Communications in Mathematics and Statistics
https://doi.org/10.1007/s40304-018-0164-7

Singular Brownian Diffusion Processes

Xicheng Zhang[1] **· Guohuan Zhao**[2]

Received: 31 July 2018 / Revised: 16 October 2018 / Accepted: 24 October 2018

Abstract
In this paper, we survey the recent progress about the SDEs with distributional drifts and generalize some well-known results about the Brownian motion with singular measure-valued drifts. In particular, we show the well-posedness of martingale problem or the existence and uniqueness of weak solutions, and obtain sharp two-sided and gradient estimates of the heat kernel associated with the above SDE. Moreover, we also study the ergodicity and global regularity of the invariant measures of the associated semigroup under some dissipative assumptions.

Keywords Singular drift · Weak solution · Heat kernel · Ergodicity · Zvonkin's transformation

Mathematics Subject Classification 60H10 · 35A08 · 37A25

1 Introduction

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

$$\mathrm{d}X_t = \sigma_t(X_t)\mathrm{d}W_t + b_t(X_t)\mathrm{d}t, \quad X_0 = x \in \mathbb{R}^d, \tag{1.1}$$

Research of X. Zhang is partially supported by NNSFC Grant 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". Research of G. Zhao is partially supported by National Postdoctoral Program for Innovative Talents (BX201600183) of China.

✉ Xicheng Zhang
XichengZhang@gmail.com

Guohuan Zhao
zhaoguohuan@gmail.com

[1] School of Mathematics and Statistics, Wuhan University, Wuhan 430072, Hubei, People's Republic of China

[2] Applied Mathematics, Chinese Academy of Science, Beijing 100081, People's Republic of China

Published online: 10 November 2018

where W is a d-dimensional standard Brownian motion on some complete filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbf{P})$, $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ is a $d \times d$-matrix-valued measurable function, and $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ is a time-dependent measurable vector field. In the theory of SDEs, there are two notions about SDE (1.1): strong solutions and weak solutions. Roughly to say, strong solution means that for given Brownian motion W, let $\mathscr{F}_t := \sigma\{W_s : s \leqslant t\}$ be the natural filtration associated with W, one needs to find an $\mathscr{F}_t$-adapted process X so that the following stochastic Itô's integral equation holds:

$$X_t = x + \int_0^t \sigma_s(X_s)\mathrm{d}W_s + \int_0^t b_s(X_s)\mathrm{d}s, \quad t \geqslant 0.$$

In other words, strong solutions can be regarded as a functional of Brownian path. While, weak solution means that we need to find a pair of processes (X, W) so that W is a Brownian motion and the above stochastic integral equation holds. Clearly, strong solution must be a weak solution. A weak solution is also simply called a solution. Related to these two notions, there are automatically two uniqueness: strong uniqueness and weak uniqueness. Strong uniqueness means that two strong solutions have the same path. Weak uniqueness means that two weak solutions have the same law in the space of continuous function spaces. It should be noticed that if two solutions are defined on the same probability space and their path coincides, we call pathwise uniqueness hold. The celebrated Yamada-Watanabe's theorem [18] tells us that weak existence plus pathwise uniqueness implies the existence and uniqueness of strong solutions.

It is a classical fact that when $b_t(x)$ and $\sigma_t(x)$ are Lipschitz continuous with respect to the spatial variable x and uniformly in t, by Picard's iteration, there is a unique strong solution to SDE (1.1). On the other hand, when σ is bounded continuous and uniformly non-degenerate, and b is bounded measurable, it is also well known that there exists a unique weak solution, or equivalently, the martingale problem associated with (1.1) is well-posed in the sense of Stroock and Varadahan [27]. Now let X be a solution of SDE (1.1) and let $f(t, x)$ be a bounded space-time function so that $\partial_t f, \nabla_x f, \nabla_x^2 f$ are bounded. By Itô's formula, we have

$$\begin{aligned} f(t, X_t) = f(0, x) &+ \int_0^t (\partial_s + \mathscr{L}_s^\sigma + b_s \cdot \nabla) f(s, X_s)\mathrm{d}s \\ &+ \int_0^t \left(\sigma_s^* \cdot \nabla f\right)(s, X_s)\mathrm{d}W_s, \end{aligned}$$

where the asterisk stands for the transpose of a matrix, $\mathscr{L}_t^\sigma$ is the time-dependent second-order differential operator defined by

$$\mathscr{L}_t^\sigma f(x) := \tfrac{1}{2}\left(\sigma_t^{ik}\sigma_t^{jk}\right)(x)\partial_i\partial_j f(x).$$

Here and below, we use the usual Einstein's convention for summation: The same index appearing in a product will be summed automatically. In particular, if we let μ_t be the probability distribution measure of X, then μ_t satisfies the following Fokker–Planck equation in the distributional sense,

$$\partial_t \mu_t = \left(\mathscr{L}_t^\sigma\right)^* \mu_t + \mathrm{div}(b_t \cdot \mu_t), \quad \mu_0 = \delta_x,$$

where δ_x is the Dirac measure concentrated at point x, and $(\mathscr{L}_t^\sigma)^*$ is the adjoint operator of $\mathscr{L}_t^\sigma$. Throughout this paper, we always assume that σ satisfies that for some $c \geqslant 1$ and $\beta \in (0, 1)$,

$$c^{-1}|\xi|^2 \leqslant |\sigma_t^*(x)\xi|^2 \leqslant c|\xi|^2, \quad \|\sigma_t(x) - \sigma_t(y)\| \leqslant c|x-y|^\beta. \qquad (\mathbf{H}_\beta^\sigma)$$

It is well known that under $(\mathbf{H}_\beta^\sigma)$ and b being bounded measurable, operator $\partial_t - \mathscr{L}_t^\sigma - b_t \cdot \nabla$ admits a fundamental solution (also called heat kernel) $p_{t,s}(x, y)$ satisfying (see [7]):

$$\partial_t p_{t,s}(x, y) + (\mathscr{L}_t^\sigma + b_t \cdot \nabla) p_{t,s}(\cdot, y)(x) = 0, \quad \lim_{t \uparrow s} p_{t,s}(x, y) = \delta_x(\mathrm{d}y), \qquad (1.2)$$

and which enjoys the following estimates:

(i) (Two-sided estimate) For any $T > 0$, there are $c_0, \kappa_0 \geqslant 1$ such that

$$c_0^{-1}|s-t|^{-d/2}\mathrm{e}^{-\frac{\kappa_0|x-y|^2}{|s-t|}} \leqslant p_{t,s}(x, y) \leqslant c_0|s-t|^{-d/2}\mathrm{e}^{-\frac{|x-y|^2}{\kappa_0|s-t|}} \quad \text{on } \mathbb{D}_0^T, \qquad (1.3)$$

where $\mathbb{D}_0^T := \{(t, x; s, y) : x, y \in \mathbb{R}^d, s, t \geqslant 0, 0 < s - t \leqslant T\}$.

(ii) (Gradient estimate) For any $T > 0$, there are $c_1, \kappa_1 \geqslant 1$ such that on $\mathbb{D}_0^T$,

$$|\nabla_x^j p_{t,s}(x, y)| \leqslant c_1|s-t|^{-(d+j)/2}\mathrm{e}^{-\frac{|x-y|^2}{\kappa_1|s-t|}}, \ j = 1, 2. \qquad (1.4)$$

(iii) (Hölder estimate in y) For any $T > 0$ and $\gamma \in (0, \beta)$, there are $c_2, \kappa_2 \geqslant 1$ such that on $\mathbb{D}_0^T$,

$$|\nabla_x^j p_{t,s}(x, y) - \nabla_x^j p_{t,s}(x, y')| \leqslant \frac{c_2|y-y'|^\gamma}{|s-t|^{(d+\gamma+j)/2}} \left(\mathrm{e}^{-\frac{|x-y|^2}{\kappa_2|s-t|}} + \mathrm{e}^{-\frac{|x-y'|^2}{\kappa_2|s-t|}}\right), \ j = 0, 1. \qquad (1.5)$$

When σ is uniformly non-degenerate and bounded Lipschitz continuous and b is uniformly Hölder continuous, Zvonkin [36] introduced a transformation of phase space to kill the drift and obtain the existence and uniqueness of strong solutions to SDE (1.1). The transformation of phase space used in [36] is now called "Zvonkin's transformation" in the literature and will be our corner stone. We will introduce it below. When $\sigma = \mathbb{I}$ and b is bounded measurable, Veretennikov [28] showed the existence and uniqueness of strong solutions to SDE (1.1). When $\sigma = \mathbb{I}$ and $b \in$

$L^q_{loc}(\mathbb{R}_+; L^p(\mathbb{R}^d))$ with $\frac{d}{p} + \frac{2}{q} < 1$, Krylov and Röckner [21] showed the strong well-posedness to SDE (1.1) under the extra assumption

$$\int_0^t |b_s(X_s)|^2 \mathrm{d}s < \infty, \quad a.s.$$

This assumption is essential in [21] because Girsanov's transformation is used therein. Such an assumption was dropped in [32] and their result was also extended to the multiplicative noise case by using Zvonkin's transformation (see [29,31,32,34]). More recent development about the strong uniqueness of SDEs with rough coefficients is referred to [6]. In one word, noise has some regularization effect in the sense that an ill-posed ODE becomes well-posed under some noise perturbations. It should be noticed that there are a lot of works to study the further properties of the strong solutions to SDE (1.1) with rough drifts such as: weak differentiability with respect to the initial values, Malliavin differentiability with respect to the sample path, stochastic flows, etc. (for examples, see [8,10,11,22,23,34]).

Now we consider SDE (1.1) with distributional drift b. Let $\mathscr{D}$ be the space of all smooth functions on $\mathbb{R}^d$ with compact supports, and $\mathscr{D}'$ the dual space of $\mathscr{D}$ called distributional function space. If $b \in \mathscr{D}'$, then the drift term $b_t(X_t)\mathrm{d}t$ in (1.1) does not make any sense in general. We call a continuous $\mathscr{F}_t$-adapted process X a solution of SDE (1.1) if

$$X_t = x + \int_0^t \sigma_s(X_s)\mathrm{d}W_s + A^b_t \text{ with } A^b_t := \lim_{n\to\infty} \int_0^t b_n(X_s)\mathrm{d}s, \tag{1.6}$$

where $(b^{(n)})_{n\in\mathbb{N}}$ is any mollifying approximation sequence of b, and the limit is taken in the sense of u.c.p (uniformly on compact subsets of time variable in probability).

In one-dimensional case, Bass and Chen [2] showed the strong well-posedness of SDE (1.6) in a special class of Dirichlet processes when σ is $\frac{1}{2}$-order Hölder continuous and bounded below by a positive constant and b is the derivative of a γ-order Hölder continuous function with $\gamma \in (\frac{1}{2}, 1)$. Therein, they used the scaling function $s(x) = \int_0^x \exp\left(\int_0^y 2b(z)/\sigma^2(z)\mathrm{d}z\right)\mathrm{d}y$ to remove the drift and then applied Yamada-Watanabe's pathwise uniqueness result about one-dimensional SDE to obtain the strong well-posedness. More results about one-dimensional SDEs driven by Brownian motion with distributional drifts are referred to [9,14–17,24]. However, in the multi-dimensional case, solving SDE (1.1) with distributional drift b becomes quite involved. Recently, when $\sigma \equiv \sqrt{2}\mathbb{I}_{d\times d}$ and

$$b \in L^\infty_{loc}(\mathbb{R}_+; H^{-\alpha,p}) \text{ with } \alpha \in \left(0, \frac{1}{2}\right) \text{ and } p \in \left(\frac{d}{1-\alpha}, \frac{d}{\alpha}\right), \tag{1.7}$$

Flandoli, Issoglio and Russo [13] showed the existence and uniqueness of "virtual" solutions (a class of special weak solutions) to SDE (1.1). More precisely, consider the following backward PDE with distributional first order term:

$$\partial_t u + (\Delta - \lambda)u + b_t \cdot \nabla u = b_t, \quad u(T) = 0, \tag{1.8}$$

where $\lambda > 0$. Under (1.7), they showed that for λ large enough, there is a solution u so that

$$|\nabla u_t(x)| \leqslant 1/2, \quad t \in [0, T], \quad x \in \mathbb{R}^d.$$

In particular, if we define $\Phi_t(x) := x + u_t(x)$, then $x \mapsto \Phi_t(x)$ is a C^1-diffeomorphism of $\mathbb{R}^d$. Using Itô's formula formally, it is easy to see that $Y_t = \Phi_t(X_t)$ solves the following new SDE:

$$Y_t = \Phi_0(x) + \int_0^t \lambda u \circ \Phi_s^{-1}(Y_s)\mathrm{d}s + \sqrt{2}\int_0^t \nabla\Phi_s \circ \Phi_s^{-1}(Y_s)\mathrm{d}W_s,$$

where $\Phi_s^{-1}(x)$ is the inverse of $x \mapsto \Phi_s(x)$. Since this new SDE has continuous and non-degenerate diffusion coefficients and the drift is Lipschitz continuous, it is well known that the above SDE admits a unique weak solution (see [27]). In [13], $X_t := \Phi_t^{-1}(Y_t)$ is in turn called "virtual" solution of SDE (1.1). The above Φ is usually called Zvonkin's transformation in the literature. Unfortunately, it is not answered whether the above constructed X really solves SDE (1.6). This question is completely answered in a recent work [35]. We will summarize the main results of [35] in Sect. 3 below.

Nevertheless, the above distribution-valued drift does not allow the measure-valued drift. In [3], Bass and Chen studied the weak well-posedness of Brownian motions with singular measure-valued drifts. That is, when $\sigma = \sqrt{2}\mathbb{I}$ and b belongs to some generalized Kato's class, they showed the well-posedness of SDE (1.6) in the class of semimartingales. In other words, $t \mapsto A_t^b$ in (1.1) has finite variation in finite time interval. In this work, we will extend Bass and Chen's result to more general case: multiplicative noise. For this aim, we introduce some new Kato's class. Our approach is still based on Zvonkin's transformation and heat kernel estimates, and looks much simpler compared with Bass and Chen's proof [3].

Another aim of this paper is to show the existence and two-sided estimate of the heat kernel and the ergodicity associated with SDE (1.6). As we shall see, Zvonkin's transformation provides a satisfactory answer. Indeed, under the homomorphism transformation, if the transformed SDE admits a density and a unique invariant probability measure, then the original SDE also admits a density and a unique invariant probability measure. Such an idea was first used in [29]. It should be noticed that for the ergodicity of SDE (1.1), we assume $b = b^{(1)} + b^{(2)}$, where $b^{(1)}$ is the dissipative part and $b^{(2)}$ is a distribution. The key observation here is that the dissipativity is preserved under Zvonkin's transformation.

This paper is organized as follows: In Sect. 2, we present some preliminary results used below. In Sect. 3, we give a short proof of Krylov and Röcker's result so that one can grasp the main points of Zvonkin's argument. In fact, Kryov's a priori estimate is the key obstacle. In Sect. 4, we survey the main results obtained in [35] when the drift is in the Bessel potential space $H^{-1/2,p}$ with $p > 2d$. In Sect. 5, we study SDE (1.6) with b in some generalized Kato's class. In particular, our results completely cover Bass and Chen's result [3].

We close this section by mentioning some conventions used throughout this paper: We use $:=$ as a way of definition. For $a, b \in \mathbb{R}$, $a \vee b := \max\{a, b\}$ and $a \wedge b := \min\{a, b\}$, $\nabla := (\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_d})$ and $\Delta := \sum_{k=1}^{d} \frac{\partial^2}{\partial x_k^2}$ denotes the gradient and Laplacian operators. The letter C with or without subscripts stands for an unimportant constant, whose value may change in different places. We use $A \asymp B$ to denote that A and B are comparable up to a constant, and use $A \lesssim B$ to denote $A \leqslant CB$ for some constant C.

2 Preliminary

In this section, we first introduce some notations and recall some basic results for later use. Let ϱ be a nonnegative smooth function in $\mathbb{R}^d$ with compact support in the unit ball and $\int \varrho = 1$. Define a family of mollifiers

$$\varrho_n(x) = n^d \varrho(nx), \quad n \in \mathbb{N}.$$

For a distribution $f \in \mathscr{D}'$, if there is no further declaration, we always use f_n to denote the mollifying approximation of f, that is,

$$f_n(x) := f * \varrho_n(x),$$

where $*$ denotes the convolution in the distributional sense. Let χ be a nonnegative smooth function with $\chi(x) = 0$ for $|x| \geqslant 2$ and $\chi(x) = 1$ for $|x| \leqslant 1$. For $R > 0$, we shall also use the following cut-off function

$$\chi_R(x) = \chi(x/R). \tag{2.1}$$

Definition 2.1 For $\alpha \in \mathbb{R}$ and $p \in [1, \infty)$, the Bessel potential space $H^{\alpha,p}$ is defined by

$$H^{\alpha,p} := (\mathbb{I} - \Delta)^{-\alpha/2}(L^p)$$

with norm

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

where $\|\cdot\|_p$ is the usual L^p-norm. We also denote by $H^{\alpha,p}_{loc}$ the space of all the distribution $f \in \mathscr{D}'$ with $f\chi_R \in H^{\alpha,p}$ for any $R > 0$, which is the local Bessel potential space.

For $\alpha \in (0, 2)$ and $p \in (1, \infty)$, by Mihlin's multiplier theorem, we have

$$\|f\|_{\alpha,p} \asymp \|(\mathbb{I} - \Delta^{\alpha/2}) f\|_p \asymp \|f\|_p + \|\Delta^{\alpha/2} f\|_p, \tag{2.2}$$

where $\Delta^{\alpha/2} := -(-\Delta)^{\alpha/2}$ is the usual fractional Laplacian, which has the following alternative expression up to a multiplying constant,

$$\Delta^{\alpha/2} f(x) = \text{P.V.} \int_{\mathbb{R}^d} \frac{f(x+y) - f(x)}{|y|^{d+\alpha}} \mathrm{d}y$$
$$= \frac{1}{2} \int_{\mathbb{R}^d} \frac{f(x+y) + f(x-y) - 2f(x)}{|y|^{d+\alpha}} \mathrm{d}y, \tag{2.3}$$

where P.V. stands for Cauchy's principle value. Notice that the following Sobolev's embedding holds:

$$H^{\alpha,p} \subset \begin{cases} \cap_{q\in[p,dp/(d-p\alpha)]} L^q, & \text{if } p\alpha < d, \\ \mathcal{C}^{\alpha-d/p} \cap (\cap_{q\geqslant p} L^q), & \text{if } p\alpha > d, \end{cases} \tag{2.4}$$

where $\mathcal{C}^{\alpha-d/p}$ is the usual Hölder space. Moreover, for any $\alpha \in (0, 1]$ and $p \in (1, \infty)$, there is a constant $C = C(\alpha, p, d) > 0$ such that for all $f \in H^{\alpha,p}$ (see [1, Theorem 2.36]),

$$\|f(\cdot + y) - f(\cdot)\|_p \leqslant C|y|^\alpha \|\Delta^{\alpha/2} f\|_p, \tag{2.5}$$

and if $p\alpha > d$, then for all $f \in H^{\alpha,p}$ and $x, y \in \mathbb{R}^d$,

$$|f(x+y) - f(x)| \leqslant C|y|^{\alpha-\frac{d}{p}} \|\Delta^{\alpha/2} f\|_p, \tag{2.6}$$

and the following Gagliardo–Nirenberg's inequality holds: for $p > 1$ and $0 < \alpha < \beta \leqslant 1$, and all $f \in H^{\beta,p} \cap L^\infty$ (see [1, Theorem 2.44]),

$$\|\Delta^{\alpha/2} f\|_{p\beta/\alpha} \leqslant C\|f\|_\infty^{1-\alpha/\beta} \|\Delta^{\beta/2} f\|_p^{\alpha/\beta}. \tag{2.7}$$

For $\alpha \in (0, 2]$ and $d \geqslant 1$, we introduce the following space-time function:

$$\rho_t^{(\alpha)}(x) := \begin{cases} (t^{1/2} + |x|)^{-d-\alpha}, & \alpha \in (0, 2), \\ t^{-d/2} \mathrm{e}^{-|x|^2/t}, & \alpha = 2. \end{cases} \tag{2.8}$$

We need the following simple lemma.

Lemma 2.2 *For any $\alpha \in (0, 2)$, there is a constant $C = C(\alpha, d) > 0$ such that*

$$|\Delta^{\alpha/2} \rho_t^{(2)}(x)| \leqslant C\rho_t^{(\alpha)}(x), \quad t > 0, x \in \mathbb{R}^d.$$

Proof By scaling, we have

$$\left(\Delta^{\alpha/2} \rho_t^{(2)}\right)(x) = t^{(-d-\alpha)/2} (\Delta^{\alpha/2} \rho_1^{(2)})(t^{-1/2} x).$$

Thus it suffices to prove the estimate for $t=1$. Suppose $|x|\geqslant 1$. By definition (2.3), we have

$$\begin{aligned}\Delta^{\alpha/2}\rho_1^{(2)}(x)\leqslant{}&\frac{1}{2}\int_{|y|\leqslant|x|/2}\frac{\rho_1^{(2)}(x+y)+\rho_1^{(2)}(x-y)-2\rho_1^{(2)}(x)}{|y|^{d+\alpha}}\mathrm{d}y\\&+\frac{1}{2}\int_{|y|>|x|/2}\frac{\rho_1^{(2)}(x+y)+\rho_1^{(2)}(x-y)-2\rho_1^{(2)}(x)}{|y|^{d+\alpha}}\mathrm{d}y\\&=:I_1(x)+I_2(x).\end{aligned}$$

Noticing that

$$|\nabla^2\rho_1^{(2)}(x)|\leqslant c_0\mathrm{e}^{-c_1|x|^2},$$

by Taylor's expansion, we have

$$I_1(x)\leqslant\int_{|y|\leqslant|x|/2}\frac{|y|^2|\nabla^2\rho_1^{(2)}(x+\theta y)|}{|y|^{d+\alpha}}\mathrm{d}y\lesssim\mathrm{e}^{-c_1|x|^2/4}|x|^{2-\alpha}\leqslant\mathrm{e}^{-c_1|x|^2/8},\quad\theta\in(0,1).$$

For $I_2(x)$, we have

$$\begin{aligned}I_2(x)\lesssim{}&|x|^{-d-\alpha}\int_{|y|>|x|/2}(\rho_1^{(2)}(x+y)+\rho_1^{(2)}(x-y))\mathrm{d}y\\&+\int_{|y|>|x|/2}\frac{\rho_1^{(2)}(x)}{|y|^{d+\alpha}}\mathrm{d}y\lesssim|x|^{-d-\alpha}.\end{aligned}$$

The proof is complete. □

The following lemma about the product of two distributions is proved in [35].

Lemma 2.3 *Let $p\in(1,\infty)$ and $\alpha\in(0,1]$ be fixed.*

(i) *For any $p_1,p_2\in[p,\infty)$ with $\frac{1}{p}\leqslant\frac{1}{p_1}+\frac{1}{p_2}<\frac{1}{p}+\frac{\alpha}{d}$, there is a constant $C>0$ such that for all $f\in H^{\alpha,p_1}$ and $g\in H^{\alpha,p_2}$,*

$$\|fg\|_{\alpha,p}\leqslant C\|f\|_{\alpha,p_1}\|g\|_{\alpha,p_2}.\tag{2.9}$$

In particular, if $p>d/\alpha$, then $H^{\alpha,p}$ is an algebra under pointwise product.

(ii) *For any $p_1\in[p,\infty)$ and $p_2\in[\frac{p_1}{p_1-1},\infty)$ with $\frac{1}{p}\leqslant\frac{1}{p_1}+\frac{1}{p_2}<\frac{1}{p}+\frac{\alpha}{d}$, there is a constant $C>0$ such that for all $f\in H^{-\alpha,p_1}$ and $g\in H^{\alpha,p_2}$,*

$$\|fg\|_{-\alpha,p}\leqslant C\|f\|_{-\alpha,p_1}\|g\|_{\alpha,p_2}.\tag{2.10}$$

Let $\mathcal{D}^0_\infty$ be the set of all C^1-diffeomorphisms on $\mathbb{R}^d$:

$$\mathcal{D}^0_\infty:=\Big\{\Phi:\mathbb{R}^d\to\mathbb{R}^d,\quad\|\Phi\|_{\mathcal{D}^0_\infty}:=\|\nabla\Phi\|_\infty+\|\nabla\Phi^{-1}\|_\infty<\infty\Big\}.$$

Clearly, $\mathcal{D}^0_\infty$ is closed under the inverse operation, that is, $\Phi \in \mathcal{D}^0_\infty$ implies $\Phi^{-1} \in \mathcal{D}^0_\infty$. For $\beta \in (0, 1]$ and $q \in (d/\beta, \infty)$, we also introduce a subclass of $\mathcal{D}^0_\infty$ as follows:

$$\mathcal{D}^\beta_q := \Big\{\Phi \in \mathcal{D}^0_\infty : \|\Phi\|_{\mathcal{D}^\beta_q} := \|\Phi\|_{\mathcal{D}^0_\infty} + \|\mathbb{I} - \nabla\Phi\|_{\beta,q} < \infty\Big\}. \tag{2.11}$$

We have the following result about the class $\mathcal{D}^\beta_q$ (see [35]).

Proposition 2.4 (i) *Let $\beta \in (0, 1]$ and $q \in (d/\beta, \infty)$. For any $\Phi \in \mathcal{D}^\beta_q$, we have $\Phi^{-1} \in \mathcal{D}^\beta_q$ and*

$$\|\det(\nabla\Phi) - 1\|_{\beta,q}, \|\det(\nabla\Phi^{-1}) - 1\|_{\beta,q} < \infty.$$

(ii) *Let $\Phi \in \mathcal{D}^0_\infty$ be a C^1-diffeomorphism. For any $\alpha \in [0, 1]$ and $p > 1$, there is a constant $C = C(\alpha, d, p, \|\Phi\|_{\mathcal{D}^0_\infty}) > 0$ such that for all $f \in H^{\alpha,p}$,*

$$\|f \circ \Phi\|_{\alpha,p} \leqslant C\|f\|_{\alpha,p}. \tag{2.12}$$

(iii) *Let $\Phi \in \mathcal{D}^\beta_q$ for some $\beta \in (0, 1)$ and $q \in (d/\beta, \infty)$. For any $\alpha \in [0, \beta]$ and $p > \frac{d}{d-\alpha}$, there is a constant $C = C(\alpha, \beta, d, p, \|\Phi\|_{\mathcal{D}^\beta_q}) > 0$ such that for all $f \in H^{-\alpha,p}$,*

$$\|f \circ \Phi\|_{-\alpha,p} \leqslant C\|f\|_{-\alpha,p}. \tag{2.13}$$

The following estimate is well known (for example, see [33]).

Lemma 2.5 *Let $b \in W^{1,1}_{loc}(\mathbb{R}^d)$. Then there exists a Lebesgue-null set $A \subset \mathbb{R}^d$ such that for all $x, y \notin A$,*

$$|b(x) - b(y)| \leqslant 2^d \int_0^{|x-y|} \fint_{B_s} |\nabla b|(x + z)\mathrm{d}z\mathrm{d}s + 2^d \int_0^{|x-y|} \fint_{B_s} |\nabla b|(y + z)\mathrm{d}z\mathrm{d}s,$$

where $B_s := \{x \in \mathbb{R}^d : |x| \leqslant s\}$. In particular, for any $R \in (0, \infty]$ and $x, y \notin A$ with $|x - y| < R$,

$$|b(x) - b(y)| \leqslant 2^d|x - y|(M_R|\nabla b|(x) + (M_R|\nabla b|(y)), \tag{2.14}$$

where $M_R f(x) := \sup_{r<R} \fint_{B_r} f(x+z)\mathrm{d}z$ is the Hardy-Littlewood maximal function.

We also need the following simple lemma.

Lemma 2.6 *Let ℓ^m_s be a family of locally integrable function and $H : [0, \infty) \to \mathbb{R}$ be a continuous function so that for any $T > 0$,*

$$\lim_{\varepsilon\to 0} \sup_{t\in[0,T]} \left|\int_0^t \ell^\varepsilon_s \mathrm{d}s - H_t\right| = 0.$$

Then for any $0 \leqslant t_0 \leqslant t_1$*, it holds that*

$$|H|_{t_0}^{t_1} \leqslant \varliminf_{m\to\infty} \int_{t_0}^{t_1} |\ell_s^m| \mathrm{d}s,$$

where $|H|_{t_0}^{t_1}$ *stands for the variation of* H *on* $[t_0, t_1]$*. If in addition,*

$$\sup_{m\in\mathbb{N}} \int_0^T |\ell_s^m| \mathrm{d}s < \infty, \quad \forall T > 0,$$

then for any bounded continuous f*,*

$$\lim_{m\to\infty} \left| \int_0^T f_s \ell_s^m \mathrm{d}s - \int_0^T f_s \mathrm{d}H_s \right| = 0.$$

Proof Let $t_0 = s_0 < s_1 < \cdots < s_n = t_1$. By the assumption we have

$$\begin{aligned}\sum_{k=0}^{n-1} |H_{s_{k+1}} - H_{s_k}| &= \sum_{k=0}^{n-1} \left| \lim_{m\to\infty} \int_{s_k}^{s_{k+1}} \ell_s^m \mathrm{d}s \right| \\ &\leqslant \varliminf_{m\to\infty} \sum_{k=0}^{n-1} \int_{s_k}^{s_{k+1}} |\ell_s^m| \mathrm{d}st = \varliminf_{m\to\infty} \int_{t_0}^{t_1} |\ell_s^m| \mathrm{d}s,\end{aligned}$$

which implies by taking supremum for partitions of $[t_0, t_1]$,

$$|H|_{t_0}^{t_1} \leqslant \varliminf_{m\to\infty} \int_{t_0}^{t_1} |\ell_s^m| \mathrm{d}s.$$

For the second conclusion, for $n \in \mathbb{N}$, letting $s_k = kT/n$, $k = 0, \ldots, n$, we have

$$\begin{aligned}\left| \int_0^T f_s \ell_s^m \mathrm{d}s - \int_0^T f_s \mathrm{d}H_s \right| \leqslant& \sup_{m\in\mathbb{N}} \left| \int_0^T f_s \ell_s^m \mathrm{d}s - \sum_{k=0}^{n-1} f_{s_k} \int_{s_k}^{s_{k+1}} \ell_s^m \mathrm{d}s \right| \\ &+ \left| \sum_{k=0}^{n-1} f_{s_k} \left[(H_{s_{k+1}} - H_{s_k}) - \int_{s_k}^{s_{k+1}} \ell_s^m \mathrm{d}s \right] \right| \\ &+ \left| \int_0^T f_s \mathrm{d}H_s - \sum_{k=0}^{n-1} f_{s_k} (H_{s_{k+1}} - H_{s_k}) \right|.\end{aligned} \tag{2.15}$$

Since f is continuous and $\sup_{m\in\mathbb{N}}\int_0^T|\ell_s^m|\mathrm{d}s<\infty$, it is easy to see that

$$\lim_{n\to\infty}\sup_{m\in\mathbb{N}}\left|\int_0^T f_s\ell_s^m\mathrm{d}s-\sum_{k=0}^{n-1}f_{s_k}\int_{s_k}^{s_{k+1}}\ell_s^m\mathrm{d}s\right|=0,$$

$$\lim_{n\to\infty}\left|\int_0^T f_s\mathrm{d}H_t-\sum_{k=0}^{n-1}f_{s_k}(H_{s_{k+1}}-H_{s_k})\right|=0.$$

Moreover, for fixed $n\in\mathbb{N}$, by the condition we also have

$$\lim_{m\to\infty}\left|\sum_{k=0}^{n-1}f_{s_k}\left[(H_{s_{k+1}}-H_{s_k})-\int_{s_k}^{s_{k+1}}\ell_s^m\mathrm{d}s\right]\right|=0.$$

Hence, by first letting $m\to\infty$ then $n\to\infty$ in (2.15), we obtain the desired limit. □

The following lemma is easy.

Lemma 2.7 *Let A_t be a continuous nonnegative adapted process and τ be any stopping time. Suppose that there is a constant $C_0>0$ such that for any stopping time $\tau'\leqslant\tau$,*

$$\mathbb{E}A_{\tau'}\leqslant C_0.$$

Then for any $q\in(0,1)$, it holds that

$$\mathbb{E}\left(\sup_{t\in[0,\tau]}A_t^q\right)\leqslant\frac{C_0^q}{1-q}.$$

Proof For $\lambda>0$, define

$$\tau'_\lambda:=\inf\{t\geqslant 0:|A_t|\geqslant\lambda\}.$$

Noticing that

$$\lambda\mathbb{P}\left(\tau'_\lambda\leqslant\tau\right)\leqslant\mathbb{E}(A_{\tau\wedge\tau'_\lambda})\leqslant C_0,$$

we have

$$\begin{aligned}\mathbb{E}\left(\sup_{t\in[0,\tau]}A_t^q\right)&=q\int_0^\infty\lambda^{q-1}\mathbb{P}\left(\sup_{t\in[0,\tau]}A_t>\lambda\right)\mathrm{d}\lambda\\&\leqslant q\int_0^\infty\lambda^{q-1}\Big(1\wedge\mathbb{P}\left(\tau'_\lambda\leqslant\tau\right)\Big)\mathrm{d}\lambda\\&\leqslant q\int_0^\infty\lambda^{q-1}\Big(1\wedge(C_0/\lambda)\Big)\mathrm{d}\lambda=\frac{C_0^q}{1-q}.\end{aligned}$$

The proof is complete. □

The following stochastic Gronwall's inequality for continuous martingales is proved by Scheutzow [25]. For general discontinuous martingales, it is due to [29].

Lemma 2.8 (Stochastic Gronwall's inequality) *Let $\xi(t)$ and $\eta(t)$ be two nonnegative càdlàg $\mathscr{F}_t$-adapted processes, A_t a continuous nondecreasing $\mathscr{F}_t$-adapted process with $A_0 = 0$, M_t a local martingale with $M_0 = 0$. Suppose that*

$$\xi(t) \leqslant \eta(t) + \int_0^t \xi(s)\mathrm{d}A_s + M_t, \quad \textit{for all } t \geqslant 0. \tag{2.16}$$

Then for any $0 < q < p < 1$ and $\tau > 0$, we have

$$\left[\mathbb{E}(\xi(\tau)^*)^q\right]^{1/q} \leqslant \left(\tfrac{p}{p-q}\right)^{1/q} \left(\mathbb{E}\mathrm{e}^{pA_\tau/(1-p)}\right)^{(1-p)/p} \mathbb{E}\big(\eta(\tau)^*\big), \tag{2.17}$$

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

Proof Without loss of generality, we may assume that the right hand side of (2.17) is finite and $\eta(t)$ is nondecreasing. Otherwise, we may replace $\eta(t)$ with $\eta(t)^* := \sup_{s\in[0,t]} \eta(s)$. Let $\bar{\xi}(t)$ be the right hand side of (2.16) and $\bar{A}_t := \int_0^t \xi(s)/\bar{\xi}(s)\mathrm{d}A_s$. Then

$$\xi(t) \leqslant \bar{\xi}(t) = \eta(t) + \int_0^t \bar{\xi}(s)\mathrm{d}\bar{A}_s + M_t.$$

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. In other words, for each $n \in \mathbb{N}$,

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

Using $\mathrm{e}^{-\bar{A}_s} \leqslant 1$, we have

$$\mathbb{E}\Big(\mathrm{e}^{-\bar{A}_{\tau\wedge\tau_R}}\bar{\xi}(\tau\wedge\tau_R)\Big) \leqslant \mathbb{E}\big(\eta(\tau\wedge\tau_R)\big) \leqslant \mathbb{E}\big(\eta(\tau)\big).$$

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

$$\mathbb{E}\Big(\mathrm{e}^{-\bar{A}_\tau}\bar{\xi}(\tau)\Big) \leqslant \mathbb{E}\big(\eta(\tau)\big),$$

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

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

Now by Lemma 2.7, we obtain (2.17). □

3 Strong Well-Posedness of SDEs with Integrable Drifts

In this section, we prove the strong well-posedness of the following SDE by using Zvonkin's method:

$$\mathrm{d}X_t = \sqrt{2}\mathrm{d}W_t + b(X_t)\mathrm{d}t, \quad X_0 = x, \tag{3.1}$$

where $b \in L^p(\mathbb{R}^d)$ for some $p > d \vee 2$. More precisely, we shall prove that

Theorem 3.1 *Assume $b \in L^{p_1}(\mathbb{R}^d)$ for some $p_1 > d \vee 2$. Then, for each $x \in \mathbb{R}^d$, there is a unique strong solution X_t to SDE (3.1) in the sense that*

$$\int_0^t |b(X_s)|\mathrm{d}s < \infty \ \ a.s.\ and\ X_t = x + \sqrt{2}W_t + \int_0^t b(X_s)\mathrm{d}s.$$

Remark 3.2 It should be noticed that in the original statement of Krylov and Röckner [21], they require $\int_0^t |b(X_s)|^2\mathrm{d}s < \infty$, in order to apply the Girsanov transformation.

Before proving this theorem, we need to first solve the following elliptic equation:

$$(\Delta - \lambda)u + b \cdot \nabla u = f, \tag{3.2}$$

where $\lambda > 0$ and $f \in L^p(\mathbb{R}^d)$. We have

Theorem 3.3 *Assume $b \in L^{p_1}(\mathbb{R}^d)$ for some $p_1 > d$. For any $p \in (d/2 \vee 1, p_1]$, there is a $\lambda_0 = \lambda_0(d, p, p_1, \|b\|_{p_1}) \geqslant 1$ such that for any $f \in L^p(\mathbb{R}^d)$ and $\lambda \geqslant \lambda_0$, there is a unique solution $u \in H^{2,p}$ to Eq. (3.2) so that*

$$\|u\|_{2,p} \leqslant C\|f\|_p, \quad \lambda^{\left(2-\alpha+\frac{d}{p'}-\frac{d}{p}\right)/2}\|u\|_{\alpha,p'} \leqslant C\|f\|_p, \tag{3.3}$$

where $\alpha \in [0, 2)$ and $p' \in [1, \infty]$ with $\frac{d}{p} < 2 - \alpha + \frac{d}{p'}$. Here, the constant C is independent of λ.

Proof We divide the proof into two steps.

(i) First of all, we assume $b = 0$ and show that for all $\lambda > 0$,

$$\|(\lambda - \Delta)^{-1} f\|_{2,p} \leqslant C_1\|f\|_p, \quad (1 \vee \lambda)^{\left(2-\alpha+\frac{d}{p'}-\frac{d}{p}\right)/2}\|(\lambda-\Delta)^{-1} f\|_{\alpha,p'} \leqslant C_2\|f\|_p, \tag{3.4}$$

where the constant C_2 does not depend on λ. The first estimate follows by Fourier's multiplier theorem (cf. [26]). We prove the second one in (3.4). Noticing that

$$u(x) := (\lambda - \Delta)^{-1} f(x) = (4\pi)^{-d/2}\int_0^\infty \mathrm{e}^{-\lambda t}\rho_t^{(2)} * f(x)\mathrm{d}t,$$

where $\rho_t^{(2)}$ is defined by (2.8), we have

$$\Delta^{\alpha/2}u(x) = (4\pi)^{-d/2}\int_0^\infty \mathrm{e}^{-\lambda t}(\Delta^{\alpha/2}\rho_t^{(2)}) * f(x)\mathrm{d}t.$$

Let $r = 1/(1 - 1/p + 1/p')$. By Lemma 2.2 and Young's inequality we have

$$\begin{aligned}\|\Delta^{\alpha/2}u\|_{p'} &\lesssim \int_0^\infty \mathrm{e}^{-\lambda t}\left\|\rho_t^{(\alpha)} * f\right\|_{p'}\mathrm{d}t \lesssim \int_0^\infty \mathrm{e}^{-\lambda t}\|\rho_t^{(\alpha)}\|_r\|f\|_p\mathrm{d}t\\ &= \left(\int_0^\infty \mathrm{e}^{-\lambda t}t^{(d/r-\alpha-d)/2}\mathrm{d}t\right)\|\rho_1^{(\alpha)}\|_r\|f\|_p \lesssim \lambda^{(\alpha+d-d/r)/2-1}\|f\|_p.\end{aligned}$$

Moreover, we also have

$$\begin{aligned}\|u\|_{p'} &\lesssim \int_0^\infty \mathrm{e}^{-\lambda t}\left\|\rho_t^{(2)} * f\right\|_{p'}\mathrm{d}t \lesssim \int_0^\infty \mathrm{e}^{-\lambda t}\|\rho_t^{(2)}\|_r\|f\|_p\mathrm{d}t\\ &= \left(\int_0^\infty \mathrm{e}^{-\lambda t}t^{(d/r-d)/2}\mathrm{d}t\right)\|\rho_1^{(2)}\|_r\|f\|_p \lesssim \lambda^{(d-d/r)/2-1}\|f\|_p.\end{aligned}$$

Combining the above two estimates, we obtain the second estimate in (3.4).

(ii) We use Picard's iteration to solve Eq. (3.2). Let $u_0 = 0$ and define for $n \in \mathbb{N}$,

$$u_n := (\Delta - \lambda)^{-1}(f - b \cdot \nabla u_{n-1}). \tag{3.5}$$

Let $p_2 := p_1 p/(p_1 - p)$, $\alpha \in [0, 2)$ and $p' \in [1, \infty]$ with $\frac{d}{p} < 2 - \alpha + \frac{d}{p'}$. By (3.4) and Hölder's inequality, we have

$$(1 \vee \lambda)^{\left(2-\alpha+\frac{d}{p'}-\frac{d}{p}\right)/2}\|u_n\|_{\alpha,p'} \lesssim \|f - b \cdot \nabla u_{n-1}\|_p \leqslant \|f\|_p + \|b\|_{p_1}\|\nabla u_{n-1}\|_{p_2} \tag{3.6}$$

and

$$(1 \vee \lambda)^{\left(2-\alpha+\frac{d}{p'}-\frac{d}{p}\right)/2}\|u_n - u_m\|_{\alpha,p'} \lesssim \|b\|_{p_1}\|\nabla u_{n-1} - \nabla u_{m-1}\|_{p_2}. \tag{3.7}$$

In particular, due to $p_1 > d$, we can take $p' = p_2$, $\alpha = 1$ and get

$$(1 \vee \lambda)^{\left(1-\frac{d}{p_1}\right)/2}\|u_n\|_{1,p_2} \leqslant C\|f\|_p + C\|b\|_{p_1}\|\nabla u_{n-1}\|_{p_2}$$

and

$$(1 \vee \lambda)^{\left(1-\frac{d}{p_1}\right)/2}\|u_n - u_m\|_{1,p_2} \leqslant C\|b\|_{p_1}\|\nabla u_{n-1} - \nabla u_{m-1}\|_{p_2}.$$

Choosing $\lambda_0 \geqslant 1$ be large enough so that $C\lambda^{(\frac{d}{p_1}-1)/2}\|b\|_{p_1} \leqslant 1/2$ for all $\lambda \geqslant \lambda_0$, we get

$$\|u_n\|_{1,p_2} \leqslant C\lambda^{\left(\frac{d}{p_1}-1\right)/2}\|f\|_p + \tfrac{1}{2}\|u_{n-1}\|_{1,p_2}$$

and for all $n \geqslant m$,

$$\|u_n - u_m\|_{1,p_2} \leqslant \tfrac{1}{2}\|u_{n-1} - u_{m-1}\|_{1,p_2}.$$

From these two estimates, by iteration, we derive that for all $\lambda \geqslant \lambda_0$,

$$\sup_n \|u_n\|_{1,p_2} \leqslant C\lambda^{\left(\frac{d}{p_1}-1\right)/2}\|f\|_p,$$

and for all $n \geqslant m$,

$$\|u_n - u_m\|_{1,p_2} \leqslant \tfrac{1}{2^m}\|u_{n-m}\|_{1,p_2} \leqslant \tfrac{C}{2^m}.$$

Substituting them into (3.6) and (3.7), we obtain

$$\lambda^{\left(2-\alpha+\frac{d}{p'}-\frac{d}{p}\right)/2}\|u_n\|_{\alpha,p'} \leqslant C\|f\|_p,$$

and for all $n \geqslant m$,

$$\lambda^{\left(2-\alpha+\frac{d}{p'}-\frac{d}{p}\right)/2}\|u_n - u_m\|_{\alpha,p'} \leqslant \tfrac{C}{2^m}.$$

Moreover, we also have

$$\|u_n\|_{2,p} \leqslant C\|f\|_p + C\|b\|_{p_1}\|\nabla u_{n-1}\|_{p_2} \leqslant C\|f\|_p.$$

Hence, there is a $u \in H^{2,p}$ such that (3.3) holds and

$$\lambda^{\left(2-\alpha+\frac{d}{p'}-\frac{d}{p}\right)/2}\|u - u_m\|_{\alpha,p'} \leqslant \tfrac{C}{2^m},$$

and u solves Eq. (3.2) by taking limits for (3.5). □

The following Krylov's estimate will play a crucial role in the proof of Theorem 3.1.

Theorem 3.4 *Let $b \in L^{p_1}(\mathbb{R}^d)$ for some $p_1 > d$. For any $p > d/2\vee 1$ and $T > 0$, there is a constant $C > 0$ such that for any solution X of SDE (3.1) and all $0 \leqslant t_0 < t_1 \leqslant T$,*

$$\mathbb{E}\left(\int_{t_0}^{t_1} |f(X_s)|\mathrm{d}s \Big| \mathscr{F}_{t_0}\right) \leqslant C(t_1 - t_0)^{1-d/(2p)}\|f\|_p. \tag{3.8}$$

Proof (i) Let $\lambda \geqslant 1$. First of all, for $f \in C_c^\infty(\mathbb{R}^d)$, let

$$u := (\Delta - \lambda)^{-1} f \in C_b^\infty(\mathbb{R}^d),$$

and for $R > 0$, define

$$\tau_R := \inf\left\{t \geqslant 0 : \int_0^t |b(X_s)|\mathrm{d}s \geqslant R\right\}.$$

By Itô's formula, we have

$$\begin{aligned}\mathbb{E}\Big(u(X_{t_1\wedge\tau_R}) - u(X_{t_0\wedge\tau_R})\Big) &= \mathbb{E}\int_{t_0\wedge\tau_R}^{t_1\wedge\tau_R} (\Delta u + b\cdot\nabla u)(X_s)\mathrm{d}s\\ &= \mathbb{E}\int_{t_0\wedge\tau_R}^{t_1\wedge\tau_R} (f + \lambda u + b\cdot\nabla u)(X_s)\mathrm{d}s.\end{aligned}$$

Therefore, for $p > d$, by (3.4) with $p' = \infty$ and $\alpha = 0, 1$, we have

$$\begin{aligned}\mathbb{E}\int_{t_0\wedge\tau_R}^{t_1\wedge\tau_R} f(X_s)\mathrm{d}s &\leqslant \lambda\|u\|_\infty(t_1 - t_0) + 2\|u\|_\infty\\ &\quad + \|\nabla u\|_\infty \mathbb{E}\int_{t_0\wedge\tau_R}^{t_1\wedge\tau_R} |b|(X_s)\mathrm{d}s\\ &\lesssim \lambda^{\frac{d}{2p}}(t_1 - t_0)\|f\|_p + \lambda^{\frac{d}{2p}-1}\|f\|_p\\ &\quad + \lambda^{\frac{d}{2p}-\frac{1}{2}}\|f\|_p \mathbb{E}\int_{t_0\wedge\tau_R}^{t_1\wedge\tau_R} |b|(X_s)\mathrm{d}s.\end{aligned} \tag{3.9}$$

By a standard monotone class argument, the above estimate still holds for any $f \in L^p(\mathbb{R}^d)$. In particular, if we take $p = p_1$ and $f = |b|$, then

$$\begin{aligned}\mathbb{E}\int_{t_0\wedge\tau_R}^{t_1\wedge\tau_R} |b|(X_s)\mathrm{d}s &\leqslant C\lambda^{\frac{d}{2p_1}-1}(\lambda(t_1 - t_0)\|b\|_{p_1} + 1)\|b\|_{p_1}\\ &\quad + C\lambda^{\frac{d}{2p_1}-\frac{1}{2}}\|b\|_{p_1}\mathbb{E}\int_{t_0\wedge\tau_R}^{t_1\wedge\tau_R} |b|(X_s)\mathrm{d}s.\end{aligned}$$

Letting $\lambda = \kappa(t_1 - t_0)^{-1}$ with κ being large enough so that $C\lambda^{(d/p_1-1)/2}\|b\|_{p_1} \leqslant 1/2$, we obtain

$$\mathbb{E}\int_{t_0\wedge\tau_R}^{t_1\wedge\tau_R} |b|(X_s)\mathrm{d}s \leqslant C(t_1 - t_0)^{1-d/(2p_1)}\|b\|_{p_1}.$$

Substituting this into (3.9) and letting $R \to \infty$, we obtain that for any $p > d$,

$$\mathbb{E}\int_{t_0}^{t_1} f(X_s)\mathrm{d}s \leqslant C(t_1 - t_0)^{1-\frac{d}{2p}}\|f\|_p.$$

(ii) Now let $p > d/2 \vee 1$. For $f \in C_c^\infty(\mathbb{R}^d)$, let $u \in H^{2,p_1}$ solve (3.2) and define $u_n(x) := u * \varrho_n(x)$. By Itô's formula, we have

$$\begin{aligned}
&\mathbb{E}((u_n(X_{t_1}) - u_n(X_{t_0}))|\mathscr{F}_{t_0}) \\
&= \mathbb{E}\left(\int_{t_0}^{t_1} (\Delta u_n + b \cdot u_n)(X_s)\mathrm{d}s \middle| \mathscr{F}_{t_0}\right) \\
&= \mathbb{E}\left(\int_{t_0}^{t_1} (b \cdot \nabla u_n - (b \cdot \nabla u) * \varrho_n + \lambda u_n + f_n)(X_s)\mathrm{d}s \middle| \mathscr{F}_{t_0}\right).
\end{aligned}$$

Letting

$$[\varrho_n, b \cdot \nabla]u := \varrho_n * (b \cdot \nabla u) - b \cdot \nabla(u * \varrho_n),$$

we have

$$\begin{aligned}
\mathbb{E}\left(\int_{t_0}^{t_1} f_n(X_s)\mathrm{d}s \middle| \mathscr{F}_{t_0}\right) &\leqslant \lambda(t_1 - t_0)\|u\|_\infty + 2\|u\|_\infty \\
&\quad + \mathbb{E}\left(\int_{t_0}^{t_1} [\varrho_n, b \cdot \nabla]u(X_s)\mathrm{d}s \middle| \mathscr{F}_{t_0}\right).
\end{aligned} \tag{3.10}$$

By (i) we have

$$\mathbb{E}\left(\int_{t_0}^{t_1} |[\varrho_n, b \cdot \nabla]u|(X_s)\mathrm{d}s\right) \leqslant C\|[\varrho_n, b \cdot \nabla]u\|_{p_1} \overset{n\to\infty}{\to} 0.$$

Thus, letting $n \to \infty$ in (3.10) and by (3.4) with $p' = p$ and $\alpha = 0$, we obtain

$$\begin{aligned}
\mathbb{E}\left(\int_{t_0}^{t_1} f(X_s)\mathrm{d}s \middle| \mathscr{F}_{t_0}\right) &\leqslant \lambda(t_1 - t_0)\|u\|_\infty + 2\|u\|_\infty \\
&\lesssim \lambda^{d/(2p)}(t_1 - t_0)\|f\|_p \\
&\quad + (t_1 - t_0)^{\alpha/2}\lambda^{\alpha/2+d/(2p)-1}\|f\|_p.
\end{aligned}$$

Taking $\lambda = \kappa(t_1 - t_0)^{-1}$ with κ large enough, we get (3.8). □

Below we give two easy applications of the above Krylov's estimate.

Corollary 3.5 (Khasminskii's estimate) *Let $p > d/2 \vee 1$. For any $f \in L^p(\mathbb{R}^d)$ and $m \in \mathbb{N}$, we have*

$$\mathbb{E}^{\mathscr{F}_{t_0}}\left(\int_{t_0}^{t_1} |f(X_s)|\mathrm{d}s\right)^m \leqslant m!(C(t_1 - t_0))^{m(1-d/(2p))}\|f\|_p^m.$$

In particular, for any $T > 0$ and $\lambda > 0$,

$$\mathbb{E}\left(\exp\left\{\lambda\int_0^T |f(X_s)|\mathrm{d}s\right\}\right) < \infty.$$

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

$$\left(\int_{t_0}^{t_1} g(s)\mathrm{d}s\right)^m = m! \int\cdots\int_{\Delta^m} g(s_1)\cdots g(s_m)\mathrm{d}s_1\cdots\mathrm{d}s_m,$$

where

$$\Delta^m := \Big\{(s_1,\ldots,s_m) : t_0 \leqslant s_1 \leqslant s_2 \leqslant \cdots \leqslant s_m \leqslant t_1\Big\},$$

by (3.8), we have

$$\begin{aligned}
&\mathbb{E}^{\mathscr{F}_{t_0}}\left(\int_{t_0}^{t_1} f(X_s)\mathrm{d}s\right)^m \\
&= m!\mathbb{E}^{\mathscr{F}_{t_0}}\left(\int\cdots\int_{\Delta^m} f(X_{s_1})\cdots f(X_{s_m})\mathrm{d}s_1\cdots\mathrm{d}s_m\right) \\
&= m!\mathbb{E}^{\mathscr{F}_{t_0}}\Bigg(\int\cdots\int_{\Delta^{m-1}} f(X_{s_1})\cdots f(X_{s_{m-1}})\mathbb{E}^{\mathscr{F}_{s_{m-1}}} \\
&\quad\times\left(\int_{s_{m-1}}^{t_1} f(X_{s_m})\mathrm{d}s_m\right)\mathrm{d}s_1\cdots\mathrm{d}s_{m-1}\Bigg) \\
&\leqslant m!\mathbb{E}^{\mathscr{F}_{t_0}}\Bigg(\int\cdots\int_{\Delta^{m-1}} f(X_{s_1})\cdots f(X_{s_{m-1}}) \\
&\quad\cdot C(t_1-t_0)^{1-d/(2p)}\|f\|_p\mathrm{d}s_1\cdots\mathrm{d}s_{m-1}\Bigg) \\
&\leqslant\cdots\leqslant m!(C(t_1-t_0)^{1-d/(2p)}\|f\|_p)^m.
\end{aligned}$$

The proof is complete. □

Corollary 3.6 (Generalized Itô's formula) *Let X_t solve SDE (3.1) and $b \in L^{p_1}(\mathbb{R}^d)$ for some $p_1 > d$. For any $f \in H^{2,p}_{loc}$ with $p > (d/2) \vee 1$, it holds that*

$$f(X_t) = f(x) + \int_0^t (\Delta f + b\cdot\nabla f)(X_s)\mathrm{d}s + \int_0^t \nabla f(X_s)\mathrm{d}W_s.$$

Proof By standard localization technique, one may assume $f \in H^{2,p}$ and $p \in (d, p_1]$. By Hölder's inequality and Sobolev's embedding (2.4), we have

$$\|\Delta f + b\cdot\nabla f\|_p \lesssim \|f\|_{2,p} + \|b\|_{p_1}\|\nabla f\|_{p_1 p/(p_1-p)} \lesssim \|f\|_{2,p}.$$

The desired formula now follows by applying Itô's formula to $f_n = f * \varrho_n$ and then taking limits by Krylov's estimate (3.8). For example, letting $p' = dp/(2(d-p))$ and by (2.4), we have

$$\mathbb{E}\left|\int_0^t (\nabla f_n(X_s) - \nabla f(X_s))\mathrm{d}W_s\right|^2 = \mathbb{E}\int_0^t |\nabla f_n(X_s) - \nabla f(X_s)|^2\mathrm{d}s$$
$$\leqslant C\| |\nabla f_n - \nabla f|^2 \|_{p'} \leqslant C\|f_n - f\|_{1,2p'}^2$$
$$\leqslant C\|f_n - f\|_{2,p}^2 \to 0.$$

The proof is complete. □

Let $b \in L^{p_1}(\mathbb{R}^d)$ and u solve the following elliptic system

$$(\Delta - \lambda)u + b \cdot \nabla u = -b.$$

By Theorem 3.3, for any $\alpha \in [0, 2)$, there are $\lambda_0 \geqslant 1$ and $C > 0$ such that for all $\lambda \geqslant \lambda_0$ and $p' \in [1, \infty]$ with $\frac{d}{p_1} < 2 - \alpha + \frac{d}{p'}$,

$$\|u\|_{2,p_1} \leqslant C\|b\|_{p_1}, \quad \lambda^{\left(2-\alpha+\frac{d}{p'}-\frac{d}{p}\right)/2}\|u\|_{\alpha,p'} \leqslant C\|b\|_{p_1}. \tag{3.11}$$

Define

$$\Phi(x) := x + u(x).$$

It is easy to see that

$$\Delta\Phi + b \cdot \nabla\Phi = \lambda u,$$

and by (3.11), there is a λ large enough so that $\|\nabla u\|_\infty \leqslant 1/2$ and

$$|x - y|/2 \leqslant |\Phi(x) - \Phi(y)| \leqslant 2|x - y|.$$

In particular, $\Phi : \mathbb{R}^d \to \mathbb{R}^d$ is a C^1-diffeomorphism and

$$\|\nabla\Phi\|_\infty \leqslant 2, \quad \|\nabla\Phi^{-1}\|_\infty \leqslant 2.$$

We have

Lemma 3.7 (Zvonkin's transformation) *X_t solves SDE (1.1) if and only if $Y_t := \Phi(X_t)$ solves*

$$\mathrm{d}Y_t = \sqrt{2}\Theta(Y_t)\mathrm{d}W_t + \lambda u \circ \Phi^{-1}(Y_t)\mathrm{d}t, \quad Y_0 = \Phi(x), \tag{3.12}$$

where $\Theta(y) := \nabla\Phi \circ \Phi^{-1}(y)$.

Proof By the generalized Itô's formula in Corollary 3.6, we have

$$\Phi(X_t) = \Phi(x) + \sqrt{2}\int_0^t \nabla\Phi(X_s)\mathrm{d}W_s + \lambda\int_0^t u(X_s)\mathrm{d}s.$$

So $Y_t = \Phi(X_t)$ solves (3.12). Similarly, one can show that if Y_t solves (3.12), then $X_t := \Phi^{-1}(Y_t)$ solves SDE (1.1) □

Now we are in the position to give

Proof of Theorem 3.1 By Lemma 3.7, it suffices to show the well-posedness of SDE (3.12). Observe that

$$\|\nabla(u \circ \Phi^{-1})\|_\infty \leqslant \|\nabla u\|_\infty \|\nabla \Phi^{-1}\|_\infty < \infty$$

and

$$\|\nabla \Theta\|_{p_1} \leqslant \|\nabla^2 \Phi\|_{p_1} \|\nabla \Phi^{-1}\|_\infty < \infty.$$

In particular, the drift in SDE (3.12) is Lipschitz continuous, while the diffusion coefficient Θ belongs to H^{1,p_1}. Since the coefficients are bounded and continuous, the existence of weak solutions follows by a standard weak convergence argument. By Yamada-Watanabe's theorem, we only need to show the pathwise uniqueness. Let Y_t and Y'_t be two solutions of SDE (3.12). By Itô's formula, we have

$$\begin{aligned}|Y_t - Y'_t|^2 = |Y_0 - Y'_0|^2 + 2\sqrt{2}\int_0^t \langle Y_s - Y'_s, \Theta(Y_s) - \Theta(Y'_s)\mathrm{d}W_s\rangle \\ + 2\int_0^t (|\Theta(Y_s) - \Theta(Y'_s)|^2 + \lambda\langle Y_s - Y'_s, \tilde{b}(Y_s) - \tilde{b}(Y'_s)\rangle)\mathrm{d}s,\end{aligned}$$

where $\tilde{b}(y) := u \circ \Phi^{-1}(y)$. Let

$$A_t := 2\int_0^t (|\Theta(Y_s) - \Theta(Y'_s)|^2 + \lambda\langle Y_s - Y'_s, \tilde{b}(Y_s) - \tilde{b}(Y'_s)\rangle)/|Y_s - Y'_s|^2 \mathrm{d}s$$

and for $R > 0$,

$$\tau_R := \inf\{s > 0 : A_s \geqslant R\}.$$

By stochastic Gronwall's inequality (2.17) with $q = 1/2$ and $p = 3/4$, we have

$$\mathbb{E}\left(\sup_{s \in [0, T\wedge\tau_R]} |Y_s - Y'_s|\right) \leqslant 3\left(\mathbb{E}|Y_0 - Y'_0|^2\right)^{1/2}\left(\mathbb{E}\mathrm{e}^{3A_{T\wedge\tau_R}}\right)^{1/6}.$$

In particular, letting $Y_0 = Y'_0$, we get

$$\mathbb{E}\left(\sup_{s \in [0, T\wedge\tau_R]} |Y_s - Y'_s|\right) = 0.$$

If we can show $\tau_R \to \infty$ as $R \to \infty$, then the uniqueness is proven. Clearly, it suffices to prove

$$A_t < \infty, \quad a.s., \quad \textit{for all } t > 0.$$

By (2.14), we have

$$A_t \leqslant C \int_0^t (M|\nabla\Theta|^2(Y_s) + M|\nabla\Theta|^2(Y_s') + M|\nabla\tilde{b}|(Y_s) + M|\nabla\tilde{b}|(Y_s'))\mathrm{d}s,$$

where $Mf(x) := \sup_{r>0} \fint_{B_r} f(x+z)\mathrm{d}z$. Since $p_1 > d \vee 2$, by Krylov's estimate and the L^p-boundedness of the maximal operator, we have

$$\mathbb{E}\int_0^t M|\nabla\Theta|^2(Y_s)\mathrm{d}s \leqslant C\|M|\nabla\Theta|^2\|_{p_1/2} \leqslant C\|\nabla\Theta\|_{p_1}^2 < \infty$$

and

$$\mathbb{E}\int_0^t M|\nabla\tilde{b}|(Y_s)\mathrm{d}s \leqslant C\|M|\nabla\tilde{b}|\,\|_\infty \leqslant C\|\nabla\tilde{b}\|_\infty < \infty.$$

Hence, $\mathbb{E}A_t < \infty$. Thus the proof is complete. □

Remark 3.8 Let $b_n = b * \varrho_n$ be the smoothing approximation of b. Consider the following approximating SDE:

$$\mathrm{d}X_t^n = \sqrt{2}\mathrm{d}W_t + b_n(X_t^n)\mathrm{d}t, \quad X_0^n = x.$$

In fact, we can show that X^n converges to the unique solution X (see [34]).

Remark 3.9 Although the above result is stated for time-independent b. The same idea also works for time-dependent b so that we can completely cover Krylov and Röckner's result (see [34]).

Remark 3.10 Here an open question is the well-posedness of SDE (3.1) when $b \in L^d(\mathbb{R}^d)$. Notice that when $d = 1$ and $b \in L^1_{loc}(\mathbb{R})$, it is well known there is a unique local strong solution (see [9]). However, when $d \geqslant 2$, it seems to be a hard problem (see [5] for some development about this problem).

4 SDEs with Distributional Drifts

In this section, we consider time-independent SDE (1.1) with drift $b \in H^{-1/2,p}$, where $p > 2d$. The main results of this section come from [35]. Let $\mathbb{C}$ be the space of all continuous functions from $\mathbb{R}_+$ to $\mathbb{R}^d$, which is endowed with the usual Borel σ-field

$\mathcal{B}(\mathbb{C})$. All the probability measures over $(\mathbb{C}, \mathcal{B}(\mathbb{C}))$ is denoted by $\mathscr{P}(\mathbb{C})$. Let w_t be the coordinate process over $\mathbb{C}$, that is,

$$w_t(\omega) = \omega_t, \quad \omega \in \mathbb{C}.$$

For $t \geqslant 0$, let $\mathcal{B}_t(\mathbb{C})$ be the natural filtration generated by $\{w_s : s \leqslant t\}$. For given $R > 0$, we shall use the following truncated $\mathcal{B}_t(\mathbb{C})$-stopping time

$$\tau_R := \inf\{t > 0 : |w_t| > R\}. \tag{4.1}$$

Notice that for each $\omega \in \mathbb{C}$, it automatically holds that

$$\lim_{R\to\infty} \tau_R(\omega) = \infty. \tag{4.2}$$

For a probability measure $\mathbb{P} \in \mathscr{P}(\mathbb{C})$, the expectation with respect to $\mathbb{P}$ will be denoted by $\mathbb{E}^{\mathbb{P}}$ or simply by $\mathbb{E}$ if there is no confusion.

We first introduce the following notion.

Definition 4.1 (*Local Krylov's estimate*) Let $\alpha \in [0, 1]$ and $p > 1$. We call a probability measure $\mathbb{P} \in \mathscr{P}(\mathbb{C})$ satisfy local Krylov's estimate with indices α, p if for any $T > 0$ and $R \geqslant 1$, there are positive constants $C_{T,R}$ and γ such that for all $f \in C^\infty$, $0 \leqslant t_0 < t_1 \leqslant T$ and $\tau \leqslant \tau_R$,

$$\mathbb{E}\left|\int_{t_0\wedge\tau}^{t_1\wedge\tau} f(w_s)\mathrm{d}s\right|^2 \leqslant C_{T,R}(t_1 - t_0)^{1+\gamma}\|f\chi_R\|^2_{-\alpha,p}. \tag{4.3}$$

If $C_{T,R}$ does not depend on R, then the above estimate will be called global Krylov's estimate. All the probability measure $\mathbb{P}$ with property (4.3) is denoted by $\mathscr{K}^\alpha_p(\mathbb{C})$.

From this definition, it is easy to see that

Proposition 4.2 *Let $\alpha \in [0, 1]$, $p > 1$ and $\mathbb{P} \in \mathscr{K}^\alpha_p(\mathbb{C})$. For any $f \in H^{-\alpha,p}_{loc}$, there is a continuous $\mathcal{B}_t(\mathbb{C})$-adapted process A^f_t such that for any mollifying approximation $f_n = f * \varrho_n$ and any $T > 0$,*

$$\lim_{n\to\infty} \mathbb{E}\left(\sup_{t\in[0,T]}\left|\int_0^t f_n(w_s)\mathrm{d}s - A^f_t\right| \wedge 1\right) = 0. \tag{4.4}$$

Moreover, for each $R \geqslant 1$, the mapping $H^{-\alpha,p} \ni f \mapsto A^f_{\cdot\wedge\tau_R} \in L^2(\mathbb{C}, \mathbb{P}; C([0, T]))$ is a bounded linear operator, where τ_R is defined in (4.1), and for all $0 \leqslant t_0 < t_1 \leqslant T$,

$$\mathbb{E}\left|A^f_{t_1\wedge\tau_R} - A^f_{t_0\wedge\tau_R}\right|^2 \leqslant C_{T,R}(t_1 - t_0)^{1+\gamma}\|f\chi_R\|^2_{-\alpha,p}, \tag{4.5}$$

where the constants $C_{T,R}$ and γ are the same as in (4.3).

Remark 4.3 (i) Estimate (4.5) implies that $t \mapsto A_t^f$ is a locally zero energy process, that is, for any $R \geqslant 1$,

$$\lim_{\delta \to 0} \sup_{\{\Pi_t : \mathrm{mesh}(\Pi_t) < \delta\}} \sum_{i=0}^{n} \mathbb{E}|A^f_{t_{i+1}\wedge\tau_R} - A^f_{t_i\wedge\tau_R}|^2 = 0,$$

where $\Pi_t := \{t_0, t_1, \dots, t_n\}$ denotes any partition of $[0, t]$.

(ii) If $f \in L^q_{loc}(\mathbb{R}^d)$ with $q \geqslant pd/(d + p\alpha)$, then $t \mapsto A_t^f$ is absolutely continuous and

$$A_t^f = \int_0^t f(w_s)\mathrm{d}s.$$

Indeed, it follows by Sobolev's embedding $L^q_{loc} \subset H^{-\alpha,p}_{loc}$.

Now we introduce the notion of martingale solutions.

Definition 4.4 (*Martingale Problem*) Let $\alpha \in [0, 1]$ and $p > 1$. We call a probability measure $\mathbb{P} \in \mathscr{K}_p^\alpha(\mathbb{C})$ a martingale solution of SDE (1.1) with starting point $x \in \mathbb{R}^d$ if for any $f \in C^\infty$,

$$M_t^f := f(w_t) - f(x) - \int_0^t (\mathscr{L}^\sigma f)(w_s)\mathrm{d}s - A_t^{b\cdot\nabla f} \tag{4.6}$$

is a continuous local $\mathcal{B}_t(\mathbb{C})$-martingale with $M_0^f = 0$ under $\mathbb{P}$, provided that $b \cdot \nabla f \in H^{-\alpha,p}_{loc}$, where $\mathscr{L}^\sigma f := \sigma^{ik}\sigma^{jk}\partial_i\partial_j f/2$. All the martingale solution $\mathbb{P} \in \mathscr{K}_p^\alpha(\mathbb{C})$ of SDE (1.1) with coefficients σ, b and starting point x is denoted by $\mathscr{M}^{\alpha,p}_{\sigma,b}(x)$.

As a direct consequence of martingale solutions and Lemma 2.3, we have

Lemma 4.5 (Generalized Itô's formula) *Let $\alpha \in (0, \frac{1}{2}]$, $p > \frac{d}{1-\alpha}$ and $\beta \in [\alpha, 1]$, $q \in (\frac{d}{\beta}, \infty)$. Suppose $\sigma \in H^{\beta,q}_{loc}$ and $b \in H^{-\alpha,p}_{loc}$. For any $f \in H^{2-\alpha,p}_{loc}$ and $\mathbb{P} \in \mathscr{M}^{\alpha,p}_{\sigma,b}(x)$,*

$$M_t^f := f(w_t) - f(x) - A_t^{(\mathscr{L}^\sigma + b\cdot\nabla)f}$$

is a continuous local $\mathcal{B}_t(\mathbb{C})$-martingale under $\mathbb{P}$.

Using this lemma and Proposition 2.4, we have the following Zvonkin's transformation as in Lemma 3.7.

Lemma 4.6 *Let $\alpha \in (0, \frac{1}{2}]$, $p > \frac{d}{1-\alpha}$ and $\beta \in [\alpha, 1]$, $q \in (\frac{d}{\beta}, \infty)$. Suppose that $\sigma \in H^{\beta,q}_{loc}$, $b \in H^{-\alpha,p}_{loc}$ and $\Phi \in \mathcal{D}_p^{1-\alpha}$. Define*

$$\tilde{\sigma} := (\nabla\Phi \cdot \sigma) \circ \Phi^{-1}, \quad \tilde{b} := (\mathscr{L}^\sigma \Phi + b \cdot \nabla\Phi) \circ \Phi^{-1}. \tag{4.7}$$

Then we have

(i) $\tilde{b} \in H_{loc}^{-\alpha,p}$ *and* $\tilde{\sigma} \in H_{loc}^{\beta',q'}$ *for* $\beta' := \beta \wedge (1-\alpha)$ *and*

$$\frac{1}{q'} := \begin{cases} \frac{1}{q} \vee \left(\frac{1}{p} - \frac{1-\alpha-\beta}{d}\right), & \beta \in [\alpha, 1-\alpha], \\ \frac{1}{p} \vee \left(\frac{1}{q} - \frac{\alpha+\beta-1}{d}\right), & \beta \in (1-\alpha, 1], \end{cases} \tag{4.8}$$

and also $q' > d/\beta'$.

(ii) *For any* $x \in \mathbb{R}^d$, *it holds that*

$$\mathbb{P} \in \mathscr{M}_{\sigma,b}^{\alpha,p}(x) \Leftrightarrow \mathbb{P} \circ \Phi^{-1} \in \mathscr{M}_{\tilde{\sigma},\tilde{b}}^{\alpha,p}(\Phi(x)). \tag{4.9}$$

Here $\mathbb{P} \circ \Phi^{-1}$ *means that for* $A \in \mathcal{B}(\mathbb{C})$, $\mathbb{P} \circ \Phi^{-1}(A) = \mathbb{P}(\{\omega : \Phi(w_\cdot(\omega)) \in A\})$.

Remark 4.7 The importance of (4.9) lies in the fact that if there is one and only one element in $\mathscr{M}_{\tilde{\sigma},\tilde{b}}^{\alpha,p}(\Phi(x))$, then there is automatically one and only one element in $\mathscr{M}_{\sigma,b}^{\alpha,p}(x)$. Moreover, the heat kernel estimates and ergodicity can also be derived by (4.9).

Next we introduce the notion of weak solutions and discuss the relationship between martingale solutions and weak solutions.

Definition 4.8 (*Weak solutions*) Let σ be locally bounded and $b \in H_{loc}^{-\alpha,p}$ for some $\alpha \in [0,1]$ and $p > 1$. Let (X, W) be two $\mathbb{R}^d$-valued continuous adapted processes on some filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbf{P})$. We call $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbf{P}; X, W)$ a weak solution of SDE (1.1) with starting point $x \in \mathbb{R}^d$ if W is an $\mathscr{F}_t$-Brownian motion and

$$X_t = x + \int_0^t \sigma(X_s)\mathrm{d}W_s + A_t^b, \quad forall t > 0, \quad \mathbf{P}-a.s., \tag{4.10}$$

where $A_t^b := \lim_{n\to\infty} \int_0^t b_n(X_s)\mathrm{d}s$ in the sense of u.c.p., and $b_n \in C^2(\mathbb{R}^d)$ is any approximation sequence of b so that for each $R > 0$,

$$\lim_{n\to\infty} \|(b_n - b)\chi_R\|_{-\alpha,p} = 0.$$

Here A_t^b does not depend on the choice of approximation sequence $b_n \in C^2(\mathbb{R}^d)$ of b.

We have the following equivalence.

Proposition 4.9 *Let* $\mathbb{P} \in \mathscr{P}(\mathbb{C})$ *satisfy that for any* $T, R > 0$ *and* $s, t \in [0, T]$,

$$\mathbb{E}|w_{t\wedge\tau_R} - w_{s\wedge\tau_R}|^2 \leqslant C_{T,R}|t-s|. \tag{4.11}$$

Let $\alpha \in [0,1]$ *and* $p > 1$. *Assume that* $b \in H_{loc}^{-\alpha,p}$ *and* σ, σ^{-1} *are locally bounded. Then* $\mathbb{P} \in \mathscr{M}_{\sigma,b}^{\alpha,p}(x)$ *if and only if there is a weak solution* $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbf{P}; X, W)$ *in the sense of Definition 4.8 so that* $\mathbf{P} \circ X^{-1} = \mathbb{P} \in \mathscr{K}_p^\alpha(\mathbb{C})$.

To state the main results, we make the following assumptions about σ and b:

($\mathbf{H}^{\sigma}_{\beta,q}$) $\|\Delta^{\beta/2}\sigma\|_q < \infty$ for some $\beta \in (0,1]$ and $q \in (\frac{d}{\beta}, \infty)$, and there is a constant $c_0 \geqslant 1$ such that

$$c_0^{-1}|\xi|^2 \leqslant |\sigma^*(x)\xi|^2 \leqslant c_0|\xi|^2, \quad \textit{for all } x, \xi \in \mathbb{R}^d. \tag{4.12}$$

($\mathbf{H}^{b}_{\alpha,p}$) $b = b^{(1)} + b^{(2)}$, where $b^{(1)}$ satisfies that for some $\vartheta \geqslant 0$ and $\kappa_0, \kappa_1, \kappa_2 > 0$,

$$\frac{\langle x, b^{(1)}(x)\rangle}{\sqrt{1+|x|^2}} \leqslant -\kappa_0|x|^{\vartheta} + \kappa_1, \quad |b^{(1)}(x)| \leqslant \kappa_2(1+|x|^{\vartheta}), \tag{4.13}$$

and $b^{(2)} \in H^{-\alpha,p}$ for some $\alpha \in (0, \frac{1}{2}]$ and $p \in (\frac{d}{1-\alpha}, \infty)$.

The following theorem is the main result in [35].

Theorem 4.10 *Let $\alpha \in (0, \frac{1}{2}]$, $p \in (\frac{d}{1-\alpha}, \infty)$ and $\beta \in [\alpha, 1]$, $q \in (\frac{d}{\beta}, \infty)$. Under ($\mathbf{H}^{\sigma}_{\beta,q}$) and ($\mathbf{H}^{b}_{\alpha,p}$), for any $x \in \mathbb{R}^d$, there exists a unique martingale solution $\mathbb{P}_x \in \mathscr{M}^{\alpha,p}_{\sigma,b}(x)$ to SDE (1.1). Moreover, letting $\mathbb{E}_x := \mathbb{E}^{\mathbb{P}_x}$, we have the following conclusions*:

(i) *For any $T > 0$ and $m \in \mathbb{N}$, there is a constant $C_T > 0$ such that for all $0 \leqslant t_0 < t_1 \leqslant T$,*

$$\mathbb{E}_x|w_{t_1} - w_{t_0}|^{2m} \leqslant C_T(t_1 - t_0)^m, \tag{4.14}$$

and for all $f \in H^{-\alpha,p}$,

$$\mathbb{E}_x\left|A^f_{t_1} - A^f_{t_0}\right|^{2m} \leqslant C_T(t_1 - t_0)^{(2-\alpha-\frac{d}{p})m}\|f\|^{2m}_{-\alpha,p}. \tag{4.15}$$

(ii) *If $\vartheta = 0$ in (4.13), then for any $\varphi \in H^{2-\alpha,p}$, $u(t,x) := P_t\varphi(x) := \mathbb{E}_x\varphi(w_t) \in L^p_{loc}(\mathbb{R}_+; H^{2-\alpha,p})$ uniquely solves the following Cauchy problem in $H^{-\alpha,p}$,*

$$\partial_t u = (\mathscr{L}^{\sigma} + b\cdot\nabla)u, \quad u(0) = \varphi. \tag{4.16}$$

Moreover, P_t admits a density $p_t(x,y)$ enjoying the following two-sided estimate: for some $c_1, c_2 \geqslant 1$ and all $t > 0$, $x, y \in \mathbb{R}^d$,

$$c_1^{-1}t^{-d/2}\mathrm{e}^{-c_2|x-y|^2/t} \leqslant p_t(x,y) \leqslant c_1 t^{-d/2}\mathrm{e}^{-c_2^{-1}|x-y|^2/t}, \tag{4.17}$$

and gradient estimate: for some $c_3, c_4 > 0$ and all $t > 0$, $x, y \in \mathbb{R}^d$,

$$|\nabla_x p_t(x,y)| \leqslant c_3 t^{-(d+1)/2}\mathrm{e}^{-c_4|x-y|^2/t}. \tag{4.18}$$

(iii) *If $\vartheta > 0$ in (4.13), then P_t admits a unique invariant probability measure $\mu(\mathrm{d}x) = \varrho(x)\mathrm{d}x$ with $\varrho \in H^{\gamma,r}$, where $\gamma \in (0, \beta\wedge(1-\alpha)]$ and $r \in (1, \frac{d}{d+\gamma-1})$.*

Proof We sketch the proof. By Lemma 2.3 and suitable freezing coefficient argument, one can show that there exists a constant $\lambda_0 > 0$ such that for all $\lambda \geqslant \lambda_0$, there is a unique $u = u_\lambda : \mathbb{R}^d \to \mathbb{R}^d$ belonging to $H^{2-\alpha,p}$ so that

$$(\mathscr{L}^\sigma - \lambda + b^{(2)} \cdot \nabla)u = -b^{(2)} \quad \text{in} \quad H^{-\alpha,p}.$$

By Sobolev's embedding (2.4), we can choose λ large enough so that

$$\|\nabla u\|_\infty \leqslant 1/2. \tag{4.19}$$

Now, define

$$\Phi(x) := x + u(x) : \mathbb{R}^d \to \mathbb{R}^d.$$

It is easy to see that

$$\begin{aligned} &\tfrac{1}{2}|x-y| \leqslant |\Phi(x) - \Phi(y)| \leqslant 2|x-y|, \\ &\|\mathbb{I} - \nabla\Phi\|_{1-\alpha,p} = \|\nabla u\|_{1-\alpha,p} \leqslant C\|b^{(2)}\|_{-\alpha,p}. \end{aligned} \tag{4.20}$$

Hence, $\Phi \in \mathcal{D}^{1-\alpha}_p$ (see (2.11) for a definition) and

$$\mathscr{L}^\sigma \Phi + b^{(2)} \cdot \nabla\Phi = \lambda u \quad \text{in } H^{-\alpha,p}. \tag{4.21}$$

Define

$$\tilde{\sigma} := (\sigma^* \cdot \nabla\Phi) \circ \Phi^{-1}, \quad \tilde{b} := (\lambda u + b^{(1)} \cdot \nabla\Phi) \circ \Phi^{-1}. \tag{4.22}$$

One can verify that for λ large enough, there are $\tilde{\kappa}_0, \tilde{\kappa}_1, \tilde{\kappa}_2 > 0$ such that for all $y \in \mathbb{R}^d$,

$$\frac{\langle y, \tilde{b}(y)\rangle}{\sqrt{1+|y|^2}} \leqslant -\tilde{\kappa}_0|y|^\vartheta + \tilde{\kappa}_1 \quad \text{and} \quad |\tilde{b}(y)| \leqslant \tilde{\kappa}_2(1 + |y|^\vartheta), \tag{4.23}$$

where ϑ is the same as in (4.13). Moreover, $\tilde{\sigma}$ satisfies $(\mathbf{H}^\sigma_{\beta',q'})$ with $\beta' = \beta \wedge (1-\alpha)$ and q' being defined by (4.8).

(i) By Lemma 4.6, it suffices to show that there is one and only one element in $\mathscr{M}^{\alpha,p}_{\tilde{\sigma},\tilde{b}}(\Phi(x))$. Since $\tilde{\sigma}$ is uniformly non-degenerate and bounded continuous, by (4.23), it is well known that there is a unique element in $\mathscr{M}^{\alpha,p}_{\tilde{\sigma},\tilde{b}}(\Phi(x))$, and (i) follows.

(ii) Since (4.17) and (4.18) are invariant under C^1-diffeomorphism transformation Φ, it follows by (1.3) and (1.4).

(iii) It follows by (4.23) and the classical Bogoliubov-Krylov's argument. More details are referred to [35]. □

As an easy corollary of Theorem 4.10 and Proposition 4.9, we have

Corollary 4.11 *Under the same assumptions of Theorem 5.16, there exists a unique weak solution* $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbf{P}; X, W)$ *for SDE (1.1) so that* $\mathbf{P}\circ X^{-1} \in \mathscr{K}_p^\alpha(\mathbb{C})$.

In the above corollary, we require that the law of weak solution satisfies the local Krylov estimate, that is, $\mathbf{P}\circ X^{-1} \in \mathscr{K}_p^\alpha(\mathbb{C})$. This is crucial when we use Zvonkin's transformation to show the uniqueness. Nevertheless, under an extra assumption, we can directly prove such a priori estimate for any weak solutions.

Theorem 4.12 *Let* $\alpha \in (0, \frac{1}{2})$, $p \in (\frac{d}{1/2-\alpha}, \infty)$ *and* $\beta \in [\alpha, 1]$, $q \in (\frac{d}{\beta}, \infty)$. *Under* $(\mathbf{H}^\sigma_{\beta,q})$ *and* $(\mathbf{H}^b_{\alpha,p})$, *for any* $x \in \mathbb{R}^d$, *there exists a unique weak solution to SDE (1.1) in the sense of Definition 4.8 so that for each* $T, R > 0$ *and* $s, t \in [0, T]$,

$$\mathbf{E}|A^b_{t\wedge\eta_R} - A^b_{s\wedge\eta_R}|^4 \leqslant C_{T,R}|t-s|^{2(2-\alpha-\frac{d}{p})}, \tag{4.24}$$

where $\eta_R := \inf\{t > 0 : |X_t| > R\}$. *Moreover,* $\mathbf{P}\circ X^{-1} \in \mathscr{K}_p^\alpha$ *and the conclusions in Theorem 4.10 still hold.*

5 SDEs with Measure-Valued Drifts

In this section, we study SDE (1.1) with drifts in some generalized Kato's class. In particular, some singular measure-valued drift is allowed, which extends the well-known results in [3]. Our proof looks much simpler than [3]. We believe that it should also work for the SDE driven by α-stable type noises.

5.1 Generalized Kato's Class of Radon Measures

In this section we introduce some generalized Kato's class of Radon measures. Let $\mathscr{R}$ be the set of signed Radon measures over $\mathbb{R}^d$, which is endowed with the vague convergence topology. For $\mu \in \mathscr{R}$, we use $|\mu|$ to denote the total variation measure of μ. Let f be a nonnegative real-valued function and $\mu \in \mathscr{R}$ a nonnegative Radon measure. We define

$$\mu * f(x) := f * \mu(x) := \int_{\mathbb{R}^d} f(x-y)\mu(\mathrm{d}y),$$

and for a measurable family of Radon measures $\mu_s : \mathbb{R} \to \mathscr{R}$ and $\delta, \lambda \geqslant 0$,

$$\begin{aligned} m^{(\alpha)}_{\lambda,\mu}(\delta) &:= \sup_{(t,x)\in\mathbb{R}^{d+1}} \int_{t-\delta}^t \mathrm{e}^{-\lambda(t-s)} \rho^{(\alpha)}_{t-s} * |\mu_s|(x)\mathrm{d}s \\ &= \sup_{(t,x)\in\mathbb{R}^{d+1}} \int_0^\delta \mathrm{e}^{-\lambda s} \rho^{(\alpha)}_s * |\mu_{t-s}|(x)\mathrm{d}s, \end{aligned}$$

where $\rho^{(\alpha)}_t$ is defined by (2.8). If $\lambda = 0$, we simply write

$$m^{(\alpha)}_\mu(\delta) := m^{(\alpha)}_{0,\mu}(\delta).$$

Below we list some important properties about $m^{(\alpha)}_{\lambda,\mu}(\delta)$ for later use.

Proposition 5.1 *Let $\alpha \in (0,2)$ and $\mu\colon \mathbb{R} \to \mathscr{R}$ be a measurable family of Radon measures.*

(i) *For any $p, q \in [1,\infty]$ with $\frac{d}{p} + \frac{2}{q} < 2 - \alpha$, there is a constant $C = C(d, p, q, \alpha) > 0$ such that for any $\mu_s(\mathrm{d}y) = f(s, y)\mathrm{d}y$ with $f \in \mathbb{L}^q_p := L^q(\mathbb{R}; L^p(\mathbb{R}^d))$,*

$$m^{(\alpha)}_{\lambda,\mu}(\delta) \leqslant C\Big(\delta^{2-\alpha-\frac{d}{p}-\frac{2}{q}} \wedge \lambda^{\left(\frac{d}{p}+\frac{2}{q}+\alpha-2\right)/2}\Big)\|f\|_{\mathbb{L}^q_p}.$$

(ii) *For any $\gamma > 0$, there is a constant $C = C(\gamma, \alpha, d) > 0$ such that*

$$t^{-\alpha/2}\rho^{(2)}_{\gamma t}(x) \leqslant C\rho^{(\alpha)}_t(x),\ t > 0, x \in \mathbb{R}^d. \tag{5.1}$$

(iii) *For any $\lambda \geqslant 0$, $\delta \mapsto m^{(\alpha)}_{\lambda,\mu}(\delta)$ is increasing on $(0,\infty)$, and for any bounded measurable f,*

$$m^{(\alpha)}_{\lambda,f\mu}(\delta) \leqslant \|f\|_\infty m^{(\alpha)}_{\lambda,\mu}(\delta),\quad m^{(\alpha)}_{\lambda,\mu*\varrho_n}(\delta) \leqslant m^{(\alpha)}_{\lambda,\mu}(\delta).$$

(iv) *For any $d > 2 - \alpha$, there is a constant $C = C(\alpha, d) \geqslant 1$ such that for any time-independent Radon measure μ and $\delta, \lambda \geqslant 0$,*

$$C^{-1}\mathrm{e}^{-\delta\lambda}\tilde m^{(\alpha)}_\mu(\delta^{1/2}) \leqslant m^{(\alpha)}_{\lambda,\mu}(\delta) \leqslant C\tilde m^{(\alpha)}_\mu(\delta^{1/2}), \tag{5.2}$$

where

$$\tilde m^{(\alpha)}_\mu(\delta) := \sup_{x\in\mathbb{R}^d}\int_{|x-y|\leqslant\delta}|x-y|^{2-\alpha-d}|\mu|(\mathrm{d}y).$$

(v) *For any $0 < \beta < \alpha < 2$, there is a constant $C = C(\alpha, \beta, d) > 0$ such that for any $\delta > 0$,*

$$m^{(\beta)}_{\lambda,\mu}(\delta) \leqslant C\delta^{(\alpha-\beta)/2}m^{(\alpha)}_{\lambda,\mu}(\delta). \tag{5.3}$$

Proof (i) Let $p' = \frac{p}{p-1}$, $q' = \frac{q}{q-1}$ and $\gamma = d + \alpha$. By Hölder's inequality we have

$$\begin{aligned}&\int_0^\delta \mathrm{e}^{-\lambda s}\rho^{(\alpha)}_s * |\mu_{t-s}|(x)\mathrm{d}s\\ &\quad\leqslant \int_0^\delta \mathrm{e}^{-\lambda s}\|\rho^{(\alpha)}_s\|_{p'}\|f_{t-s}\|_p\mathrm{d}s \leqslant \left(\int_0^\delta \|\mathrm{e}^{-\lambda s}\rho^{(\alpha)}_s\|^{q'}_{p'}\mathrm{d}s\right)^{1/q'}\|f\|_{\mathbb{L}^q_p}\end{aligned}$$

$$\leqslant C\left[\int_0^\delta \mathrm{e}^{-\lambda q' s}\left(s^{\frac{d-\gamma p'}{2}} + \int_{|x|>s^{1/2}} |x|^{-\gamma p'}\mathrm{d}x\right)^{q'/p'}\mathrm{d}s\right]^{1/q'} \|f\|_{\mathbb{L}_p^q}$$
$$\leqslant C\left(\delta^{2-\alpha-\frac{d}{p}-\frac{2}{q}} \wedge \lambda^{\left(\frac{d}{p}+\frac{2}{q}+\alpha-2\right)/2}\right)\|f\|_{\mathbb{L}_p^q}.$$

(ii) It follows by

$$t^{-(d+\alpha)/2}\mathrm{e}^{-|x|^2/(\gamma t)}(t^{1/2}+|x|)^{d+\alpha} \leqslant \sup_{r>0}(1+r)^{d+\alpha}\mathrm{e}^{-r^2/\gamma}.$$

(iii) The increasing of $\delta \mapsto m^{(\alpha)}_{\lambda,\mu}(\delta)$ and $m^{(\alpha)}_{\lambda,f\mu}(\delta) \leqslant \|f\|_\infty m^{(\alpha)}_{\lambda,\mu}(\delta)$ are direct by definition. Moreover, by the definition of convolution, it follows that

$$\left\|\int_0^\delta \mathrm{e}^{-\lambda s}\rho_s^{(\alpha)} * |\mu^n_{t-s}|\mathrm{d}s\right\|_\infty \leqslant \left\|\varrho_n * \int_0^\delta \mathrm{e}^{-\lambda s}\rho_s^{(\alpha)} * |\mu_{t-s}|\mathrm{d}s\right\|_\infty$$
$$\leqslant \left\|\int_0^\delta \mathrm{e}^{-\lambda s}\rho_s^{(\alpha)} * |\mu_{t-s}|\mathrm{d}s\right\|_\infty.$$

(iv) Since $\mathrm{e}^{-\lambda\delta} \leqslant \mathrm{e}^{-\lambda s} \leqslant 1$ for $s \in [0,\delta]$, without loss of generality we assume $\lambda = 0$. Notice that

$$\int_0^\delta \rho_s^{(\alpha)}(x)\mathrm{d}s = |x|^{2-\alpha-d}\int_0^{\delta/|x|^2}(s^{1/2}+1)^{-d-\alpha}\mathrm{d}s =: |x|^{2-\alpha-d}g_\delta(|x|).$$

Since $g_\delta(r) \leqslant g_\infty(1) < \infty$ for $r^2 < \delta$ and $g_\delta(r) \leqslant g_\gamma(1) \leqslant \gamma$ for $r^2 > \delta/\gamma$, we have

$$\int_0^\delta \rho_s^{(\alpha)} * |\mu|(x)\mathrm{d}s \leqslant \int_{\mathbb{R}^d} |x-y|^{2-\alpha-d}g_\delta(|x-y|)\mu(\mathrm{d}y)$$
$$= \int_{|x-y|^2\leqslant\delta} \frac{g_\delta(|x-y|)}{|x-y|^{d+\alpha-2}}\mu(\mathrm{d}y)$$
$$+ \sum_{k=0}^\infty \int_{2^k\delta<|x-y|^2\leqslant 2^{k+1}\delta} \frac{g_\delta(|x-y|)}{|x-y|^{d+\alpha-2}}\mu(\mathrm{d}y)$$
$$\leqslant g_\infty(1)\tilde{m}^{(\alpha)}_\mu(\delta^{1/2})$$
$$+ \sum_{k=0}^\infty 2^{-k}(2^k\delta)^{(2-\alpha-d)/2}\int_{2^k\delta<|x-y|^2\leqslant 2^{k+1}\delta}\mu(\mathrm{d}y).$$

For the annulus $\mathcal{C}^k_\delta := \{y : 2^k\delta < |x-y|^2 \leqslant 2^{k+1}\delta\}$, there are at most $N2^{kd/2}$ balls with radius $\sqrt{\delta}$ and centers $\{x_i, i = 1, \ldots, N2^{kd/2}\}$, where $N = N(d) \in \mathbb{N}$, such that

$$\mathcal{C}^k_\delta \subset \bigcup_{i=1}^{N2^{kd/2}} \{y : |y-x_i|^2 \leqslant \delta\}.$$

Hence,

$$\mu(\mathcal{C}_\delta^k) \leqslant \sum_{i=1}^{N2^{kd/2}} \int_{|y-x_i|^2\leqslant\delta} \mu(\mathrm{d}y) \leqslant N2^{kd/2}\delta^{(d+\alpha-2)/2}\tilde{m}_\mu^{(\alpha)}(\delta^{1/2}).$$

Thus we get

$$m_\mu^{(\alpha)}(\delta) \leqslant C\tilde{m}_\mu^{(\alpha)}(\delta^{1/2}) + \sum_{k=0}^{\infty} 2^{-k}2^{k(2-\alpha)/2}\tilde{m}_\mu^{(\alpha)}(\delta^{1/2}) \leqslant C\tilde{m}_\mu^{(\alpha)}(\delta^{1/2}).$$

On the other hand, we clearly have

$$1_{\{r^2\leqslant\delta\}} \leqslant \int_0^{\delta/r^2} (s^{1/2}+1)^{-d-\alpha}\mathrm{d}s \Big/ g_1(1),$$

which implies that

$$\tilde{m}_\mu^{(\alpha)}(\delta^{1/2}) \leqslant Cm_\mu^{(\alpha)}(\delta).$$

(v) Clearly, by definition,

$$\int_0^\delta\int_{|x-y|^2\leqslant\delta} \mathrm{e}^{-\lambda s}\rho_s^{(\beta)}(x-y)\mu_{t-s}(\mathrm{d}y)\mathrm{d}s \leqslant (2\delta)^{(\alpha-\beta)/2}m_{\lambda,\mu}^{(\alpha)}(\delta). \tag{5.4}$$

We now estimate

$$\begin{aligned}
&\int_0^\delta\int_{|x-y|^2>\delta} \mathrm{e}^{-\lambda s}\rho_s^{(\beta)}(x-y)\mu_{t-s}(\mathrm{d}y)\mathrm{d}s \\
&\quad= \sum_{k=0}^{\infty}\int_0^\delta\int_{2^k\delta<|x-y|^2\leqslant 2^{k+1}\delta} \mathrm{e}^{-\lambda s}\rho_s^{(\beta)}(x-y)\mu_{t-s}(\mathrm{d}y)\mathrm{d}s \\
&\quad\leqslant \sum_{k=0}^{\infty}(2^k\delta)^{-(d+\beta)/2}\int_0^\delta \mathrm{e}^{-\lambda s}\int_{2^k\delta<|x-y|^2\leqslant 2^{k+1}\delta} \mu_{t-s}(\mathrm{d}y)\mathrm{d}s.
\end{aligned}$$

As above, we have

$$\begin{aligned}
&\int_0^\delta \mathrm{e}^{-\lambda s}\int_{2^k\delta<|x-y|^2\leqslant 2^{k+1}\delta} \mu_{t-s}(\mathrm{d}y)\mathrm{d}s \\
&\quad\leqslant \sum_{i=0}^{N2^{kd/2}}\int_0^\delta \mathrm{e}^{-\lambda s}\int_{|y-x_i|^2\leqslant\delta} \mu_{t-s}(\mathrm{d}y)\mathrm{d}s
\end{aligned}$$

$$\leqslant \delta^{(d+\alpha)/2} \sum_{i=0}^{N2^{kd/2}} \int_0^\delta \int_{|y-x_i|^2 \leqslant \delta} e^{-\lambda s} \rho_s^{(\alpha)}(x_i - y)\mu_{t-s}(dy)ds$$
$$\leqslant \delta^{(d+\alpha)/2} N2^{kd/2} m_{\lambda,\mu}^{(\alpha)}(\delta).$$

Hence,

$$\int_0^\delta \int_{|x-y|^2 > \delta} e^{-\lambda s} \rho_s^{(\beta)}(x-y)\mu_{t-s}(dy)ds \leqslant \delta^{(\alpha-\beta)/2} N \sum_{k=0}^{\infty} 2^{-k\beta/2} m_{\lambda,\mu}^{(\alpha)}(\delta),$$

which together with (5.4) yields (5.3). □

Definition 5.2 For $\alpha \in (0, 2]$, the generalized Kato's class of Radon measures is defined by

$$\mathbb{K}_\alpha := \left\{ \mu : \mathbb{R} \to \mathscr{R} \text{ satisfies } \lim_{\delta \to 0} m_\mu^{(\alpha)}(\delta) = 0 \right\}.$$

Moreover, we also introduce

$$\mathbb{K}'_\alpha := \left\{ \mu : \mathbb{R} \to \mathscr{R} \text{ satisfies } m_\mu^{(\alpha)}(\delta) < \infty \text{ for some } \delta > 0 \right\}$$

and

$$\mathbb{K}''_\alpha := \left\{ \mu \in \mathbb{K}_\alpha \text{ satisfies } \lim_{\lambda \to \infty} m_{\lambda,\mu}^{(\alpha)}(\delta) = 0 \text{ for some } \delta > 0 \right\}.$$

Clearly,

$$\mathbb{K}''_\alpha \subset \mathbb{K}_\alpha \subset \mathbb{K}'_\alpha.$$

Examples of singular measures in $\mathbb{K}_\alpha$ are referred to [3]. The following lemma is easy.

Lemma 5.3 *Let $\mu \in \mathbb{K}'_\alpha$ be time-independent and $\mu_n(x) = \mu * \varrho_n(x)$. Then for any $m \in \mathbb{N}$, it holds that*

$$\|\nabla_x^m \mu_n\|_\infty \leqslant 2^{d+\alpha} \|\nabla_x^m \varrho_n\|_\infty m_\mu^{(\alpha)}(1).$$

Proof By definition, we have

$$|\nabla_x^m \mu_n|(x) \leqslant \int_{\mathbb{R}^d} |\nabla_x^m \varrho_n(x-y)|\, |\mu|(dy)$$
$$= \int_{\mathbb{R}^d} |\nabla_x^m \varrho_n(x-y)| \frac{\rho_s^{(\alpha)}(x-y)}{\rho_s^{(\alpha)}(x-y)} |\mu|(dy)$$
$$\leqslant 2^{d+\alpha} \|\nabla_x^m \varrho_n\|_\infty \int_{\mathbb{R}^d} \rho_s^{(\alpha)}(x-y)\, |\mu|(dy), \;\; s \in (0,1).$$

Integrating both sides with respect to s from 0 to 1, we obtain the desired estimate. □

5.2 Solvability of PDEs with Measure-Valued Datas

In the following, we always assume that σ and b are time-independent. For a Radon measure μ and $\lambda \geqslant 0$, we define

$$P_t^\lambda \mu(x) := \int_{\mathbb{R}^d} \mathrm{e}^{-\lambda t} p_t(x, y)\mu(\mathrm{d}y), \tag{5.5}$$

where $p_t(x, y)$ is the fundamental solution of operator $\mathscr{L}^\sigma$ (see (1.2)). The following lemma plays a crucial role in solving the PDE:

Lemma 5.4 *For any $\mu \in \mathbb{K}'_1$, there is a constant $C = C(\sigma, d) > 0$ such that for any $\lambda, t \geqslant 0$,*

$$\left\|\nabla u_\mu^\lambda(t)\right\|_\infty \leqslant C\, m_{\lambda,\mu}^{(1)}(t), \tag{5.6}$$

where

$$u_\mu^\lambda(t, x) := \int_0^t P_{t-s}^\lambda \mu_s(x)\mathrm{d}s.$$

For $\alpha \in [1, 2)$ and $\mu \in \mathbb{K}'_\alpha$, there is a constant $C = C(\alpha, d, \sigma) > 0$ such that for all $t, \delta, \lambda \geqslant 0$ and $x, y \in \mathbb{R}^d$,

$$\begin{aligned}&\left|\nabla u_\mu^\lambda(t, x) - \nabla u_\mu^\lambda(t, y)\right| \\ &\quad\leqslant C\left((\delta \wedge t)^{\frac{\alpha-1}{2}} m_{\lambda,\mu}^{(\alpha)}((\delta \wedge t)) + |x - y|(\delta \wedge t)^{\frac{\alpha-2}{2}} m_{\lambda,\mu}^{(\alpha)}(t)\right).\end{aligned} \tag{5.7}$$

Proof (i) By (1.4) and (5.1), we have

$$|\nabla u_\mu^\lambda(t, x)| = \left|\int_0^t \nabla P_{t-s}^\lambda \mu_s(x)\mathrm{d}s\right| \lesssim \left|\int_0^t \mathrm{e}^{-\lambda(t-s)} \rho_{t-s}^{(1)} * |\mu_s|(x)\mathrm{d}s\right|. \tag{5.8}$$

Thus (5.6) follows by definition.
(ii) First of all, if $t \leqslant \delta$, then by (5.6) and (5.3),

$$\begin{aligned}&|\nabla u_\mu^\lambda(t, x) - \nabla u_\mu^\lambda(t, y)| \\ &= \left|\int_0^t \nabla P_{t-s}^\lambda \mu_s(x)\mathrm{d}s - \int_0^t \nabla P_{t-s}^\lambda \mu_s(y)\mathrm{d}s\right| \lesssim m_{\lambda,\mu}^{(1)}(t) \lesssim t^{\frac{\alpha-1}{2}} m_{\lambda,\mu}^{(\alpha)}(t).\end{aligned} \tag{5.9}$$

Suppose now $\delta < t$. We write

$$
\begin{aligned}
&\left|\int_0^t \nabla P^\lambda_{t-s}\mu_s(x)\mathrm{d}s - \int_0^t \nabla P^\lambda_{t-s}\mu_s(y)\mathrm{d}s\right| \\
&\quad \leqslant \left|\int_{t-\delta}^t \nabla P^\lambda_{t-s}\mu_s(x)\mathrm{d}s\right| + \left|\int_{t-\delta}^t \nabla P^\lambda_{t-s}\mu_s(y)\mathrm{d}s\right| \\
&\qquad + \left|\int_0^{t-\delta} \nabla P^\lambda_{t-s}\mu_s(x)\mathrm{d}s - \int_0^{t-\delta} \nabla P^\lambda_{t-s}\mu_s(y)\mathrm{d}s\right|.
\end{aligned}
\tag{5.10}
$$

As in (5.8), by (1.4), (5.1) and (5.3), we have

$$
\left|\int_{t-\delta}^t \nabla P^\lambda_{t-s}\mu_s(x)\mathrm{d}s\right| \lesssim \left|\int_{t-\delta}^t \mathrm{e}^{-\lambda(t-s)}\rho^{(1)}_{t-s} * |\mu_s|(x)\mathrm{d}s\right| \leqslant m^{(1)}_{\lambda,\mu}(\delta) \lesssim \delta^{\frac{\alpha-1}{2}} m^{(\alpha)}_{\lambda,\mu}(\delta)
\tag{5.11}
$$

and

$$
\begin{aligned}
&\left|\int_0^{t-\delta} \nabla P^\lambda_{t-s}\mu_s(x)\mathrm{d}s - \int_0^{t-\delta} \nabla P^\lambda_{t-s}\mu_s(y)\mathrm{d}s\right| \\
&\quad \leqslant |x-y| \left\|\nabla \int_0^{t-\delta} \nabla P^\lambda_{t-s}\mu_s \mathrm{d}s\right\|_\infty \\
&\quad \leqslant |x-y| \left\|\int_0^{t-\delta} \mathrm{e}^{-\lambda(t-s)}(t-s)^{-1}\rho^{(2)}_{\kappa_2(t-s)} * |\mu_s|\mathrm{d}s\right\|_\infty \\
&\quad \lesssim |x-y|\,\delta^{\frac{\alpha-2}{2}} \left\|\int_0^{t-\delta} \mathrm{e}^{-\lambda(t-s)}\rho^{(\alpha)}_{t-s} * |\mu_s|\mathrm{d}s\right\|_\infty \leqslant |x-y|\delta^{\frac{\alpha-2}{2}} m^{(\alpha)}_{\lambda,\mu}(t).
\end{aligned}
\tag{5.12}
$$

Combining (5.9)–(5.12), we obtain (5.7). □

For $b, f \in \mathbb{K}_1$, we consider the following PDE:

$$
\partial_t u = (\mathscr{L}^\sigma - \lambda)u + b\cdot\nabla u + f,\ u(0) = 0,
\tag{5.13}
$$

where $b\cdot\nabla u$ is understood as the measure $\sum_{i=1}^d \partial_i u(x) b^i(\mathrm{d}x)$. We introduce the following Banach space of continuous functions

$$
\mathbf{C}^{0,1}_T := \{f, \nabla f \text{ are continuous on } [0,T]\times\mathbb{R}^d \to \mathbb{R}^d\},
$$

which is endowed with the uniform norm:

$$
\|f\|_{\mathbf{C}^{0,1}_T} := \|f\|_{\mathbb{L}^\infty_T} + \|\nabla f\|_{\mathbb{L}^\infty_T},
$$

where for a space-time function f,

$$
\|f\|_{\mathbb{L}^\infty_T} := \sup_{t\in[0,T]} \sup_{x\in\mathbb{R}^d} |f(t,x)|.
$$

We call a function $u \in \mathbf{C}^{0,1}_T$ a mild solution of PDE (5.13) if u satisfies

$$u(t,x) = \int_0^t P^\lambda_{t-s}(b\cdot\nabla u)(s,x)\mathrm{d}s + \int_0^t P^\lambda_{t-s} f(s,x)\mathrm{d}s, \quad \forall (t,x)\in[0,T]\times\mathbb{R}^d, \tag{5.14}$$

where P^λ_t is defined by (5.5). We now establish the following main result of this section.

Theorem 5.5 *Under* ($\mathbf{H}^\sigma_\beta$) *and $b \in \mathbb{K}_1$ being time-independent, for any $f \in \mathbb{K}'_1$ and $\lambda \geqslant 0$, there are $T = T(\sigma, d, m^{(1)}_b)$ small enough and a unique mild solution $u \in \mathbf{C}^{0,1}_T$ to* (5.14) *with*

$$\|u\|_{\mathbf{C}^{0,1}_T} \leqslant C m^{(1)}_{\lambda,f}(T), \tag{5.15}$$

where $C = C(\sigma, d) > 0$ is independent of λ, T. Moreover, we have the following conclusions:

(a) *There is a constant $C = C(\sigma, d, m^{(1)}_b(T), m^{(1)}_f(T)) > 0$ such that for all $t \in [0,T]$ and $x, y \in \mathbb{R}^d$,*

$$|\nabla u(t,x) - \nabla u(t,y)| \leqslant C\left(m^{(1)}_{\lambda,b}(|x-y|) + m^{(1)}_{\lambda,f}(|x-y|) + |x-y|^{\frac{1}{2}}\right). \tag{5.16}$$

(b) *let b_n and f_n be the mollifying approximation of b and f respectively. We have*

$$\lim_{n\to\infty} \|u_n - u\|_{\mathbf{C}^{0,1}_T} = 0. \tag{5.17}$$

(c) *There is a constant $C = C(\sigma, d, m^{(1)}_b(T)) > 0$ such that for all $n \in \mathbb{N}$ and $\delta \in (0,T)$,*

$$\|u_{f_n} - u_f\|_{\mathbf{C}^{0,1}_T} \leqslant C\left(m^{(1)}_{\lambda,f}(\delta) + n^{-\gamma}\delta^{-\gamma/2} m^{(1)}_{\lambda,f}(T)\right), \tag{5.18}$$

where u_f stands for the solution of (5.14) *with nonhomogeneous term f.*

Proof It suffices to prove (5.15)–(5.17) since the existence and uniqueness are easily derived from the above a priori estimates. Let $T > 0$ be fixed and whose value will be determined below.

(i) By (5.6) we have for all $t \in [0,T]$,

$$\|\nabla u(t)\|_\infty \leqslant C_0\left(m^{(1)}_{\lambda, b\cdot\nabla u}(T) + m^{(1)}_{\lambda,f}(T)\right) \leqslant C_0\|\nabla u\|_{\mathbb{L}^\infty_T} m^{(1)}_{0,b}(T) + C_0 m^{(1)}_{\lambda,f}(T).$$

Since $b \in \mathbb{K}_1$, we can choose T small enough so that

$$C_0 m^{(1)}_{0,b}(T) \leqslant 1/2 \tag{5.19}$$

and

$$\|\nabla u\|_{\mathbb{L}^\infty_T} \leqslant 2C_0 m^{(1)}_{\lambda,f}(T).$$

Thus we obtain (5.15).

(ii) By (5.7) with $\alpha = 1$ and $\delta = |x-y|$, there is a $C = C(\sigma, d) > 0$ such that for all $t \in [0, T]$ and $x, y \in \mathbb{R}^d$,

$$|\nabla u(t,x) - \nabla u(t,y)| \leqslant C\left(m^{(1)}_{\lambda, b\cdot\nabla u+f}(|x-y|) + |x-y|^{\frac{1}{2}} m^{(1)}_{\lambda, b\cdot\nabla u+f}(T)\right).$$

Notice that

$$m^{(1)}_{\lambda, b\cdot\nabla u+f}(\delta) \leqslant \|\nabla u\|_{\mathbb{L}^\infty_T} m^{(1)}_{\lambda,b}(\delta) + m^{(1)}_{\lambda,f}(\delta),\ \ \delta > 0.$$

The estimate (5.16) now follows by (5.15).

(iii) Let u_n satisfy the following integral equation

$$u_n(t,x) = \int_0^t P^\lambda_{t-s}(b_n \cdot \nabla u_n)(s,x)\mathrm{d}s + \int_0^t P^\lambda_{t-s} f_n(s,x)\mathrm{d}s.$$

We have

$$\begin{aligned}\|\nabla u_n(t) - \nabla u(t)\|_\infty \leqslant &\left\|\int_0^t \nabla P^\lambda_{t-s}(b_n\cdot\nabla(u_n-u))(s)\mathrm{d}s\right\|_\infty \\ &+ \left\|\int_0^t \nabla P^\lambda_{t-s}((b_n-b)\cdot\nabla u)(s)\mathrm{d}s\right\|_\infty \\ &+ \left\|\int_0^t \nabla P^\lambda_{t-s}(f_n-f)(s)\mathrm{d}s\right\|_\infty =: I_1 + I_2 + I_3.\end{aligned}$$

For I_1, by (5.6) we obviously have

$$I_1 \leqslant C_0 m^{(1)}_{\lambda,b_n}(T)\|\nabla u_n - \nabla u\|_{\mathbb{L}^\infty_T} \leqslant C_0 m^{(1)}_{0,b}(T)\|\nabla u_n - \nabla u\|_{\mathbb{L}^\infty_T}.$$

Next we treat I_2. For $\delta \in (0,t)$, we make the following decomposition:

$$\begin{aligned}&\int_0^t \nabla P^\lambda_{t-s}((b^i_n - b^i)\cdot\partial_i u)(s,x)\mathrm{d}s \\ &= \int_{t-\delta}^t \nabla P^\lambda_{t-s}((b^i_n - b^i)\cdot\partial_i u)(s,x)\mathrm{d}s \\ &\quad + \int_0^{t-\delta}\left(\mathrm{e}^{-\lambda(t-s)}\int_{\mathbb{R}^d}(\nabla_x p_{t-s}(x,y) - \nabla_x p_{t-s}(x,z))\partial_i u(s,y)\varrho_n(y-z)\mathrm{d}y\right) b^i(\mathrm{d}z)\mathrm{d}s \\ &\quad + \int_0^{t-\delta}\left(\mathrm{e}^{-\lambda(t-s)}\int_{\mathbb{R}^d}\nabla_x p_{t-s}(x,z)(\partial_i u(s,y) - \partial_i u(s,z))\varrho_n(y-z)\mathrm{d}y\right) b^i(\mathrm{d}z)\mathrm{d}s \\ &=: J^{(1)}_{\delta,n}(t,x) + J^{(2)}_{\delta,n}(t,x) + J^{(3)}_{\delta,n}(t,x).\end{aligned}$$

For $J^{(1)}_{\delta,n}$, by (1.4) and (5.1), we have

$$\|J^{(1)}_{\delta,n}(t)\|_\infty \lesssim \|\nabla u\|_{\mathbb{L}^\infty_T} \left\| \int_{t-\delta}^t \mathrm{e}^{-\lambda(t-s)} \rho^{(1)}_{t-s} * (|b_n| + |b|)\mathrm{d}s \right\|_\infty \lesssim \|\nabla u\|_{\mathbb{L}^\infty_T} m^{(1)}_{\lambda,b}(\delta).$$

For $J^{(2)}_{\delta,n}$, noticing that by (1.5) and (5.1),

$$\begin{aligned}&\int_{\mathbb{R}^d} |\nabla_x p_{t-s}(x,y) - \nabla_x p_{t-s}(x,z)|\,|\nabla u(s,y)|\varrho_n(y-z)\mathrm{d}y \\ &\quad \lesssim n^{-\gamma}(t-s)^{-\gamma/2}\|\nabla u(s)\|_\infty \int_{\mathbb{R}^d} (\rho^{(1)}_{t-s}(x-y) + \rho^{(1)}_{t-s}(x-z))\varrho_n(y-z)\mathrm{d}y,\end{aligned}$$

we have

$$\begin{aligned}\|J^{(2)}_{\delta,n}(t)\|_\infty &\lesssim \|\nabla u\|_{\mathbb{L}^\infty_T} n^{-\gamma}\delta^{-\gamma/2} \left\| \int_0^{t-\delta} \mathrm{e}^{-\lambda(t-s)} \Big(\rho^{(1)}_{t-s} * \varrho_n * |b| + \rho^{(1)}_{t-s} * |b|\Big)\mathrm{d}s \right\|_\infty \\ &\leqslant 2\|\nabla u\|_{\mathbb{L}^\infty_T} n^{-\gamma}\delta^{-\gamma/2} m^{(1)}_{\lambda,b}(t).\end{aligned}$$

For $J^{(3)}_{\delta,n}$, we have

$$\|J^{(3)}_{\delta,n}(t)\|_\infty \lesssim \sup_{s\in[0,T]} \sup_{|y-z|\leqslant 1/n} |\nabla u(s,y) - \nabla u(s,z)| \cdot m^{(1)}_{\lambda,b}(t).$$

Combining the above calculations, we obtain

$$\begin{aligned}I_2 &\leqslant \|\nabla u\|_{\mathbf{C}^0_T} \left(m^{(1)}_{\lambda,b}(\delta) + n^{-\gamma}\delta^{-\gamma/2}\right) \\ &\quad + \sup_{s\in[0,T]} \sup_{|y-z|\leqslant 1/n} |\nabla u(s,y) - \nabla u(s,z)| \cdot m^{(1)}_{\lambda,b}(T).\end{aligned}$$

For I_3, we similarly have

$$I_3 \leqslant C\left(m^{(1)}_{\lambda,f}(\delta) + n^{-\gamma}\delta^{-\gamma/2} m^{(1)}_{\lambda,f}(T)\right).$$

Hence, by (5.19),

$$\begin{aligned}\|\nabla u_n - \nabla u\|_{\mathbb{L}^\infty_T} &\lesssim \left(m^{(1)}_f(\delta) + n^{-\gamma}\delta^{-\gamma/2} m^{(1)}_{\lambda,f}(T)\right) \\ &\quad + \|\nabla u\|_{\mathbb{L}^\infty_T} \left(m^{(1)}_{\lambda,b}(\delta) + n^{-\gamma}\delta^{-\gamma/2}\right) \\ &\quad + \sup_{s\in[0,T]} \sup_{|y-z|\leqslant 1/n} |\nabla u(s,y) - \nabla u(s,z)| \cdot m^{(1)}_{\lambda,b}(T),\end{aligned}$$

which gives (5.17) by first letting $n \to \infty$ and then $\delta \to 0$. Moreover, it is easier to show

$$\lim_{n\to\infty} \|u_n - u\|_{\mathbb{L}^\infty_T} = 0.$$

(iv) Finally, for (5.18), it is similar. □

Remark 5.6 If we assume $b, f \in \mathbb{K}'_\alpha$ with $\alpha \in (1,2)$, then by using (5.7) with $\delta = |x-y|^2$, we can improve the estimate (5.16) as

$$|\nabla u(t,x) - \nabla u(t,y)| \leqslant C|x-y|^{\alpha-1}, \quad |x-y| \leqslant 1. \tag{5.20}$$

We also have the following solvability of elliptic equation.

Theorem 5.7 *Under* $(\mathbf{H}^\sigma_\beta)$ *and* $b \in C^1_b(\mathbb{R}^d) \cap \mathbb{K}_1$ *there is a* $\lambda_0 = \lambda_0(\sigma, d, m^{(1)}_b) \geqslant 1$ *large enough such that for all* $\lambda \geqslant \lambda_0$ *and* $f \in C^1_b(\mathbb{R}^d) \cap \mathbb{K}_1$, *there is a unique* $u \in C^2_b(\mathbb{R}^d)$ *such that*

$$(\mathscr{L}^\sigma - \lambda + b\cdot\nabla)u = f, \tag{5.21}$$

and

$$\|u\|_{C^1_b} \leqslant C m^{(1)}_{\lambda,f}(T), \tag{5.22}$$

where $C = C(\sigma, d) > 0$ *and* $T = T(\sigma, d, m^{(1)}_b) > 0$ *are independent of* λ. *Moreover,*

$$|\nabla u(x) - \nabla u(y)| \leqslant \ell(|x-y|), \tag{5.23}$$

where $\ell : \mathbb{R}_+ \to \mathbb{R}_+$ *with* $\ell(0) = 0$ *is a continuous increasing function only depending on* $\lambda, \sigma, d, m^{(1)}_{\lambda,b}, m^{(1)}_{\lambda,f}$.

Proof Since $b, f \in C^1_b(\mathbb{R}^d)$ and $(\mathbf{H}^\sigma_\beta)$ is satisfied, for any $\lambda > 0$, the existence and uniqueness of $u \in C^2_b(\mathbb{R}^d)$ to PDE (5.21) is classical. We only need to prove estimates (5.22) and (5.23). Let $T > 0$ and $\phi : \mathbb{R} \to \mathbb{R}$ be a nonzero smooth function with compact support in $(0, T)$. Let $u \in C^2_b(\mathbb{R}^d)$ solve (5.21). Then $\bar u(t,x) := u(x)\phi(t)$ satisfies the following parabolic equation in $[0, T]$,

$$\partial_t \bar u = (\mathscr{L}^\sigma - \lambda + b\cdot\nabla)\bar u - f\phi + u\phi'.$$

Since $b \in \mathbb{K}_1$, by (5.15), there are $T = T(\sigma, d, m^{(1)}_b) > 0$ small enough and $C = C(\sigma, d) > 0$ such that for all $\lambda \geqslant 0$,

$$\|u\|_{C^1_b}\|\phi\|_\infty = \|\bar u\|_{\mathbf{C}^{0,1}_T} \leqslant C m^{(1)}_{\lambda, f\phi+u\phi'}(T) \leqslant C\Big(\|\phi\|_\infty m^{(1)}_{\lambda,f}(T) + \|u\|_\infty m^{(1)}_{\lambda,\phi'}(T)\Big).$$

Since $\lim_{\lambda\to\infty} m^{(1)}_{\lambda,\phi'}(T) = 0$, one may choose λ large enough so that $Cm^{(1)}_{\lambda,\phi'}(T) = \frac{\|\phi\|_\infty}{2}$. Thus, we obtain the desired estimate (5.22). On the other hand, by (5.16) we have

$$
\begin{aligned}
|\phi(t)|\,|\nabla u(x) - \nabla u(y)| &= |\nabla u(t,x) - \nabla u(t,y)| \\
&\leqslant C\left(m^{(1)}_{\lambda,b}(|x-y|) + m^{(1)}_{\lambda,f\phi+u\phi'}(|x-y|) + |x-y|^{\frac{1}{2}}\right),
\end{aligned}
$$

which implies (5.23). □

Remark 5.8 If we assume $b, f \in C^1_b(\mathbb{R}^d) \cap \mathbb{K}'_\alpha$ for some $\alpha \in (1,2)$, then by (5.20) we have

$$
|\nabla u(x) - \nabla u(y)| \leqslant C|x-y|^{\alpha-1}, \quad |x-y| \leqslant 1. \tag{5.24}
$$

5.3 SDE with Measure-Valued Drifts

Now we consider SDE (1.1) with $b \in \mathbb{K}_1$ being a Radon measure. We first introduce the following definition of martingale solutions.

Definition 5.9 (*Martingale solution*) Given $x \in \mathbb{R}^d$, we call a probability measure $\mathbb{P}_x \in \mathscr{P}(\mathbb{C})$ a martingale solution of SDE (1.1) with starting point $x \in \mathbb{R}^d$ if

(i) There is a continuous finite variation process A defined on $\mathbb{C}$ such that for any $t > 0$,

$$
\lim_{n\to\infty} \mathbb{E}_x\left(\sup_{s\in[0,t]}\left|\int_0^s b_n(w_r)\mathrm{d}r - A_s\right| \wedge 1\right) = 0,
$$

where $b_n = b * \varrho_n$ is any mollifying approximation of b.

(ii) For any function $f \in C^2(\mathbb{R}^d)$, it holds that

$$
M^f_t := f(w_t) - f(x) - \int_0^t \mathscr{L}^\sigma f(w_s)\mathrm{d}s - \int_0^t \nabla f(w_s)\mathrm{d}A_s \tag{5.25}
$$

is a continuous local martingale under $\mathbb{P}_x$ with $\mathbb{P}_x(M^f_0 = 0) = 1$.

All the martingale solutions $\mathbb{P}_x$ of SDE (1.1) with starting point x will be denoted by $\mathscr{M}^x_{\sigma,b}$.

We also introduce the following notion of weak solutions.

Definition 5.10 (*Weak solution*) Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbf{P})$ be a complete filtration probability space. Let (X, W) be a pair of continuous $\mathscr{F}_t$-adapted process. We call $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbf{P}; X, W)$ a weak solution of SDE (1.1) with starting point x if

(i) W is a standard d-dimensional $\mathscr{F}_t$-Brownian motion.

(ii) $A_t := \lim_{n\to\infty} \int_0^t b_n(X_s)\mathrm{d}s$ is a finite variation process, where $b_n = b * \varrho_n$ is any mollifying approximation of b and the limit is taken in the sense of u.c.p.
(iii) $\mathbf{P}(X_0 = x) = 1$ and it holds that

$$X_t = x + \int_0^t \sigma(X_s)\mathrm{d}W_s + A_t, \quad t \geqslant 0, \ a.s.$$

The following lemma is standard (see [27]). For the reader's convenience, we provide a detailed proof here.

Proposition 5.11 *Suppose that σ and σ^{-1} are locally bounded. For fixed $x \in \mathbb{R}^d$, there is a martingale solution $\mathbb{P}_x \in \mathscr{M}^x_{\sigma,b}$ if and only if there is a weak solution $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbf{P}; X, W)$ with starting point x in the sense of Definition 5.10 so that $\mathbb{P}_x = \mathbf{P} \circ X^{-1}$.*

Proof (i) Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbf{P}; X, W)$ be a weak solution of SDE (1.1) with starting point x in the sense of Definition 5.10. By Itô's formula, it follows that $\mathbb{P}_x = \mathbf{P} \circ X^{-1}$ is a martingale solution of SDE (1.1) in the sense of Definition 5.9.
(ii) Suppose that $\mathbb{P} \in \mathscr{M}^x_{\sigma,b}$. By choosing $f(x) = x_i$ in (5.25), one sees that $M^i_t := w^i_t - x^i - A^i_t$ is a continuous local martingale under $\mathbb{P}$. By Itô's formula, we have

$$w^i_t w^j_t - x^i x^j = \int_0^t (w^j_s \mathrm{d}M^i_s + w^i_s \mathrm{d}M^j_s) + \int_0^t (w^j_s \mathrm{d}A^i_s + w^i_s \mathrm{d}A^j_s) + [M^i, M^j]_t.$$

On the other hand, for any $i, j = 1, \ldots, d$, if we choose $f(x) = x_i x_j$ in (5.25), then

$$w^i_t w^j_t - x^i x^j - \int_0^t w^j_s \mathrm{d}A^i_s - \int_0^t w^i_s \mathrm{d}A^j_s - \int_0^t (\sigma^{ik}\sigma^{jk})(w_s)\mathrm{d}s$$

is also a continuous local martingale. Hence,

$$[M^i, M^j]_t = \int_0^t (\sigma^{ik}\sigma^{jk})(w_s)\mathrm{d}s.$$

Now we define

$$W_t := \int_0^t \sigma^{-1}(w_s)\mathrm{d}M_s, \quad t \geqslant 0.$$

Since σ^{-1} is locally bounded, W is a continuous $\mathcal{B}_t(\mathbb{C})$-local martingale under $\mathbb{P}$ and by definition,

$$[W^i, W^j]_t = \delta_{ij}\, t, \ i, j = 1, \ldots, d.$$

By Lévy's characterization, W is a $\mathcal{B}_t(\mathbb{C})$-Brownian motion under $\mathbb{P}$. Moreover,

$$w_t = x + A_t + \int_0^t \sigma(w_s)\mathrm{d}W_s, \quad \mathbb{P} - a.s.$$

Thus $(\mathbb{C}, \mathcal{B}(\mathbb{C}), (\mathcal{B}_t(\mathbb{C}))_{t\geqslant 0}, \mathbb{P}; w, W)$ is a weak solution in the sense of Definition 4.8. □

Lemma 5.12 (Krylov type estimate) *Under* $(\mathbf{H}^\sigma_\beta)$ *and* $b \in \mathbb{K}_1$, *there are* $T, C > 0$ *only depending on* $\sigma, d, m^{(1)}_b$ *such that for any* $x \in \mathbb{R}^d$ *and* $\mathbb{P} \in \mathscr{M}^x_{\sigma,b}$, $f \in C^1_b(\mathbb{R}^d)$ *and all* $0 \leqslant t_0 < t_1 \leqslant T$,

$$\mathbb{E}\left(\int_{t_0}^{t_1} |f(w_s)|\mathrm{d}s\Big|\mathscr{F}_{t_0}\right) \leqslant Cm^{(1)}_f(t_1 - t_0).$$

Proof Let $T > 0$ be fixed, whose value will be determined below. Fix $0 \leqslant t_0 < t_1 \leqslant T$. Let $f \in C^1_b(\mathbb{R}^d)$ and $u \in \mathbf{C}^{0,2}_{t_1}$ solve the following backward PDE:

$$\partial_t u + \mathscr{L}^\sigma u = f, \quad u(t_1) = 0.$$

Without loss of generality we assume $\mathbb{E}|A|^{t_1}_{t_0} < \infty$. Otherwise, one just needs to replace w_s below by $w_{s\wedge\tau_n}$, where $\tau_n = \{s > t_0 : |A|^s_{t_0} > n\}$, and then let $n \to \infty$. By (ii) of Definition 5.9 and the optional stopping theorem, we have

$$\mathbb{E}(u(t_1, w_{t_1})|\mathscr{F}_{t_0}) = u(t_0, w_{t_0}) + \mathbb{E}\left(\int_{t_0}^{t_1} f(w_s)\mathrm{d}s\Big|\mathscr{F}_{t_0}\right) + \mathbb{E}\left(\int_{t_0}^{t_1} \nabla u(s, w_s)\mathrm{d}A_s\Big|\mathscr{F}_{t_0}\right).$$

Hence,

$$\begin{aligned}\mathbb{E}\left(\int_{t_0}^{t_1} f(w_s)\mathrm{d}s\Big|\mathscr{F}_{t_0}\right) &\leqslant 2\|u\|_{\mathbb{L}^\infty(t_0,t_1)} + \|\nabla u\|_{\mathbb{L}^\infty(t_0,t_1)}\mathbb{E}\left(|A|^{t_1}_{t_0}|\mathscr{F}_{t_0}\right)\\ &\leqslant c_1 m^{(1)}_f(t_1 - t_0) + c_2 m^{(1)}_f(t_1 - t_0)\mathbb{E}\left(|A|^{t_1}_{t_0}|\mathscr{F}_{t_0}\right).\end{aligned} \tag{5.26}$$

Taking $f = |b_\varepsilon|$, we obtain

$$\mathbb{E}\left(\int_{t_0}^{t_1} |b_\varepsilon|(w_s)\mathrm{d}s\Big|\mathscr{F}_{t_0}\right) \leqslant c_1 m^{(1)}_b(t_1 - t_0) + c_2 m^{(1)}_b(t_1 - t_0)\mathbb{E}\left(|A|^{t_1}_{t_0}|\mathscr{F}_{t_0}\right).$$

By Lemma 2.6 and Fatou's lemma, we get

$$\mathbb{E}\left(|A|^{t_1}_{t_0}|\mathscr{F}_{t_0}\right) \leqslant c_1 m^{(1)}_b(t_1 - t_0) + c_2 m^{(1)}_b(t_1 - t_0)\mathbb{E}\left(|A|^{t_1}_{t_0}|\mathscr{F}_{t_0}\right).$$

Letting T be small enough so that

$$c_2 m^{(1)}_b(T) \leqslant 1/2.$$

Thus we obtain

$$\mathbb{E}\left(|A|^{t_1}_{t_0}|\mathscr{F}_{t_0}\right) \leqslant 2c_1 m^{(1)}_b(t_1 - t_0).$$

Substituting this into (5.26) yields the desired estimate. □

Remark 5.13 As a corollary of the above Krylov type estimate (see Corollary 3.5), for any $f \in \mathbb{K}_1$ we have

$$\sup_n \mathbb{E}\left(\mathrm{e}^{\lambda \int_0^T |f_n(w_s)|\mathrm{d}s}\right) < \infty, \quad \forall \lambda > 0.$$

Theorem 5.14 *Under* $(\mathbf{H}^\sigma_\beta)$ *and* $b \in \mathbb{K}_1$, *for any* $x \in \mathbb{R}^d$, *there is a unique martingale solution* $\mathbb{P} \in \mathscr{M}^x_{\sigma,b}$, *equivalently, a unique weak solution in the sense of Definition 5.10.*

Proof We first show the existence and uniqueness of weak solutions $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbf{P}; X, W)$ with starting point x in a short time $T > 0$, where T is independent of x, whose value will be determined below, and in fact only depends on $\sigma, d, m^{(1)}_b$ as above.

(**Uniqueness**) Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbf{P}; X, W)$ be a weak solution defined on the time interval $[0, T]$. For $n \in \mathbb{N}$, let $b_n := b * \varrho_n$ be the mollifying approximation of b. Since $b_n \in C^\infty_b(\mathbb{R}^d)$ by Lemma 5.3, it is well known that there is a unique $u_n : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ in $\mathbf{C}^{0,2}_T(\mathbb{R}^d)$ solving the following backward PDE

$$\partial_t u_n + \mathscr{L}^\sigma u_n + b_n \cdot \nabla u_n = -b_n, \quad u_n(T) = 0.$$

By Theorem 5.5, there are $T = T(\sigma, d, m^{(1)}_b) > 0$ small enough such that for all $n \in \mathbb{N}$,

$$\|u_n\|_{\mathbf{C}^{0,1}_T} \leqslant 1/2, \tag{5.27}$$

and for some $C > 0$ independent of n,

$$|\nabla u_n(t,x) - \nabla u_n(t,y)| \leqslant C m^{(1)}_{\lambda,b_n}(|x-y|) \leqslant C m^{(1)}_{\lambda,b}(|x-y|) \tag{5.28}$$

and

$$\lim_{n\to\infty} \|u_n - u\|_{\mathbf{C}^{0,1}_T} = 0,$$

where $u \in \mathbf{C}^{0,1}_T$ is the unique mild solution of PDE (5.14) with $f = b$.

Define

$$\Phi^n_t(x) := x + u_n(t,x), \quad \Phi_t(x) := x + u(t,x).$$

Then we have

$$\partial_t \Phi^n + \mathscr{L}^\sigma \Phi^n + b_n \cdot \nabla \Phi^n = 0.$$

By Itô's formula, we have for each $n \in \mathbb{N}$,

$$\Phi^n_t(X_t) = \Phi^n_0(x) + \int_0^t \nabla \Phi^n_s(X_s)\Big[\mathrm{d}A_s - b_n(X_s)\mathrm{d}s\Big] + \int_0^t (\sigma^* \cdot \nabla \Phi^n_s)(X_s)\mathrm{d}W_s.$$

By BDG's inequality and the dominated convergence theorem, we clearly have

$$\begin{aligned}&\lim_{n\to\infty}\mathbb{E}\left(\sup_{t\in[0,T]}\left|\int_0^t(\sigma^*\cdot\nabla\Phi^n_s-\sigma^*\cdot\nabla\Phi_s)(X_s)\mathrm{d}W_s\right|^2\right)\\&\leqslant 4\|\sigma\|_\infty\lim_{n\to\infty}\mathbb{E}\left(\int_0^t|\nabla\Phi^n_s-\nabla\Phi_s|^2(X_s)\mathrm{d}s\right)\\&\leqslant 4\|\sigma\|_\infty\lim_{n\to\infty}\|\nabla u_n-\nabla u\|^2_{\mathbb{L}^\infty_t}=0.\end{aligned}$$

Moreover, for each $t\in[0,T]$, by Lemma 5.12, we also have

$$\begin{aligned}&\lim_{n\to\infty}\mathbb{E}\left|\int_0^t(\nabla\Phi^n_s(X_s)-\nabla\Phi_s(X_s))\Big[\mathrm{d}A_s-b_n(X_s)\mathrm{d}s\Big]\right|\\&\leqslant\lim_{n\to\infty}\|\nabla u_n-\nabla u_\infty\|_{\mathbb{L}^\infty_T}\mathbb{E}\left(|A|^t_0+\int_0^t|b_n(X_s)|\mathrm{d}s\right)=0.\end{aligned}$$

On the other hand, due to $\nabla\Phi=\mathbb{I}+\nabla u\in\mathbf{C}^0_T$, by Lemmas 2.6 and 5.12, we get

$$\lim_{n\to\infty}\mathbb{E}\left|\int_0^t\nabla\Phi_s(X_s)\Big(\mathrm{d}A_s-b_n(X_s)\mathrm{d}s\Big)\right|=0.$$

Combining the above limits, we arrive at

$$\Phi_t(X_t)=\Phi_0(x)+\int_0^t(\sigma^*\cdot\nabla\Phi_s)(X_s)\mathrm{d}W_s.$$

By (5.27), one sees that

$$\tfrac{1}{2}|x-y|\leqslant|\Phi_t(x)-\Phi_t(y)|\leqslant 2|x-y|,$$

and $x\mapsto\Phi_t(x)$ is a C^1-diffeomorphism. Let Φ_t^{-1} be the inverse of Φ_t and define $Y_t:=\Phi_t(X_t)$. Then Y_t solves the following SDE

$$Y_t=\Phi_0(x)+\int_0^t\tilde{\sigma}_s(Y_s)\mathrm{d}W_s,\qquad(5.29)$$

where

$$\tilde{\sigma}_s(y):=[\sigma^*\cdot\nabla\Phi_s]\circ\Phi_s^{-1}(y).$$

Clearly, by definition and (5.28) we have

$$\lim_{|y-y'|\to 0}\sup_{s\in[0,T]}|\tilde{\sigma}_s(y)-\tilde{\sigma}_s(y')|=0,$$

which, together with $\tilde{\sigma}$ being bounded and uniformly non-degenerate, yields that SDE (5.29) admits a unique weak solution (cf. [27]). Thus, the uniqueness of the original SDE follows.

(**Existence**) Next, we show the existence. For each $n \in \mathbb{N}$ and $x \in \mathbb{R}^d$, since $b_n := b * \varrho_n \in C_b^1(\mathbb{R}^d)$, it is well known that there is a unique martingale solution $\mathbb{P}_x^n \in \mathscr{M}_{\sigma,b_n}^x$ so that $(\mathbb{P}_x^n)_{x\in\mathbb{R}^d}$ forms a family of strong Markov processes. For any stopping time $\tau \leqslant T$, by the strong Markov property and Lemma 5.12, we have

$$\mathbb{E}_x^n\left(\int_\tau^{\tau+\delta} |b_n(w_s)|\mathrm{d}s\right) \leqslant \sup_{y\in\mathbb{R}^d} \mathbb{E}_y^n\left(\int_0^{\delta} |b_n(w_s)|\mathrm{d}s\right) \leqslant Cm_{b_n}^{(1)}(\delta) \leqslant Cm_b^{(1)}(\delta),$$

where $\mathbb{E}_x^n$ denotes the expectation with respect to the probability measure $\mathbb{P}_x^n$, and C is independent of x, n. Hence, by Lemma 2.7, we have

$$\mathbb{E}_x^n\left(\sup_{t\in[0,T]}\left(\int_t^{t+\delta} |b_n(w_s)|\mathrm{d}s\right)^{1/2}\right) \leqslant 2(Cm_b^{(1)}(\delta))^{1/2}.$$

In particular, since $b \in \mathbb{K}_1$, by Chebyschev's inequality, for any $\varepsilon > 0$,

$$\lim_{\delta\to 0}\sup_n \mathbb{P}_x^n\left(\sup_{\substack{0\leqslant t'<t\leqslant T\\ |t-t'|\leqslant\delta}}\left|\int_{t'}^t b_n(w_s)\mathrm{d}s\right| > \varepsilon\right) = 0. \tag{5.30}$$

Moreover, it is easy to see that

$$\lim_{\delta\to 0}\sup_n \mathbb{P}_x^n\left(\sup_{\substack{0\leqslant t'<t\leqslant T\\ |t-t'|\leqslant\delta}}\left|\int_{t'}^t \sigma(w_s)\mathrm{d}W_s\right| > \varepsilon\right) = 0. \tag{5.31}$$

On the other hand, by the equivalence between martingale solutions and weak solutions, there exists a weak solution $(\Omega^n, \mathscr{F}^n, (\mathscr{F}_t^n)_{t\geqslant 0}, \mathbf{P}^n; X^n, W^n)$ so that $\mathbf{P}^n \circ (X^n)^{-1} = \mathbb{P}_x^n$ and

$$X_t^n = x + \int_0^t \sigma(X_s^n)\mathrm{d}W_s^n + \int_0^t b_n(X_s^n)\mathrm{d}s (=: A_t^n),\ t \in [0, T]. \tag{5.32}$$

Let $\mathbb{Q}^n$ be the law of (X^n, W^n, A^n) in $\mathbb{C}\times\mathbb{C}\times\mathbb{C}$. By (5.30) and (5.31), one sees that $(\mathbb{Q}^n)_{n\in\mathbb{N}}$ is tight. Hence, there is a subsequence still denoted by n so that $\mathbb{Q}^n$ weakly converges to some probability measure $\mathbb{Q}$. By Skorokhod's representation theorem, there are probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathbf{P}})$ and random variables $(\tilde{X}^n, \tilde{W}^n, \tilde{A}^n)$ and $(\tilde{X}, \tilde{W}, \tilde{A})$ defined on it such that

$$(\tilde{X}^n, \tilde{W}^n, \tilde{A}^n) \to (\tilde{X}, \tilde{W}, \tilde{A}), \quad \tilde{\mathbf{P}} - a.s. \tag{5.33}$$

and

$$\tilde{\mathbf{P}} \circ (\tilde{X}^n, \tilde{W}^n, \tilde{A}^n)^{-1} = \mathbb{Q}^n, \quad \tilde{\mathbf{P}} \circ (\tilde{X}, \tilde{W}, \tilde{A})^{-1} = \mathbb{Q}. \tag{5.34}$$

Define $\tilde{\mathscr{F}}_t^n := \sigma(\tilde{W}_s^n; s \leqslant t)$. Notice that

$$\mathbf{P}^n(W_t^n - W_s^n \in \cdot | \mathscr{F}_s^n) = \mathbf{P}^n(W_t^n - W_s^n \in \cdot)$$
$$\Rightarrow \tilde{\mathbf{P}}(\tilde{W}_t^n - \tilde{W}_s^n \in \cdot | \tilde{\mathscr{F}}_s^n) = \tilde{\mathbf{P}}(\tilde{W}_t^n - \tilde{W}_s^n \in \cdot).$$

In other words, $\tilde{W}^n$ is an $\tilde{\mathscr{F}}_t^n$-Brownian motion. Thus, by (5.32) and (5.34) we have

$$\tilde{X}_t^n = x + \int_0^t \sigma(\tilde{X}_s^n)\mathrm{d}\tilde{W}_s^n + \tilde{A}_t^n, \quad \tilde{A}_t^n := \int_0^t b_n(\tilde{X}_s^n)\mathrm{d}s.$$

By taking limits $n \to \infty$ and (5.33),

$$\tilde{X}_t = x + \int_0^t \sigma(\tilde{X}_s)\mathrm{d}\tilde{W}_s + \tilde{A}_t,$$

where $\tilde{A}$ is a finite variation process. Indeed, by Lemmas 2.6 and 5.12,

$$\tilde{\mathbf{E}}|\tilde{A}|_0^T \leqslant \tilde{\mathbf{E}} \varliminf_{n\to\infty} |\tilde{A}^n|_0^T \leqslant \varliminf_{n\to\infty} \tilde{\mathbf{E}}|\tilde{A}^n|_0^T = \varliminf_{n\to\infty} \mathbb{E}_x^n \left(\int_0^T |b_n(w_s)|\mathrm{d}s \right) < \infty.$$

To complete the proof, it suffices to show that

$$\lim_{n\to\infty} \tilde{\mathbf{E}} \left(\sup_{t\in[0,T]} \left| \tilde{A}_t - \int_0^t b_n(\tilde{X}_s)\mathrm{d}s \right| \wedge 1 \right) = 0. \tag{5.35}$$

Below we drop the tilde for simplicity. For $n, k \in \mathbb{N}$, let $u_{n,k}$ solve the following PDE

$$\partial_t u_{n,k} + \mathscr{L}^\sigma u_{n,k} + b_n \cdot \nabla u_{n,k} = b_k, \quad u_{n,k}(T) = 0.$$

By Itô's formula, we have

$$\int_0^t b_k(\tilde{X}_s^n)\mathrm{d}s = u_{n,k}(t, \tilde{X}_t^n) - u_{n,k}(0, x) + \int_0^t (\sigma^* \cdot \nabla u_{n,k})(s, \tilde{X}_s^n)\mathrm{d}\tilde{W}_s.$$

Hence, for any stopping time $\tau \leqslant T$,

$$\begin{aligned}\tilde{\mathbf{E}}\left|\int_0^\tau (b_k - b_l)(\tilde{X}_s^n)\mathrm{d}s\right|^2 &\leqslant 4\|u_{n,k} - u_{n,l}\|^2_{\mathbb{L}^\infty_T} + 2\tilde{\mathbf{E}}\left|\int_0^\tau (\sigma^* \cdot (\nabla u_{n,k} - \nabla u_{n,l}))(\tilde{X}_s^n)\mathrm{d}W_s\right|^2 \\ &\leqslant 4\|u_{n,k} - u_{n,l}\|^2_{\mathbb{L}^\infty_T} + 2T\|\sigma\|_\infty\|\nabla u_{n,k} - \nabla u_{n,l}\|^2_{\mathbb{L}^\infty_T},\end{aligned}$$

which implies by Lemma 2.7 that

$$\tilde{\mathbf{E}}\left(\sup_{t\in[0,T]}\left|\int_0^t (b_k - b_l)(\tilde{X}_s^n)\mathrm{d}s\right|\right) \leqslant C\|u_{n,k} - u_{n,l}\|_{\mathbf{C}^{0,1}_T},$$

where C is independent of n, k, l. By (5.18) we get

$$\lim_{l,k\to\infty}\sup_n \tilde{\mathbf{E}}\left(\sup_{t\in[0,T]}\left|\int_0^t (b_k(\tilde{X}_s^n) - b_l(\tilde{X}_s^n))\mathrm{d}s\right|\right) = 0. \tag{5.36}$$

On the other hand, for fixed $l \in \mathbb{N}$, by the dominated convergence theorem, we have

$$\lim_{n\to\infty}\tilde{\mathbf{E}}\left(\int_0^T |b_l(\tilde{X}_s^n) - b_l(\tilde{X}_s)|\mathrm{d}s\right) = 0. \tag{5.37}$$

Since $\tilde{A}^n = \int_0^\cdot b_n(\tilde{X}_s^n)\mathrm{d}s$ converges to $\tilde{A}$ a.s., by (5.36) and (5.37) we obtain (5.35).

Finally, we need to extend the solution from the short time T to the arbitrary time T'. This can be done by a standard patching up technique. We left it to the readers. □

We have the following easy corollary (see [4]).

Corollary 5.15 *Suppose $d = 1$ and $b \in \mathbb{K}_{3/2}$. Under $(\mathbf{H}^\sigma_{1/2})$, for any $x \in \mathbb{R}^d$, SDE (1.6) has a unique strong solution.*

Proof Choosing $\alpha = 3/2$ in Remark 5.6, by (5.3), we get

$$\|\tilde{\sigma}\|_{C^{1/2}} \leqslant \|\nabla\Phi\|_{C^{1/2}} \cdot \|\sigma\|_{C^{1/2}} < \infty.$$

By Yamada-Watanabe's theorem, one sees that (5.29) has a unique strong solution. Hence, the pathwise uniqueness holds for the original SDE. □

To state the ergodicity, we make the following assumption on b:

$(\hat{\mathbf{H}}^b)$ $b = b^{(1)} + b^{(2)}$, where $b^{(2)} \in \mathbb{K}''_\alpha$ for some $\alpha \in (1, 2)$ and $b^{(1)}$ satisfies that for some $\vartheta \geqslant 0$ and $\kappa_0, \kappa_1, \kappa_2 > 0$,

$$\frac{\langle x, b^{(1)}(x)\rangle}{\sqrt{1 + |x|^2}} \leqslant -\kappa_0|x|^\vartheta + \kappa_1, \quad |b^{(1)}(x)| \leqslant \kappa_2(1 + |x|^\vartheta). \tag{5.38}$$

We have the following ergodicity result.

Theorem 5.16 *Under* $(\mathbf{H}^\sigma_\beta)$ *and* $(\hat{\mathbf{H}}^b)$, *for any* $x \in \mathbb{R}^d$, *there exists a unique martingale solution* $\mathbb{P}_x \in \mathscr{M}^x_{\sigma,b}$ *to SDE* (1.1). *Moreover, letting* $\mathbb{E}_x := \mathbb{E}^{\mathbb{P}_x}$, *we have the following conclusions*:

(i) *If* $\vartheta = 0$ *in* (5.38), *then* $\mathbb{P}_x^{-1} \circ w_t$ *admits a density* $p_t(x, y)$ *and for fixed* $T > 0$, $p_t(x, y)$ *enjoys the following two-sided estimate*: *for some* $c_1, c_2 \geqslant 1$ *and all* $t \in (0, T]$, $x, y \in \mathbb{R}^d$,

$$c_1^{-1} t^{-d/2} \mathrm{e}^{-c_2|x-y|^2/t} \leqslant p_t(x, y) \leqslant c_1 t^{-d/2} \mathrm{e}^{-c_2^{-1}|x-y|^2/t}, \tag{5.39}$$

and gradient estimate: *for some* $c_3, c_4 > 0$ *and all* $t \in (0, T]$, $x, y \in \mathbb{R}^d$,

$$|\nabla_x p_t(x, y)| \leqslant c_3 t^{-(d+1)/2} \mathrm{e}^{-c_4|x-y|^2/t}. \tag{5.40}$$

(ii) *If* $\vartheta > 0$ *in* (5.38), *then* $P_t\varphi(x) := \mathbb{E}_x\varphi(w_t)$ *admits a unique invariant probability measure* $\mu(\mathrm{d}x) = \varrho(x)\mathrm{d}x$ *with* $\varrho \in H^{\gamma,r}$, *where* $\gamma \in (0, \beta \wedge (\alpha - 1)]$ *and* $r \in (1, \frac{d}{d+\gamma-1})$.

Proof The proof is essentially the same as in [35, Theorem 5.1]. We sketch the key point: global Zvonkin's transformation. For $n \in \mathbb{N}$, let $b^{(2)}_n := b^{(2)} * \varrho_n$ be the mollifying approximation of $b^{(2)}$. For $\lambda > 0$, let $u_n \in C^2_b(\mathbb{R}^d)$ solve the following elliptic PDE

$$(\mathscr{L}^\sigma - \lambda)u_n + b^{(2)}_n \cdot \nabla u_n = -b^{(2)}_n.$$

By Theorem 5.5, there are $\lambda_0, T, C > 0$ depending only on $\sigma, d, m^{(1)}_{b^{(2)}}$ and a continuous function $\ell : \mathbb{R}_+ \to \mathbb{R}_+$ with $\ell(0) = 0$ such that for all $n \in \mathbb{N}$ and $\lambda \geqslant \lambda_0$,

$$\|u_n\|_{C^1_b} \leqslant C m^{(1)}_{\lambda, b^{(2)}_n}(T) \leqslant C m^{(1)}_{\lambda, b^{(2)}}(T) \tag{5.41}$$

and

$$|\nabla u_n(x) - \nabla u_n(y)| \leqslant \ell(|x - y|). \tag{5.42}$$

Now by Ascolli-Arzela's lemma, there is a subsequence still denoted by n and $u \in C^1_b(\mathbb{R}^d)$ such that

$$\|u\|_{C^1_b} \leqslant C m^{(1)}_{\lambda, b^{(2)}}(T), \ \lim_{n\to\infty} \sup_{|x|\leqslant R} |\nabla^j u_n(x) - \nabla^j u(x)| = 0, \ \forall R > 0, j = 0, 1. \tag{5.43}$$

Define

$$\Phi_n(x) := x + u_n(x), \quad \Phi(x) := x + u(x).$$

Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbf{P}; X, W)$ be the unique weak solution of (1.1), i.e.,

$$X_t = x + A_t + \int_0^t b^{(1)}(X_s)\mathrm{d}s + \int_0^t \sigma(X_s)\mathrm{d}W_s,\ A_t := \lim_{n\to\infty}\int_0^t b_n^{(2)}(X_s)\mathrm{d}s.$$

By Itô's formula, we have for each $n \in \mathbb{N}$,

$$\Phi_n(X_t) = \Phi_n(x) + \int_0^t \nabla\Phi_n(X_s)\Big[\mathrm{d}A_s - b_n^{(2)}(X_s)\mathrm{d}s + b^{(1)}(X_s)\mathrm{d}s\Big] + \lambda\int_0^t u_n(X_s)\mathrm{d}s + \int_0^t (\sigma^* \cdot \nabla\Phi_n)(X_s)\mathrm{d}W_s.$$

As in the proof of Theorem 5.14, by standard stopping technique and taking limits, we obtain

$$\Phi(X_t) = \Phi(x) + \int_0^t (b^{(1)} \cdot \nabla\Phi + \lambda u)(X_s)\mathrm{d}s + \int_0^t (\sigma^* \cdot \nabla\Phi)(X_s)\mathrm{d}W_s.$$

Since $b^{(2)} \in \mathbb{K}''_\alpha$ for some $\alpha \in (1, 2)$, by (5.43), letting λ be large enough, one sees that

$$\tfrac{1}{2}|x - y| \leqslant |\Phi(x) - \Phi(y)| \leqslant 2|x - y|,$$

and by (5.24),

$$|\nabla\Phi(x) - \nabla\Phi(y)| \leqslant C|x - y|^{\alpha-1}.$$

So, $x \mapsto \Phi(x)$ is a C^1-diffemorphism. Let Φ^{-1} be the inverse of Φ and define $Y_t := \Phi(X_t)$. Then Y_t solves the following SDE

$$Y_t = \Phi(x) + \int_0^t \tilde{b}(Y_s)\mathrm{d}s + \int_0^t \tilde{\sigma}(Y_s)\mathrm{d}W_s, \tag{5.44}$$

where

$$\tilde{b}(y) := (b^{(1)} \cdot \nabla\Phi + \lambda u) \circ \Phi^{-1}(y),\ \ \tilde{\sigma}(y) := [\sigma^* \cdot \nabla\Phi] \circ \Phi^{-1}(y).$$

Since $\tilde{b}$ still satisfies (5.38) for large λ (see [35, Lemma 5.9]) and $\tilde{\sigma}$ is Hölder continuous, we can use Theorem 4.10 to conclude the desired results. □

Remark 5.17 When $\sigma = \mathbb{I}$ and $b \in \mathbb{K}_1$, Kim and Song in [19] proved estimates (5.39) and (5.40) by direct perturbation argument.

References

1. Bahouri, H., Chemin, J.Y., Danchin, R.: Fourier Analysis and Nonlinear Partial Differential Equations. Springer, Berlin (2011)
2. Bass, R.F., Chen, Z.Q.: Stochastic differential equations for Dirichlet processes. Probab. Theory Rel. Fields **121**(3), 422–446 (2001)
3. Bass, R.F., Chen, Z.Q.: Brownian motion with singular drift. Ann. Probab. **31**(2), 791–817 (2003)
4. Bass, R.F., Chen, Z.Q.: One-dimensional stochastic differential equations with singular and degenerate coefficients. Sankhyā **67**(1), 19–45 (2005)
5. Beck L., Flandoli F., Gubinelli M., Maurelli M.: Stochastic ODEs and stochastic linear PDEs with critical drift: regularity, duality and uniqueness. arXiv:1401.1530v1
6. Champagnat, N., Jabin, P.E.: Strong solutions to stochastic differential equations with rough coefficients. Ann. Probab. **46**(3), 1498–1541 (2018)
7. Chen, Z.Q., Hu, E., Xie, L., Zhang, X.: Heat kernels for non-symmetric diffusion operators with jumps. J. Differ. Equs. **263**, 6576–6634 (2017)
8. Chen, X., Li, X.M.: Strong completeness for a class of stochastic differential equations with irregular coefficients. Electron. J. Probab. **19**(91), 1–34 (2014)
9. Cherny A.S., Engelbert H.J.: Singular Stochastic Differential Equations. Lecture Notes in Mathematics. **1858**, Springer (2005)
10. Fedrizzi, E., Flandoli, F.: Noise prevents singularities in linear transport equations. J. Funct. Anal. **264**, 1329–1354 (2013)
11. Fedrizzi, E., Flandoli, F.: Hölder flow and differentiability for SDEs with non regular drift. Stoch. Anal. Appl. **31**, 708–736 (2013)
12. Fedrizzi, E., Flandoli, F., Priola, E., Vovelle, J.: Regularity of stochastic kinetic equations. Electron. J. Probab. **22**, 42 (2017). paper no. 48
13. Flandoli, F., Issoglio, E., Russo, F.: Multidimensional stochastic differential equations with distributional drift. Trans. Am. Math. Soc. **369**(3), 1665–1688 (2017)
14. Flandoli, F., Russo, F., Wolf, J.: Some SDEs with distributional drift. I. General calculas. Osaka J. Math. **40**(2), 493–542 (2003)
15. Flandoli, F., Russo, F., Wolf, J.: Some SDEs with distributional drift. II. Lyons–Zheng structure, Itô's formula and semimartingale characterization. Random Oper. Stoch. Equ. **12**(2), 145–184 (2004)
16. He, K., Zhang, X.: One dimensional stochastic differential equations with distributional drifts. Acta Mathematicae Applicatae Sinica **23**(3), 501–512 (2007). English Series
17. Hu, Y., Lê, K., Mytnik, L.: Stochastic differential equation for Brox diffusion. Stoch. Process. Appl. **127**, 2281–2315 (2017)
18. Ikeda, N., Watanabe, S.: Stochastic Differential Equations and Diffusion Processes, 2nd edn. North-Holland Publishing Company, Amsterdam (1989)
19. Kim, P., Song, R.: Two-sided estimates on the density of Brownian motion with singular drift. Ill. J. Math. **50**, 635–688 (2006)
20. Krylov, N.V.: Controlled Diffusion Processes. Springer, Berlin (1980)
21. Krylov, N.V., Röckner, M.: Strong solutions of stochastic equations with singular time dependent drift. Probab. Theory Rel. Fields **131**, 154–196 (2005)
22. Menoukeu-Pamen, O., Meyer-Brandis, T., Nilssen, T., Proske, F., Zhang, T.: A variational approach to the construction and Malliavin differentiability of strong solutions of SDE's. Math. Ann. **357**(2), 761–799 (2013)
23. Mohammed, S.E.A., Nilssen, T., Proske, F.: Sobolev differentiable stochastic flows for SDE0s with singular coefficients: applications to the transport equation. Ann. Probab. **43**(3), 1535–1576 (2015)
24. Russo, F., Trutnau, G.: Some parabolic PDEs whose drift is an irregular random noise in space. Ann. Probab. **35**(6), 2213–2262 (2007)
25. Scheutzow, M.: A stochastic Gronwall's lemma. Infin. Dimens. Anal. Quantum Probab. Rel. Top. **16**(2), 1350019 (2013). (4 pages)
26. Stein, E.M.: Singular Integrals and Differentiability Properties of Functions. Princeton University Press, Princeton (1970)
27. Stroock, D.W., Varadhan, S.S.: Multidimensional Diffusion Processes. Springer, Berlin (1979)
28. Veretennikov, AJu: On the strong solutions of stochastic differential equations. Theory Probab. Appl. **24**, 354–366 (1979)

29. Xie, L., Zhang X.: Ergodicity of Stochastic Differential Equations with Jumps and Singular Coefficients. arXiv:1705.07402
30. Zhang, Q.S.: Gaussian bounds for the fundamental solutions of $\nabla(A\nabla u)+B\nabla u-u_t=0$. Manuscripta Math. **93**(3), 381–390 (1997)
31. Zhang, X.: Strong solutions of SDEs with singular drift and Sobolev diffusion coefficients. Stoch. Proc. Appl. **115**, 1805–1818 (2005)
32. Zhang, X.: Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients. Electron. J. Probab. **16**, 1096–1116 (2011)
33. Zhang, X.: Well-posedness and large deviation for degenerate SDEs with Sobolev coefficients. Rev. Mate. Ibv. **29**(1), 25–52 (2013)
34. Zhang, X.: Stochastic differential equations with Sobolev diffusion and singular drift. Ann. Appl. Probab. **26**(5), 2697–2732 (2016)
35. Zhang, X., Zhao G.H.: Heat Kernel and Ergodicity of SDEs with Distributional Drifts. arXiv:1710.10537
36. Zvonkin, A.K.: A transformation of the phase space of a diffusion process that removes the drift. Math. Sb. **93**(135), 129–149 (1974)