho^{(\vartheta)}_{-\vartheta}(t-s,\cdot)*f\right\|_{p'}\\ &\leq \left\|\varrho^{(\vartheta)}_{-\vartheta}(t-s,\cdot)\right\|_r\|f\|_p \overset{(4.18)}{\lesssim} (t-s)^{d/(\alpha r)-(\vartheta+d)/\alpha}\|f\|_p,\end{aligned}$$

and for $j=0,1$, by (4.15),

$$\begin{aligned}\left\|\nabla_x^j P^\kappa_{s,t}f\right\|_{p'} &= \left\|\int_{\mathbb{R}^d}\nabla_x\rho_\kappa(s,\cdot;t,y)f(y)\,\mathrm{d}y\right\|_{p'} \lesssim \left\|\varrho^{(\alpha)}_{-j}(t-s,\cdot)*f\right\|_{p'}\\ &\leq \left\|\varrho^{(\alpha)}_{-j}(t-s,\cdot)\right\|_r\|f\|_p \overset{(4.18)}{\lesssim} (t-s)^{d/(\alpha r)-(j+d)/\alpha}\|f\|_p.\end{aligned}$$

Let $h_\lambda(t):=\mathrm{e}^{-\lambda t}t^{d/(\alpha r)-(\vartheta+d)/\alpha}1_{\{t>0\}}$. By Young's convolution inequality again,

$$\begin{aligned}\left\|\Delta^{\vartheta/2}u\right\|_{\mathbb{L}^{q'}_{p'}(S,T)} &\lesssim \left\|h_\lambda*\left\|f(\cdot)\right\|_p\right\|_{L^{q'}(S,T)} \leq \|h_\lambda\|_{L^{1/(1+1/q'-1/q)}(0,T-S)}\|f\|_{\mathbb{L}^q_p(S,T)}\\ &\lesssim \lambda^{\frac{1}{\alpha}(\vartheta-\alpha-\frac{d}{p'}-\frac{\alpha}{q'}+\frac{d}{p}+\frac{\alpha}{q})}\|f\|_{\mathbb{L}^q_p(S,T)}.\end{aligned}$$

Thus we get (4.23).

(*Step 2*) We show (4.24) for $b=0$ and $\mathcal{K}=0$. For $\vartheta \in [\alpha,(\alpha+\beta)\wedge 2)$, since

$$\int_{\mathbb{R}^d}\Delta_x^{\vartheta/2}\rho_\kappa(s,x;t,y)\,\mathrm{d}y=\Delta_x^{\vartheta/2}1=0,$$

by the definition and (4.16), we have

$$\left|\Delta_x^{\vartheta/2}P^\kappa_{s,t}f(x)\right|=\left|\int_{\mathbb{R}^d}\Delta_x^{\vartheta/2}\rho_\kappa(s,x;t,y)\big(f(y)-f(x)\big)\,\mathrm{d}y\right|$$

$$\lesssim \int_{\mathbb{R}^d} \varrho^{(\alpha)}_{-\vartheta}(t-s, x-y)\big|f(y)-f(x)\big|\,\mathrm{d}y$$
$$= \int_{\mathbb{R}^d} \varrho^{(\alpha)}_{-\vartheta}(t-s, y)\big|f(x-y)-f(x)\big|\,\mathrm{d}y.$$

By Minkovskii's inequality, (4.3) and (4.18), we get

$$\big\|\Delta^{\vartheta/2}P^\kappa_{s,t}f\big\|_p \lesssim \|f\|_{\theta,p}\int_{\mathbb{R}^d} \varrho^{(\alpha)}_{-\vartheta}(t-s,y)|y|^\theta\,\mathrm{d}y$$
$$\lesssim \|f\|_{\theta,p}\int_{\mathbb{R}^d} \varrho^{(\alpha-\theta)}_{\theta-\vartheta}(t-s,y)\,\mathrm{d}y \lesssim \|f\|_{\theta,p}(t-s)^{(\theta-\vartheta)/\alpha}.$$

Hence,

$$\big\|\Delta^{\vartheta/2}u(s)\big\|_p \lesssim \int_s^T \mathrm{e}^{-\lambda(t-s)}(t-s)^{(\theta-\vartheta)/\alpha}\big\|f(t)\big\|_{\theta,p}\,\mathrm{d}t,$$

which in turn gives (4.24) by Hölder's inequality.

(*Step 3*) We prove the a priori estimate (4.23). First of all, for any $\vartheta\in[0,\alpha)$ and $p'\in[p,\infty]$, $q'\in[q,\infty]$ satisfying (4.22), by Step 1, we have

$$\lambda^{\frac{1}{\alpha}(\alpha-\vartheta+\frac{d}{p'}+\frac{\alpha}{q'}-\frac{d}{p}-\frac{\alpha}{q})}\|u\|_{\mathbb{H}^{\vartheta,q'}_{p'}(S,T)} \le c_1\big\|f+\mathscr{L}^b_1u+\mathscr{K}u\big\|_{\mathbb{L}^q_p(S,T)}. \tag{4.25}$$

Letting $q_2:=qq_1/(q_1-q)$ and $p_2:=pp_1/(p_1-p)$, by Hölder's inequality, we have

$$\big\|\mathscr{L}^b_1u\big\|_{\mathbb{L}^q_p(S,T)} \le \|b_1\|_{\mathbb{L}^{q_1}_{p_1}(S,T)}\|u\|_{\mathbb{H}^{1,q_2}_{p_2}(S,T)}+\|b_2\|_{\mathbb{L}^\infty(S,T)}\|u\|_{\mathbb{H}^{1,q}_p(S,T)}. \tag{4.26}$$

In particular, in (4.25), taking $\vartheta=1$, $(q',p')=(q_2,p_2)$ and $(q',p')=(q,p)$ respectively, and by (4.21), we obtain

$$\lambda^{\frac{1}{\alpha}(\alpha-1-\frac{d}{p_1}-\frac{\alpha}{q_1})}\|u\|_{\mathbb{H}^{1,q_2}_{p_2}(S,T)}+\lambda^{1-\frac{1}{\alpha}}\|u\|_{\mathbb{H}^{1,q}_p(S,T)}$$
$$\le c\|f\|_{\mathbb{L}^q_p(S,T)}+c\|b_1\|_{\mathbb{L}^{q_1}_{p_1}(S,T)}\|u\|_{\mathbb{H}^{1,q_2}_{p_2}(S,T)}+c\big(\|b_2\|_{\mathbb{L}^\infty(S,T)}+1\big)\|u\|_{\mathbb{H}^{1,q}_p(S,T)}.$$

Choosing λ_0 large enough so that

$$\lambda_0^{(\alpha-1-\frac{d}{p_1}-\frac{\alpha}{q_2})/\alpha}\ge 2c\|b\|_{\mathbb{L}^{q_1}_{p_1}(T)},\qquad \lambda_0^{1-\frac{1}{\alpha}}\ge 2c\big(\|b_2\|_{\mathbb{L}^\infty(T)}+1\big),$$

we obtain that for all $\lambda\ge\lambda_0$,

$$\lambda^{\frac{1}{\alpha}(\alpha-1-\frac{d}{p_1}-\frac{\alpha}{q_1})}\|u\|_{\mathbb{H}^{1,q_2}_{p_2}(S,T)}+\lambda^{1-\frac{1}{\alpha}}\|u\|_{\mathbb{H}^{1,q}_p(S,T)}\le 2c\|f\|_{\mathbb{L}^q_p(S,T)}.$$

(*Step 4*) We prove the estimate (4.24). Since $\theta\le\vartheta-1$, by Lemma 4.4, we have

$$\big\|\mathscr{L}^b_1u\big\|_{\mathbb{H}^{\theta,q}_p(T)}\le c\big(\|b_1\|_{\mathbb{H}^{\theta,q}_{p_1}(T)}+\|b_2\|_{\mathbb{H}^{\theta,q}_\infty(T)}\big)\|u\|_{\mathbb{H}^{\vartheta,\infty}_p(T)}.$$

Thus, by Step 2 and (4.21), we can get

$$\lambda^{1-\frac{1}{q}-\frac{\vartheta-\theta}{\alpha}}\|u\|_{\mathbb{H}^{\vartheta,\infty}_p(T)}\le c\big\|f+\mathscr{L}^b_1u+\mathscr{K}u\big\|_{\mathbb{H}^{\theta,q}_p(T)}$$
$$\le c\big(\|b_1\|_{\mathbb{H}^{\theta,q}_{p_1}(T)}+\|b_2\|_{\mathbb{H}^{\theta,q}_\infty(T)}\big)\|u\|_{\mathbb{H}^{\vartheta,\infty}_p(T)}+c\|f\|_{\mathbb{H}^{\theta,q}_p(T)}+c\|u\|_{\mathbb{H}^{1,\infty}_p(T)}.$$

Choosing λ_0 large enough, we obtain the desired estimate (4.24). □

Remark 4.6. It should be noticed that in the above case (ii), if $\vartheta>\alpha$, then u solves (4.19) because u is in the domain of $\mathscr{L}^\kappa_\alpha$, $\mathscr{L}^b_1$ and $\mathscr{K}$.

5. Krylov's estimates for semimartingales

This section is devoted to the study of Krylov's estimates for discontinuous semimartingales, which can be regarded as a priori estimates for the solution of SDE (1.2).

5.1. *General discontinuous semimartingales*

The classical Krylov's estimate on the distribution of continuous martingales is well known, see [29] or [22, Lemma 3.1]. Below, we generalize it to discontinuous semimartingales.

The following important result on the existence of a solution for a partial differential inequality comes from Krylov [30, Chapter III, p. 55, Theorem 4].

Lemma 5.1. *Given a nonnegative smooth function f on $\mathbb{R}_+\times\mathbb{R}^d$ with compact support and $\lambda>0$, there exists a nonnegative smooth function $u(t,x)$ such that for all nonnegative definite symmetric matrices $a=(a^{ij})_{d\times d}$ and $\beta\geq 0$,*

$$\beta\partial_t u+a^{ij}\partial_i\partial_j u-\lambda(\beta+\operatorname{tr}a)u+(\beta\det a)^{1/(d+1)}f\leq 0, \tag{5.1}$$

and

$$|\nabla u|\leq\sqrt{\lambda}u,\qquad u\leq K_d\lambda^{-d/(2(d+1))}\|f\|_{\mathbb{L}^{d+1}(T)}, \tag{5.2}$$

where $K_d>0$ depends only on the dimension d.

Using this lemma, we show the following Krylov estimate for general discontinuous semimartingales.

Theorem 5.2. *Let $m=m(t)$ be an $\mathbb{R}^d$-valued continuous local martingale, $V=V(t)$ an $\mathbb{R}^d$-valued continuous adapted process with finite variation on finite time intervals, $N(\mathrm{d}t,\mathrm{d}z)$ a Poisson random measure with compensator $\mathrm{d}t\nu(\mathrm{d}z)$, where ν is a Lévy measure, and $G:\mathbb{R}_+\times\Omega\times\mathbb{R}^d\to\mathbb{R}^d$ a predictable process with*

$$\int_0^t\int_{|z|\leq R}\big|G(s,z)\big|^2\,\mathrm{d}s\nu(\mathrm{d}z)<\infty,\quad a.s.,$$

where $R>0$. Suppose that

$$m(0)=V(0)=0,\qquad \mathrm{d}\langle m^i,m^j\rangle_t\ll\mathrm{d}t.$$

Let $a^{ij}(t):=\frac{\mathrm{d}\langle m^i,m^j\rangle_t}{2\,\mathrm{d}t}$ and

$$X(t):=X_0+m(t)+V(t)+\int_0^t\int_{|z|\leq R}G(s,z)\tilde N(\mathrm{d}s,\mathrm{d}z)+\int_0^t\int_{|z|>R}G(s,z)N(\mathrm{d}s,\mathrm{d}z).$$

Then for any $T>0$, $p\geq d+1$ and $\alpha\in[1,2]$, there is a constant $c=c(T,p,d,\alpha)>0$ such that for any stopping time τ and nonnegative $f\in\mathbb{L}^p(T)$,

$$\mathbb{E}\bigg(\int_0^{T\wedge\tau}\big(\det a(t)\big)^{\frac1p}f(t,X_t)\,\mathrm{d}t\bigg)\leq c\big(1+\mathbb{V}^2+\mathbb{A}+\mathbb{G}_\alpha^{\frac2\alpha}\big)^{\frac{d}{2p}}\|f\|_{\mathbb{L}^p(T)}, \tag{5.3}$$

where

$$\mathbb{V}:=\mathbb{E}\bigg(\int_0^{T\wedge\tau}\big|\mathrm{d}V(t)\big|\bigg),\qquad \mathbb{A}:=\mathbb{E}\bigg(\int_0^{T\wedge\tau}\operatorname{tr}a(t)\,\mathrm{d}t\bigg),$$

$$\mathbb{G}_\alpha:=\mathbb{E}\bigg(\int_0^{T\wedge\tau}\int_{|z|<R}\big|G(t,z)\big|^\alpha\,\mathrm{d}t\nu(\mathrm{d}z)\bigg).$$

Proof. By standard approximation, we may assume that $f \in C_0^\infty(\mathbb{R}_+ \times \mathbb{R}^d)$ and $\mathbb{V}$, $\mathbb{A}$, $\mathbb{G}_\alpha$ are finite. For a given constant $\lambda > 0$ whose precise value will be decided latter, let u be the nonnegative smooth function given by Lemma 5.1 corresponding to λ and f. By Itô's formula, we have

$$\begin{aligned} Z_t &:= u(t, X_t) - u(0, X_0) - \int_0^t \big(\partial_s u + a^{ij}\partial_{ij}u + \mathcal{L}_\nu^G u\big)(s, X_s)\,\mathrm{d}s - \int_0^t \partial_i u(s, X_s)\,\mathrm{d}V_s^i \\ &= \int_0^t \partial_i u(s, X_s)\,\mathrm{d}m_s^i + \int_0^t \int_{\mathbb{R}^d} \big(u\big(s, X_{s-} + G(s,z)\big) - u(s, X_{s-})\big)\tilde{N}(\mathrm{d}s, \mathrm{d}z) \end{aligned}$$

is a local martingale, where

$$\mathcal{L}_\nu^G u(t,x) := \int_{\mathbb{R}^d} \big[u\big(t, x + G(t,z)\big) - u(t,x) - 1_{|z|\le R} G^i(t,z)\partial_i u(t,x)\big]\nu(\mathrm{d}z).$$

Observing that for $|z| < R$,

$$\begin{aligned} \Sigma_t(x,z) &:= u\big(t, x + G(t,z)\big) - u(t,x) - G^i(t,z)\partial_i u(t,x) \\ &= G^i(t,z)\int_0^1 \big(\partial_i u\big(t, x + s_1 G(t,z)\big) - \partial_i u(t,x)\big)\,\mathrm{d}s_1 \\ &= G^i(t,z)G^j(t,z)\int_0^1\int_0^1 s_1 \partial_i\partial_j u\big(t, x + s_1 s_2 G(t,z)\big)\,\mathrm{d}s_1\,\mathrm{d}s_2, \end{aligned}$$

by (5.1) with $\beta = 0$, we have

$$G^i G^j \partial_i\partial_j u \le \lambda |G|^2 u \quad \Rightarrow \quad \Sigma_t(x,z) \le \lambda \big|G(t,z)\big|^2 \|u\|_{\mathbb{L}^\infty(T)},$$

and by (5.2),

$$G^i \partial_i u \le \sqrt{\lambda}|G|u \quad \Rightarrow \quad \big|\Sigma_t(x,z)\big| \le \sqrt{\lambda}\big|G(t,z)\big|\|u\|_{\mathbb{L}^\infty(T)}.$$

Hence, for any $\alpha \in [1,2]$,

$$\mathcal{L}_\nu^G u(t,x) \le \left(2\nu\big(B_R^c\big) + \lambda^{\frac{\alpha}{2}} \int_{|z|<R} \big|G(t,z)\big|^\alpha \nu(\mathrm{d}z)\right)\|u\|_{\mathbb{L}^\infty(T)}. \tag{5.4}$$

For $n \in \mathbb{N}$, if we define the stopping time

$$\tau_n := \tau \wedge \inf\left\{t \ge 0 : |m_t| + \int_0^t \int_{|z|\le R} \big|G(s,z)\big|^2 \,\mathrm{d}s\,\nu(\mathrm{d}z) \ge n\right\},$$

then $t \mapsto Z_{t\wedge\tau_n}$ is a martingale. Thus, by the definition of Z_t, (5.1) with $\beta = 1$, (5.2) and (5.4), we have

$$\begin{aligned} \mathbb{E}u(t\wedge\tau_n, X_{t\wedge\tau_n}) - \mathbb{E}u(0, X_0) &\le -\mathbb{E}\left(\int_0^{t\wedge\tau_n} \big(\det a(s)\big)^{\frac{1}{d+1}} f(s, X_s)\,\mathrm{d}s\right) \\ &\quad + \mathbb{E}\left(\sqrt{\lambda}\int_0^{t\wedge\tau_n} \mathrm{d}|V_s| + \lambda\int_0^{t\wedge\tau_n} \big(\operatorname{tr} a(s) + 1\big)\,\mathrm{d}s \right. \\ &\quad \left. + 2\nu\big(B_R^c\big) + \lambda^{\frac{\alpha}{2}} \int_0^{t\wedge\tau_n} \int_{|z|<R} \big|G(t,z)\big|^\alpha \nu(\mathrm{d}z)\,\mathrm{d}s\right)\|u\|_{\mathbb{L}^\infty(T)} \\ &\le -\mathbb{E}\left(\int_0^{t\wedge\tau_n} \big(\det a(s)\big)^{\frac{1}{d+1}} f(s, X_s)\,\mathrm{d}s\right) \\ &\quad + \big(\sqrt{\lambda}\mathbb{V} + \lambda(\mathbb{A} + t) + 2\nu\big(B_R^c\big) + \lambda^{\frac{\alpha}{2}}\mathbb{G}_\alpha\big)\|u\|_{\mathbb{L}^\infty(T)}. \end{aligned}$$

Taking into account (5.2), we get

$$\mathbb{E}\left(\int_0^{t\wedge\tau_n}\big(\det a(s)\big)^{\frac{1}{d+1}}f(s,X_s)\,\mathrm{d}s\right)$$
$$\lesssim\big(\sqrt{\lambda}\mathbb{V}+\lambda(\mathbb{A}+1)+\lambda^{\frac{\alpha}{2}}\mathbb{G}_\alpha+1\big)\lambda^{-d/(2(d+1))}\|f\|_{\mathbb{L}^{d+1}(T)},$$

which, by taking $\lambda^{-1}=\mathbb{V}^2\vee\mathbb{A}\vee\mathbb{G}_\alpha^{\frac{2}{\alpha}}\vee 1$ and letting $n\to\infty$, implies (5.3) for $p=d+1$. Finally, for $p>d+1$, by Hölder's inequality, we have

$$\mathbb{E}\left(\int_0^{T\wedge\tau}\big(\det a(t)\big)^{\frac{1}{p}}f(t,X_t)\,\mathrm{d}t\right)\lesssim\left(\mathbb{E}\int_0^{T\wedge\tau}\big(\det a(t)\big)^{\frac{1}{d+1}}\big|f(t,X_t)\big|^{\frac{p}{d+1}}\,\mathrm{d}t\right)^{\frac{d+1}{p}}$$
$$\lesssim\big(1+\mathbb{V}^2+\mathbb{A}+\mathbb{G}_\alpha^{\frac{2}{\alpha}}\big)^{\frac{d}{2p}}\|f\|_{\mathbb{L}^p(T)}.$$

The proof is finished. □

Remark 5.3. A similar result can be found in [45, Theorem 165]. However, the right hand side of (5.3) in our result is more precise, which is important for us below.

5.2. *Non-degenerate diffusion SDEs with jumps*

Below, for the moment we suppose that X_t satisfies the following equation:

$$X_t=X_0+\int_0^t\sigma_s(X_s)\,\mathrm{d}W_s+\int_0^t\int_{|z|\le R}g_s(X_{s-},z)\tilde{N}(\mathrm{d}s,\mathrm{d}z)$$
$$+\int_0^t\int_{|z|>R}g_s(X_{s-},z)N(\mathrm{d}s,\mathrm{d}z)+\int_0^t\xi(s)\,\mathrm{d}s,\tag{5.5}$$

where $\xi(t)$ is a measurable $\mathcal{F}_t$-adapted process. The reason of considering this form X_t is that we have more flexibility of choosing the drift $\xi(s)$.

The following lemma is an easy consequence of Theorem 5.2.

Lemma 5.4. *Let X_t be of the form (5.5). Suppose that $\sigma\sigma^*$ is bounded and uniformly positive definite, and for some $\alpha\in[1,2]$ and $q\ge d+1$, $\Gamma^{0,\alpha}_{0,R}(g)\in\mathbb{L}^q(T)$, where $\Gamma^{0,\alpha}_{0,R}(g)$ is defined by (2.1). Then for any $p\ge d+1$ and $\delta>0$, there is a constant $c_\delta>0$ such that for any stopping time τ and $f\in\mathbb{L}^p(T)$,*

$$\mathbb{E}\left(\int_0^{T\wedge\tau}f(s,X_s)\,\mathrm{d}s\right)\le\left(c_\delta+\delta\mathbb{E}\left(\int_0^{T\wedge\tau}\big|\xi(s)\big|\,\mathrm{d}s\right)\right)\|f\|_{\mathbb{L}^p(T)}.\tag{5.6}$$

Proof. Without loss of generality, we assume $\mathbb{E}(\int_0^{T\wedge\tau}|\xi(s)|\,\mathrm{d}s)<\infty$. In order to use Theorem 5.2, we take

$$m(t):=\int_0^t\sigma_s(X_s)\,\mathrm{d}W_s,\qquad V(t):=\int_0^t\big|\xi(s)\big|\,\mathrm{d}s,\qquad G(t,z):=g_t(X_{t-},z).$$

Thus, by the assumption on σ, for any $p\ge d+1$, by (5.3), there is a constant $c>0$ such that for all $f\in\mathbb{L}^p(T)$,

$$\mathbb{E}\left(\int_0^{T\wedge\tau}f(t,X_t)\,\mathrm{d}t\right)\le c\big(1+\mathbb{V}^2+\mathbb{G}_\alpha^{\frac{2}{\alpha}}\big)^{\frac{d}{2p}}\|f\|_{\mathbb{L}^p(T)}.\tag{5.7}$$

Here, $\mathbb{V}:=\mathbb{E}(\int_0^{T\wedge\tau}|\xi(s)|\,\mathrm{d}s)$ and

$$\mathbb{G}_\alpha:=\mathbb{E}\left(\int_0^{T\wedge\tau}\int_{|z|<R}\big|g_t(X_t,z)\big|^\alpha\,\mathrm{d}t\,\nu(\mathrm{d}z)\right)=\mathbb{E}\left(\int_0^{T\wedge\tau}\Gamma^{0,\alpha}_{0,R}(g_t)(X_t)\,\mathrm{d}t\right).$$

By (5.7) with $f=\Gamma^{0,\alpha}_{0,R}(g)$ and the assumption, we have

$$\mathbb{G}_\alpha \le c\big(1+\mathbb{V}^2+\mathbb{G}_\alpha^{\frac{2}{\alpha}}\big)^{\frac{d}{2q}}\big\|\Gamma^{0,\alpha}_{0,R}(g)\big\|_{\mathbb{L}^q(T)} \le c\big(1+\mathbb{V}^{d/q}\big)+\frac{1}{2}\mathbb{G}_\alpha,$$

which implies $\mathbb{G}_\alpha \le c(1+\mathbb{V}^{d/q})$. Thus, we get (5.6) by (5.7) and Young's inequality. □

In the above estimate, it is required $p\ge d+1$, which is too strong for our purpose. Below we use Theorem 4.3 to obtain better integrability index p. The price we have to pay is to strengthen the assumption on $\Gamma^{0,\alpha}_{0,R}(g)$.

Lemma 5.5. *Let X be of the form (5.5) and $\Gamma^{0,2}_{0,R}(g)$ be defined by (2.1). Let $T>0$. Suppose that $(\mathbf{H}^\sigma_\beta)$ holds and*

$$\Gamma^{0,2}_{0,R}(g)\in\mathbb{L}^\infty(T) \quad \text{and} \quad \lim_{\varepsilon\to 0}\big\|\Gamma^{0,2}_{0,\varepsilon}(g)\big\|_{\mathbb{L}^\infty(T)}=0.$$

Then for any $p,q\in(1,\infty)$ with $\frac{d}{p}+\frac{2}{q}<1$ and each $\delta>0$, there is a constant $c_\delta>0$ such that for any stopping time τ and $0\le t_0\le t_1\le T$ and $f\in\mathbb{L}^q_p(t_0,t_1)$,

$$\mathbb{E}\bigg(\int_{t_0\wedge\tau}^{t_1\wedge\tau} f(s,X_s)\,\mathrm{d}s\bigg|\mathcal{F}_{t_0\wedge\tau}\bigg)\le \|f\|_{\mathbb{L}^q_p(t_0,t_1)}\bigg[c_\delta+\delta\mathbb{E}\bigg(\int_{t_0\wedge\tau}^{t_1\wedge\tau}\big|\xi(s)\big|\,\mathrm{d}s\bigg|\mathcal{F}_{t_0\wedge\tau}\bigg)\bigg]. \tag{5.8}$$

Moreover, if $\xi\equiv 0$, then we can relax p,q to satisfy $\frac{d}{p}+\frac{2}{q}<2$.

Proof. We may assume without loss of generality that $f\in C_0^\infty(\mathbb{R}^{d+1})$ and

$$\mathbb{E}\bigg(\int_0^{T\wedge\tau}\big|\xi(s)\big|\,\mathrm{d}s\bigg)<+\infty.$$

Let r be large enough so that

$$\frac{d}{r}+\frac{2}{r}\le\frac{d}{p}+\frac{2}{q}<1.$$

Let λ_0 be the constant in Theorem 4.3. For $\lambda\ge\lambda_0$ and $t_1\in(0,T]$, since $f\in\mathbb{L}^q_p(t_1)\cap\mathbb{L}^r(t_1)$, by Theorem 4.3, there exists a unique solution $u\in\mathbb{H}^{2,q}_p(t_1)\cap\mathbb{H}^{2,r}_r(t_1)$ with $\partial_t u\in\mathbb{L}^r(t_1)$ to the following backward equation:

$$\partial_t u+\big(\mathscr{L}^a_2-\lambda\big)u+\mathscr{L}^g_{\nu,R}u=f,\quad u(t_1,x)=0,$$

where $a=\sigma\sigma^*/2$. Let ϕ be a non-negative smooth function on $\mathbb{R}^{d+1}$ with support in $\{(t,x)\in\mathbb{R}^{d+1}:|(t,x)|\le 1\}$ and $\int_{\mathbb{R}^{d+1}}\phi(t,x)\,\mathrm{d}t\,\mathrm{d}x=1$. Set

$$\phi_n(t,x):=n^{d+1}\phi(nt,nx),$$

and extend $u(t,x)$ to $\mathbb{R}$ by setting $u(t,x)=0$ for $t\ge t_1$ and $u(t,x)=u(0,x)$ for $t\le 0$. Define

$$u_n(t,x):=u*\phi_n(t,x):=\int_{\mathbb{R}^{d+1}}u(s,y)\phi_n(t-s,x-y)\,\mathrm{d}s\,\mathrm{d}y \tag{5.9}$$

and

$$f_n:=\partial_t u_n+\big(\mathscr{L}^a_2-\lambda\big)u_n+\mathscr{L}^g_{\nu,R}u_n, \tag{5.10}$$

where $\mathscr{L}^g_{\nu,R}$ is defined by (1.5). Since $\frac{d}{p}+\frac{2}{q}<1$, by the property of convolution and using (4.9) with $\gamma=1$ and $p'=q'=\infty$, there is a constant $c>0$ independent of n such that for all $\lambda\ge 1$ and $t_0\in[0,t_1]$,

$$\|u_n\|_{\mathbb{H}^{1,\infty}_\infty(t_0,t_1)}\le\|u\|_{\mathbb{H}^{1,\infty}_\infty(t_0,t_1)}\le c\lambda^{\frac{1}{2}(\frac{d}{p}+\frac{2}{q}-1)}\|f\|_{\mathbb{L}^q_p(t_0,t_1)}, \tag{5.11}$$

and

$$
\begin{aligned}
\|f_n - f\|_{\mathbb{L}^r_r(t_1)} &\le \lambda\|u_n - u\|_{\mathbb{L}^r_r(t_1)} + \big\|\partial_t(u_n - u)\big\|_{\mathbb{L}^r_r(t_1)} \\
&\quad + c\big\|\nabla^2(u_n - u)\big\|_{\mathbb{L}^r_r(t_1)} + \big\|\mathscr{L}^g_{\nu,R}(u_n - u)\big\|_{\mathbb{L}^r_r(t_1)} \\
&\le \big\|\partial_t(u_n - u)\big\|_{\mathbb{L}^r_r(t_1)} + c\|u_n - u\|_{\mathbb{H}^{2,r}_r(t_1)} \to 0, \quad n \to \infty,
\end{aligned}
$$

where we have used the same estimate as in (4.11). Therefore, by the Krylov estimate (5.6), we have

$$
\lim_{n\to\infty} \mathbb{E}\bigg(\int_0^{T\wedge\tau} \big|f_n(s, X_s) - f(s, X_s)\big|\, \mathrm{d}s\bigg) \le c \lim_{n\to\infty} \|f_n - f\|_{\mathbb{L}^r_r(T)} = 0. \tag{5.12}
$$

Now, applying Itô's formula to $u_n(t, x)$, we have

$$
\begin{aligned}
u_n(t, X_t) = u_n(0, X_0) &+ \int_0^t \big(\partial_s u_n + \mathscr{L}^a_2 u_n + \mathscr{L}^g_\nu u_n\big)(s, X_s)\, \mathrm{d}s \\
&+ \int_0^t \xi(s) \cdot \nabla u_n(s, X_s)\, \mathrm{d}s + \text{a martingale}.
\end{aligned}
$$

Thus, by Doob's optional stopping theorem and (5.10), we obtain

$$
\begin{aligned}
&\mathbb{E}\big(u_n(t_1 \wedge \tau, X_{t_1\wedge\tau})|\mathcal{F}_{t_0\wedge\tau}\big) - u_n(t_0 \wedge \tau, X_{t_0\wedge\tau}) \\
&\quad = \mathbb{E}\bigg(\int_{t_0\wedge\tau}^{t_1\wedge\tau} \big(\partial_s u_n + \mathscr{L}^a_2 u_n + \mathscr{L}^g_\nu u_n\big)(s, X_s)\, \mathrm{d}s\bigg|\mathcal{F}_{t_0\wedge\tau}\bigg) \\
&\qquad + \mathbb{E}\bigg(\int_{t_0\wedge\tau}^{t_1\wedge\tau} \xi(s) \cdot \nabla u_n(s, X_s)\, \mathrm{d}s\bigg|\mathcal{F}_{t_0\wedge\tau}\bigg) \\
&\quad \ge \mathbb{E}\bigg(\int_{t_0\wedge\tau}^{t_1\wedge\tau} \big(\lambda u_n(s, X_s) + f_n(s, X_s)\big)\, \mathrm{d}s\bigg|\mathcal{F}_{t_0\wedge\tau}\bigg) \\
&\qquad - 2\|u_n\|_{\mathbb{L}^\infty(t_0,t_1)}\nu\big(B^c_R\big)t_1 - \|\nabla u_n\|_{\mathbb{L}^\infty(t_0,t_1)}\mathbb{E}\bigg(\int_{t_0\wedge\tau}^{t_1\wedge\tau} \big|\xi(s)\big|\, \mathrm{d}s\bigg|\mathcal{F}_{t_0\wedge\tau}\bigg),
\end{aligned}
$$

which implies that by (5.11),

$$
\begin{aligned}
\mathbb{E}\bigg(\int_{t_0\wedge\tau}^{t_1\wedge\tau} f_n(s, X_s)\, \mathrm{d}s\bigg|\mathcal{F}_{t_0\wedge\tau}\bigg) &\le \big(2 + \lambda T + 2\nu\big(B^c_R\big)T\big)\|u_n\|_{\mathbb{L}^\infty(t_0,t_1)} \\
&\quad + \|\nabla u_n\|_{\mathbb{L}^\infty(t_0,t_1)}\mathbb{E}\bigg(\int_{t_0\wedge\tau}^{t_1\wedge\tau} \big|\xi(s)\big|\, \mathrm{d}s\bigg|\mathcal{F}_{t_0\wedge\tau}\bigg) \\
&\le \bigg[c_\lambda + c\lambda^{\frac{1}{2}(\frac{d}{p}+\frac{2}{q}-1)}\mathbb{E}\bigg(\int_{t_0\wedge\tau}^{t_1\wedge\tau} \big|\xi(s)\big|\, \mathrm{d}s\bigg|\mathcal{F}_{t_0\wedge\tau}\bigg)\bigg]\|f\|_{\mathbb{L}^q_p(t_0,t_1)}.
\end{aligned}
$$

Letting $n \to \infty$ and λ be large enough, by (5.12) we get (5.8). If $\xi \equiv 0$, then we only need to control $\|u\|_{\mathbb{L}^\infty(t_0,t_1)}$, which follows by (4.9) with $\vartheta = 0$ and $p' = q' = \infty$. □

Remark 5.6. Lemma 5.5 will be used to derive Krylov's estimate for SDE with polynomial growth drift in the proof of ergodicity for SDEs with singular drifts.

Now we show the following important Krvlov's estimate for SDE (1.2).

Theorem 5.7. *Let $T > 0$. Assume that $(\mathbf{H}^\sigma_\beta)$ holds and for some $p_1, q_1 \in (2, \infty]$ with $\frac{d}{p_1} + \frac{2}{q_1} < 1$,*

$$
b \in \mathbb{L}^{q_1}_{p_1}(T), \Gamma^{0,2}_{0,R}(g) \in \mathbb{L}^\infty(T), \quad \lim_{\varepsilon\to 0}\big\|\Gamma^{0,2}_{0,\varepsilon}(g)\big\|_{\mathbb{L}^\infty(T)} = 0.
$$

Then for any $p, q \in (1, \infty)$ with $\frac{d}{p} + \frac{2}{q} < 2$, the solution X of SDE (1.2) satisfies Krylov's estimate with index p, q.

Proof. (i) First of all, we show that X satisfies Krylov's estimate for all $p, q \in (1, \infty)$ with $\frac{d}{p} + \frac{2}{q} < 1$. By Lemma 5.5, it suffices to show that for all $0 \le t_0 \le t_1 \le T$,

$$\mathbb{E}\left(\int_{t_0}^{t_1} \left|b_s(X_s)\right| \mathrm{d}s \middle| \mathcal{F}_{t_0}\right) \le c\|b\|_{\mathbb{L}_{p_1}^{q_1}(t_0,t_1)}. \tag{5.13}$$

For $n \in \mathbb{N}$, define a stopping time

$$\tau_n := \inf\left\{t > 0 : \int_0^t \left|b_s(X_s)\right| \mathrm{d}s \ge n\right\}.$$

Taking $\xi(s) = b_s(X_s)$ and $f = |b|$ in Lemma 5.5, we get that for every $\delta > 0$ and $0 \le t_0 \le t_1 \le T$,

$$\mathbb{E}\left(\int_{t_0\wedge\tau_n}^{t_1\wedge\tau_n} \left|b_s(X_s)\right| \mathrm{d}s \middle| \mathcal{F}_{t_0\wedge\tau_n}\right) \le \left[c_\delta + \delta\mathbb{E}\left(\int_{t_0\wedge\tau_n}^{t_1\wedge\tau_n} \left|b_s(X_s)\right| \mathrm{d}s \middle| \mathcal{F}_{t_0\wedge\tau_n}\right)\right]\|b\|_{\mathbb{L}_{p_1}^{q_1}(t_0,t_1)}.$$

Choosing δ be small enough such that

$$\delta\|b\|_{\mathbb{L}_{p_1}^{q_1}(T)} \le \frac{1}{2},$$

we obtain that for all $0 \le t_0 \le t_1 \le T$,

$$\mathbb{E}\left(\int_{t_0\wedge\tau_n}^{t_1\wedge\tau_n} \left|b_s(X_s)\right| \mathrm{d}s \middle| \mathcal{F}_{t_0\wedge\tau_n}\right) \le c\|b\|_{\mathbb{L}_{p_1}^{q_1}(t_0,t_1)},$$

where c is independent of n. Letting $n \to \infty$, we get (5.13).

(ii) In this step we show that X satisfies the Krylov estimate for $p = p_1/2$ and $q = q_1/2$. Without loss of generality, we assume $p_1, q_1 \in (2, \infty)$ and $f \in C_0^\infty(\mathbb{R}^{d+1})$. Let λ_0 be the constant in Theorem 4.3. For $\lambda \ge \lambda_0$ and $t_1 \in (0, T]$, since $f \in \mathbb{L}_p^q(t_1) \cap \mathbb{L}_{p_1}^{q_1}(t_1)$, by Theorem 4.3, there exists a unique solution $u \in \mathbb{H}_p^{2,q}(t_1) \cap \mathbb{H}_{p_1}^{2,q_1}(t_1)$ with $\partial_t u \in \mathbb{L}_{p_1}^{q_1}(t_1)$ to the following backward equation:

$$\partial_t u + \left(\mathcal{L}_2^a - \lambda\right)u + \mathcal{L}_{\nu,R}^g u + \mathcal{L}_1^b u = f, \quad u(t_1, x) = 0.$$

Let $u_n := u * \phi_n$ be defined as in (5.9), and

$$f_n := \partial_t u_n + \left(\mathcal{L}_2^a - \lambda\right)u_n + \mathcal{L}_{\nu,R}^g u_n + \mathcal{L}_1^b u_n.$$

As in the proof of Lemma 5.5 we have

$$\mathbb{E}\left(u_n(t_1, X_{t_1}) \middle| \mathcal{F}_{t_0}\right) - u_n(t_0, X_{t_0}) = \mathbb{E}\left(\int_{t_0}^{t_1} (f_n + \lambda u_n)(s, X_s)\, \mathrm{d}s \middle| \mathcal{F}_{t_0}\right),$$

which implies by (4.9) with $\vartheta = 0$ and $p' = q' = \infty$ that

$$\mathbb{E}\left(\int_{t_0}^{t_1} f_n(s, X_s)\, \mathrm{d}s \middle| \mathcal{F}_{t_0}\right) \le (\lambda T + 2)\|u_n\|_{\mathbb{L}^\infty(t_1)} \le c\|f\|_{\mathbb{L}_p^q(t_1)}. \tag{5.14}$$

Noticing that

$$\lim_{n\to\infty} \|f_n - f\|_{\mathbb{L}_{p_1}^{q_1}(t_1)} = 0,$$

by step (i) and taking limits $n \to \infty$ for (5.14), we get

$$\mathbb{E}\left(\int_{t_0}^{t_1} f(s, X_s)\, \mathrm{d}s \middle| \mathcal{F}_{t_0}\right) \le c\|f\|_{\mathbb{L}_p^q(t_1)} = c\|f\|_{\mathbb{L}_{p_1/2}^{q_1/2}(t_1)}. \tag{5.15}$$

(iii) By (5.15), we have for all $0 \le t_0 \le t_1 \le T$,

$$\mathbb{E}\left(\int_{t_0}^{t_1} \left|b_s(X_s)\right|^2 \mathrm{d}s \middle| \mathcal{F}_{t_0}\right) \le c\|b\|_{\mathbb{L}_{p_1}^{q_1}(t_0,t_1)}.$$

By Lemma 3.5, for any $\lambda > 0$, there is a constant $c > 0$ such that for all $0 \le t_0 < t_1 \le T$,

$$\mathbb{E}\bigg(\exp\bigg\{\lambda \int_{t_0}^{t_1} \big|b_s(X_s)\big|^2 \,\mathrm{d}s\bigg\}\bigg|\mathscr{F}_{t_0}\bigg) \le c. \tag{5.16}$$

Define for $\gamma \in \mathbb{R}$,

$$\mathcal{E}^{(\gamma)}_{t_0,t_1} := \exp\bigg\{\gamma \int_{t_0}^{t_1} \big(\sigma_s^{-1} b_s\big)(X_s)\,\mathrm{d}W_s - \frac{\gamma^2}{2}\int_{t_0}^{t_1} \big|\sigma_s^{-1} b_s\big|^2(X_s)\,\mathrm{d}s\bigg\}.$$

By Novikov's criterion, $t \mapsto \mathcal{E}^{(\gamma)}_{0,t}$ is an exponential martingale. Hence, by (5.16) and Hölder's inequality,

$$\mathbb{E}\big(\big(\mathcal{E}^{(1)}_{t_0,t_1}\big)^{\gamma}|\mathscr{F}_{t_0}\big) \le \bigg(\mathbb{E}\bigg(\exp\bigg\{\big(2\gamma^2-\gamma\big)\int_{t_0}^{t_1} \big|\sigma_s^{-1} b_s\big|^2(X_s)\,\mathrm{d}s\bigg\}\bigg|\mathscr{F}_{t_0}\bigg)\bigg)^{1/2} \le c. \tag{5.17}$$

Define a new probability $\mathbb{Q}_{t_0,t_1} := \mathcal{E}^{(1)}_{t_0,t_1}\mathbb{P}$. By Girsanov's theorem, under the probability measure $\mathbb{Q}_{t_0,t_1}$, after time t_0, $\tilde{W}_t := W_t + \int_{t_0}^t (\sigma_s^{-1} b_s)(X_s)\,\mathrm{d}s$ is still a Brownian motion and $N(\mathrm{d}t, \mathrm{d}z)$ is still a Poisson random measure with the same compensator $\mathrm{d}t\,\nu(\mathrm{d}z)$. Moreover, X_t satisfies

$$X_t = X_{t_0} + \int_{t_0}^t \sigma_s(X_s)\,\mathrm{d}\tilde{W}_s + \int_{t_0}^t \int_{|z|<R} g_s(X_{s-}, z)\tilde{N}(\mathrm{d}s, \mathrm{d}z) + \int_{t_0}^t \int_{|z|>R} g_s(X_{s-}, z)N(\mathrm{d}s, \mathrm{d}z).$$

Hence, by Lemma 5.5 with $\xi \equiv 0$, for any $p, q \in (1, \infty)$ with $\frac{d}{p} + \frac{2}{q} < 2$,

$$\mathbb{E}^{\mathbb{Q}_{t_0,t_1}}\bigg(\int_{t_0}^{t_1} f(s, X_s)\,\mathrm{d}s\bigg|\mathscr{F}_{t_0}\bigg) \le c\|f\|_{\mathbb{L}^q_p(t_0,t_1)}. \tag{5.18}$$

Noticing that for any nonnegative random variable ζ,

$$\mathbb{E}\big(\zeta\,\mathcal{E}^{(1)}_{t_0,t_1}|\mathscr{F}_{t_0}\big) = \mathbb{E}^{\mathbb{Q}_{t_0,t_1}}(\zeta|\mathscr{F}_{t_0})\mathbb{E}\big(\mathcal{E}^{(1)}_{t_0,t_1}|\mathscr{F}_{t_0}\big),$$

by (5.18), (5.17) and suitable Hölder's inequality, we can get the desired Krylov's estimate. □

5.3. *SDEs driven by pure jump Lévy noises*

In this subsection we assume $\nu(\mathrm{d}z) = \mathrm{d}z/|z|^{d+\alpha}$ for some $\alpha \in (1, 2)$ and $(\mathbf{H}^g_\beta)$ holds. We proceed to show Krylov's estimate for pure jump cases. In order to use Theorem 4.5, we need to write $\mathscr{L}^g_\nu$ defined in (1.5) in the form of $\mathscr{L}^\kappa_\alpha$ defined in (4.13). First of all, under (2.3), the map $z \mapsto g_t(x, z)$ admits an inverse denoted by $g_t^{-1}(x, z)$. By the change of variables, it allows us to write

$$\mathscr{L}^g_\nu u(x) = \int_{\mathbb{R}^d} \big[u(x+z) - u(x) - z\cdot\nabla u(x)\big]\frac{\kappa(t,x,z)}{|z|^{d+\alpha}}\,\mathrm{d}z + \bigg(\int_{|z|\ge R} g_t(x,z)\nu(\mathrm{d}z)\bigg)\cdot\nabla u(x), \tag{5.19}$$

where

$$\kappa(t, x, z) := \frac{|z|^{d+\alpha}\det(\nabla_z g_t^{-1}(x, z))}{|g_t^{-1}(x, z)|^{d+\alpha}}. \tag{5.20}$$

Lemma 5.8. *Under $(\mathbf{H}^g_\beta)$, there is a constant $\kappa_0 \ge 1$ such that for all $t \ge 0$ and $x, y, z \in \mathbb{R}^d$,*

$$\kappa_0^{-1} \le \kappa(t, x, z) \le \kappa_0, \qquad \big|\kappa(t, x, z) - \kappa(t, y, z)\big| \le \kappa_0|x - y|^\beta\big(1 + |z|\big). \tag{5.21}$$

Proof. By (2.3) and $g_t(x, 0) = 0$, one sees that

$$g_t^{-1}(x, 0) = 0, \qquad c_1^{-1}\big|z - z'\big| \le \big|g_t^{-1}(x, z) - g_t^{-1}\big(x, z'\big)\big| \le c_1\big|z - z'\big|.$$

In particular,

$$c_1^{-1}|z| \le \left|g_t^{-1}(x,z)\right| \le c_1|z|, \qquad \left\|\nabla_z g^{-1}\right\|_\infty \le c_1. \tag{5.22}$$

Moreover, for $x, y \in \mathbb{R}^d$, letting $\tilde{z} := g_t^{-1}(x,z)$, we have

$$\begin{aligned}\left|g_t^{-1}(x,z) - g_t^{-1}(y,z)\right| &= \left|g_t^{-1}\big(y, g_t(y,\tilde{z})\big) - g_t^{-1}\big(y, g_t(x,\tilde{z})\big)\right| \\ &\le c_1\left|g_t(y,\tilde{z}) - g_t(x,\tilde{z})\right| \le c_1^2|x-y|^\beta|\tilde{z}| \le c_1^3|x-y|^\beta|z|.\end{aligned} \tag{5.23}$$

Noticing that

$$\nabla_z g_t^{-1}(x,z) = [\nabla_z g_t]^{-1}\big(x, g_t^{-1}(x,z)\big),$$

by (2.4) and (2.5), we have

$$\begin{aligned}\left|\nabla_z g_t^{-1}(x,z) - \nabla_z g_t^{-1}(y,z)\right| &\le \left\|\nabla_z g^{-1}\right\|_\infty^2\left|\nabla_z g_t\big(x, g_t^{-1}(x,z)\big) - \nabla_z g_t\big(y, g_t^{-1}(y,z)\big)\right| \\ &\lesssim |x-y|^\beta\big(1+|z|\big) + \left|g_t^{-1}(x,z) - g_t^{-1}(y,z)\right| \\ &\lesssim |x-y|^\beta\big(1+|z|\big),\end{aligned}$$

which together with (5.20), (5.22) and (5.23) yields (5.21). □

Notice that κ defined in (5.20) is not Hölder continuous uniformly in z. In order to use Theorem 4.5 in Section 4.2, we need to write the operator $\mathcal{L}_\nu^g$ as follows:

$$\mathcal{L}_\nu^g u(x) = \mathcal{L}_\alpha^{\kappa'} u(x) + \bar{\mathcal{L}}_\alpha^{\kappa''} u(x) + \bar{b}_t^g(x)\cdot\nabla u(x), \tag{5.24}$$

where

$$\kappa'(t,x,z) := \big(\kappa(t,x,z)1_{|z|<1} + \kappa_0 1_{|z|\ge 1}\big), \qquad \kappa''(t,x,z) := \big(\kappa(t,x,z) - \kappa_0\big)1_{|z|\ge 1}$$

and

$$\begin{aligned}\bar{\mathcal{L}}_\alpha^{\kappa''} u(x) &:= \int_{|z|\ge 1}\big(u(x+z) - u(x)\big)\frac{\kappa''(t,x,z)}{|z|^{d+\alpha}}\,\mathrm{d}z, \\ \bar{b}_t^g(x) &:= \int_{|z|\ge R} g_t(x,z)\nu(\mathrm{d}z) - \int_{|z|\ge 1}\frac{\kappa''(t,x,z)}{|z|^{d+\alpha}}\,\mathrm{d}z.\end{aligned} \tag{5.25}$$

By (5.21), one sees that κ' satisfies $(\mathbf{H}_\beta^\kappa)$, and due to assumption $(\mathbf{H}_\beta^g)$, we have for some constant $c_0 > 0$,

$$\left|\bar{b}_t^g(x) - \bar{b}_t^g(y)\right| \le c_0\big(|x-y|^\beta \wedge 1\big). \tag{5.26}$$

Now, using Theorem 4.5 and as in the proof of Lemma 5.5, we have

Lemma 5.9. *Suppose that $(\mathbf{H}_\beta^g)$ holds and X_t satisfies*

$$X_t = X_0 + \int_0^t\int_{|z|<R} g_s(X_{s-},z)\tilde{N}(\mathrm{d}s,\mathrm{d}z) + \int_0^t\int_{|z|\ge R}\eta_s(z)N(\mathrm{d}s,\mathrm{d}z) + \int_0^t \xi(s)\,\mathrm{d}s,$$

where $\eta : \mathbb{R}_+ \times \Omega \times \mathbb{R}^d \to \mathbb{R}^d$ is a predictable process, and ξ is a measurable adapted process. For any $p, q \in (1,\infty)$ with $\frac{d}{p} + \frac{\alpha}{q} < \alpha - 1$ and each $\delta > 0$, there is a constant $c_\delta > 0$ such that for all $0 \le t_0 < t_1 \le T$, any stopping time τ and $f \in \mathbb{L}_p^q(T)$,

$$\mathbb{E}\left(\int_{t_0\wedge\tau}^{t_1\wedge\tau} f(s,X_s)\,\mathrm{d}s\,\middle|\,\mathcal{F}_{t_0\wedge\tau}\right) \le \left[c_\delta + \delta\mathbb{E}\left(\int_{t_0\wedge\tau}^{t_1\wedge\tau}\left|\xi(s)\right|\mathrm{d}s\,\middle|\,\mathcal{F}_{t_0\wedge\tau}\right)\right]\|f\|_{\mathbb{L}_p^q(T)}. \tag{5.27}$$

Proof. Without loss of generality, we assume $f \in C_c^\infty(\mathbb{R}_+ \times \mathbb{R}^d)$. By (5.21), (5.26), Theorem 4.5 and Remark 4.6, for some $\varepsilon \in (0,1)$ small, there is a unique $u \in \mathbb{H}_\infty^{\alpha+\varepsilon,\infty}(T)$ solving the following equation

$$\partial_t u + \big(\mathscr{L}_\alpha^{\kappa'} - \lambda\big)u + \bar{b}_t^g \cdot \nabla u = f, \quad u(t_1, x) = 0.$$

By Itô's formula and Doob's optional stopping theorem, we have

$$\begin{aligned}
&\mathbb{E}\big(u(t_1 \wedge \tau, X_{t_1\wedge\tau})|\mathcal{F}_{t_0\wedge\tau}\big) - u(t_0 \wedge \tau, X_{t_0\wedge\tau}) \\
&\quad = \mathbb{E}\bigg(\int_{t_0\wedge\tau}^{t_1\wedge\tau} \big(\partial_s u + \mathscr{L}_{\nu,R}^g u + \bar{\mathscr{L}}_{\nu,R}^\eta u + \xi(s)\cdot\nabla u\big)(s, X_s)\,\mathrm{d}s \bigg| \mathcal{F}_{t_0\wedge\tau}\bigg) \\
&\quad \overset{(5.24)}{=} \mathbb{E}\bigg(\int_{t_0\wedge\tau}^{t_1\wedge\tau} \big(\lambda u + f + \mathscr{L}_\alpha^{\kappa''} u + \bar{\mathscr{L}}_{\nu,R}^g u + \bar{\mathscr{L}}_{\nu,R}^\eta u\big)(s, X_s)\,\mathrm{d}s \bigg| \mathcal{F}_{t_0\wedge\tau}\bigg) \\
&\qquad + \mathbb{E}\bigg(\int_{t_0\wedge\tau}^{t_1\wedge\tau} \xi(s)\cdot\nabla u(s, X_s)\,\mathrm{d}s \bigg| \mathcal{F}_{t_0\wedge\tau}\bigg),
\end{aligned}$$

where $\bar{\mathscr{L}}_{\nu,R}^g$ and $\bar{\mathscr{L}}_{\nu,R}^\eta$ are defined as in (1.5). Notice that by definitions (5.25) and (1.5),

$$\big\|\mathscr{L}_\alpha^{\kappa''} u + \bar{\mathscr{L}}_{\nu,R}^g u + \bar{\mathscr{L}}_{\nu,R}^\eta u\big\|_\infty \le c_0 \|u\|_\infty.$$

Hence, by (4.23) with $\vartheta = 0, 1$ and $q' = p' = \infty$, we have for $\lambda \ge 1$,

$$\begin{aligned}
\mathbb{E}\bigg(\int_{t_0\wedge\tau}^{t_1\wedge\tau} f(s, X_s)\,\mathrm{d}s \bigg| \mathcal{F}_{t_0\wedge\tau}\bigg) &\le \big(2 + (t_1 - t_0)(\lambda + c_0)\big)\|u\|_{\mathbb{L}^\infty(T)} \\
&\quad + \|\nabla u\|_{\mathbb{L}^\infty(T)} \mathbb{E}\bigg(\int_{t_0\wedge\tau}^{t_1\wedge\tau} \big|\xi(s)\big|\,\mathrm{d}s \bigg| \mathcal{F}_{t_0\wedge\tau}\bigg) \\
&\le \|f\|_{\mathbb{L}_p^q(T)} \bigg(c_\lambda + c_1 \lambda^{\frac{1}{q} + \frac{d}{\alpha p} + \frac{1}{\alpha} - 1} \mathbb{E}\bigg(\int_{t_0\wedge\tau}^{t_1\wedge\tau} \big|\xi(s)\big|\,\mathrm{d}s \bigg| \mathcal{F}_{t_0\wedge\tau}\bigg)\bigg),
\end{aligned}$$

where c_1 is independent of $\lambda \ge 1$. The desired estimate now follows by letting λ large enough since $\frac{d}{p} + \frac{\alpha}{q} < \alpha - 1$. □

The above Krylov estimate requires $\frac{d}{p} + \frac{\alpha}{q} < \alpha - 1$, which is too strong for later use. As we shall see below, for proving the well-posedness of SDE (1.2), we need to relax it to $\frac{d}{p} + \frac{\alpha}{q} < \alpha$. The following result is similar to Theorem 5.7.

Theorem 5.10. *Let $T > 0$ and $p_1, q_1 \in (1, \infty)$ with $\frac{d}{p_1} + \frac{\alpha}{q_1} < \alpha - 1$ and $b \in \mathbb{L}_{p_1}^{q_1}(T)$. Suppose that* $(\mathbf{H}_\beta^g)$ *holds for some $\beta \in (0,1)$ and X_t satisfies*

$$X_t = X_0 + \int_0^t \int_{|z|<R} g_s(X_{s-}, z)\tilde{N}(\mathrm{d}s, \mathrm{d}z) + \int_0^t \int_{|z|\ge R} \eta_s(z) N(\mathrm{d}s, \mathrm{d}z) + \int_0^t b_s(X_s)\,\mathrm{d}s,$$

where $\eta : \mathbb{R}_+ \times \Omega \times \mathbb{R}^d \to \mathbb{R}^d$ is a predictable process. Then for any $p, q \in (1, \infty)$ with $\frac{d}{p} + \frac{\alpha}{q} < \alpha$, the Krylov estimate holds for X with index p, q.

Proof. First of all, we show that for all $p, q \in (1, \infty)$ with $\frac{d}{p} + \frac{\alpha}{q} < \alpha - 1$,

$$\mathbb{E}\bigg(\int_{t_0}^{t_1} f(s, X_s)\,\mathrm{d}s \bigg| \mathcal{F}_{t_0}\bigg) \le c\|f\|_{\mathbb{L}_p^q(T)}, \quad 0 \le t_0 < t_1 \le T. \tag{5.28}$$

For $n > 0$, define

$$\tau_n := \inf\bigg\{t \ge 0 : \int_0^t |b_s|(X_s)\,\mathrm{d}s \ge n\bigg\}.$$

In (5.27), if we take $f = |b|$, $\xi(s) = |b_s|(X_s)$ and $\delta = 1/(1 \vee (2\|b\|_{\mathbb{L}^{q_1}_{p_1}(T)}))$, then

$$\mathbb{E}\left(\int_{t_0\wedge\tau_n}^{t_1\wedge\tau_n} |b_s|(X_s)\,\mathrm{d}s\Big|\mathscr{F}_{t_0\wedge\tau_n}\right) \le c\|b\|_{\mathbb{L}^{q_1}_{p_1}(T)}.$$

Letting $n \to \infty$, we further have

$$\mathbb{E}\left(\int_{t_0}^{t_1} |b_s|(X_s)\,\mathrm{d}s\Big|\mathscr{F}_{t_0}\right) \le c\|b\|_{\mathbb{L}^{q_1}_{p_1}(T)}.$$

Substituting this into (5.27) with $\tau = T$, we get (5.28).

Below, without loss of generality, we assume $f \in C_0^\infty(\mathbb{R}_+ \times \mathbb{R}^d)$. Let $b^n := b * \phi_n$ be defined as in (5.9). Since $b_n \in \mathbb{H}^{1,\infty}_\infty(T)$, by (5.21), (5.26), (ii) of Theorem 4.5 and Remark 4.6, for ε small enough and λ large enough, there exists a unique $u_n \in \mathbb{H}^{\alpha+\varepsilon,\infty}_\infty(T)$ solving the following equation

$$\partial_t u_n + \left(\mathscr{L}^{\kappa'}_\alpha - \lambda\right)u_n + \mathscr{L}^{\bar{b}^g}_1 u_n + \mathscr{L}^{b^n}_1 u_n = f. \tag{5.29}$$

By Itô's formula and (5.29), we have

$$\begin{aligned}
&\mathbb{E}\big(u_n(t_1, X_{t_1})|\mathscr{F}_{t_0}\big) - u_n(t_0, X_{t_0})\\
&\quad= \mathbb{E}\left(\int_{t_0}^{t_1} \big(\partial_s u_n + \mathscr{L}^g_{\nu,R}u_n + \bar{\mathscr{L}}^\eta_{\nu,R}u_n + \mathscr{L}^b_1 u_n\big)(s, X_s)\,\mathrm{d}s\Big|\mathscr{F}_{t_0}\right)\\
&\quad\overset{(5.24)}{=} \mathbb{E}\left(\int_{t_0}^{t_1} \big(f + \big(\lambda + \mathscr{L}^{\kappa''}_\alpha\big)u_n + \bar{\mathscr{L}}^g_{\nu,R}u_n + \bar{\mathscr{L}}^\eta_{\nu,R}u_n + \big(b - b^n\big)\cdot\nabla u_n\big)(s, X_s)\,\mathrm{d}s\Big|\mathscr{F}_{t_0}\right),
\end{aligned}$$

where $\bar{\mathscr{L}}^\eta_{\nu,R}$ is defined as in (1.5) in terms of η. Hence, by (5.28) and (4.23) with $\vartheta = 0, 1$ and $p' = q' = \infty$, we have

$$\begin{aligned}
\mathbb{E}\left(\int_{t_0}^{t_1} f(s, X_s)\,\mathrm{d}s\Big|\mathscr{F}_{t_0}\right) &\le c\|u_n\|_{\mathbb{L}^\infty(t_0,t_1)} + c\|\nabla u_n\|_{\mathbb{L}^\infty(T)}\big\|b - b^n\big\|_{\mathbb{L}^{q_1}_{p_1}(T)}\\
&\le c\|f\|_{\mathbb{L}^q_p(t_0,t_1)} + c\|f\|_{\mathbb{L}^{q_1}_{p_1}(t_0,t_1)}\big\|b - b^n\big\|_{\mathbb{L}^{q_1}_{p_1}(T)},
\end{aligned}$$

where c is independent of n due to $\|b_n\|_{\mathbb{L}^{q_1}_{p_1}(T)} \le \|b\|_{\mathbb{L}^{q_1}_{p_1}(T)}$. Letting $n \to \infty$, we obtain the desired estimate. □

6. Strong well-posedness of SDEs with jumps

Now, the Kolmogorov equations associated with SDE (1.2) have been studied in Section 4, and the desired Krylov estimates were obtained in Theorem 5.7 and Theorem 5.10. Combining these with the results obtained in Section 3, we give the proofs of the strong well-posedness of SDE (1.2).

6.1. *Proof of Theorem* 2.1

Below we fix $T > 0$ and assume that $(\mathbf{H}^\sigma_\beta)$ holds and for some $p, q \in (2, \infty)$ with $\frac{d}{p} + \frac{2}{q} < 1$,

$$|\nabla\sigma|, b, \left(\Gamma^{1,2}_{0,R}(g)\right)^{1/2} \in \mathbb{L}^q_p(T),$$

and

$$\Gamma^{0,2}_{0,R}(g) \in \mathbb{L}^\infty(T), \quad \lim_{\varepsilon\to 0}\big\|\Gamma^{0,2}_{0,\varepsilon}(g)\big\|_{\mathbb{L}^\infty(T)} = 0,$$

where $\Gamma^{j,2}_{0,R}(g)$ is defined by (2.1).

Consider the following backward second order partial integro-differential equation:

$$\partial_t u + \left(\mathscr{L}^a_2 - \lambda\right)u + \mathscr{L}^b_1 u + \mathscr{L}^g_{\nu,R}u + b = 0, \quad u(T, x) = 0. \tag{6.1}$$

Since $\frac{d}{p}+\frac{2}{q}<1$, by Theorem 4.3, for λ large enough, there is a unique solution $u\in\mathbb{H}^{2,q}_p(T)$ to the above equation with

$$\|u\|_{\mathbb{L}^\infty(T)}+\|\nabla u\|_{\mathbb{L}^\infty(T)}\le\frac{1}{2}.$$

Let $u_\infty(t,x):=u(t,x)$ and u_n be defined as in (5.9). Define for $n\in\mathbb{N}\cup\{\infty\}$,

$$\Phi_n(t,x):=x+u_n(t,x).$$

Since for each $t\in[0,T]$,

$$\frac{1}{2}|x-y|\le\big|\Phi_n(t,x)-\Phi_n(t,y)\big|\le\frac{3}{2}|x-y|,\tag{6.2}$$

the map $x\to\Phi_n(t,x)$ forms a C^1-diffeomorphism and

$$1/2\le\|\nabla\Phi_n\|_{\mathbb{L}^\infty(T)},\ \big\|\nabla\Phi_n^{-1}\big\|_{\mathbb{L}^\infty(T)}\le 2,\tag{6.3}$$

where $\Phi_n^{-1}(t,\cdot)$ is the inverse of $\Phi_n(t,\cdot)$ and

$$\Phi_n^{-1}(t,y)=y-u_n\big(t,\Phi_n^{-1}(t,y)\big).$$

The following limits are easily verified by the definition, $u\in\mathbb{H}^{2,q}_p(T)$ and (6.3):

$$\begin{aligned}&\lim_{n\to\infty}\big\|\nabla^j\Phi_n-\nabla^j\Phi_\infty\big\|_{\mathbb{L}^\infty(T)}=0,\qquad\lim_{n\to\infty}\big\|\nabla^j\Phi_n^{-1}-\nabla^j\Phi_\infty^{-1}\big\|_{\mathbb{L}^\infty(T)}=0,\quad j=0,1,\\&\lim_{n\to\infty}\big\|(\Phi_n-\Phi_\infty)\chi_m\big\|_{\mathbb{H}^{2,q}_p(T)}=0,\qquad\lim_{n\to\infty}\big\|\big(\Phi_n^{-1}-\Phi_\infty^{-1}\big)\chi_m\big\|_{\mathbb{H}^{2,q}_p(T)}=0,\end{aligned}\tag{6.4}$$

where χ_m is defined by (2.8). Now let us define $\Phi_t(x):=\Phi_\infty(t,x)$ and

$$\begin{aligned}&\tilde\sigma_t(y):=(\nabla\Phi_t\cdot\sigma_t)\circ\Phi_t^{-1}(y),\qquad\tilde b_t(y):=\lambda u\big(t,\Phi_t^{-1}(y)\big),\\&\tilde g_t(y,z):=\Phi_t\big(\Phi_t^{-1}(y)+g_t\big(\Phi_t^{-1}(y),z\big)\big)-y.\end{aligned}\tag{6.5}$$

The following lemma is proven in the Appendix.

Lemma 6.1.

(i) *$\tilde\sigma$ satisfies* $(\mathbf{H}^\sigma_\beta)$ *and* $|\nabla\tilde\sigma|\in\mathbb{L}^q_p(T)$, $\tilde b\in\mathbb{H}^{1,\infty}_\infty(T)$ *and*

$$\big(\Gamma^{1,2}_{0,R}(\tilde g)\big)^{1/2}\in\mathbb{L}^q_p(T),\ \Gamma^{0,2}_{0,R}(\tilde g)\in\mathbb{L}^\infty(T),\quad\lim_{\varepsilon\to0}\big\|\Gamma^{0,2}_{0,\varepsilon}(\tilde g)\big\|_{\mathbb{L}^\infty(T)}=0.$$

(ii) $\lim_{n\to\infty}\|(\partial_s+\mathscr{L}^\sigma_2+\mathscr{L}^b_1+\mathscr{L}^g_{\nu,R})\Phi_n-\lambda u\|_{\mathbb{L}^q_p(T)}=0$.

(iii) $\lim_{n\to\infty}\|((\partial_s+\mathscr{L}^{\tilde\sigma}_2+\mathscr{L}^{\tilde b}_1+\mathscr{L}^{\tilde g}_{\nu,R})\Phi_n^{-1}-b\circ\Phi^{-1})\chi_m\|_{\mathbb{L}^q_p(T)}=0$, *where χ_m is defined by* (2.8).

Now, as a consequence of Theorem 3.10, Lemma 6.1 and Theorem 5.7, we have

Lemma 6.2. *Let $\Phi_t(x)$ be defined as above. Then X_t solves SDE* (1.2) *if and only if $Y_t:=\Phi_t(X_t)$ solves the following SDE*:

$$\mathrm{d}Y_t=\tilde\sigma_t(Y_t)\,\mathrm{d}W_t+\tilde b_t(Y_t)\,\mathrm{d}t+\int_{|z|<R}\tilde g_t(Y_{t-},z)\tilde N(\mathrm{d}t,\mathrm{d}z)+\int_{|z|\ge R}\tilde g_t(Y_{t-},z)N(\mathrm{d}t,\mathrm{d}z),\tag{6.6}$$

where $\tilde\sigma$, $\tilde b$ and $\tilde g$ are defined by (6.5).

We are now in the position to give:

Proof of Theorem 2.1. By Lemma 6.2, it suffices to prove Theorem 2.1 for SDE (6.6). For the sake of simplicity, we drop the tilde over $\tilde\sigma$, $\tilde b$ and $\tilde g$.

(i) Define

$$\sigma_t^{(n)}(y) := \sigma_t * \phi_n(y), \qquad g_t^{(n)}(y,z) := g_t(\cdot,z) * \phi_n(y),$$

where ϕ_n is the mollifiers in $\mathbb{R}^d$. Since σ satisfies $(\mathbf{H}^\sigma_\beta)$, there is a n_0 large enough such that for all $n \geq n_0$,

$\sigma^{(n)}$ satisfies $(\mathbf{H}^\sigma_\beta)$ uniformly with respect to n,

and

$$\|\nabla\sigma^{(n)}\|_{\mathbb{L}^q_p(T)} \leq \|\nabla\sigma\|_{\mathbb{L}^q_p(T)}, \qquad \|(\Gamma^{1,2}_{0,R}(g^{(n)}))^{1/2}\|_{\mathbb{L}^q_p(T)} \leq \|(\Gamma^{1,2}_{0,R}(g))^{1/2}\|_{\mathbb{L}^q_p(T)},$$
$$\|\Gamma^{0,2}_{0,R}(g^{(n)})\|_{\mathbb{L}^\infty(T)} \leq \|\Gamma^{0,2}_{0,R}(g)\|_{\mathbb{L}^\infty(T)}, \qquad \lim_{\varepsilon\to 0}\sup_n \|\Gamma^{0,2}_{0,\varepsilon}(g^{(n)})\|_{\mathbb{L}^\infty(T)} = 0.$$

Let $Y^{(n)}$ solve the following SDE with no big jumps:

$$Y_t^{(n)} = y + \int_0^t \sigma_s^{(n)}(Y_s^{(n)})\,\mathrm{d}W_s + \int_0^t b_s(Y_s^{(n)})\,\mathrm{d}s + \int_0^t\int_{|z|<R} g_s^{(n)}(Y_{s-}^{(n)},z)\tilde{N}(\mathrm{d}s,\mathrm{d}z). \tag{6.7}$$

Since $\sigma^{(n)}$, $g^{(n)}$ satisfy the assumptions of Theorem 5.7 uniformly with respect to n, by Theorem 5.7, $Y^{(n)}$ satisfies the Krylov estimate for all p', q' with $\frac{d}{p'} + \frac{2}{q'} < 2$ and the Krylov constant is independent of n. Thus, by Theorem 3.9 with $r = 1$, we have for any $\theta \in (0,1)$,

$$\mathbb{E}\Big(\sup_{t\in[0,T]}|Y_t^{(n)} - Y_t^{(m)}|^{2\theta}\Big) \lesssim \|\sigma^{(n)} - \sigma^{(m)}\|^{2\theta}_{\mathbb{L}^2_\infty(T)} + \|\Gamma^{0,2}_{0,R}(g^{(n)} - g^{(m)})\|^{\theta}_{\mathbb{L}^1_\infty(T)}.$$

Here and below, the constant contained in $\lesssim$ is independent of n. Since $p > d$, by (A.2) with $\mathbb{B} = L^2(B_R;\nu)$, we have

$$\begin{aligned}
\Gamma^{0,2}_{0,R}(g_t^{(n)} - g_t)(y) &= \int_{|z|<R}\left|\int_{\mathbb{R}^d}\big(g_t(y-y',z) - g_t(y,z)\big)\phi_n(y')\,\mathrm{d}y'\right|^2 \nu(\mathrm{d}z) \\
&\leq \Big(\int_{\mathbb{R}^d}\|g(y-y',\cdot) - g(y,\cdot)\|_{L^2(B_R;\nu)}\phi_n(y')\,\mathrm{d}y'\Big)^2 \\
&\lesssim \Big(\int_{\mathbb{R}^d}|y'|^{1-d/p}\phi_n(y')\,\mathrm{d}y'\Big)^2 \|(\Gamma^{1,2}_{0,R}(g_t))^{1/2}\|_p^2 \\
&\leq n^{-2+2d/p}\|(\Gamma^{1,2}_{0,R}(g_t))^{1/2}\|_p^2,
\end{aligned} \tag{6.8}$$

and by (A.2) with $\mathbb{B} = \mathbb{R}^d \otimes \mathbb{R}^d$,

$$\|\sigma_t^{(n)}(y) - \sigma_t(y)\| \lesssim n^{-1+d/p}\|\nabla\sigma_t\|_p. \tag{6.9}$$

Therefore,

$$\mathbb{E}\Big(\sup_{t\in[0,T]}|Y_t^{(n)} - Y_t^{(m)}|^{2\theta}\Big) \lesssim \big(n^{-1+d/p} + m^{-1+d/p}\big)^{2\theta} \to 0 \quad \text{as } n,m\to\infty,$$

and there exists a càdlàg $\mathcal{F}_t$-adapted process Y such that

$$\lim_{n\to\infty}\mathbb{E}\Big(\sup_{t\in[0,T]}|Y_t^{(n)} - Y_t|^{2\theta}\Big) = 0.$$

By Remark 3.4, Y_t also satisfies the Krylov estimate with index p, q. By taking limits for (6.7), one finds that Y_t solves

$$Y_t = y + \int_0^t \sigma_s(Y_s)\,\mathrm{d}W_s + \int_0^t b_s(Y_s)\,\mathrm{d}s + \int_0^t\int_{|z|<R} g_s(Y_{s-},z)\tilde{N}(\mathrm{d}s,\mathrm{d}z). \tag{6.10}$$

For example, letting $Y_t^\infty := Y_t$, by (6.8), we have

$$\sup_{n\in\mathbb{N}\cup\{\infty\}} \mathbb{E}\left|\int_0^t \int_{|z|<R} \big(g_s^{(m)}\big(Y_{s-}^{(n)}, z\big) - g_s\big(Y_{s-}^{(n)}, z\big)\big)\tilde{N}(\mathrm{d}s, \mathrm{d}z)\right|^2$$
$$= \sup_{n\in\mathbb{N}\cup\{\infty\}} \mathbb{E}\int_0^t \int_{|z|<R} \left|g_s^{(m)}\big(Y_s^{(n)}, z\big) - g_s\big(Y_s^{(n)}, z\big)\right|^2 \nu(\mathrm{d}z)\,\mathrm{d}s$$
$$\le \big\|\Gamma_{0,R}^{0,2}\big(g^{(m)} - g\big)\big\|_{\mathbb{L}^1_\infty(T)} \to 0, \quad m\to\infty,$$

and for each $m \in \mathbb{N}$,

$$\lim_{n\to\infty} \mathbb{E}\left|\int_0^t \int_{|z|<R} g_s^{(m)}\big(Y_{s-}^{(n)}, z\big)\tilde{N}(\mathrm{d}s, \mathrm{d}z) - \int_0^t \int_{|z|<R} g_s^{(m)}(Y_{s-}, z)\tilde{N}(\mathrm{d}s, \mathrm{d}z)\right|^2 = 0.$$

Combining the above two estimates, we obtain

$$\lim_{n\to\infty} \mathbb{E}\left|\int_0^t \int_{|z|<R} g_s^{(n)}\big(Y_{s-}^{(n)}, z\big)\tilde{N}(\mathrm{d}s, \mathrm{d}z) - \int_0^t \int_{|z|<R} g_s(Y_{s-}, z)\tilde{N}(\mathrm{d}s, \mathrm{d}z)\right|^2 = 0.$$

(ii) To show (2.2), we first consider SDE (6.10). By the classical Bismut–Elworthy–Li's formula (see [52]), we have for any $h \in \mathbb{R}^d$,

$$\nabla_h \mathbb{E}\varphi\big(Y_t^{(n)}(y)\big) = \frac{1}{t}\mathbb{E}\left[\varphi\big(Y_t^{(n)}(y)\big)\int_0^t \big[\sigma\big(Y_s^{(n)}(y)\big)\big]^{-1}\nabla_h Y_s^{(n)}(y)\,\mathrm{d}W_s\right], \tag{6.11}$$

where $\nabla_h Y_t^{(n)}(y) := \lim_{\varepsilon\to 0}[Y_t^{(n)}(y+\varepsilon h) - Y_t^{(n)}(y)]/\varepsilon$ is the derivative flow of $Y_t^{(n)}(y)$ with respect to the initial value y.

Now by Theorem 3.9, we have for any $\theta \in (0, 1)$,

$$\mathbb{E}\big|Y_t^{(n)}(y) - Y_t^{(n)}\big(y'\big)\big|^{2\theta} \le c\big|y - y'\big|^{2\theta},$$

where c is independent of n. Let $\theta \in (1/2, 1)$ and $T > 0$. By Theorem A.2 in the Appendix with $p = 2\theta$ and $q = r = \infty$, we get

$$\sup_n \sup_y \mathbb{E}\Big(\sup_{t\in[0,T]}\big|\nabla Y_t^{(n)}(y)\big|^{2\theta}\Big) \le c.$$

Hence, by (6.11) and Burkholder's inequality, for $t \in [0, T]$,

$$\sup_y\big|\nabla\mathbb{E}\varphi\big(Y_t^{(n)}(y)\big)\big| \le \frac{\|\varphi\|_\infty\|\sigma^{-1}\|_\infty}{t}\sup_y \mathbb{E}\left[\int_0^t \big|\nabla Y_s^{(n)}(y)\big|^2\,\mathrm{d}s\right]^{1/2}$$
$$\le \frac{\|\varphi\|_\infty\|\sigma^{-1}\|_\infty}{\sqrt{t}}\sup_y \mathbb{E}\Big[\sup_{s\in[0,t]}\big|\nabla Y_s^{(n)}(y)\big|\Big] \le c\|\varphi\|_\infty t^{-1/2},$$

which means that

$$\big|\mathbb{E}\varphi\big(Y_t^{(n)}(y)\big) - \mathbb{E}\varphi\big(Y_t^{(n)}\big(y'\big)\big)\big| \le c_T\|\varphi\|_\infty t^{-1/2}\big|y - y'\big|.$$

By taking limits $n \to \infty$ we get

$$\mathrm{Var}\big(P_t(y,\cdot) - P_t\big(y',\cdot\big)\big) = \sup_{\varphi\in C_b(\mathbb{R}^d), \|\varphi\|_\infty\le 1}\big|\mathbb{E}\varphi\big(Y_t(y)\big) - \mathbb{E}\varphi\big(Y_t\big(y'\big)\big)\big| \le c_T t^{-1/2}\big|y - y'\big|,$$

where $P_t(y,\cdot)$ denotes the law of $Y_t(y)$.

(iii) To allow the large jump in the equation, we shall use the interlacing technique. More precisely, let p_s be a point function on $\mathbb{R}_+$ with values in B_R^c, μ the associated counting measure, i.e.,

$$\mu\big([0, t] \times A\big) := \#\big\{\mathrm{p}_s \in A : s \in [0, t]\big\}, \quad A \in \mathscr{B}\big(B_R^c\big).$$

Let $\tau_n^{\mathrm{p}} := \inf\{t > 0 : \mu([0,t] \times B_R^c) = n\}$ be the n-th jump time of $t \mapsto \mu([0,t] \times B_R^c)$. Let $Y_{s,t}(y)$ solve SDE (6.10) starting from y at time $t = s$. Let $\tau_0^{\mathrm{p}} = 0$ and $Y_0^{\mathrm{p}}(y) = y$. Define $Y_t^{\mathrm{p}}(y)$ recursively by

$$Y_t^{\mathrm{p}} := Y_t^{\mathrm{p}}(y) := \begin{cases} Y_{\tau_{n-1}^{\mathrm{p}},t}(Y_{\tau_{n-1}^{\mathrm{p}}}^{\mathrm{p}}(y)), & t \in [\tau_{n-1}^{\mathrm{p}}, \tau_n^{\mathrm{p}}), \\ Y_{\tau_n^{\mathrm{p}}-}^{\mathrm{p}}(y) + g_{\tau_n^{\mathrm{p}}}(Y_{\tau_n^{\mathrm{p}}-}^{\mathrm{p}}(y), \mathrm{p}_{\tau_n^{\mathrm{p}}}), & t = \tau_n^{\mathrm{p}}. \end{cases}$$

It is easy to see that Y_t^{p} solves the following SDE with starting point $Y_0^{\mathrm{p}} = y$:

$$\mathrm{d}Y_t^{\mathrm{p}} = \sigma_t\big(Y_t^{\mathrm{p}}\big)\,\mathrm{d}W_t + b_t\big(Y_t^{\mathrm{p}}\big)\,\mathrm{d}t + \int_{|z|<R} g_t\big(Y_{t-}^{\mathrm{p}}, z\big)\tilde{N}(\mathrm{d}t, \mathrm{d}z) + \int_{|z|\geq R} g_t\big(Y_{t-}^{\mathrm{p}}, z\big)\mu(\mathrm{d}t, \mathrm{d}z).$$

In particular, if we let p_s^N be the Poisson point process with values in B_R^c associated to the Poisson random measure $N(\mathrm{d}t, \mathrm{d}z)$, i.e.,

$$N\big([0,t] \times A\big) = \#\big\{\mathrm{p}_s^N \in A : s \in [0,t]\big\}, \quad A \in \mathscr{B}\big(B_R^c\big),$$

then p^N is independent of Y. Therefore, $\tilde{Y}_t := Y_t^{\mathrm{p}^N}$ solves SDE (1.2) with $Y_0 = y$.

Next we show (2.2) for $\tilde{Y}_t(y)$. We adopt the same argument as used in [52]. We first look at it for $Y_t^{\mathrm{p}}(y)$. Observing that

$$Y_t^{\mathrm{p}}(y) = \begin{cases} Y_t(y), & t < \tau_1^{\mathrm{p}}, \\ Y_{\tau_1^{\mathrm{p}}-}(y) + g_{\tau_1^{\mathrm{p}}}(Y_{\tau_1^{\mathrm{p}}-}^{\mathrm{p}}(y), \mathrm{p}_{\tau_1^{\mathrm{p}}}), & t = \tau_1^{\mathrm{p}}, \\ Y_{\tau_1^{\mathrm{p}},t}(Y_{\tau_1^{\mathrm{p}}}^{\mathrm{p}}(y)), & t \in [\tau_1^{\mathrm{p}}, \tau_2^{\mathrm{p}}), \\ \cdots, \end{cases}$$

by what we have proved in step (ii), and since $Y_{s,t}(\cdot)$ and $Y_{0,s}(\cdot)$ are independent, one sees that

$$\big|\mathbb{E}\varphi\big(Y_t^{\mathrm{p}}(y)\big) - \mathbb{E}\varphi\big(Y_t^{\mathrm{p}}\big(y'\big)\big)\big| \leq c_T \|\varphi\|_\infty \big(t \wedge \tau_1^{\mathrm{p}}\big)^{-1/2}\big|y - y'\big|, \quad t \in [0,T].$$

Hence,

$$\big|\mathbb{E}\varphi\big(\tilde{Y}_t(y)\big) - \mathbb{E}\varphi\big(\tilde{Y}_t\big(y'\big)\big)\big| \leq c_T \|\varphi\|_\infty \mathbb{E}\big(t \wedge \tau_1^{\mathrm{p}^N}\big)^{-1/2}\big|y - y'\big|, \quad t \in [0,T].$$

Since the random variable $\tau_1^{\mathrm{p}^N} = \inf\{t > 0 : N([0,t] \times B_R^c) = 1\}$ obeys the exponential distribution with parameter $\nu(B_R^c)$, by easy calculations, we have

$$\mathbb{E}\big(t \wedge \tau_1^{\mathrm{p}^N}\big)^{-1/2} \leq ct^{-1/2}.$$

Thus, we get (2.2) for $\tilde{Y}_t(y)$. The proof is complete. □

6.2. *Proof of Theorem* 2.4

Let $T > 0$ and $\nu(\mathrm{d}z) = |z|^{-d-\alpha}\,\mathrm{d}z$ for some $\alpha \in (1,2)$. Below, we assume $(\mathbf{H}_\beta^g)$ holds with $2 - \alpha > \beta > 1 - \frac{\alpha}{2}$, and for some $\theta \in (1 - \frac{\alpha}{2}, 1)$, $p \in (\frac{2d}{\alpha} \vee 2, \infty)$ and $q \in (\frac{\alpha}{\alpha-1}, \infty)$,

$$\big(\Gamma_{0,R}^{1,2}(g)\big)^{1/2} \in \mathbb{L}_p^q(T), \quad b \in \mathbb{H}_p^{\theta,q}(T).$$

We also fix

$$\vartheta \in \big((1 + \alpha/2) \vee (1 + \theta) \vee (1 + d/p), (\alpha + \beta) \wedge (\theta + \alpha - \alpha/q)\big).$$

Consider the following backward nonlocal equation:

$$\partial_t u + \big(\mathscr{L}_\nu^g - \lambda\big)u + \mathscr{L}_1^b u + b = 0, \quad u(T,x) = 0. \tag{6.12}$$

Recalling (5.24), we can rewrite the above equation as

$$\partial_t u + \big(\mathscr{L}^{\kappa'}_{\nu} - \lambda\big)u + \mathscr{K}u + \mathscr{L}^{\bar b^g}_1 u + \mathscr{L}^b_1 u + b = 0, \quad u(T,x) = 0, \tag{6.13}$$

where $\mathscr{K}u := \bar{\mathscr{L}}^{\kappa''}_{\alpha} u$, and $\bar{\mathscr{L}}^{\kappa''}_{\alpha}$, $\bar b^g$ are defined by (5.25). The following lemma is obvious by Lemma 4.4.

Lemma 6.3. *For any $p > 1$ and $\theta \in [0, \beta)$, we have $\|\mathscr{K}u\|_{\theta,p} \le c\|u\|_{1,p}$.*

By Lemma 6.3 and Theorem 4.5, for any $\lambda \ge \lambda_0$, there is a unique solution $u \in \mathbb{H}^{\vartheta,\infty}_p(T)$ to equation (6.13), and so does for equation (6.12). Moreover, by Sobolev's embedding (4.2),

$$\sup_{t\in[0,T]} \|u(t)\|_{C_b^{\vartheta-d/p}} \le \|u\|_{\mathbb{H}^{\vartheta-d/p,\infty}_{\infty}(T)} \le \frac{1}{2}. \tag{6.14}$$

As in Section 6.1, we introduce u_∞, u_n and Φ_n so that (6.2) and (6.3) hold, and also define $\Phi_t(x) := \Phi_\infty(t,x)$ and

$$\begin{aligned} \tilde b_t(y) &:= \lambda u\big(t, \Phi_t^{-1}(y)\big) - \big(\bar{\mathscr{L}}^g_{\nu,R}\Phi_t\big) \circ \Phi_t^{-1}(y), \\ \tilde g_t(y,z) &:= \Phi_t\big(\Phi_t^{-1}(y) + g_t\big(\Phi_t^{-1}(y), z\big)\big) - y. \end{aligned} \tag{6.15}$$

Then by (6.14), we also have

$$\nabla\Phi_n, \nabla\Phi_n^{-1} \text{ are Hölder continuous uniformly with respect to } t, n. \tag{6.16}$$

The following lemma is proven in the Appendix.

Lemma 6.4.

(i) $\tilde b \in \mathbb{L}^\infty(T) \cap \mathbb{H}^{1,q}_p(T)$ *and* $\tilde g$ *satisfies* $(\mathbf{H}^g_\beta)$, $(\Gamma^{1,2}_{0,R}(\tilde g))^{1/2} \in \mathbb{L}^q_p(T)$.

(ii) $\lim_{n\to\infty} \|(\partial_s + \mathscr{L}^b_1 + \mathscr{L}^g_{\nu,R})\Phi_n - \tilde b \circ \Phi\|_{\mathbb{L}^q_p(T)} = 0$.

(iii) $\lim_{n\to\infty} \|((\partial_s + \mathscr{L}^{\tilde b}_1 + \mathscr{L}^{\tilde g}_{\nu,R})\Phi_n^{-1} - b \circ \Phi^{-1})\chi_m\|_{\mathbb{L}^q_p(T)} = 0$, *where χ_m is defined by* (2.8).

Let $p_1 := dp/(d - \theta p)$. Since $\frac{d}{p_1} + \frac{\alpha}{q} < \alpha - 1$ and $b \in \mathbb{L}^q_{p_1}(T)$, by Theorem 5.10, any solution X of SDE (1.2) with $\sigma \equiv 0$ satisfies the Krylov estimate for all $p', q' \in (1,\infty)$ with $\frac{d}{p'} + \frac{\alpha}{q'} < \alpha$. As in Lemma 6.2, by Lemma 6.4, Theorem 5.10 and Theorem 3.10, we have the following lemma.

Lemma 6.5. *Let $\Phi_t(x)$ be defined as above. Then X_t solves SDE*

$$\mathrm{d}X_t = b_t(X_t)\,\mathrm{d}t + \int_{|z|<R} g_t(X_{t-}, z)\tilde N(\mathrm{d}t, \mathrm{d}z) + \int_{|z|\ge R} g_t(X_{t-}, z) N(\mathrm{d}t, \mathrm{d}z)$$

if and only if $Y_t := \Phi_t(X_t)$ solves the following SDE:

$$\mathrm{d}Y_t = \tilde b_t(Y_t)\,\mathrm{d}t + \int_{|z|<R} \tilde g_t(Y_{t-}, z)\tilde N(\mathrm{d}t, \mathrm{d}z) + \int_{|z|\ge R} \tilde g_t(Y_{t-}, z) N(\mathrm{d}t, \mathrm{d}z), \tag{6.17}$$

where $\tilde b$ and $\tilde g$ are defined by (6.15).

Now, we are in the position to give:

Proof of Theorem 2.4. By Lemma 6.5, it suffices to prove the theorem for SDE (6.17).

(i) Let $\tilde b^{(n)}_t(y) := \tilde b_t * \phi_n(y)$ and $\tilde g^{(n)}_t(y,z) := \tilde g_t(\cdot, z) * \phi_n(y)$. By (i) of Lemma 6.4, there is a n_0 large enough such that for all $n \ge n_0$,

$$\tilde g^{(n)} \text{ satisfies } (\mathbf{H}^g_\beta) \text{ with constant } c_1 \text{ independent of } n,$$

and

$$\big\|\big(\Gamma^{1,2}_{0,R}\big(\tilde g^{(n)}\big)\big)^{1/2}\big\|_{\mathbb{L}^q_p(T)} \le \big\|\big(\Gamma^{1,2}_{0,R}(\tilde g)\big)^{1/2}\big\|_{\mathbb{L}^q_p(T)}.$$

Let $Y^{(n)}$ satisfy

$$Y_t^{(n)} = y + \int_0^t \tilde{b}_s^{(n)}\big(Y_s^{(n)}\big)\,\mathrm{d}s + \int_0^t \int_{|z|<R} \tilde{g}_s^{(n)}\big(Y_{s-}^{(n)}, z\big)\tilde{N}(\mathrm{d}s, \mathrm{d}z). \tag{6.18}$$

By Theorem 5.10, for any p', q' with $\frac{d}{p'} + \frac{\alpha}{q'} < \alpha$, $Y^{(n)}$ satisfies Krylov's estimate with index p', q' and the Krylov constant is independent of n. Thus, by Theorem 3.9 with $r = 1$, for any $\theta \in (0, 1)$, we have

$$\mathbb{E}\Big(\sup_{t\in[0,T]}\big|Y_t^{(n)} - Y_t^{(m)}\big|^\theta\Big) \lesssim \big\|b^{(n)} - b^{(m)}\big\|_{\mathbb{L}^2_\infty(T)}^{2\theta} + \big\|\Gamma_{0,R}^{0,2}\big(\tilde{g}^{(n)} - \tilde{g}^{(m)}\big)\big\|_{\mathbb{L}^1_\infty(T)}^{\theta},$$

which converges to zero as $n, m \to \infty$ by (6.8) and similar (6.9). Therefore, there exists a càdlàg $\mathcal{F}_t$-adapted process Y such that

$$\lim_{n\to\infty} \mathbb{E}\Big(\sup_{t\in[0,T]}\big|Y_t^{(n)} - Y_t\big|^\theta\Big) = 0,$$

and by Remark 3.4, Y_t also satisfies the Krylov estimate with index p, q. By taking limits for (6.18), one finds that Y_t solves

$$Y_t = y + \int_0^t \tilde{b}_s(Y_s)\,\mathrm{d}s + \int_0^t \int_{|z|<R} \tilde{g}_s(Y_{s-}, z)\tilde{N}(\mathrm{d}s, \mathrm{d}z).$$

The uniqueness follows by Theorem 3.9. For the large jump, we use the same technique as used in the proof of Theorem 2.1.

(ii) To show the existence and the estimates of the distribution density of of X_t, we use the results obtained in [14]. In view of (5.19), we have

$$\mathcal{L} := \mathcal{L}_\nu^g + \mathcal{L}_1^b = \mathcal{L}_\alpha^\kappa + \mathcal{L}_1^{\bar{b}^g} + \mathcal{L}_1^b,$$

where k is defined by (5.20) and $\bar{b}^g := \int_{|z|\ge R} g_t(x, z)\nu(\mathrm{d}z)$. By (2.3), $\bar{b}^g$ is bounded. Thus by [14, Theorem 1.5], the operator $\mathcal{L}$ admits a fundamental solution $\tilde{\rho}(s, x; t, y)$ so that for ε small enough, and for all $0 \le s < t \le T$ and $x, y \in \mathbb{R}^d$,

$$c_0\varrho_0^{(\alpha)}(s - t, x - y) \le \tilde{\rho}(s, x; t, y) \le c_1\big(\varrho_0^{(\alpha)} + \varrho_\varepsilon^{(\alpha-\varepsilon)}\big)(s - t, x - y), \tag{6.19}$$

where $\varrho_\gamma^{(\beta)}$ is defined by (4.17), and

$$\big|\nabla_x\tilde{\rho}(s, x; t, y)\big| \le c_2\big(\varrho_{-1}^{(\alpha)} + \varrho_{\varepsilon-1}^{(\alpha-\varepsilon)}\big)(s - t, x - y). \tag{6.20}$$

Moreover, $\tilde{\rho}(s, x; t, y)$ is a family of transition probability density functions, which determines a Feller process

$$\big(\Omega, \mathcal{F}, (\mathbb{P}_{s,x})_{(s,x)\in\mathbb{R}_+\times\mathbb{R}^d}; (X_t)_{t\ge0}\big),$$

with the property that

$$\mathbb{P}_{s,x}(X_t = x, 0 \le t \le s) = 1,$$

and for $r \in [s, t]$ and $E \in \mathcal{B}(\mathbb{R}^d)$,

$$\mathbb{E}_{s,x}(X_t \in E|X_r) = \int_E \tilde{\rho}(r, X_r; s, y)\,\mathrm{d}y.$$

In particular, for any $f \in C_b^2(\mathbb{R}^d)$, it follows from the Markov property of X that under $\mathbb{P}_{s,x}$, with respect to the filtration $\mathcal{F}_t := \sigma\{X_r, r \le t\}$,

$$M_s^f := f(X_s) - f(X_t) - \int_t^s \mathcal{L}f(X_r)\,\mathrm{d}r \text{ is a martingale.}$$

In other words, $\mathbb{P}_{s,x}$ solves the martingale problem for $(\mathcal{L}, C_b^2(\mathbb{R}^d))$. On the other hand, by [1] or [15], we know that the martingale problem for $\mathcal{L}$ is well-posed, and by Itô's formula, any solution of SDE (1.2) is a martingale solution of $\mathcal{L}$. Therefore, the strong solution $X_t(x)$ admits a density $\rho(t, x, y) = \tilde{\rho}(0, x; t, y)$, and the desired estimates (2.6) and (2.7) follow by (6.19) and (6.20). □

7. Ergodicity of SDEs with jumps

This section is devoted to the study of the existence and uniqueness of invariant probability measures associated with the time-independent SDE (2.10). To prove Theorem 2.9 and Theorem 2.12, we shall first establish a general ergodicity result for SDE (2.10) with dissipative drifts in Section 7.1. Then, we shall use Zvonkin's method to transform SDE (2.10) with singular dissipative drifts into a new SDE, and verify that the new SDE satisfies the conditions in Theorem 7.4. Thus, the conclusions in Theorem 2.9 and Theorem 2.12 follow by Proposition 2.8 and Theorem 7.4.

7.1. *SDEs with dissipative drifts*

For each $m \in \mathbb{N}$, let $\chi_m(x)$ be the cutoff function in (2.8). Let

$$\sigma_m(x) := \sigma\big(x\chi_m(x)\big), \qquad b_m(x) := b(x)\chi_m(x), \qquad g_m(x,z) := g\big(x\chi_m(x), z\big).$$

Below, we always assume that one of the following conditions holds:

(C1) For each $m \in \mathbb{N}$, (σ_m, b_m, g_m) satisfies the assumptions of Theorem 2.1.
(C2) For each $m \in \mathbb{N}$, $(0, b_m, g_m)$ satisfies the assumptions of Theorem 2.4.

Under **(C1)** or **(C2)**, by Corollary 2.6, there exists a unique local strong solution to SDE (2.10). To show the non-explosion and ergodicity, we make the following assumptions:

(C3) For some $r > -1$ and $\kappa_1, \kappa_2, \kappa_3 > 0$, it holds that

$$2\langle x, b(x)\rangle + \|\sigma(x)\|^2 \le -\kappa_1|x|^{2+r} + \kappa_2, \qquad \big|b(x)\big| \le \kappa_3\big(1 + |x|^{1+r}\big), \tag{7.1}$$

and for any $\varepsilon > 0$ and $\lambda \ge R$, there is a $c_{\varepsilon,\lambda} > 0$ such that

$$\Gamma^{0,2}_{0,\lambda}(g)(x) + \Gamma^{0,1}_{\lambda,\infty}(g)(x) \le \varepsilon|x|^{1+r} + c_{\varepsilon,\lambda}. \tag{7.2}$$

We first show the non-explosion and some moment estimates of the unique strong solution.

Lemma 7.1. *Under* **(C3)**, *there is no explosion to SDE* (2.10). *Moreover, for any $\vartheta \in (0,1)$, there is a constant $c > 0$ such that for all $t > 0$ and $x \in \mathbb{R}^d$*,

$$\int_0^t \mathbb{E}\big|X_s(x)\big|^{1+r}\,\mathrm{d}s + \Big[\mathbb{E}\Big(\sup_{s\in[0,t]}\big|X_s(x)\big|^{\vartheta}\Big)\Big]^{1/\vartheta} \le c\big(|x| + t + 1\big), \tag{7.3}$$

and

$$\mathbb{E}\big|X_t(x)\big| \le \begin{cases} c\mathrm{e}^{-t/c}|x| + c, & r = 0, \\ c(1 + t^{-1/2}), & r > 0. \end{cases} \tag{7.4}$$

Proof. Let $h(x) := \sqrt{1 + |x|^2}$. By Itô's formula, we have

$$\mathrm{d}h(X_t) = \big[\mathscr{L}^{\sigma}_2 h + \mathscr{L}^{b}_1 h + \mathscr{L}^{g}_{\nu} h\big](X_t)\,\mathrm{d}t + \mathrm{d}M_t,$$

where M_t is a local martingale. Noticing that for $i, j = 1, \dots, d$,

$$\partial_i h(x) = x_i\big(1 + |x|^2\big)^{-1/2}/2$$

and

$$\partial_i\partial_j h(x) = \big(1 + |x|^2\big)^{-1/2}\delta_{ij}/2 - 3x_i x_j\big(1 + |x|^2\big)^{-3/2}/4,$$

we have

$$\mathscr{L}^{\sigma}_2 h(x) + \mathscr{L}^{b}_1 h(x) \le \big(\|\sigma(x)\|^2 + 2\langle x, b(x)\rangle\big)\big(1 + |x|^2\big)^{-1/2}/4. \tag{7.5}$$

On the other hand, observing that

$$\big|h(x+y)-h(x)\big| \le |y|\int_0^1 \big|\nabla h(x+sy)\big|\,\mathrm{d}s \le |y|/2,$$
$$h(x+y)-h(x)-y\cdot\nabla h(x) \le |y|^2/2,$$

we have

$$\begin{aligned}\mathscr{L}^g_\nu h(x) &= \int_{\mathbb{R}^d}\big[h\big(x+g(x,z)\big)-h(x)-1_{|z|<R}\,g(x,z)\cdot\nabla h(x)\big]\nu(\mathrm{d}z)\\ &\le \frac{1}{2}\int_{|z|<R}\big|g(x,z)\big|^2\nu(\mathrm{d}z)+\frac{1}{2}\int_{|z|\ge R}\big|g(x,z)\big|\nu(\mathrm{d}z)\\ &= \big(\Gamma^{0,2}_{0,R}(g)(x)+\Gamma^{0,1}_{R,\infty}(g)(x)\big)/2. \end{aligned} \tag{7.6}$$

By (7.5), (7.6) and (7.1), (7.2), there are $c_1, c_2 > 0$ only depending on κ_i in (7.1) such that

$$\big[\mathscr{L}^\sigma_2 h+\mathscr{L}^b_1 h+\mathscr{L}^g_\nu h\big](x) \le -c_1\big(1+|x|^2\big)^{(1+r)/2}+c_2.$$

Hence,

$$\mathrm{d}h(X_t) \le -c_1 h(X_t)^{1+r}\,\mathrm{d}t + c_2\,\mathrm{d}t + \mathrm{d}M_t.$$

Letting $\tau_n := \inf\{t>0 : |X_t| \ge n\}$, we have

$$c_1\mathbb{E}\bigg(\int_0^{t\wedge\tau_n} h(X_s)^{1+r}\,\mathrm{d}s\bigg) \le h(x)+c_2 t,$$

and by Lemma 3.7, for any $\vartheta\in(0,1)$,

$$\mathbb{E}\Big(\sup_{s\in[0,t\wedge\tau_n]} h(X_s)^\vartheta\Big) \le c_\vartheta\big(h(x)+c_2 t\big)^\vartheta,$$

which yields by a contradiction method that $\tau_n\to\infty$ as $n\to\infty$. By taking limits $n\to\infty$, we then obtain (7.3). Moreover, we also have

$$\mathrm{d}\mathbb{E}h(X_t)/\mathrm{d}t \le \begin{cases} -c_1\mathbb{E}(h(X_t))+c_2, & r=0,\\ -c_1(\mathbb{E}h(X_t))^{1+r}+c_2, & r>0.\end{cases}$$

Solving this differential inequality, we get (7.4). □

The following lemma is useful for showing the irreducibility in the non-degenerate diffusion case.

Lemma 7.2. *For given $x_0 \ne y_0 \in \mathbb{R}^d$ and $m \ge 1$, let Z_t solve the following SDE*:

$$\begin{aligned}\mathrm{d}Z_t &= \big[b(Z_t)-m(Z_t-y_0)/2\big]\mathrm{d}t+\sigma(Z_t)\,\mathrm{d}W_t\\ &\quad+\int_{|z|<R} g(Z_{t-},z)\tilde{N}(\mathrm{d}t,\mathrm{d}z)+\int_{|z|\ge R} g(Z_{t-},z)N(\mathrm{d}t,\mathrm{d}z),\quad Z_0=x_0.\end{aligned}\tag{7.7}$$

Under **(C3)**, *for any $0<a<|x_0-y_0|$ and $T>0$, there exists an m large enough such that*

$$\mathbb{P}\big(\big|Z_T(x_0)-y_0\big|>a\big)<1/2, \tag{7.8}$$

and for any $\vartheta\in(0,1)$,

$$\mathbb{E}\Big(\sup_{t\in[0,T]}|Z_t|^\vartheta\Big)<\infty. \tag{7.9}$$

Proof. First of all, by using the same argument as in estimating (7.3), we have (7.9). Let us show (7.8). For $\lambda > 0$, define

$$\tau_\lambda := \inf\{t \geq 0 : N([0,t] \times B_\lambda^c) = 1\}.$$

Let $T > 0$ be fixed. Since τ_λ obeys the exponential distribution with parameter $\nu(B_\lambda^c)$, one can choose $\lambda \geq R$ large enough so that

$$\mathbb{P}(\tau_\lambda \leq T) = 1 - \mathrm{e}^{-T\nu(B_\lambda^c)} \leq 1/4. \tag{7.10}$$

For this λ, let Z_t^λ solve the following SDE with starting point $Z_0^\lambda = x_0$,

$$\mathrm{d}Z_t^\lambda = \left[b_\lambda(Z_t^\lambda) - m(Z_t^\lambda - y_0)/2\right]\mathrm{d}t + \sigma(Z_t^\lambda)\,\mathrm{d}W_t + \int_{|z|<\lambda} g(Z_{t-}^\lambda, z)\tilde{N}(\mathrm{d}t, \mathrm{d}z),$$

where $b_\lambda(x) := b(x) + \int_{R \leq |z| < \lambda} g(x,z)\nu(\mathrm{d}z)$. Clearly,

$$Z_t = Z_t^\lambda, \quad t \in [0, \tau_\lambda). \tag{7.11}$$

By Itô's formula and (7.1), (7.2), we have

$$\begin{aligned}
\mathrm{e}^{mt}\mathbb{E}|Z_t^\lambda - y_0|^2 &= |x_0 - y_0|^2 + \mathbb{E}\int_0^t \mathrm{e}^{ms}\left(2\langle Z_s^\lambda - y_0, b_\lambda(Z_s^\lambda)\rangle + \|\sigma(Z_s^\lambda)\|^2\right)\mathrm{d}s \\
&\quad + \mathbb{E}\int_0^t \mathrm{e}^{ms}\int_{|z|\leq\lambda} |g(Z_s^\lambda, z)|^2\nu(\mathrm{d}z)\,\mathrm{d}s \\
&= |x_0 - y_0|^2 + \mathbb{E}\int_0^t \mathrm{e}^{ms}\left(2\langle Z_s^\lambda, b(Z_s^\lambda)\rangle + \|\sigma(Z_s^\lambda)\|^2 - 2\langle y_0, b_\lambda(Z_s^\lambda)\rangle\right)\mathrm{d}s \\
&\quad + \mathbb{E}\int_0^t \mathrm{e}^{ms}\left(2\left\langle Z_s^\lambda, \int_{R\leq|z|<\lambda} g(Z_s^\lambda, z)\nu(\mathrm{d}z)\right\rangle + \Gamma_{0,\lambda}^{0,2}(g)(Z_s^\lambda)\right)\mathrm{d}s \\
&\leq |x_0 - y_0|^2 + \mathbb{E}\int_0^t \mathrm{e}^{ms}\left(-\kappa_1|Z_s^\lambda|^{2+r} + \kappa_2\right)\mathrm{d}s \\
&\quad + 2|y_0|\mathbb{E}\int_0^t \mathrm{e}^{ms}\left(\kappa_3\left(|Z_s^\lambda|^{1+r} + 1\right) + \Gamma_{R,\lambda}^{0,1}(g)(Z_s^\lambda)\right)\mathrm{d}s \\
&\quad + \mathbb{E}\int_0^t \mathrm{e}^{ms}\left(2|Z_s^\lambda|\Gamma_{R,\lambda}^{0,1}(g)(Z_s^\lambda) + \Gamma_{0,\lambda}^{0,2}(g)(Z_s^\lambda)\right)\mathrm{d}s \\
&\leq |x_0 - y_0|^2 + c\left(\mathrm{e}^{mt} - 1\right)/m,
\end{aligned}$$

where $c > 0$ is independent of m. From this we derive that for m large enough,

$$\mathbb{P}\left(|Z_T^\lambda(x_0) - y_0| > a\right) \leq \frac{\mathbb{E}|Z_T^\lambda(x_0) - y_0|^2}{a^2} \leq \frac{\mathrm{e}^{-mT}|x_0 - y_0|^2}{a^2} + \frac{c(1 - \mathrm{e}^{-mT})}{ma^2} \leq 1/4,$$

which together with (7.10) and (7.11) yields that

$$\mathbb{P}\left(|Z_T(x_0) - y_0| > a\right) \leq \mathbb{P}\left(|Z_T(x_0) - y_0| > a, T < \tau_\lambda\right) + \mathbb{P}(T \geq \tau_\lambda) \leq 1/2.$$

The proof is complete. □

Let $P_t\varphi(x) := \mathbb{E}\varphi(X_t(x))$. We have

Lemma 7.3. *Under* **(C1)** *or* **(C2)**, *and* **(C3)**, *the semigroup P_t has the C_b-strong Feller property and irreducibility.*

Proof. (i) Let $X_t^m(x)$ be the solution of SDE (2.10) corresponding to (σ_m, b_m, g_m). In the case of **(C1)**, by (2.2), for any bounded measurable function f and $t > 0$,

$$x \mapsto \mathbb{E}f\left(X_t^m(x)\right) \text{ is continuous.} \tag{7.12}$$

In the case of **(C2)**, by the gradient estimate (2.7), we still have (7.12).

Now fix $K > 0$. For $x \in \mathbb{R}^d$ and $m > K$, define a stopping time

$$\tau_m^x := \{t \geq 0 : |X_t(x)| \geq m\}.$$

By Chebyshev's inequality and (7.3), we have

$$\lim_{m\to\infty} \sup_{|x|\leq K} \mathbb{P}(t \geq \tau_m^x) \leq \lim_{m\to\infty} \sup_{|x|\leq K} \mathbb{E}\Big(\sup_{s\in[0,t]} |X_s(x)|^{\vartheta}\Big)/m^{\vartheta} = 0. \tag{7.13}$$

Moreover, by the local uniqueness of solutions to SDE (2.10), we have

$$X_t(x) = X_t^m(x), \quad t \in [0, \tau_m^x), |x| \leq K.$$

Let f be a bounded measurable function. For any $x, y \in B_K$, we have

$$\begin{aligned}
&\big|\mathbb{E}\big(f(X_t(x)) - f(X_t(y))\big)\big| \\
&\quad \leq \big|\mathbb{E}\big(f(X_t(x)) - f(X_t(y))1_{\{t<\tau_m^x\wedge\tau_m^y\}}\big)\big| + 2\|f\|_\infty \mathbb{P}(t \geq \tau_m^x \wedge \tau_m^y) \\
&\quad = \big|\mathbb{E}\big(f(X_t^m(x)) - f(X_t^m(y))1_{\{t<\tau_m^x\wedge\tau_m^y\}}\big)\big| + 2\|f\|_\infty \mathbb{P}(t \geq \tau_m^x \wedge \tau_m^y) \\
&\quad \leq \big|\mathbb{E}\big(f(X_t^m(x)) - f(X_t^m(y))\big)\big| + 4\|f\|_\infty \mathbb{P}(t \geq \tau_m^x \wedge \tau_m^y) \\
&\quad \leq \big|\mathbb{E}\big(f(X_t^m(x)) - f(X_t^m(y))\big)\big| + 4\|f\|_\infty \big(\mathbb{P}(t \geq \tau_m^x) + \mathbb{P}(t \geq \tau_m^y)\big),
\end{aligned}$$

which together with (7.12), (7.13) yields the continuity of $x \mapsto \mathbb{E}(f(X_t(x)))$.

(ii) For the irreducibility, it suffices to prove that for any $T, a > 0$ and $x_0, y_0 \in \mathbb{R}^d$,

$$\mathbb{P}(|X_T(x_0) - y_0| \leq a) > 0. \tag{7.14}$$

In the case of **(C1)**, we use Lemma 7.2 and Girsanov's transformation to show (7.14), see [43]. Let $Z_t(x_0)$ solve SDE (7.7) and set for $K > 0$,

$$\tau_K := \inf\{t : |Z_t(x_0)| \geq K\}.$$

By (7.8) and (7.9), we may fix K and m large enough so that

$$\mathbb{P}(\tau_K \leq T) + \mathbb{P}(|Z_T(x_0) - y_0| > a) < 1. \tag{7.15}$$

Define

$$U_t := -m\sigma(Z_t)^{-1}(Z_t - y_0), \qquad \tilde{W}_t := W_t + \int_0^{t\wedge\tau_K} U_s\,\mathrm{d}s,$$

and

$$\mathcal{E}_T := \exp\bigg(\int_0^{T\wedge\tau_K} U_s\,\mathrm{d}W_s - \frac{1}{2}\int_0^{T\wedge\tau_K} |U_s|^2\,\mathrm{d}s\bigg).$$

Since $|U_t 1_{\{t<\tau_K\}}|^2$ is bounded, we have $\mathbb{E}[\mathcal{E}_T] = 1$. By Girsanov's theorem (see [45, Theorem 132]), under the new probability measure $\mathbb{Q} := \mathcal{E}_T\mathbb{P}$, $\tilde{W}_t$ is still a Brownian motion, and $N(\mathrm{d}t, \mathrm{d}z)$ is a Poisson random measure with the same compensator $\mathrm{d}t\,\nu(\mathrm{d}z)$. In view of (7.15), we also have

$$\mathbb{Q}(\{\tau_K \leq T\} \cup \{|Z_T(x_0) - y_0| > a\}) < 1.$$

Note that the solution Z_t of (7.7) also solves the following SDE:

$$\begin{aligned}
Z_{t\wedge\tau_K} &= x_0 + \int_0^{t\wedge\tau_K} b(Z_s)\,\mathrm{d}s + \int_0^{t\wedge\tau_K} \sigma(Z_s)\,\mathrm{d}\tilde{W}_s \\
&\quad + \int_0^{t\wedge\tau_K}\int_{|z|<R} g(Z_{s-}, z)\tilde{N}(\mathrm{d}s, \mathrm{d}z) + \int_0^{t\wedge\tau_K}\int_{|z|\geq R} g(Z_{s-}, z)N(\mathrm{d}s, \mathrm{d}z).
\end{aligned}$$

Set

$$\theta_K := \inf\{t : |X_t| \geq K\}.$$

Then the law uniqueness for (2.10) yields that the law of $\{(X_t 1_{\{t<\theta_K\}})_{t\in[0,T]}, \theta_K\}$ under $\mathbb{P}$ is the same as that of $\{(Z_t 1_{\{t<\tau_K\}})_{t\in[0,T]}, \tau_K\}$ under $\mathbb{Q}$. Hence

$$\begin{aligned}\mathbb{P}\big(\big|X_T(x_0) - y_0\big| > a\big) &\leq \mathbb{P}\big(\{\theta_K \leq T\} \cup \big\{\theta_K > T, \big|X_T(x_0) - y_0\big| > a\big\}\big) \\ &= \mathbb{Q}\big(\{\tau_K \leq T\} \cup \big\{\tau_K > T, \big|Z_T(x_0) - y_0\big| > a\big\}\big) \\ &\leq \mathbb{Q}\big(\{\tau_K \leq T\} \cup \big\{\big|Z_T(x_0) - y_0\big| > a\big\}\big) < 1,\end{aligned}$$

which implies (7.14).

In the case of **(C2)**, we shall use the positivity of the Dirichlet heat kernel, which is proved in Theorem A.3 in the Appendix. Let $D_m := \{x : |x| < m\}$ be a ball containing x_0 and $B_a(y_0)$. We have

$$\begin{aligned}\mathbb{P}\big(\big|X_T(x_0) - y_0\big| \leq a\big) &\geq \mathbb{P}\big(X_T(x_0) \in B_a(y_0); T < \tau_{D_m}\big) \\ &= \mathbb{P}\big(X_T^m(x_0) \in B_a(y_0); T < \tau_{D_m}\big),\end{aligned}$$

where $X_T^m(x_0)$ is the solution of SDE (2.10) corresponding to $(0, b_m, g_m)$. By (2.6), we can check that the functions

$$\varrho_1(t,r) := c_1 t\big(t^{1/\alpha} + r\big)^{-d-\alpha}, \qquad \varrho_2(t,r) := c_2 t\big(t^{1/\alpha} + r\big)^{-d-\alpha}\big(1 + \big(t^{1/\alpha} + r\big)^{\varepsilon}\big)$$

satisfy $(\mathbf{H}^{\varrho})$ in the Appendix. Thus, by (A.10) with $D = D_m$ and using Theorem A.3, we get (7.14). The proof is complete. □

Now we show the following ergodicity result.

Theorem 7.4. *Under* **(C1)** *or* **(C2)**, *and* **(C3)**, *P_t admits a unique invariant probability measure μ. Moreover, if $r = 0$ in* **(C3)**, *then P_t is V-uniformly exponential ergodic with $V(x) = 1 + |x|$; if $r > 0$, then P_t is uniformly exponential ergodic.*

Proof. By (7.4), the existence of invariant probability measures for P_t follows by the standard Bogoliov–Krylov's argument. The uniqueness is a direct consequence of the C_b-strong Feller property and the irreducibility. Moreover, still by the C_b-strong Feller property and the irreducibility, we can derive easily that for any $y \in \mathbb{R}^d$ and $r, t > 0$,

$$\inf_{x\in B_r} \mathbb{P}\big(X_t(x) \in B_r(y)\big) > 0.$$

Combing this with (7.4) and [21, Theorem 2.5], we get the desired exponential ergodicity. □

7.2. *SDEs with singular and dissipative drifts*

In this subsection we study the ergodicity of SDE (2.10) with singular and dissipative drifts. The main idea is to use Zvonkin's transformation to kill the singular part. We point out that Krylov's estimates obtained in Theorem 5.7 and Theorem 5.10 are not applicable for solutions of SDE (2.10) due to the dissipative part in the drift.

First of all, we consider the case of non-degenerate diffusion, and show the following non explosion and Krylov's estimate.

Lemma 7.5. *Under* $(\mathbf{H}^{\sigma}_{\beta})$, $(\mathbf{H}^{b})$ *and* (2.12), *any solution $X_t(x)$ to SDE* (2.10) *does not explode. Moreover, for any $T > 0$ and $f \in L^{p'}(\mathbb{R}^d)$ with $p' > d$,*

$$\mathbb{E}\bigg(\int_0^T f\big(X_s(x)\big)\,\mathrm{d}s\bigg) \leq c\big(|x| + 1\big)\|f\|_{p'}, \tag{7.16}$$

where $c > 0$ is independent of x.

Proof. For $n>0$, let $\tau_n:=\inf\{t\geq 0:|X_t|\geq n\}$. By Lemma 5.5, for any $T>0$, $p'>d$ and $\delta>0$, there exists a constant $c_\delta>0$ such that for any $f\in L^{p'}(\mathbb{R}^d)$,

$$\mathbb{E}\bigg(\int_0^{T\wedge\tau_n} f(X_s)\,\mathrm{d}s\bigg)\leq\bigg(c_\delta+\delta\mathbb{E}\bigg(\int_0^{T\wedge\tau_n}|b_1+b_2|(X_s)\,\mathrm{d}s\bigg)\bigg)\|f\|_{p'}. \tag{7.17}$$

Since $b_1\in L^p(\mathbb{R}^d)$ with $p>d$, for every $\delta_0>0$, we can take $f=|b_1|$ and choose δ small enough such that $\delta\|b_1\|_p<\delta_0$ in the above inequality to get

$$\begin{aligned}\mathbb{E}\bigg(\int_0^{T\wedge\tau_n} b_1(X_s)\,\mathrm{d}s\bigg) &\leq c_{\delta_0}+\delta_0\mathbb{E}\bigg(\int_0^{T\wedge\tau_n}\big|b_2(X_s)\big|\,\mathrm{d}s\bigg)\\ &\overset{(2.11)}{\leq} c_{\delta_0}+\kappa_3\delta_0\mathbb{E}\bigg(\int_0^{T\wedge\tau_n}\big(1+|X_s|^2\big)^{(1+r)/2}\,\mathrm{d}s\bigg).\end{aligned} \tag{7.18}$$

On the other hand, let $h(x):=\sqrt{1+|x|^2}$. By Itô's formula, we have

$$\mathbb{E}h(X_{t\wedge\tau_n})=h(x)+\mathbb{E}\int_0^{t\wedge\tau_n}\big[\mathcal{L}_2^\sigma h+\mathcal{L}_1^b h+\mathcal{L}_\nu^g h\big](X_s)\,\mathrm{d}s. \tag{7.19}$$

As the calculations in Lemma 7.1, by the assumptions, we have

$$\begin{aligned}\mathcal{L}_2^\sigma h(x) &\leq\frac{1}{2}\big(\sigma^{ik}\sigma^{ik}\big)(x)\big(1+|x|^2\big)^{-1/2}\leq c,\\ \mathcal{L}_1^b h(x) &\leq\big(-\kappa_1|x|^{2+r}+\kappa_2\big)\big(1+|x|^2\big)^{-1/2}+\big|b_1(x)\big|\\ &\leq-\kappa_1\big(1+|x|^2\big)^{(1+r)/2}/2+c+\big|b_1(x)\big|,\end{aligned}$$

and

$$\begin{aligned}\mathcal{L}_\nu^g h(x) &=\int_{\mathbb{R}^d}\big[h\big(x+g(x,z)\big)-h(x)-1_{|z|\leq R}g(x,z)\cdot\nabla h(x)\big]\nu(\mathrm{d}z)\\ &\leq\int_{|z|<R}\big|g(x,z)\big|^2\nu(\mathrm{d}z)+\int_{|z|\geq R}\big|g(x,z)\big|\nu(\mathrm{d}z)\leq c.\end{aligned}$$

Hence, by (7.19) and (7.18) with δ_0 small enough, we obtain

$$\begin{aligned}\mathbb{E}\big(1+|X_{T\wedge\tau_n}|^2\big)^{1/2} &\leq\big(1+|x|^2\big)^{1/2}-\frac{\kappa_1}{2}\mathbb{E}\int_0^{T\wedge\tau_n}\big(1+|X_s|^2\big)^{(1+r)/2}\,\mathrm{d}s\\ &\quad+\mathbb{E}\int_0^{T\wedge\tau_n}\big|b_1(X_s)\big|\,\mathrm{d}s+ct\\ &\leq\big(1+|x|^2\big)^{1/2}-\frac{\kappa_1}{4}\mathbb{E}\int_0^{T\wedge\tau_n}\big(1+|X_s|^2\big)^{(1+r)/2}\,\mathrm{d}s+c_T,\end{aligned}$$

which implies that $\lim_{n\to\infty}\tau_n=\infty$ and

$$\mathbb{E}\big(1+|X_T|^2\big)^{1/2}+\frac{\kappa_1}{4}\mathbb{E}\int_0^T\big(1+|X_s|^2\big)^{(1+r)/2}\,\mathrm{d}s\leq\big(1+|x|^2\big)^{1/2}+c_T.$$

Substituting this into (7.17) and (7.18), we obtain (7.16). □

To perform Zvonkin's transformation, we need to solve a related elliptic equation, which is a consequence of Theorem 4.3.

Theorem 7.6. *Suppose that* $(\mathbf{H}_\beta^\sigma)$ *holds and* $b\in L^p(\mathbb{R}^d)$ *for some* $p>d$, *and*

$$\Gamma_{0,R}^{0,2}(g)\in L^\infty\big(\mathbb{R}^d\big),\qquad\lim_{\varepsilon\to 0}\big\|\Gamma_{0,\varepsilon}^{0,2}(g)\big\|_\infty=0.$$

Then for some $\lambda_1 \geq 1$ *large enough and for all* $\lambda \geq \lambda_1$ *and* $f \in L^p(\mathbb{R}^d)$, *there exists a unique solution* $u \in H_p^2$ *to the following elliptic equation*:

$$\big(\mathscr{L}_2^\sigma - \lambda\big)u + \mathscr{L}_{\nu,R}^g u + \mathscr{L}_1^b u = f, \tag{7.20}$$

and for any $p' \in [p, \infty]$ *and* $\vartheta \in (0, 2)$ *with* $\frac{d}{p} < 2 - \vartheta + \frac{d}{p'}$,

$$\lambda^{\frac{1}{2}(2-\vartheta+\frac{d}{p'}-\frac{d}{p})} \|u\|_{\vartheta,p'} + \big\|\nabla^2 u\big\|_p \leq c\|f\|_p. \tag{7.21}$$

Proof. As usual, it suffices to show the a priori estimate (7.21). Let $u \in H_p^2$ solve (7.20). Let $T > 0$ and $\phi(t)$ be a nonnegative and nonzero smooth function with support in $(0, T)$. Let $\bar{u}(t, x) := u(x)\phi(t)$. It is easy to see that $\bar{u}$ satisfies the following parabolic equation:

$$\partial_t \bar{u} + \big(\mathscr{L}_2^\sigma - \lambda\big)\bar{u} + \mathscr{L}_{\nu,R}^g \bar{u} + \mathscr{L}_1^b \bar{u} = u\phi' + f\phi.$$

Thus, by Theorem 4.3, there is a $\lambda_0 \geq 1$ depending on $\|b\|_p$ and $\|\Gamma_{0,R}^{0,2}(g)\|_\infty$ such that for all $\lambda \geq \lambda_0$, $p' \in [p, \infty]$ and $\vartheta \in (0, 2)$ with $\frac{d}{p} < 2 - \vartheta + \frac{d}{p'}$,

$$\lambda^{\frac{1}{2}(2-\vartheta+\frac{d}{p'}-\frac{d}{p})} \|\bar{u}\|_{\mathbb{H}_{p'}^{\vartheta,\infty}(T)} + \big\|\nabla^2 \bar{u}\big\|_{\mathbb{L}_p^\infty(T)} \leq c\big\|u\phi' + f\phi\big\|_{\mathbb{L}_p^\infty},$$

which implies that

$$\lambda^{\frac{1}{2}(2-\vartheta+\frac{d}{p'}-\frac{d}{p})} \|u\|_{\vartheta,p'} + \big\|\nabla^2 u\big\|_p \leq c\|\phi\|_\infty^{-1}\big(\|u\|_p\big\|\phi'\big\|_\infty + \|f\|_p\|\phi\|_\infty\big). \tag{7.22}$$

Letting $p' = p$ and $\vartheta = 0$ in (7.22) and choosing $\lambda_1 \geq \lambda_0$ large enough, we get

$$\|u\|_p \leq c\|f\|_p.$$

Finally, substituting this into (7.22), we obtain the desired estimate (7.21). □

Below we assume that $(\mathbf{H}_\beta^\sigma)$ holds and for some $p > d$,

$$b_1, |\nabla\sigma|, \big(\Gamma_{0,R}^{1,2}(g)\big)^{1/2} \in L^p\big(\mathbb{R}^d\big), \Gamma_{0,R}^{0,2}(g) \in L^\infty\big(\mathbb{R}^d\big), \quad \lim_{\varepsilon\to 0}\big\|\Gamma_{0,\varepsilon}^{0,2}(g)\big\|_\infty = 0.$$

Now consider the following elliptic equation system:

$$\big(\mathscr{L}_2^\sigma - \lambda\big)u + \mathscr{L}_{\nu,R}^g u + \mathscr{L}_1^{b_1} u = b_1.$$

Note that only the first part b_1 in the drift of SDE (2.10) is involved. By (7.21), there are $c, \lambda_1 \geq 1$ such that for all $\lambda \geq \lambda_1$,

$$\|u\|_\infty + \|\nabla u\|_\infty \leq c\lambda^{\frac{1}{2}(\frac{d}{p}-1)}. \tag{7.23}$$

Define

$$\Phi(x) := x + u(x).$$

By (7.23) with λ large enough, the map $x \to \Phi(x)$ forms a C^1-diffeomorphism and

$$1/2 \leq \|\nabla\Phi\|_\infty, \big\|\nabla\Phi^{-1}\big\|_\infty \leq 2,$$

where Φ^{-1} is the inverse of Φ.

By Lemma 7.5 and Theorem 7.6, the following result can be shown in the same way as in Lemma 6.2. We omit the details.

Lemma 7.7. *X_t solves SDE (2.10) if and only if $Y_t := \Phi(X_t)$ solves*

$$\mathrm{d}Y_t = \tilde{\sigma}(Y_t)\,\mathrm{d}W_t + \tilde{b}(Y_t)\,\mathrm{d}t + \int_{|z|<R} \tilde{g}(Y_{t-}, z)\tilde{N}(\mathrm{d}t, \mathrm{d}z) + \int_{|z|\geq R} \tilde{g}(Y_{t-}, z)N(\mathrm{d}t, \mathrm{d}z), \tag{7.24}$$

where $y := \Phi(x)$ and

$$\tilde{\sigma}(y) := (\nabla\Phi \cdot \sigma) \circ \Phi^{-1}(y), \qquad \tilde{b}(y) := (\lambda u + \nabla\Phi \cdot b_2) \circ \Phi^{-1}(y),$$
$$\tilde{g}(y, z) := \Phi\big(\Phi^{-1}(y) + g\big(\Phi^{-1}(y), z\big)\big) - y.$$

The following proposition is the key observation, which shows that the dissipativity (2.11) is preserved under Zvonkin's transformation.

Proposition 7.8. *Under (2.11), for λ large enough, there are $\tilde{\kappa}_1, \tilde{\kappa}_2, \tilde{\kappa}_3 > 0$ such that for all $y \in \mathbb{R}^d$,*

$$\big\langle y, \tilde{b}(y)\big\rangle \leq -\tilde{\kappa}_1 |y|^{2+r} + \tilde{\kappa}_2 \quad \textit{and} \quad \big|\tilde{b}(y)\big| \leq \tilde{\kappa}_3\big(1 + |y|^{1+r}\big).$$

Proof. Noticing that

$$y = \Phi^{-1}(y) + u\big(\Phi^{-1}(y)\big), \qquad \nabla\Phi(x) = \mathbb{I} + \nabla u(x),$$

by the definition of $\tilde{b}$ and (2.11), we have

$$\begin{aligned}
\big\langle y, \tilde{b}(y)\big\rangle &= \lambda\big\langle y, u\big(\Phi^{-1}(y)\big)\big\rangle + \big\langle y, b_2\big(\Phi^{-1}(y)\big)\big\rangle + \big\langle y, (b_2\nabla u)\big(\Phi^{-1}(y)\big)\big\rangle \\
&\leq \lambda\|u\|_\infty |y| + \big\langle \Phi^{-1}(y), b_2\big(\Phi^{-1}(y)\big)\big\rangle + \|u\|_\infty \cdot \big|b_2\big(\Phi^{-1}(y)\big)\big| \\
&\quad + \|\nabla u\|_\infty |y| \cdot \big|b_2\big(\Phi^{-1}(y)\big)\big| \\
&\leq \lambda\|u\|_\infty |y| - \kappa_1 \big|\Phi^{-1}(y)\big|^{2+r} + \kappa_2 \\
&\quad + \kappa_3\big(1 + \big|\Phi^{-1}(y)\big|^2\big)^{(1+r)/2}\big(\|u\|_\infty + \|\nabla u\|_\infty |y|\big) \\
&\leq \lambda\|u\|_\infty |y| - \kappa_1\big(|y| - \|u\|_\infty\big)^{2+r} + \kappa_2 \\
&\quad + \kappa_3\big(1 + \big(|y| + \|u\|_\infty\big)^2\big)^{(1+r)/2}\big(\|u\|_\infty + \|\nabla u\|_\infty |y|\big) \\
&\leq \left(c_1\|\nabla u\|_\infty - \frac{\kappa_1}{2}\right)|y|^{2+r} + c_\lambda,
\end{aligned}$$

where c_1 only depends on κ_3 and r. By (7.23) with λ large enough so that $c_1\|\nabla u\|_\infty \leq \frac{\kappa_1}{4}$, we get the first estimate. The second estimate is easy. □

Now we can give

Proof of Theorem 2.9. For the first part of the result, by Proposition 2.8, we only need to prove that the the conclusions in Theorem 7.4 hold for SDE (7.24). To this end, it suffices to check that the new coefficients of SDE (7.24) satisfy the requirement in Theorem 7.4. The fact that $\tilde{\sigma}$, $\tilde{b}$, $\tilde{g}$ satisfy the local condition and (7.2) can be proved by direct computations, we omit the details. Since $\tilde{\sigma}$ is bounded and by Proposition 7.8, we have (7.1) is true. The desired result follows. To finish the proof, it remains to show that the invariant probability measure μ has a density $\rho \in L^q(\mathbb{R}^d)$ with $q < d/(d-1)$. By Zvonkin's transformation Lemma 7.7, we may assume $b_1 = 0$. Let $f \in C_0^\infty(\mathbb{R}^d)$, and for $p > d$ let $u \in H_p^2$ solve the following elliptic equation:

$$\big(\mathscr{L}_2^\sigma - \lambda\big)u + \mathscr{L}_{\nu,R}^g u = f.$$

Let $u_n = u * \phi_n$ be the mollifying approximation of u and define

$$f_n := \big(\mathscr{L}_2^\sigma - \lambda\big)u_n + \mathscr{L}_{\nu,R}^g u_n.$$

By Itô's formula, we have

$$\mathbb{E}u_n(X_T)=u_n(x)+\mathbb{E}\left(\int_0^T (f_n+\lambda u_n+b\cdot\nabla u_n)(X_t)\,\mathrm{d}t\right).$$

Noticing that

$$\|f_n-f\|_p\le c\|u_n-u\|_{2,p},$$

by Krylov's estimate (7.16) and (2.11) we have

$$\begin{aligned}\mathbb{E}\left(\int_0^T f(X_t)\,\mathrm{d}t\right)&=\lim_{n\to\infty}\mathbb{E}\left(\int_0^T f_n(X_t)\,\mathrm{d}t\right)\\&\le(\lambda+2)\|u_n\|_\infty+\|\nabla u_n\|_\infty\mathbb{E}\left(\int_0^T |b|(X_t)\,\mathrm{d}t\right)\\&\le(\lambda+2)\|u\|_\infty+\kappa_3\|\nabla u\|_\infty\mathbb{E}\left(\int_0^T \left(1+|X_t|^{1+r}\right)\mathrm{d}t\right),\end{aligned}$$

which yields by (7.3) and (7.21) that

$$\mathbb{E}\left(\int_0^T f(X_t)\,\mathrm{d}t\right)\le c\big(1+|x|+T\big)\|f\|_p,$$

where c is independent of T and x. By (2.9) we get for any $p>d$,

$$\mu(f)\le c\|f\|_p,\quad f\in C_0^\infty\big(\mathbb{R}^d\big),$$

which in turn implies by Riesz's representation theorem that μ has a density $\rho\in L^{p/(p-1)}(\mathbb{R}^d)$. The proof is complete. □

The proof of Theorem 2.12 is entirely similar to Theorem 2.9. As in Lemma 7.5, the following lemma can be proven by Lemma 5.9.

Lemma 7.9. *Under* $(\mathbf{H}^g_\beta)$, $(\mathbf{H}^b)$ *and* (2.12), *any solution* $X_t(x)$ *to SDE* (2.10) *does not explode. Moreover, for any* $T>0$ *and* $f\in L^{p'}(\mathbb{R}^d)$ *with* $p'>d/(\alpha-1)$,

$$\mathbb{E}\left(\int_0^T f\big(X_s(x)\big)\,\mathrm{d}s\right)\le c\big(|x|+1\big)\|f\|_{p'},$$

where $c>0$ *is independent of* x.

We also have the following solvability of nonlocal elliptic equations.

Theorem 7.10. *Let* $\alpha\in(1,2)$ *and* $\mathscr{L}^\kappa_\alpha$ *be defined by* (4.13), *where* κ *satisfies* (4.14). *Let* $\vartheta\in[\alpha,(\alpha+2)\wedge 2)$ *and* $\theta\in(\vartheta-\alpha,(\vartheta-1)\wedge\beta)$. *Suppose that* $b=b_1+b_2$ *with* $b_1\in H^\theta_p(\mathbb{R}^d)$ *for some* $p>2d/\alpha$. *Then for some* $\lambda_1\ge 1$ *large enough and for all* $\lambda\ge\lambda_1$ *and* $f\in H^\theta_p(\mathbb{R}^d)$, *there exists a unique solution* $u\in H^\vartheta_p$ *to the following nonlocal elliptic equation*:

$$\big(\mathscr{L}^\kappa_\alpha-\lambda\big)u+\mathscr{K}u+\mathscr{L}^b_1u=f,\tag{7.25}$$

so that

$$\lambda^{1-\frac{\vartheta-\theta}{\alpha}}\|u\|_{\vartheta,p}\le c\|f\|_{\theta,p}.\tag{7.26}$$

Moreover, for any $\gamma\in[0,\alpha)$ *and* $p'\in[p,\infty]$ *with* $\frac{d}{p}<\alpha-\gamma+\frac{d}{p'}$,

$$\lambda^{\frac{1}{\alpha}(\alpha-\gamma+\frac{d}{p'}-\frac{d}{p})}\|u\|_{\gamma,p'}\le c\|f\|_p.\tag{7.27}$$

Proof. We show the apriori estimate (7.26) and (7.27). Suppose $u \in H^\vartheta_p$ satisfies (7.25). Let $T > 0$ and $\phi(t)$ be a non-negative and nonzero smooth function with support in $(0, T)$. Let $\bar u(t, x) := u(x)\phi(t)$. Then

$$\partial_t \bar u + \big(\mathcal{L}^g_\nu - \lambda + \mathcal{K} + \mathcal{L}^b_1\big)\bar u = u\phi' + f\phi.$$

By Theorem 4.5, we have

$$\lambda^{1-\frac{\vartheta-\theta}{\alpha}} \|\bar u\|_{\mathbb{H}^{\vartheta,\infty}_p(T)} \le c\big\|u\phi' + f\phi\big\|_{\mathbb{H}^{\theta,\infty}_p(T)},$$

which implies that

$$\lambda^{1-\frac{\vartheta-\theta}{\alpha}} \|u\|_{\vartheta,p}\|\phi\|_\infty \le c\big(\|u\|_{\theta,p}\big\|\phi'\big\|_\infty + \|f\|_{\theta,p}\|\phi\|_\infty\big).$$

Letting λ be large enough, we get (7.26). On the other hand, by (4.23) we also have

$$\lambda^{\frac{1}{\alpha}(\alpha-\gamma+\frac{d}{p'}-\frac{d}{p})} \|\bar u\|_{\mathbb{H}^{\gamma,\infty}_{p'}(T)} \le c\big\|u\phi' + f\phi\big\|_{\mathbb{L}^p(T)},$$

which also implies (7.27) as above. □

Below we assume that $\sigma \equiv 0$, $\nu(\mathrm{d}z) = |z|^{-d-\alpha}\,\mathrm{d}z$, $(\mathbf{H}^g_\beta)$ holds with $\beta > 1 - \alpha/2$, and for some $\theta \in (1 - \alpha/2, 1)$ and $p > 2d/\alpha$,

$$(I - \Delta)^{\theta/2} b_1, \big(\Gamma^{1,2}_{0,R}(g)\big)^{1/2}, \Gamma^{1,1}_{R,\infty}(g) \in L^p\big(\mathbb{R}^d\big), \qquad \Gamma^{0,2}_{0,R}(g) \in L^\infty\big(\mathbb{R}^d\big).$$

Consider the following nonlocal elliptic equation system:

$$\big(\mathcal{L}^g_\nu - \lambda\big)u + \mathcal{L}^{\bar b^g}_1 u + \mathcal{L}^{b_1}_1 u = b_1.$$

By (7.27), there are $c, \lambda_1 \ge 1$ such that for all $\lambda \ge \lambda_1$,

$$\|u\|_\infty + \|\nabla u\|_\infty \le c\lambda^{\frac{1}{\alpha}(\frac{d}{p}+1-\alpha)}. \tag{7.28}$$

Define

$$\Phi(x) := x + u(x).$$

By (7.28) with λ large enough, the map $x \to \Phi(x)$ forms a C^1-diffeomorphism and

$$1/2 \le \|\nabla\Phi\|_\infty, \big\|\nabla\Phi^{-1}\big\|_\infty \le 2,$$

where Φ^{-1} is the inverse of Φ.

By Lemma 7.9 and Theorem 7.10, the following result can be shown in the same way as in Lemma 6.2. We omit the details.

Lemma 7.11. *X_t solves SDE (2.10) with $\sigma = 0$ if and only if $Y_t := \Phi(X_t)$ solves*

$$\mathrm{d}Y_t = \tilde b(Y_t)\,\mathrm{d}t + \int_{|z|<R} \tilde g(Y_{t-}, z)\tilde N(\mathrm{d}t, \mathrm{d}z) + \int_{|z|\ge R} \tilde g(Y_{t-}, z) N(\mathrm{d}t, \mathrm{d}z),$$

where $y := \Phi(x)$ and

$$\tilde b(y) := \big(\lambda u - \bar{\mathcal{L}}^g_{\nu,R}\Phi + \nabla\Phi \cdot b_2\big) \circ \Phi^{-1}(y),$$
$$\tilde g(y, z) := \Phi\big(\Phi^{-1}(y) + g\big(\Phi^{-1}(y), z\big)\big) - y.$$

Proof of Theorem 2.12. By Lemma 7.11, Proposition 2.8 and Proposition 7.8, the result follows by Theorem 7.4. As for the conclusion that μ has a density $\rho \in L^q(\mathbb{R}^d)$ with $q < d/(d - \alpha + 1)$, it follows by Theorem 7.10 and the same argument as used in the proof of Theorem 2.9. □

Appendix

A.1. *Maximal functions*

Let f be a locally integrable function on $\mathbb{R}^d$. The Hardy–Littlewood maximal function of f is defined by

$$\mathcal{M}f(x) := \sup_{r>0} ⨍_{B_r} \big|f(x+y)\big|\,\mathrm{d}y,$$

where $⨍_{B_r} := \frac{1}{|B_r|}\int_{B_r}$ and $|B_r|$ denotes the Lebesgue measure of ball $B_r := \{x : |x| < r\}$. We have

Lemma A.1.

(i) *Let $\mathbb{B}$ be a Banach space and $f : \mathbb{R}^d \to \mathbb{B}$ a locally integrable function with $\nabla f \in L^1_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{B}^d)$. There is a Lebesgue zero set E such that for all $x, y \notin E$,*

$$\big\|f(x) - f(y)\big\|_{\mathbb{B}} \le 2^d \int_0^{|x-y|} ⨍_{B_s} \big[\|\nabla f\|_{\mathbb{B}}(x+w) + \|\nabla f\|_{\mathbb{B}}(y+w)\big]\,\mathrm{d}w\,\mathrm{d}s. \tag{A.1}$$

In particular, if $\nabla f \in L^p(\mathbb{R}^d; \mathbb{B}^d)$ for some $p > d$, then

$$\big\|f(x) - f(y)\big\|_{\mathbb{B}} \le c_{d,p}|x-y|^{1-d/p}\|\nabla f\|_p. \tag{A.2}$$

(ii) *For $p \in (1, \infty]$, there is a constant $c_{d,p} > 0$ such that for all $f \in L^p(\mathbb{R}^d)$,*

$$\|\mathcal{M}f\|_p \le c_{d,p}\|f\|_p. \tag{A.3}$$

Proof. The estimates (A.1) and (A.3) can be found in [55, Lemma 5.4] and [46, p. 5, Theorem 1]. We prove (A.2). For $\alpha \in (\frac{d}{p} - \frac{1}{p}, 1 - \frac{1}{p})$, by Hölder's inequality, we have

$$\begin{aligned}\int_0^{|x-y|} ⨍_{B_s} \|\nabla f\|_{\mathbb{B}}(x+w)\,\mathrm{d}w\,\mathrm{d}s &\le \left(\int_0^{|x-y|} s^{-\alpha p^*}\,\mathrm{d}s\right)^{\frac{1}{p^*}} \left(\int_0^{|x-y|} s^{\alpha p} ⨍_{B_s} \|\nabla f\|^p_{\mathbb{B}}(x+w)\,\mathrm{d}w\,\mathrm{d}s\right)^{\frac{1}{p}} \\ &\lesssim \left(\int_0^{|x-y|} s^{-\alpha p^*}\,\mathrm{d}s\right)^{\frac{1}{p^*}} \left(\int_0^{|x-y|} s^{\alpha p - d}\,\mathrm{d}s\right)^{\frac{1}{p}} \|\nabla f\|_p \\ &\lesssim |x-y|^{1-d/p}\|\nabla f\|_p.\end{aligned}$$

Substituting this into (A.1), we obtain (A.2). □

For $p, q, r \in [1, \infty]$ and $T > 0$, let $L^r(T) := L^r([0, T])$ and define

$$H^1_q\big(\mathbb{R}^d; L^p\big(\Omega; L^r(T)\big)\big) := \big\{f(x, \omega, t) : f, \nabla f \in L^q\big(\mathbb{R}^d; L^p\big(\Omega; L^r(T)\big)\big)\big\},$$

and

$$\|f\|_{H^1_q(\mathbb{R}^d;L^p(\Omega;L^r(T)))} := \|f\|_{L^q(\mathbb{R}^d;L^p(\Omega;L^r(T)))} + \|\nabla f\|_{L^q(\mathbb{R}^d;L^p(\Omega;L^r(T)))}.$$

The following characterization about the Sobolev differentiability of random fields can be found in [53, Theorem 1.1], which is used to prove the Sobolev differentiability of the strong solution to SDEs with respect to the initial value.

Theorem A.2. *Let $f \in L^q(\mathbb{R}^d; L^p(\Omega; L^r(T)))$ for some $p \in (1, \infty)$ and $q, r \in (1, \infty]$. Then $f \in H^1_q(\mathbb{R}^d; L^p(\Omega; L^r(T)))$ if and only if there exists a nonnegative function $g \in L^q(\mathbb{R}^d)$ such that for Lebesgue-almost all $x, y \in \mathbb{R}^d$,*

$$\big\|f(x, \cdot) - f(y, \cdot)\big\|_{L^p(\Omega;L^r(T))} \le |x-y|\big(g(x) + g(y)\big). \tag{A.4}$$

Moreover, if (A.4) holds, then for Lebesgue-almost all $x \in \mathbb{R}^d$,

$$\big\|\partial_i f(x, \cdot)\big\|_{L^p(\Omega;L^r(T))} \le 2g(x), \quad i = 1, \ldots, d,$$

where $\partial_i f$ is the weak partial derivative of f with respect to the i-th spacial variable.

A.2. Proofs of Lemmas 4.4, 6.1 *and* 6.4

Proof of Lemma 4.4. Let $p_0 := dp_1/(d - p_1\gamma_1) \geq p$. First of all, by Hölder's inequality and Sobolev's embedding (4.2), we have

$$\|fg\|_p \leq \|f\|_{p_0}\|g\|_{pp_0/(p_0-p)} \lesssim \|f\|_{\gamma_1,p_1}\|g\|_{\alpha+\gamma_2,p_2}. \tag{A.5}$$

Notice that by (4.1),

$$\Delta^{\alpha/2}(fg) = \int_{\mathbb{R}^d} \frac{(f(\cdot+y)-f(\cdot))(g(\cdot+y)-g(\cdot))}{|y|^{d+\alpha}}\,\mathrm{d}y + \big(\Delta^{\alpha/2}f\big)g + f\Delta^{\alpha/2}g.$$

Hence,

$$\begin{aligned}\big\|\Delta^{\alpha/2}(fg)\big\|_p \leq{}& \int_{\mathbb{R}^d} \frac{\|(f(\cdot+y)-f(\cdot))(g(\cdot+y)-g(\cdot))\|_p}{|y|^{d+\alpha}}\,\mathrm{d}y \\ &+ \big\|\big(\Delta^{\alpha/2}f\big)g\big\|_p + \big\|f\Delta^{\alpha/2}g\big\|_p.\end{aligned} \tag{A.6}$$

As above, by Hölder's inequality and Sobolev's embedding, we have

$$\big\|\big(\Delta^{\alpha/2}f\big)g\big\|_p \leq \|f\|_{\alpha,p_0}\|g\|_{p_0p/(p_0-p)} \lesssim \|f\|_{\alpha+\gamma_1,p_1}\|g\|_{\alpha+\gamma_2,p_2}, \tag{A.7}$$

and by symmetry,

$$\big\|f\big(\Delta^{\alpha/2}g\big)\big\|_p \lesssim \|f\|_{\alpha+\gamma_1,p_1}\|g\|_{\alpha+\gamma_2,p_2}.$$

Moreover, for $\varepsilon \in (0, \gamma_1+\gamma_2+\alpha-\frac{d}{p_1}-\frac{d}{p_2}+\frac{d}{p})$, by Hölder's inequality, Sobolev's embedding and (4.3), we have

$$\begin{aligned}&\big\|\big(f(\cdot+y)-f(\cdot)\big)\big(g(\cdot+y)-g(\cdot)\big)\big\|_p \\ &\quad\leq \big\|f(\cdot+y)-f(\cdot)\big\|_{p_0}\big\|g(\cdot+y)-g(\cdot)\big\|_{pp_0/(p_0-p)} \\ &\quad\lesssim \big\|f(\cdot+y)-f(\cdot)\big\|_{\gamma_1,p_1}\big\|g(\cdot+y)-g(\cdot)\big\|_{\alpha+\gamma_2-\varepsilon,p_2} \\ &\quad\lesssim \big(\big(|y|^\alpha\|f\|_{\alpha+\gamma_1,p_1}\big)\wedge\big(2\|f\|_{\gamma_1,p_1}\big)\big)\big(\big(|y|^\varepsilon\|g\|_{\alpha+\gamma_2,p_2}\big)\wedge\big(2\|g\|_{\alpha+\gamma_2-\varepsilon,p_2}\big)\big) \\ &\quad\lesssim \big(|y|^{\alpha+\varepsilon}\wedge 1\big)\|f\|_{\alpha+\gamma_1,p_1}\|g\|_{\alpha+\gamma_2,p_2}.\end{aligned} \tag{A.8}$$

Substituting (A.7)–(A.8) into (A.6), we obtain

$$\big\|\Delta^{\alpha/2}(fg)\big\|_p \lesssim \|f\|_{\alpha+\gamma_1,p_1}\|g\|_{\alpha+\gamma_2,p_2},$$

which together with (A.5) yields the desired estimate. □

Proof of Lemma 6.1. (i) We only show $(\Gamma^{1,2}_{0,R}(\tilde{g}))^{1/2} \in \mathbb{L}^q_p(T)$. The others are direct by definition. Let

$$\bar{g}_t(x,z) := \Phi_t\big(x+g_t(x,z)\big) - \Phi_t(x). \tag{A.9}$$

By (6.3), it suffices to show $(\Gamma^{1,2}_{0,R}(\bar{g}))^{1/2} \in \mathbb{L}^q_p(T)$. Noticing that

$$\begin{aligned}\big|\nabla_x\bar{g}_t(x,z)\big| &= \big|(\nabla\Phi_t)\big(x+g_t(x,z)\big)\cdot\big(\mathbb{I}+\nabla_x g_t(x,z)\big)-\nabla\Phi_t(x)\big| \\ &\leq \sup_y |y|^{-1}\big|(\nabla\Phi_t)(x+y)-\nabla\Phi_t(x)\big|\cdot\big|g_t(x,z)\big| + 2\big|\nabla_x g_t(x,z)\big|,\end{aligned}$$

in view of $p > d$, by Lemma 4.1, we have

$$\big\|\big(\Gamma^{1,2}_{0,R}(\bar{g})\big)^{1/2}\big\|_{\mathbb{L}^q_p(T)} \leq \Big\|\sup_y |y|^{-1}\big|\nabla\Phi(\cdot+y)-\nabla\Phi(\cdot)\big|\Big\|_{\mathbb{L}^q_p(T)}$$

$$\times \big\| \big(\Gamma^{0,2}_{0,R}(g)\big)^{1/2} \big\|_{\mathbb{L}^\infty(T)} + 2\big\| \big(\Gamma^{1,2}_{0,R}(g)\big)^{1/2} \big\|_{\mathbb{L}^q_p(T)}$$

$$\lesssim \|\nabla u\|_{\mathbb{H}^{1,q}_p(T)} \big\|\Gamma^{0,2}_{0,R}(g)\big\|^{1/2}_{\mathbb{L}^\infty(T)} + 2\big\| \big(\Gamma^{1,2}_{0,R}(g)\big)^{1/2} \big\|_{\mathbb{L}^q_p(T)}.$$

(ii) By (6.1) and easy calculations, we have

$$\big(\partial_s + \mathscr{L}^\sigma_2 + \mathscr{L}^b_1 + \mathscr{L}^g_{\nu,R}\big)\Phi = \lambda u.$$

Hence, as in (4.11), one sees that as $n \to \infty$,

$$\big\| \big(\partial_s + \mathscr{L}^\sigma_2 + \mathscr{L}^b_1 + \mathscr{L}^g_{\nu,R}\big)\Phi_n - \lambda u \big\|_{\mathbb{L}^q_p(T)} \lesssim \|\Phi_n - \Phi\|_{\mathbb{H}^{2,q}_p(T)} = \|u_n - u\|_{\mathbb{H}^{2,q}_p(T)} \to 0.$$

(iii) Notice that in the generalized sense,

$$\partial_s \Phi^{-1} = -\nabla\Phi^{-1} \cdot (\partial_s \Phi) \circ \Phi^{-1}, \qquad \nabla\Phi^{-1} = (\nabla\Phi)^{-1} \circ \Phi^{-1}$$

$$\nabla^2\Phi^{-1} = -\big[(\nabla\Phi)^{-1} \cdot \nabla^2\Phi \cdot (\nabla\Phi)^{-1}\big] \circ \Phi^{-1} \cdot \nabla\Phi^{-1}.$$

By cumbersome calculations (see [54,55]), we have

$$\big(\partial_s + \mathscr{L}^{\tilde\sigma}_2 + \mathscr{L}^{\tilde b}_1 + \mathscr{L}^{\tilde g}_{\nu,R}\big)\Phi^{-1} = b \circ \Phi^{-1}.$$

The limit in (iii) now follows by (6.4). □

Proof of Lemma 6.4. (i) It is clear that $\tilde b \in \mathbb{H}^{1,\infty}_\infty(T)$ by definition. Let $\bar g$ be defined as in (A.9). We show that $\bar g$ satisfies $(\mathbf{H}^g_\beta)$ and $(\Gamma^{1,2}_{0,R}(\bar g))^{1/2} \in \mathbb{L}^q_p(T)$. Clearly, $\bar g_t(x,0) = 0$ by $g_t(x,0) = 0$, and by (6.2),

$$(2c_1)^{-1}\big|z - z'\big| \le \big|\bar g_t(x,z) - \bar g_t\big(x,z'\big)\big| \le 2c_1\big|z - z'\big|.$$

Moreover, notice that

$$\nabla_z \bar g_t(x,z) = (\nabla\Phi_t)\big(x + g_t(x,z)\big) \cdot \nabla_z g_t(x,z).$$

Since g satisfies $(\mathbf{H}^g_\beta)$, by (6.3) and (6.16), it is easy to see that $\bar g$ also satisfies $(\mathbf{H}^g_\beta)$, and so does $\tilde g$. On the other hand, note that

$$\big|\nabla_x \bar g_t(x,z)\big| \le \big|\nabla\Phi_t\big(x + g(x,z)\big) - \nabla\Phi_t(x)\big| + 2\big|\nabla_x g_t(x,z)\big|$$

$$\le U_t(x)\big|g(x,z)\big|^{\vartheta-1} + 2\big|\nabla_x g_t(x,z)\big|,$$

where

$$U_t(x) := \sup_y |y|^{1-\vartheta}\big|\nabla u_t(x+y) - \nabla u_t(x)\big|.$$

We have by $(\mathbf{H}^g_\beta)$,

$$\Gamma^{1,2}_{0,R}(\bar g)(x) = \int_{|z|<R} \big|\nabla \bar g_t(x,z)\big|^2 \nu(\mathrm{d}z) \lesssim U^2(x) \int_{|z|<R} |z|^{2(\vartheta-1)-d-\alpha}\,\mathrm{d}z + \Gamma^{1,2}_{0,R}(g)(x).$$

By Lemma 4.1, since $p(\vartheta - 1) > d$, we have

$$\|U\|_{\mathbb{L}^q_p(T)} \lesssim \|\nabla u\|_{\mathbb{H}^{\vartheta-1,q}_p(T)} \le \|u\|_{\mathbb{H}^{\vartheta,q}_p(T)}.$$

Since $\vartheta > 1 + \frac{\alpha}{2}$, we get $(\Gamma^{1,2}_{0,R}(\bar g))^{1/2} \in \mathbb{L}^q_p(T)$ and so $(\Gamma^{1,2}_{0,R}(\tilde g))^{1/2} \in \mathbb{L}^q_p(T)$.
(ii) and (iii) are the same as in the proof of Lemma 6.1. □

A.3. Positivity of Dirichlet heat kernel

Let $\rho(t,x,y)$ be a family of jointly continuous transition probability density functions in $\mathbb{R}^d$. Let $(X,\mathbb{P}_x)_{x\in\mathbb{R}^d}$ be the associated homogeneous Markov processes, that is, $\mathbb{P}_x(X_0=x)=1$ and for any $t>0$,

$$\int_A \rho(t,x,y)\,\mathrm{d}y=\mathbb{P}_x(X_t\in A),\quad A\in\mathscr{B}(\mathbb{R}^d).$$

Let D be a domain (bounded open subset of $\mathbb{R}^d$), and $\tau_D:=\{t>0:X_t\notin D\}$ be the exit time of X from D. Let X^D be the killed Markov process outside D, and P^D_t the transition probability of X^D, that is,

$$P^D_t(x,A):=\mathbb{P}_x(t<\tau_D;X_t\in A),\quad A\in\mathscr{B}(D). \tag{A.10}$$

Define the Dirichlet heat kernel by

$$\rho^D(t,x,y):=\rho(t,x,y)-r^D(t,x,y),$$

where

$$r^D(t,x,y):=\mathbb{E}^x\big[\tau_D<t;\rho\big(t-\tau_D;X(\tau_D),y\big)\big].$$

Let $\varrho_i(t,r):\mathbb{R}_+\times\mathbb{R}_+\to\mathbb{R}_+$, $i=1,2$ be two continuous functions and satisfy that

($\mathbf{H}^\varrho$) For each $t>0$, the map $r\mapsto\varrho_i(t,r)$ is decreasing, and for each $\delta>0$,

$$\sup_{t>0,r>\delta}\varrho_2(t,r)<\infty,$$

and there are $t_0=t_0(\delta)$ and $R=R(\delta)>0$ such that $t\mapsto\varrho_i(t,\delta)$ is increasing on $(0,t_0)$ and

$$\varrho_1(t,\delta/R)>\varrho_2(t,\delta),\quad t\in(0,t_0).$$

The following result is essentially due to Hunt (cf. [16, Theorem 2.4]).

Theorem A.3. *Let ϱ_1 and ϱ_2 satisfy* ($\mathbf{H}^\varrho$). *Suppose that*

$$\varrho_1\big(t,|x-y|\big)\le\rho(t,x,y)\le\varrho_2\big(t,|x-y|\big). \tag{A.11}$$

Then ρ^D is the transition probability density function of X^D, i.e., for any $t>0$,

$$P^D_t(x,A)=\int_A\rho^D(t,x,y)\,\mathrm{d}y,\quad x\in\mathbb{R}^d,A\in\mathscr{B}(D).$$

Moreover, ρ^D is continuous and strictly positive on $\mathbb{R}_+\times D\times D$ and for $0<s<t<\infty$ and $x,y\in\mathbb{R}^d$,

$$\rho^D(t,x,y)=\int_D\rho^D(s,x,z)\rho^D(t-s,z,y)\,\mathrm{d}z. \tag{A.12}$$

Proof. We only show the strict positivity of $\rho^D(t,x,y)$. The others are completely same as in [16, Theorem 2.4]. Fix $x,y\in D$ and let $d(y,\partial D)$ be the distance of y to the boundary ∂D. Let $\delta\in(0,d(y,\partial D))$ be given. By the assumption on ϱ_i, there are $t_0=t_0(\delta)>0$ and $R>0$ such that

$$\varrho_1(t,\delta/R)>\varrho_2(t,\delta),\quad t\in(0,t_0).$$

Hence, by the definition of r^D, (A.11) and the assumptions of ϱ_2, we have for $t\in(0,t_0)$,

$$r^D(t,x,y)\le\mathbb{E}^x\big[\tau_D<t;\varrho_2\big(t-\tau_D,\big|X(\tau_D)-y\big|\big)\big]\le\varrho_2(t,\delta).$$

Consequently, if $|x-y|\le\delta/R\le\delta<\rho(y,\partial D)$, then

$$\rho^D(t,x,y)\ge\varrho_1\big(t,|x-y|\big)-\varrho_2(t,\delta)\ge\varrho_1(t,\delta/R)-\varrho_2(t,\delta)>0. \tag{A.13}$$

Now for any $t > 0$ and $x, y \in D$. Let Γ be a curve in D connecting x and y. Let $\delta := \rho(\Gamma, \partial D)$. Let n be large enough such that $t \le nt_0(\delta)$ and there are points $a_0, a_1, \ldots, a_{n+1}$ on Γ with $a_0 = x$, $a_{n+1} = y$ and $a_i \in B(a_{i-1}, \delta/(3R))$. Notice that for $x_{i-1} \in B(a_{i-1}, \delta/(3R))$ and $x_i \in B(a_i, \delta/(3R))$,

$$|x_i - x_{i-1}| \le |x_i - a_i| + |a_{i-1} - x_{i-1}| + |a_i - a_{i-1}| \le \delta/R.$$

By C-K equation (A.12) and (A.13), we have

$$\begin{aligned}\rho^D(t,x,y) &= \int_D \cdots \int_D \rho^D\left(\frac{t}{n}, x, x_1\right) \cdots \rho^D\left(\frac{t}{n}, x_n, y\right) \mathrm{d}x_1 \cdots \mathrm{d}x_n \\ &\ge \int_{B(a_1,\delta/(3R))} \cdots \int_{B(a_n,\delta/(3R))} \rho^D\left(\frac{t}{n}, x, x_1\right) \cdots \rho^D\left(\frac{t}{n}, x_n, y\right) \mathrm{d}x_1 \cdots \mathrm{d}x_n > 0.\end{aligned}$$

The proof is complete. □

Acknowledgements

The authors would like to thank Zhen-Qing Chen, Zenghu Li, Renming Song, Feng-Yu Wang, Yinchao Xie and Guohuan Zhao for their quite useful conversations. The referee's useful suggestions are also acknowledged.

References

[1] H. Abels and M. Kassmann. The Cauchy problem and the martingale problem for integro-differential operators with non-smooth kernels. *Osaka J. Math.* **46** (2009) 661–683. MR2583323

[2] A. Arapostathis, A. Biswas and L. Caffarelli. The Dirichlet problem for stable-like operators and related probabilistic representations. *Comm. Partial Differential Equations* **41** (9) (2016) 1472–1511. MR3551465 https://doi.org/10.1080/03605302.2016.1207084

[3] R. F. Bass, K. Burdzy and Z. Chen. Stochastic differential equations driven by stable processes for which pathwise uniqueness fails. *Stochastic Process. Appl.* **111** (2004) 1–15. MR2049566 https://doi.org/10.1016/j.spa.2004.01.010

[4] J. Bergh and J. Löfström. *An Introduction to Interpolation Spaces*. Springer-Verlag, Berlin, 1970. MR0482275

[5] V. I. Bogachev, N. V. Krylov and M. Röckner. On regularity of transition probabilities and invariant measures of singular diffusions under minimal conditions. *Comm. Partial Differential Equations* **26** (11–12) (2001) 2037–2080. MR1876411 https://doi.org/10.1081/PDE-100107815

[6] V. I. Bogachev, N. V. Krylov and M. Röckner. Elliptic and parabolic equations for measures. *Uspekhi Mat. Nauk* **64** (6) (2009) 5–116 [in Russian]. English transl.: *Russian Math. Surveys* **64** (6) (2009) 973–1078. MR2640966 https://doi.org/10.1070/RM2009v064n06ABEH004652

[7] V. I. Bogachev, G. D. Prato and M. Röckner. Existence of solutions to weak parabolic equations for measures. *Proc. Lond. Math. Soc.* **88** (2004) 753–774. MR2044056 https://doi.org/10.1112/S0024611503014540

[8] V. I. Bogachev, M. Röckner and S. V. Shaposhnikov. On parabolic inequalities for generators of diffusions with jumps. *Probab. Theory Related Fields* **158** (2014) 465–476. MR3152788 https://doi.org/10.1007/s00440-013-0485-0

[9] V. I. Bogachev and P. A. Yu. Strong solutions of stochastic equations with Lévy noise and a discontinuous drift coefficient. *Dokl. Math.* **92** (1) (2015) 471–475. MR3444003 https://doi.org/10.1134/s1064562415040213

[10] V. I. Bogachev and P. A. Yu. Strong solutions to stochastic equations with a Lévy noise and a non-constant diffusion coefficient. *Dokl. Math.* **94** (1) (2016) 438–440. MR3561342 https://doi.org/10.1134/s1064562416040244

[11] Z. Chen, E. Hu, L. Xie and X. Zhang. Heat kernels for non-symmetric diffusions operators with jumps. *J. Differential Equations* **263** (2017) 6576–6634. MR3693184 https://doi.org/10.1016/j.jde.2017.07.023

[12] Z. Chen, R. Song and X. Zhang. Stochastic flows for Lévy processes with Hölder drift. *Rev. Mat. Iberoam.* **34** (2018) 1755–1788. MR3896248 https://doi.org/10.4171/rmi/1042

[13] Z. Chen and X. Zhang. Heat kernels and analyticity of non-symmetric jump diffusion semigroups. *Probab. Theory Related Fields* **165** (2016) 267–312. MR3500272 https://doi.org/10.1007/s00440-015-0631-y

[14] Z. Chen and X. Zhang. Heat kernels for time-dependent non-symmetric stable-like operators. *J. Math. Anal. Appl.* **465** (2018) 1–21. MR3806688 https://doi.org/10.1016/j.jmaa.2018.03.054

[15] Z. Chen and X. Zhang. Uniqueness of stable like processes. Available at arXiv:1604.02681.

[16] K. L. Chung and Z. Zhao. *From Brownian Motion to Schrödinger's Equation*. Springer-Verlag, Berlin, 1995. MR1329992 https://doi.org/10.1007/978-3-642-57856-4

[17] S. D. Eidelman, S. D. Ivasyshen and A. N. Kochubei. *Analytic Methods in the Theory of Differential and Pseudo-Differential Equations of Parabolic Type*. Birkhäuser, Basel, 2004. MR2093219 https://doi.org/10.1007/978-3-0348-7844-9

[18] E. Fedrizzi and F. Flandoli. Pathwise uniqueness and continuous dependence of SDEs with non-regular drift. *Stochastics* **83** (3) (2011) 241–257. MR2810591 https://doi.org/10.1080/17442508.2011.553681

[19] E. Fedrizzi and F. Flandoli. Hölder flow and differentiability for SDEs with nonregular drift. *Stoch. Anal. Appl.* **31** (2013) 708–736. MR3175794 https://doi.org/10.1080/07362994.2012.628908

[20] F. Flandoli, M. Gubinelli and E. Priola. Well-posedness of the transport equation by stochastic perturbation. *Invent. Math.* **180** (1) (2010) 1–53. MR2593276 https://doi.org/10.1007/s00222-009-0224-4

[21] B. Goldys and B. Maslowski. Exponential ergodicity for stochastic reaction-diffusion equations. In *Stochastic Partial Differential Equations and Applications, XVII* 115–131. *Lect. Notes Pure Appl. Math.* **245**. Chapman Hall/CRC, Boca Raton, FL, 2006. MR2227225 https://doi.org/10.1201/9781420028720.ch12

[22] I. Gyöngy and T. Martinez. On stochastic differential equations with locally unbounded drift. *Czechoslovak Math. J.* **51** (4) (2001) 763–783. MR1864041 https://doi.org/10.1023/A:1013764929351

[23] S. Haadem and F. Proske. On the construction and Malliavin differentiability of solutions of Lévy noise driven SDE's with singular coefficients. *J. Funct. Anal.* **266** (2014) 5321–5359. MR3177338 https://doi.org/10.1016/j.jfa.2014.02.009

[24] M. Hairer. An introduction to stochastic PDEs. Available at http://www.hairer.org/notes/SPDEs.pdf.

[25] R. Z. Hasminskii. *Stochastic Stability of Differential Equations*. Sijthoff & Noordhoff, Rockville, 1980. MR0600653

[26] V. A. Ju. On the strong solutions of stochastic differential equations. *Theory Probab. Appl.* **24** (1979) 354–366. MR0532447

[27] K.-H. Kim. $L_q(L_p)$-Theory of parabolic PDEs with variable coefficients. *Bull. Korean Math. Soc.* **45** (2008) 169–190. MR2391465 https://doi.org/10.4134/BKMS.2008.45.1.169

[28] R. Kruse and M. Scheutzow. A discrete stochastic Gronwall lemma. *Math. Comput. Simulation* **143** (2018) 149–157. MR3698223 https://doi.org/10.1016/j.matcom.2016.07.002

[29] N. V. Krylov. *Controlled Diffusion Processes*. *Applications of Mathematics* **14**. Springer-Verlag, New York, 1980. Translated from the Russian by A. B. Aries. MR0601776

[30] N. V. Krylov. *Nonlinear Elliptic and Parabolic Equations of Second Order*. Nauka, Moscow, 1985. MR0815513

[31] N. V. Krylov and M. Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields* **131** (2) (2005) 154–196. MR2117951 https://doi.org/10.1007/s00440-004-0361-z

[32] A. Kulik. Exponential ergodicity of the solutions to SDE's with a jump noise. *Stochastic Process. Appl.* **119** (2) (2009) 602–632. MR2494006 https://doi.org/10.1016/j.spa.2008.02.006

[33] V. P. Kurenok. Stochastic equations with time-dependent drift driven by Lévy processes. *J. Theoret. Probab.* **20** (2007) 859–869. MR2359059 https://doi.org/10.1007/s10959-007-0086-x

[34] H. Masuda. Ergodicity and exponential β-mixing bounds for multidimensional diffusions with jumps. *Stochastic Process. Appl.* **117** (2007) 35–56. MR2287102 https://doi.org/10.1016/j.spa.2006.04.010

[35] P. O. Menoukeu, B. T. Meyer, T. Nilssen, F. Proske and T. Zhang. A variational approach to the construction and Malliavin differentiability of strong solutions of SDEs. *Math. Ann.* **357** (2013) 761–799. MR3096525 https://doi.org/10.1007/s00208-013-0916-3

[36] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Springer-Verlag, Berlin, 1993. MR1287609 https://doi.org/10.1007/978-1-4471-3267-7

[37] R. Mikulevicius and H. Pragarauskas. On the Cauchy problem for certain integro-differential operators in Sobolev and Hölder spaces. *Lith. Math. J.* **32** (2) (1992) 377–396. MR1246036 https://doi.org/10.1007/BF02450422

[38] S. E. A. Mohammed, T. Nilssen and F. Proske. Sobolev differentiable stochastic flows of SDE's with singular coefficients: Applications to the transport equation. *Ann. Probab.* **43** (3) (2015) 1535–1576. MR3342670 https://doi.org/10.1214/14-AOP909

[39] H. Pragarauskas. On L^p-estimates of stochastic integrals. In *Probab. Theory and Math. Stat* 579–588, 1999.

[40] E. Priola. Pathwise uniqueness for singular SDEs driven by stable processes. *Osaka J. Math.* **49** (2012) 421–447. MR2945756

[41] E. Priola. Stochastic flow for SDEs with jumps and irregular drift term. *Banach Center Publ.* **105** (2015) 193–210. MR3445537 https://doi.org/10.4064/bc105-0-12

[42] E. Priola. Davie's type uniqueness for a class of SDEs with jumps. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** (2018) 694–725. MR3795063 https://doi.org/10.1214/16-AIHP818

[43] J. Ren, J. Wu and X. Zhang. Exponential ergodicity of multi-valued stochastic differential equations. *Bull. Sci. Math.* **134** (2010) 391–404. MR2651898 https://doi.org/10.1016/j.bulsci.2009.01.003

[44] M. Scheutzow. A stochastic Gronwall's lemma. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **16** (2) (2013) 1350019 (4 pages). MR3078830 https://doi.org/10.1142/S0219025713500197

[45] R. Situ. *Theory of Stochastic Differential Equations with Jumps and Applications*. Springer, Berlin, 2005. MR2160585

[46] E. M. Stein. *Singular Integrals and Differentiability Properties of Functions*. *Princeton Mathematical Series* **30**. Princeton University Press, Princeton, NJ, 1970. MR0290095

[47] H. Tanaka, M. Tsuchiya and S. Watanabe. Perturbation of drift-type for Lévy processes. *J. Math. Kyoto Univ.* **14** (1974) 73–92. MR0368146 https://doi.org/10.1215/kjm/1250523280

[48] H. Triebel. *Interpolation Theory, Function Spaces, Differential Operators*. North-Holland Publishing Company, Amsterdam, 1978. MR0503903

[49] F. Y. Wang. Gradient estimates and applications for SDEs in Hilbert space with multiplicative noise and Dini continuous drift. *J. Differential Equations* **3** (2016) 2792–2829. MR3427683 https://doi.org/10.1016/j.jde.2015.10.020

[50] F. Y. Wang. Integrability conditions for SDEs and semi-linear SPDEs. *Ann. Probab.* **45** (2017) 3223–3265. MR3706742 https://doi.org/10.1214/16-AOP1135

[51] F. Y. Wang and X. Zhang. Degenerate SDE with Hölder–Dini drift and non-Lipschitz noise coefficient. *SIAM J. Math. Anal.* **48** (3) (2016) 2189–2222. MR3511355 https://doi.org/10.1137/15M1023671

[52] L. Wang, L. Xie and X. Zhang. Derivative formulae for SDEs driven by multiplicative α-stable-like processes. *Stochastic Process. Appl.* **125** (3) (2015) 867–885. MR3303960 https://doi.org/10.1016/j.spa.2014.10.011

[53] L. Xie and X. Zhang. Sobolev differentiable flows of SDEs with local Sobolev and super-linear growth coefficients. *Ann. Probab.* **44** (6) (2016) 3661–3687. MR3572321 https://doi.org/10.1214/15-AOP1057

[54] X. Zhang. Strong solutions of SDEs with singular drift and Sobolev diffusion coefficients. *Stochastic Process. Appl.* **115** (2005) 1805–1818. MR2172887 https://doi.org/10.1016/j.spa.2005.06.003

[55] X. Zhang. Stochastic homemomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients. *Electron. J. Probab.* **16** (2011) 1096–1116. MR2820071 https://doi.org/10.1214/EJP.v16-887

[56] X. Zhang. Stochastic differential equations with Sobolev drifts and driven by α-stable processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **49** (2013) 915–1231. MR3127913 https://doi.org/10.1214/12-AIHP476

[57] X. Zhang. Stochastic differential equations with Sobolev coefficients and applications. *Ann. Appl. Probab.* **26** (5) (2016) 2697–2732. MR3563191 https://doi.org/10.1214/15-AAP1159

[58] X. Zhang. Multidimensional singular stochastic differential equations. In *Stochastic Partial Differential Equations and Related Fields* 391–403. *Springer Proc. Math. Stat.* **229**. Springer, Cham, 2018. MR3828184