SDES WITH SUPERCRITICAL DISTRIBUTIONAL DRIFTS

ZIMO HAO AND XICHENG ZHANG

Abstract. Let $d \geqslant 2$. In this paper, we investigate the following stochastic differential equation (SDE) in $\mathbb{R}^d$ driven by Brownian motion

$$\mathrm{d}X_t = b(t, X_t)\mathrm{d}t + \sqrt{2}\mathrm{d}W_t,$$

where b belongs to the space $\mathbb{L}^q_T\mathbf{H}^\alpha_p$ with $\alpha \in [-1, 0]$ and $p, q \in [2, \infty]$, which is a distribution-valued and divergence-free vector field. In the subcritical case $\frac{d}{p} + \frac{2}{q} < 1 + \alpha$, we establish the existence and uniqueness of a weak solution to the integral equation:

$$X_t = X_0 + \lim_{n\to\infty} \int_0^t b_n(s, X_s)\mathrm{d}s + \sqrt{2}W_t.$$

Here, $b_n := b*\phi_n$ represents the mollifying approximation, and the limit is taken in the L^2-sense. In the critical and supercritical case $1+\alpha \leqslant \frac{d}{p}+\frac{2}{q} < 2+\alpha$, assuming the initial distribution has an L^2-density, we show the existence of weak solutions and associated Markov processes. Moreover, under the additional assumption that $b = b_1 + b_2 + \mathrm{div}a$, where $b_1 \in \mathbb{L}^\infty_T\mathbf{B}^{-1}_{\infty,2}$, $b_2 \in \mathbb{L}^2_T L^2$, and a is a bounded antisymmetric matrix-valued function, we establish the convergence of mollifying approximation solutions without the need to subtract a subsequence. To illustrate our results, we provide examples of Gaussian random fields and singular interacting particle systems, including the two-dimensional vortex models.

Contents

Keywords: Supercritical SDEs, distributional drifts, Krylov's estimate, Schauder's estimate.

This work is partially supported by NNSFC grants of China (Nos. 12131019), and the German Research Foundation (DFG) through the Collaborative Research Centre(CRC) 1283/2 2021 - 317210226 "Taming uncertainty and profiting from randomness and low regularity in analysis, stochastics and their applications".

1. Introduction

Throughout this paper we fix $d \geqslant 2$. Let $(W_t)_{t\geqslant 0}$ be a standard d-dimensional Brownian motion on a stochastic basis $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbf{P})$. Consider the following stochastic differential equation (SDE) in $\mathbb{R}^d$ driven by Brownian motion W:

$$\mathrm{d}X_t = b(t, X_t)\mathrm{d}t + \sqrt{2}\mathrm{d}W_t, \quad t \geqslant 0, \tag{1.1}$$

where the drift b is a time-dependent distribution. Since b is a distribution, the drift term does not make a meaning in the classical sense as we cannot assign a value to a distribution at the point X_t. To formulate a solution, one natural approach is to employ mollifying approximation. Let $\phi_n(x) = n^d\phi(nx)$ be a family of modifiers, where $\phi \in C_c^\infty(\mathbb{R}^d)$ is a smooth probability density function with compact support. Define the smooth approximation of b by convolution as:

$$b_n(t, x) := b(t, \cdot) * \phi_n(x). \tag{1.2}$$

We begin by introducing the concept of weak solutions for the above SDE with a distributional drift. Let $\mathscr{P}(\mathbb{R}^d)$ be the space of all probability measures on $\mathbb{R}^d$.

Definition 1.1 (Weak solutions). *Let $\mathfrak{F} := (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbf{P})$ be a stochastic basis and (X, W) be a pair of $\mathbb{R}^d$-valued continuous $\mathscr{F}_t$-adapted processes on $\mathfrak{F}$. We call $(\mathfrak{F}, X, W)$ a weak solution of SDE (1.1) with initial distribution $\mu \in \mathscr{P}(\mathbb{R}^d)$ if W is an $\mathscr{F}_t$-Brownian motion, $\mathbf{P} \circ X_0^{-1} = \mu$, and for all $t \in [0, T]$,*

$$X_t = X_0 + A_t^b + \sqrt{2}W_t, \quad a.s.,$$

where $A_t^b := \lim_{n\to\infty} \int_0^t b_n(s, X_s)\mathrm{d}s$ exists in the L^2-sense.

It is noted that the definition of a weak solution in Definition 1.1 depends on the choice of mollifiers ϕ_n. The goal of this paper is to find conditions on b as weak as possible so that SDE (1.1) can be solved in the sense of Definition 1.1 and to construct the associated Markov process.

To present our results, we commence by conducting a basic scale analysis, which enables us to distinguish between subcritical, critical, and supercritical drifts. Let $\mathbf{H}_p^\alpha$ (resp. $\dot{\mathbf{H}}_p^\alpha$) be the (resp. homogenous) Bessel potential space, where $\alpha \in \mathbb{R}$ and $p \in [1, \infty]$ (see [5] or Subsection 2.1 below for precise definitions). Suppose that for some $\alpha \in \mathbb{R}$ and $q, p \in [1, \infty]$,

$$b \in L^q(\mathbb{R}_+; \dot{\mathbf{H}}_p^\alpha),$$

and SDE (1.1) admits a solution denoted by X. For $\lambda > 0$, define

$$X_t^\lambda := \lambda^{-1}X_{\lambda^2 t}, \quad W_t^\lambda := \lambda^{-1}W_{\lambda^2 t}, \quad b_\lambda(t, x) := \lambda b(\lambda^2 t, \lambda x).$$

Formally, one sees that

$$\mathrm{d}X_t^\lambda = b_\lambda(t, X_t^\lambda)\mathrm{d}t + \sqrt{2}\mathrm{d}W_t^\lambda.$$

Moreover, by the change of variable we have

$$\|b_\lambda\|_{L^q(\mathbb{R}_+;\dot{\mathbf{H}}_p^\alpha)} = \lambda^{1+\alpha-\frac{d}{p}-\frac{2}{q}}\|b\|_{L^q(\mathbb{R}_+;\dot{\mathbf{H}}_p^\alpha)}.$$

As $\lambda \to 0$, we categorize SDE (1.1) into the following three cases:

Subcritical: $\frac{d}{p} + \frac{2}{q} < 1 + \alpha$; **Critical**: $\frac{d}{p} + \frac{2}{q} = 1 + \alpha$; **Supercritical**: $\frac{d}{p} + \frac{2}{q} > 1 + \alpha$.

To solve the SDE (1.1), a crucial step involves establishing a more precise regularity estimate for the associated Kolmogorov equation:

$$\partial_t u = \Delta u + b \cdot \nabla u + f. \tag{1.3}$$

Let's perform a straightforward analysis of the differentiability index α when $b \in \mathbf{C}^\alpha$ with $\alpha < 0$. According to the Schauder theory of the heat equation, u belongs to at most $\mathbf{C}^{2+\alpha}$. To make the product $b \cdot \nabla u$ meaningful, we need to stipulate that $1 + 2\alpha > 0$, which implies $\alpha > -\frac{1}{2}$.

We summarize part of well-known results regarding the SDE (1.1) in the three cases, categorized based on the value of α. For clarity, we will utilize the following abbreviations in the subsequent table:

SEU: Strong existence-uniqueness; WEU: Weak existence-uniqueness;
WE: Weak existence; EUE: Existence-uniqueness of energy solution.

Value of α	Subcritical	Critical	Supercritical
$\alpha = 0$	SEU: V$^{79}_{[41]}$, KR$^{05}_{[29]}$, Z$^{05,10}_{[43, 44]}$	WEU&SEU: BFGM$^{19}_{[4]}$, K$^{21}_{[28]}$, RZ$^{21}_{[33, 34]}$, KM$^{23}_{[24]}$	WE: ZZ$^{21}_{[49]}$
$\alpha \in [-\frac{1}{2}, 0)$	WEU: BC$^{01}_{[2]}$, FIR$^{17}_{[14]}$, ZZ$^{17}_{[48]}$	–	–
$\alpha \in [-1, -\frac{1}{2})$	EUE: GP$^{23}_{[20]}$	–	EUE: GDJP$^{23}_{[19]}$

- When $\alpha = 0$ and $p = q = \infty$, Veretenikov [41] firstly obtained the strong well-posedness of SDE (1.1) for any starting point X_0.
- In the subcritical case with $\alpha = 0$, Krylov and Röckner [29] demonstrated the strong well-posedness of SDE (1.1) using Girsanov's technique. Related results obtained through Zvonkin's transformation can be found in [43, 44].
- In the critical case with $\alpha = 0$, the weak and the strong well-posedness are studied in recent work [4, 24, 33, 34]. See also Krylov's series of works [26, 27, 28].
- In the supercritical case, when $\alpha = 0$ and $\mathrm{div} b = 0$, a weak solution is constructed in [49] using the maximum principle proven by De-Giorgi's iteration technique method. For the case of multiplicative and possibly degenerate noise, see [46].
- When $\alpha \in (-\frac{1}{2}, 0)$ and b is a time-independent vector field belonging to the space $\mathbf{C}^\alpha$, the existence and uniqueness of a solution termed as "virtual solution" was proved in [14]. When $\alpha \in [-\frac{1}{2}, 0)$ and $b \in \mathbf{H}^\alpha_p$ with $\alpha \in [-\frac{1}{2}, 0]$ and $p > \frac{d}{1+\alpha}$, the authors in [48] showed that there is a unique weak solution $(\mathfrak{F}, X, W)$ to SDE (1.1) in the class that the following Krylov estimate holds: for some $\theta, T > 0$ and any $m \in \mathbb{N}$, $f \in C_c(\mathbb{R}^d)$ and $0 \leqslant t_0 < t_1 \leqslant T$,
$$\left\| \int_{t_0}^{t_1} f(X_s) \mathrm{d}s \right\|_{L^m(\Omega)} \leqslant C(t_1 - t_0)^{\frac{1+\theta}{2}} \|f\|_{\mathbf{H}^\alpha_p}, \tag{1.4}$$
where the constant $C = C(m, T, d, \alpha, p, \|b\|_{\mathbf{H}^\alpha_p}) > 0$. In particular, the above Krylov estimate implies that $A^b_t := \lim_{n\to\infty} \int_0^t b_n(X_s)\mathrm{d}s$ exists and $A^b_\cdot$ is a zero energy process. In the one-dimensional case, the zero energy solution was explored by Bass and Chen in [2].
- In the subcritical case, with $\alpha \in (-1, 0]$ and a time-independent, divergence-free b, a unique energy solution on the torus is established in [20] for SDE (1.1) with initial data whose probability distribution has an L^2-density with respect to the Lebesgue measure.
- In the critical and supercritical cases, when $b = b(x) \in \mathbf{H}^{-1}_p$ with some $p > 2$ and is divergence-free, the uniqueness of the energy solution is established in [19] for any initial data whose law has a bounded density with respect to the Lebesgue measure.
- Without assuming that b is divergence-free, when $b \in \mathbf{C}^\alpha$ for some $\alpha \in (-\frac{2}{3}, -\frac{1}{2})$, representing a certain Gaussian noise, the authors in [11] and [7] utilized the rough path and paracontrolled theory to independently establish the existence of a unique martingale solution. Additional references to this scenario can be found in [25] for the Lévy process case and [22] for the degenerate kinetic case.

1.1. **Main results.** In this subsection, we present our main results in two cases: the subcritical case and the supercritical case. Our first result establishes the well-posedness of SDE (1.1) in the subcritical case when b is divergence-free.

Theorem 1.2 (Subcritical case)**.** *Let $\alpha_b\in(-1,-\frac12]$ and $p_b,q_b\in[2,\infty]$ satisfy $\frac{d}{p_b}+\frac{2}{q_b}<1+\alpha_b$. Suppose that $b\in\cap_{T>0}\mathbb{L}^{q_b}_T\mathbf{H}^{\alpha_b}_{p_b}$ is divergence-free. Then for any $\mu\in\mathcal{P}(\mathbb{R}^d)$, there is a weak solution to SDE (1.1) starting from initial distribution μ in the sense of Definition 1.1, which is also unique in the class that the following Krylov estimate holds: for any $(\alpha,p,q)\in[\alpha_b,0]\times[2,\infty]^2$ with*

$$\tfrac{q_b}{2}\leqslant q\leqslant q_b,\quad p\leqslant p_b,\quad \alpha-\tfrac{d}{p}-\tfrac{2}{q}\geqslant\alpha_b-\tfrac{d}{p_b}-\tfrac{2}{q_b},\tag{1.5}$$

for any $T>0$, there is a $\theta>0$ such that for any $m\in\mathbb{N}$, $f\in\mathbb{L}^q_T\mathbf{H}^\alpha_p\cap\mathbb{L}^q_TC^\infty_b$ and $0\leqslant t_0<t_1\leqslant T$,

$$\left\|\int_{t_0}^{t_1}f(s,X_s)\mathrm{d}s\right\|_{L^m(\Omega)}\lesssim_C(t_1-t_0)^{\frac{1+\theta}{2}}\|f\|_{\mathbb{L}^q_T\mathbf{H}^\alpha_p},\tag{1.6}$$

where the constant C only depends on $\theta,m,T,d,\alpha,p,q,\alpha_b,p_b,q_b$ and $\|b\|_{\mathbb{L}^{q_b}_T\mathbf{H}^{\alpha_b}_{p_b}}$. Moreover, for each $t>0$ and $X_0=x\in\mathbb{R}^d$, the law of X_t admits a density $\rho_t(x,x')$ called the heat kernel of $\Delta+b\cdot\nabla$, which enjoys the following two sides Gaussian estimate: for fixed $T>0$, all $x,x'\in\mathbb{R}^d$ and $t\in(0,T]$,

$$C_0t^{-d/2}\mathrm{e}^{-c_0|x-x'|^2/t}\leqslant\rho_t(x,x')\leqslant C_1t^{-d/2}\mathrm{e}^{-c_1|x-x'|^2/t},$$

where $C_0,C_1,c_0,c_1>0$ only depend on the parameters T,d,α_b,p_b,q_b and $\|b\|_{\mathbb{L}^{q_b}_T\mathbf{H}^{\alpha_b}_{p_b}}$.

Remark 1.3. *In the subcritical case, our result improves upon* [20, Theorem 2.10] *since our initial distribution is not required to have an L^2-density; it can be a Dirac measure. Moreover, we are working in the whole space, not in the torus. Our proof is based on the Schauder estimate of heat equation in Besov spaces and Zvonkin's transformation. Moreover, by (1.6), one easily sees that the solution does not depend on the choice of the mollifiers.*

To present our main result in the supercritical case, we introduce the following class of distributions for later use:

$$\mathcal{B}:=\left\{b\in\mathscr{S}'(\mathbb{R}^d;\mathbb{R}^d):\|b\|_{\mathcal{B}}:=\sup_{\varphi\in C^\infty_c(\mathbb{R}^d)}\|b\cdot\nabla\varphi\|_{\mathbf{H}^{-1}_2}/\|\varphi\|_{\mathbf{H}^1_2}<\infty\right\},\tag{1.7}$$

where $\mathscr{S}'(\mathbb{R}^d;\mathbb{R}^d)$ stands for the class of $\mathbb{R}^d$-valued Schwartz distributions over $\mathbb{R}^d$.

Examples: (i) Suppose that $b=\mathrm{div}a$, where $a:\mathbb{R}^d\to\mathbb{R}^d\otimes\mathbb{R}^d$ is an antisymmetric matrix-valued bounded measurable function. Then $b\in\mathcal{B}$. In fact, for any $h\in\mathbf{H}^1_2$ and $\varphi\in C^\infty_c(\mathbb{R}^d)$,

$$\langle b\cdot\nabla\varphi,h\rangle=\langle\mathrm{div}a\cdot\nabla\varphi,h\rangle=-\langle a,\nabla\varphi\otimes\nabla h\rangle\leqslant\|a\|_\infty\|\nabla\varphi\|_2\|\nabla h\|_2,$$

where we have used that $\sum_{i,j}a_{ij}\partial_i\partial_j\varphi=0$.

(ii) $\mathbf{B}^{-1}_{\infty,2}\subset\mathcal{B}$. We shall prove it after introducing Besov spaces (see Lemma 2.7 below).

In the following we fix a time level $T>0$. Let $\mathbb{C}_T:=C([0,T];\mathbb{R}^d)$ be the Banach space of all continuous functions which is endowed with the natural filtration $\mathscr{B}_t:=\sigma(\omega_s,s\leqslant t)$. A path in $\mathbb{C}_T$ is denoted by ω. The canonical process is denoted by

$$w_t(\omega)=\omega(t).$$

Our second main result of this paper is:

Theorem 1.4 (Supercritical case)**.** *Let $\alpha_b\in[-1,0]$ and $p_b,q_b\in[2,\infty]$ satisfy $\frac{2}{q_b}+\frac{d}{p_b}<2+\alpha_b$. Suppose that $b\in\mathbb{L}^{q_b}_T\mathbf{H}^{\alpha_b}_{p_b}$ is divergence-free. For any initial distribution $\mu\in\mathcal{P}(\mathbb{R}^d)$ having L^2-density ρ_0, there is a weak solution (X,W) to SDE (1.1) in the sense of Definition 1.1, which has a density $\rho_t\in\mathbb{L}^\infty_TL^2\cap\mathbb{L}^2_T\mathbf{H}^1_2$ that satisfies the Fokker-Planck equation in the distributional sense*

$$\partial_t\rho=\Delta\rho-\mathrm{div}(b\rho).$$

Moreover, consider the approximation SDE

$$X^n_t=X_0+\int_0^tb_n(s,X^n_s)\mathrm{d}s+\sqrt{2}W_t,$$

where $b_n\in\mathbb{L}^{q_b}_TC^\infty_b$ is defined by (1.2). Let $\mathbb{P}_n$ be the law of $X^n_\cdot$ in $\mathbb{C}_T$.

(A) *For any subsequence n_k, there is a subsubsequence n'_k such that for* **any** *initial distribution $\mu \in \mathscr{P}(\mathbb{R}^d)$ with L^2-density, $\mathbb{P}_{n'_k}$ weakly converges to a solution of SDE (1.1) starting from μ.*

(B) *Let $\mathbb{P}_\mu$ be the law of the solution constructed in* **(A)**. *The following Markov property holds:*

$$\mathbb{E}^{\mathbb{P}_\mu}(f(w_t)|\mathscr{B}_s) = \mathbb{E}^{\mathbb{P}_\mu}(f(w_t)|w_s),\ 0 \leqslant s \leqslant t \leqslant T,\ f \in C_b(\mathbb{R}^d).$$

(C) *If in addition that $b = b_1 + b_2$, where $b_1 \in \mathbb{L}^\infty_T \mathscr{B}$ and $b_2 \in \mathbb{L}^2_T L^2$, then without subtracting subsequence, $\mathbb{P}_n$ weakly converges to $\mathbb{P}_\mu$ in $\mathbb{C}_T$ as $n \to \infty$.*

Remark 1.5. *In the supercritical case, we are unable to demonstrate the uniqueness of weak solutions a prioril. However, the assertion* **(A)** *affirms that we can identify a subsequence such that, for any initial value, the corresponding approximation solution converges weakly to a solution. This enables us to select a Markov process. In the case where $b = b_1 + b_2$, with $b_1 \in \mathbb{L}^\infty_T \mathscr{B}$ and $b_2 \in \mathbb{L}^2_T L^2$, it is unique in the approximation sense since it is not necessary to choose a subsequence.*

Remark 1.6. *One of the motivations for exploring the supercritical case arises from the observation of the super-diffusion phenomenon, as discussed in [8, 9]. In this context, the drift term is defined as the Leray projection of the 2-dimensional white noise, specifically represented as:*

$$\mathbf{E}\big(b(f)b(g)\big) = \int_{\mathbb{R}^2} \hat{f}(\xi)\left(\mathbb{I}_{2\times 2} - \frac{\xi \otimes \xi}{|\xi|^2}\right)\hat{g}(-\xi)\mathrm{d}\xi,$$

where $\hat{f}$ is the Fourier transform of Schwartz function $f : \mathbb{R}^2 \to \mathbb{R}^2$. Given the absence of well-posedness results for the related stochastic differential equations as presented in [9], modifications were introduced to the drift term by truncating its Fourier transform, i.e.

$$\mathbf{E}\big(b_\varepsilon(f)b_\varepsilon(g)\big) = \int_{\mathbb{R}^2} 1_{\{|\xi|<1/\varepsilon\}}\hat{f}(\xi)\left(\mathbb{I}_{2\times 2} - \frac{\xi \otimes \xi}{|\xi|^2}\right)\hat{g}(-\xi)\mathrm{d}\xi,$$

where ε is taken 1 in [9]. Note that estimate (6.14) and Lemma 5.3 below allows us to establish the tightness of solutions to the SDE with the modified drift term $b_{\varepsilon_n}/\sqrt{\ln(1/\varepsilon_n)}$, where $\varepsilon_n \to 0$.

The proof of Theorem 1.4 depends on the solvability of PDE. Specifically, we consider the following parabolic equation:

$$\partial_t u = \Delta u + b \cdot \nabla u + f. \tag{1.8}$$

When $f = 0$ and $b = \mathrm{div} a$, where $a : \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ is an antisymmetric matrix-valued measurable function, the Harnack inequality was established in [32] for $a \in L^\infty$ and in [35] for $a \in BMO$. In the case where $u_0 = 0$ and $b \in L^p$ with $\mathrm{div} b = 0$, through De-Giorgi's argument, the maximum estimate was obtained in [49] for $p > d/2$ and in [46, Corollary 1.5] for $p > (d-1)/2$. In this work, for drifts that are distribution-valued and under certain conditions on the divergence of b, we show the weak well-posedness of PDE (1.8) using the energy method. To establish the existence of a weak solution to SDE (1.1), we rely on two types of Krylov's estimates. One is used to prove tightness (see Lemma 5.1), while the other is utilized to take limits (see Lemma 5.5).

To establish conclusions **(A)** and **(C)**, we will make use of the notion of generalized martingale solutions. In cases involving distributional drift, it's important to note that the weak solution is not generally equivalent to the generalized martingale solutions. However, in our proofs, both of these solutions are the limits of smooth approximation solutions. It's worth emphasizing that in the supercritical case, even without uniqueness, we have the flexibility to select a Markov process.

As an application, we consider the following singular interaction particle system in $\mathbb{R}^{Nd}$:

$$\mathrm{d}X^{N,i}_t = \sum_{j\neq i} \gamma_j K(X^{N,i}_t - X^{N,j}_t)\mathrm{d}t + \sqrt{2}\mathrm{d}W^{N,i}_t,\quad i = 1, \cdots, N, \tag{1.9}$$

where $K \in \mathbf{H}^{-1}_\infty(\mathbb{R}^d;\mathbb{R}^d)$ is divergence free, $W^{N,i}_t, i = 1,\cdots,N$ are N-independent standard d-dimensional Brownian motions, $\gamma_j \in \mathbb{R}$ and initial value $\mathbf{X}^N_0 := (X^{N,1}_0,\cdots,X^{N,N}_0)$ has an L^2-density ρ^N_0. Define $b(\mathbf{x})$ for $\mathbf{x} = (x_1,\cdots,x_N)$ as

$$b(\mathbf{x}) := \left(\sum_{j\neq 1}\gamma_j K(x_1 - x_j),\cdots,\sum_{j\neq N}\gamma_j K(x_N - x_j)\right).$$

It can be verified that $b \in \mathbf{H}^{-1}_\infty(\mathbb{R}^{Nd};\mathbb{R}^{Nd})$ is divergence-free. Consequently, by Theorem 1.4 above, the SDE (1.9) has a solution $\mathbf{X}^N_t$ with a density ρ^N_t that satisfies the Liouville equation:

$$\partial_t\rho^N_t = \Delta\rho^N_t - \sum_{i\neq j}\gamma_j K(x_i - x_j)\partial_{x_i}\rho^N_t.$$

In [23, Proposition 1], the entropy solution of the above Liouville equation is directly constructed. However, the existence of a solution to the SDE (1.9) appears to be an open question. In the context of the two-dimensional vortex model, where

$$K(x) = \frac{(x_2,-x_1)}{|x|^2} = \Big(-\partial_{x_1}\arctan\big(\frac{x_2}{x_1}\big), \partial_{x_2}\arctan\big(\frac{x_1}{x_2}\big)\Big) \in \mathbf{H}^{-1}_\infty(\mathbb{R}^2;\mathbb{R}^2)$$

known as the Biot-Savart law, and when γ_j has the same sign, Takanobu [40] established, through a purely probabilistic argument, the existence and uniqueness of a solution that avoids collisions starting from a point $\mathbf{x} = (x_1,\cdots,x_N)$ with $x_i \neq x_j$ (see also [40] for some extensions). For general $\gamma_j \in \mathbb{R}$, Osada [31] showed the same result based on heat kernel estimates for generators in a generalized divergence form, obtained in [32], and on potential theoretical results. In our construction, two particles of the solution are allowed to meet, and as a result, the singular point can be touched.

1.2. **Comparison with related works.** In this subsection we make a detailed comparison with the related well-known results.

Subcritical case: $\alpha \in (-\frac{1}{2},0]$**.** Our result, Theorem 4.8 below, not only encompasses but also extends the work presented in [14, 48] for the range $\alpha \in (-\frac{1}{2},0]$. In particular, when $\alpha = \frac{1}{2}$, our conditions for the coefficient b are not compared to the conditions stated in [48], where $b \in \mathbf{H}^{-\frac{1}{2}}_p$ with $p > 2d$ is time-independent. Instead, our result allows b to belong to $\mathbb{L}^q_T\mathbf{B}^{-\frac{1}{2}}_{p,1}$ with $\frac{d}{p} + \frac{2}{q} < \frac{1}{2}$, which encompasses cases with time integrability index $q \in (4,\infty]$.

Energy method comparison for $\alpha \in (-1,-\frac{1}{2}]$**.** In contrast to the energy method employed in [20], our approach differs significantly. We will outline these differences within the context of the following three aspects.

- **Dimensional dependence:** Our supercritical condition, which requires $b \in \mathbb{L}^q_T\mathbf{H}^{-1}_p$ with $\frac{2}{q} + \frac{d}{p} < 1$, is dimension-dependent. This is in contrast to the assumption made in [20, 19], where they require $b \in \mathbf{H}^{-1}_{2+}$, and this assumption is independent of dimension.
- **Initial conditions:** In Lemma 5.3 below, our method achieves the tightness without any assumption on the initial data. Furthermore, for the existence of a weak solution and the uniqueness of the generalized martingale solution, we only require that the law of the initial data has an L^2-density with respect to the Lebesgue measure.
- **Methodology:** Our methodologies differ fundamentally from each other. In [20], the focus is on establishing the existence of a unique energy solution on the torus $\mathbb{T}^d$. The key aspect of the energy method is to find a stationary distribution μ and show the following Itô trick:

$$\mathbb{E}_\mu\left[\sup_{t\in[0,T]}\left|\int_0^t f(X_s)\mathrm{d}s\right|^m\right] \lesssim_C T^{m/2}\|(-\mathscr{L}_s)^{-1/2}f\|^p_{L^p(\mathbb{T}^d)},$$

where $\mathscr{L}_s := \frac{1}{2}(\mathscr{L} + \mathscr{L}^*)$ is the symmetric part of the generator $\mathscr{L} := \Delta + b \cdot \nabla$, which is exactly the Laplacian Δ when $\mathrm{div} b = 0$. On the other hand, in order to consider the whole space $\mathbb{R}^d$, the authors in [19] introduce an additional term $-\frac{1}{n}X_t$ to construct the stationary distribution, followed by taking the limit as $n \to \infty$. This approach makes it challenging to address the time-dependent case. In our work, we make use of a classical Krylov-type estimate, where the condition $\frac{2}{q} + \frac{d}{p} < 1$ is sharp due to scaling. Reducing the assumption to $p > 2$ is difficult within our framework. Importantly, our method does not necessitate the existence of a stationary distribution, which can be challenging in unbounded spaces. Additionally, our approach allows us to handle time-dependent drifts efficiently. Furthermore, our method could be used to deal with the multiplicative noise as well as the degenerate kinetic SDEs, which seems not possible by the energy method developed from [20].

Rough path method. Recently, numerous studies have been dedicated to investigating singular SDEs driven by fractional Brownian motions. Due to the absence of Markov and martingale properties for fractional Brownian motions, traditional stochastic analysis arguments are no longer applicable. In [6, 18], path-by-path well-posedness is established by using the Young integral in rough path theory and the stochastic sewing lemma. It's important to note that this method only works with subcritical drifts.

1.3. **Organization of the paper.** In Section 2, we provide a brief overview of classical Besov and Bessel potential spaces and establish a precise Schauder estimate for classical heat equations.

Section 3 is dedicated to the study of the solvability of the linear PDE (1.8) with a distributional drift b. In the subcritical case, we utilize Duhamel's formulation and fixed point theorem, while in the supercritical case, we apply maximum estimates and the energy method.

In Section 4, we focus on proving Theorem 1.2, primarily through the Zvonkin transformation. A crucial step in this process is to establish the following uniform Krylov estimate for the approximation solution X^n: for any $m \in \mathbb{N}$, $0 \leqslant t_0 < t_1 \leqslant T$ and $f \in \mathbb{L}^q_T \mathbf{H}^\alpha_p$,

$$\sup_n \left\| \int_{t_0}^{t_1} f(s, X^n_s) \mathrm{d}s \right\|_{L^m(\Omega)} \leqslant C(t_1 - t_0)^{\frac{1+\theta_0}{2}} \|f\|_{\mathbb{L}^q_T \mathbf{H}^\alpha_p},$$

where $\theta_0 > 0$ and (α, p, q) satisfies (1.5). This estimate ensures that A^b_t is a zero-energy process, which, in turn, enables us to apply the generalized Itô's formula and derive the effectiveness of Zvonkin's transformation. Subsequently, the uniqueness and heat kernel estimates follow by the corresponding results of the transformed SDEs.

Section 5 is dedicated to the proof of Theorem 1.4. To establish the existence of a weak solution, we need to establish two types of Krylov's estimates (as presented in Lemma 5.1 and Lemma 5.5). For the well-posedness of the associated generalized martingale problem, which is crucial for **(A)**-**(C)**, our approach primarily revolves around solving PDE (1.8)

Finally in Section 6, we apply the results to the diffusion in random environments. In particular, vector-valued Gaussian random fields are provided to illustrate our results.

Throughout this paper, we use C with or without subscripts to denote constants, whose values may change from line to line. We also use := to indicate a definition and set

$$a \wedge b := \max(a, b), \quad a \vee b := \min(a, b).$$

By $A \lesssim_C B$ or simply $A \lesssim B$, we mean that for some constant $C \geqslant 1$, $A \leqslant CB$.

2. Prelimiaries

In this section, we recall the definition and some properties of Besov and Bessel potential spaces. Additionally, we establish a Schauder estimate for the heat equation in Besov spaces, which appears to be new. This estimate will be used to solve the Kolmogorov equation in the subcritical case.

Let $\mathscr{S}(\mathbb{R}^d)$ be the Schwartz space of all rapidly decreasing functions on $\mathbb{R}^d$, and $\mathscr{S}'(\mathbb{R}^d)$ the dual space of $\mathscr{S}(\mathbb{R}^d)$ called Schwartz generalized function (or tempered distribution) space. Given $f \in \mathscr{S}(\mathbb{R}^d)$, the Fourier transform $\hat{f}$ and inverse Fourier transform $\check{f}$ are defined, respectively, by

$$\hat{f}(\xi) := \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}\xi\cdot x} f(x)\mathrm{d}x, \quad \xi \in \mathbb{R}^d,$$
$$\check{f}(x) := \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}\xi\cdot x} f(\xi)\mathrm{d}\xi, \quad x \in \mathbb{R}^d.$$

For $q \in [1, \infty]$ and a normed space $\mathbb{B}$, we shall simply denote

$$\mathbb{L}^q_T\mathbb{B} := L^q([0,T];\mathbb{B}), \quad \mathbb{L}^q_T := \mathbb{L}^q_T L^q,$$

where $(L^p, \|\cdot\|_p)$ is the usual L^p-space in $\mathbb{R}^d$. Moreover, we use $C_b = C_b(\mathbb{R}^d)$ to denote the Banach space of all bounded continuous functions on $\mathbb{R}^d$, and $C^\infty_b = C^\infty_b(\mathbb{R}^d)$ the space of all smooth functions with bounded derivatives of all orders.

2.1. **Besov and Bessel potential spaces.** To introduce the Besov space, we first introduce the dyadic partition of unity. Let ϕ_{-1} be a symmetric nonnegative $C^\infty-$function on $\mathbb{R}^d$ with

$$\phi_{-1}(\xi) = 1 \text{ for } \xi \in B_{1/2} \text{ and } \phi_{-1}(\xi) = 0 \text{ for } \xi \notin B_{2/3}.$$

For $j \geqslant 0$, we define

$$\phi_j(\xi) := \phi_{-1}(2^{-(j+1)}\xi) - \phi_{-1}(2^{-j}\xi). \tag{2.1}$$

By definition, one sees that for $j \geqslant 0$, $\phi_j(\xi) = \phi_0(2^{-j}\xi)$ and

$$\operatorname{supp}\phi_j \subset B_{2^{j+2}/3} \setminus B_{2^{j-1}}, \quad \sum_{j=-1}^{n} \phi_j(\xi) = \phi_{-1}(2^{-(n+1)}\xi) \to 1, \quad n \to \infty.$$

Definition 2.1. *For given $j \geqslant -1$, the block operator $\mathcal{R}_j$ is defined on $\mathscr{S}'(\mathbb{R}^d)$ by*

$$\mathcal{R}_j f(x) := (\phi_j \hat{f})\check{}(x) = \check{\phi}_j * f(x),$$

with the convention $\mathcal{R}_j \equiv 0$ for $j \leqslant -2$. In particular, for $j \geqslant 0$,

$$\mathcal{R}_j f(x) = 2^{jd} \int_{\mathbb{R}^d} \check{\phi}_0(2^j y) f(x-y)\mathrm{d}y. \tag{2.2}$$

For $j \geqslant -1$, by definition it is easy to see that

$$\mathcal{R}_j = \mathcal{R}_j \widetilde{\mathcal{R}}_j, \quad \text{where} \quad \widetilde{\mathcal{R}}_j := \mathcal{R}_{j-1} + \mathcal{R}_j + \mathcal{R}_{j+1}, \tag{2.3}$$

and $\mathcal{R}_j$ is symmetric in the sense that

$$\langle g, \mathcal{R}_j f\rangle = \langle f, \mathcal{R}_j g\rangle, \quad f, g \in \mathscr{S}'(\mathbb{R}^d),$$

where $\langle\cdot,\cdot\rangle$ stands for the dual pair between $\mathscr{S}'(\mathbb{R}^d)$ and $\mathscr{S}(\mathbb{R}^d)$.

Now we recall the definition of Besov spaces (see [1]).

Definition 2.2. *Let $p, q \in [1, \infty]$ and $s \in \mathbb{R}$. The Besov space $\mathbf{B}^s_{p,q}$ is defined by*

$$\mathbf{B}^s_{p,q} := \left\{ f \in \mathscr{S}'(\mathbb{R}^d) : \|f\|_{\mathbf{B}^s_{p,q}} := \left(\sum_{j\geqslant -1} 2^{sjq} \|\mathcal{R}_j f\|_p^q \right)^{1/q} < \infty \right\}.$$

Remark 2.3. *For $s \in (0,1)$, an equivalent characterization of $\mathbf{B}^s_{p,q}$ is given by (see [1, p74, Theorem 2.36] or [22, Theorem 2.7])*

$$\|f\|_{\mathbf{B}^s_{p,q}} \asymp \left(\int_{|h|\leqslant 1} \left(\frac{\|f(\cdot+h) - f(\cdot)\|_p}{|h|^s} \right)^q \frac{\mathrm{d}h}{|h|^d} \right)^{1/q} + \|f\|_p.$$

In particular, for any $s\in(0,1)$ and $p\in[1,\infty]$, there is a constant $C=C(s,d,p)>0$ such that

$$\|f(\cdot+h)-f(\cdot)\|_p\leqslant C\|f\|_{\mathbf{B}^s_{p,\infty}}(|h|^s\wedge 1),$$

and for any $s_0\in\mathbb{R}$,

$$\|f(\cdot+h)-f(\cdot)\|_{\mathbf{B}^{s_0}_{p,\infty}}\leqslant C\|f\|_{\mathbf{B}^{s_0+s}_{p,\infty}}(|h|^s\wedge 1). \tag{2.4}$$

For $(\alpha,p)\in(\mathbb{R}\setminus\mathbb{Z})\times[1,\infty]$, let $\mathbf{H}^\alpha_p$ be the Bessel potential space defined by:

$$\mathbf{H}^\alpha_p:=\big\{f\in\mathscr{S}'(\mathbb{R}^d):\|f\|_{\mathbf{H}^\alpha_p}:=\|(\mathbb{I}-\Delta)^{\alpha/2}f\|_p<\infty\big\},$$

where $(\mathbb{I}-\Delta)^{\alpha/2}f$ is defined through Fourier's transform

$$(\mathbb{I}-\Delta)^{\alpha/2}f:=\big((1+|\cdot|^2)^{\alpha/2}\hat f\big)\check{}.$$

For $\alpha=0,1,2,\cdots$ and $p\in[1,\infty]$, we define

$$\mathbf{H}^\alpha_p:=\big\{f\in\mathscr{S}'(\mathbb{R}^d):\|f\|_{\mathbf{H}^\alpha_p}:=\|f\|_p+\|\nabla^\alpha f\|_p<\infty\big\},$$

and for $\alpha=-1,-2,\cdots$ and $p\in[1,\infty]$,

$$\mathbf{H}^\alpha_p:=(\mathbf{H}^{-\alpha}_{p/(p-1)})'.$$

Note that we do not use the space $\mathbf{H}^\alpha_1$ for $\alpha<0$. Moreover, for integer $\alpha\in\mathbb{N}$ and $p\in(1,\infty)$, an equivalent norm in $\mathbf{H}^\alpha_p$ is given by (cf. [37, p135, Theorem 3])

$$\|f\|_{\mathbf{H}^\alpha_p}\asymp\|(I-\Delta)^{\alpha/2}f\|_p. \tag{2.5}$$

Below we recall some well-known facts about Besov spaces and Bessel potential spaces, where (2.6) and (2.11) below are easily derived by definition.

Lemma 2.4. *(i) For any $p\in[1,\infty]$, $s'>s$ and $q\in[1,\infty]$, it holds that*

$$\mathbf{B}^0_{p,1}\hookrightarrow L^p\hookrightarrow\mathbf{B}^0_{p,\infty},\quad \mathbf{B}^{s'}_{p,\infty}\hookrightarrow\mathbf{B}^s_{p,1}\hookrightarrow\mathbf{B}^s_{p,q}. \tag{2.6}$$

(ii) For any $\alpha\in\mathbb{R}$ and $p\in(1,\infty)$, we have the following embedding (see [5, Theorem 6.4.4])

$$\begin{cases}\mathbf{B}^\alpha_{p,2}\hookrightarrow\mathbf{H}^\alpha_p, & p\in[2,\infty),\\ \mathbf{H}^\alpha_p\hookrightarrow\mathbf{B}^\alpha_{p,2}, & p\in(1,2],\end{cases} \tag{2.7}$$

and for $\alpha\in\mathbb{R}$ and $p\in(1,\infty]$ (see [5, Theorem 6.2.4]),

$$\mathbf{B}^\alpha_{p,1}\subset\mathbf{H}^\alpha_p\subset\mathbf{B}^\alpha_{p,\infty}.$$

(iii) For $1\leqslant p_1\leqslant p\leqslant\infty$, $q\in[1,\infty]$ and $\alpha=\alpha_1-\frac{d}{p_1}+\frac{d}{p}$, it holds that (see [5, Theorem 6.5.1])

$$\|f\|_{\mathbf{B}^\alpha_{p,q}}\lesssim_C\|f\|_{\mathbf{B}^{\alpha_1}_{p_1,q}}, \tag{2.8}$$

and for $1<p_1\leqslant p<\infty$,

$$\|f\|_{\mathbf{H}^\alpha_p}\lesssim_C\|f\|_{\mathbf{H}^{\alpha_1}_{p_1}}. \tag{2.9}$$

(iv) Let $\alpha\in[-1,0)$ and $p\in(1,\infty)$. For any $p_1,p_2\in(1,\infty)$ satisfying

$$p_1\geqslant p,\ p_2\geqslant\tfrac{p_1}{p_1-1},\quad \tfrac1p\leqslant\tfrac1{p_1}+\tfrac1{p_2}<\tfrac1p+\tfrac{\alpha}{d},$$

there is a constant $C>0$ such that for all $f\in\mathbf{H}^\alpha_{p_1}$ and $g\in\mathbf{H}^{-\alpha}_{p_2}$ (see [48, Lemma 2.1]),

$$\|fg\|_{\mathbf{H}^\alpha_p}\lesssim_C\|f\|_{\mathbf{H}^\alpha_{p_1}}\|g\|_{\mathbf{H}^{-\alpha}_{p_2}}, \tag{2.10}$$

and

$$\|fg\|_{\mathbf{H}^{-1}_\infty}\leqslant\|f\|_{\mathbf{H}^{-1}_\infty}\|g\|_{\mathbf{H}^1_\infty}. \tag{2.11}$$

Next we recall the well-known Bony decomposition and related paraproduct estimates. Let S_k be the cut-off low frequency operator defined by

$$S_k f := \sum_{j=-1}^{k-1} \mathcal{R}_j f \to f, \quad k \to \infty. \tag{2.12}$$

For $f, g \in \mathscr{S}'(\mathbb{R}^d)$, define

$$f \prec g := \sum_{k \geqslant -1} S_{k-1} f \mathcal{R}_k g, \quad f \circ g := \sum_{|i-j| \leqslant 1} \mathcal{R}_i f \mathcal{R}_j g.$$

The Bony decomposition of fg is formally given by (cf. [1])

$$fg = f \prec g + f \circ g + g \prec f =: f \preceq g + f \succ g. \tag{2.13}$$

We recall the following paraproduct estimates (cf. [1, Theorem 2.82 and Theorem 2.85]).

Lemma 2.5. *Let $p, p_1, p_2, q, q_1, q_2 \in [1, \infty]$ with $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$ and $\frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2}$ and $\alpha, \beta \in \mathbb{R}$.*

(i) If $\beta < 0$, then there is a constant $C = C(\alpha, \beta, d, p, q, p_1, q_1, p_2, q_2) > 0$ such that

$$\|f \prec g\|_{\mathbf{B}^{\alpha+\beta}_{p,q}} \lesssim_C \|f\|_{\mathbf{B}^{\beta}_{p_1,q_1}} \|g\|_{\mathbf{B}^{\alpha}_{p_2,q_2}}.$$

Moreover, for $\beta = 0$, we have

$$\|f \prec g\|_{\mathbf{B}^{\alpha}_{p,q}} \lesssim_C \|f\|_{p_1} \|g\|_{\mathbf{B}^{\alpha}_{p_2,q}}.$$

(ii) If $\alpha + \beta > 0$, then there is a constant $C = C(\alpha, \beta, d, p, q, p_1, q_1, p_2, q_2) > 0$ such that

$$\|f \circ g\|_{\mathbf{B}^{\alpha+\beta}_{p,q}} \lesssim_C \|f\|_{\mathbf{B}^{\beta}_{p_1,q_1}} \|g\|_{\mathbf{B}^{\alpha}_{p_2,q_2}}.$$

Moreover, when $\alpha + \beta = 0$ and $q = 1$, we have

$$\|f \circ g\|_{\mathbf{B}^{0}_{p,\infty}} \lesssim_C \|f \circ g\|_p \lesssim_C \|f\|_{\mathbf{B}^{-\alpha}_{p_1,q_1}} \|g\|_{\mathbf{B}^{\alpha}_{p_1,q_2}}. \tag{2.14}$$

Proof. We only prove (2.14). By definition and Hölder's inequality, we have

$$\begin{aligned} \|f \circ g\|_{\mathbf{B}^{0}_{p,\infty}} &\lesssim \|f \circ g\|_p \leqslant \sum_{|i-j| \leqslant 1} \|\mathcal{R}_i f \mathcal{R}_j g\|_p \leqslant \sum_{|i-j| \leqslant 1} \|\mathcal{R}_i f\|_{p_1} \|\mathcal{R}_j g\|_{p_2} \\ &\lesssim \left(\sum_i 2^{-q_1 \alpha i} \|\mathcal{R}_i f\|_{p_1}^{q_1} \right)^{1/q_1} \left(\sum_j 2^{q_2 \alpha j} \|\mathcal{R}_i f\|_{p_2}^{q_2} \right)^{1/q_2} = \|f\|_{\mathbf{B}^{-\alpha}_{p_1,q_1}} \|g\|_{\mathbf{B}^{\alpha}_{p_1,q_2}}. \end{aligned}$$

The proof is complete. □

Now we discuss a decomposition of $b \cdot \nabla u$, that arises in the study of PDE (1.3), where $b \in \mathscr{S}'(\mathbb{R}^d)$ is a distribution-valued vector field and $u \in \mathscr{S}(\mathbb{R}^d)$. By Bony's decomposition, we can write

$$b \cdot \nabla u = \{\mathrm{div}(b \preceq u) + b \succ \nabla u\} - \mathrm{div} b \preceq u =: b \odot \nabla u - \mathrm{div} b \preceq u. \tag{2.15}$$

In particular, if $\mathrm{div} b = 0$, then

$$b \cdot \nabla u = b \odot \nabla u.$$

We have the following regularity estimate.

Lemma 2.6. *For any $\alpha \in (-1, 0]$ and $p, q \in [1, \infty]$, there is a $C = C(d, \alpha, p, q) > 0$ such that*

$$\|b \odot \nabla u\|_{\mathbf{B}^{\alpha}_{p,q}} \lesssim_C \|b\|_{\mathbf{B}^{\alpha}_{p,q}} \|u\|_{\mathbf{B}^{1+d/p}_{p,1}} \tag{2.16}$$

and

$$\|\mathrm{div} b \preceq u\|_{\mathbf{B}^{0}_{p/2,\infty}} \lesssim_C \|\mathrm{div} b\|_{\mathbf{B}^{-\alpha-2}_{p,q/(q-1)}} \|u\|_{\mathbf{B}^{2+\alpha}_{p,q}}. \tag{2.17}$$

Proof. Since $\alpha\in(-1,0]$, by Lemma 2.5 we have

$$\|\mathrm{div}(b\preceq u)\|_{\mathbf{B}^{\alpha}_{p,q}}\lesssim\|b\preceq u\|_{\mathbf{B}^{1+\alpha}_{p,q}}\lesssim\|b\|_{\mathbf{B}^{\alpha}_{p,q}}\|u\|_{\mathbf{B}^{1}_{\infty,\infty}}$$

and

$$\|b\succ\nabla u\|_{\mathbf{B}^{\alpha}_{p,q}}\lesssim\|b\|_{\mathbf{B}^{\alpha}_{p,q}}\|\nabla u\|_\infty\lesssim\|b\|_{\mathbf{B}^{\alpha}_{p,q}}\|u\|_{\mathbf{B}^{1}_{\infty,1}}.$$

Hence, by embedding (2.8),

$$\|b\odot\nabla u\|_{\mathbf{B}^{\alpha}_{p,q}}\lesssim\|b\|_{\mathbf{B}^{\alpha}_{p,q}}\|u\|_{\mathbf{B}^{1}_{\infty,1}}\lesssim_C\|b\|_{\mathbf{B}^{\alpha}_{p,q}}\|u\|_{\mathbf{B}^{1+d/p}_{p,1}}.$$

For the second estimate (2.17), it follows by Lemma 2.5. □

We also have the following easy result.

Lemma 2.7. *If $b,\mathrm{div}b\in\mathbf{B}^{-1}_{\infty,2}$, then $b\in\mathcal{B}$, where $\mathcal{B}$ is defined in (1.7).*

Proof. By definition (2.15), (2.7) and Lemma 2.5, we have

$$\begin{aligned}\|b\odot\nabla u\|_{\mathbf{H}^{-1}_2}&\leqslant\|\mathrm{div}(b\preceq u)\|_{\mathbf{H}^{-1}_2}+\|b\succ\nabla u\|_{\mathbf{H}^{-1}_2}\lesssim\|b\preceq u\|_2+\|b\succ\nabla u\|_{\mathbf{B}^{-1}_{2,2}}\\&\lesssim\|b\prec u\|_{\mathbf{B}^{0}_{2,1}}+\|b\circ u\|_2+\|b\|_{\mathbf{B}^{-1}_{\infty,2}}\|\nabla u\|_2\lesssim\|b\|_{\mathbf{B}^{-1}_{\infty,2}}\big(\|u\|_{\mathbf{B}^{1}_{2,2}}+\|u\|_{\mathbf{H}^1_2}\big),\end{aligned}$$

and

$$\|\mathrm{div}b\preceq u\|_{\mathbf{H}^{-1}_2}\lesssim\|\mathrm{div}b\prec u\|_{\mathbf{B}^{-1}_{2,2}}+\|\mathrm{div}b\circ u\|_{\mathbf{B}^{-1}_{2,2}}\lesssim\|\mathrm{div}b\|_{\mathbf{B}^{-1}_{\infty,2}}\|u\|_{\mathbf{B}^{1}_{2,2}}.$$

Combining the above two estimates and by the definition of $\mathcal{B}$ and (2.7), we get $b\in\mathcal{B}$. □

2.2. **Schauder's estimate for heat equation in Besov spaces.** In this subsection we study the regularity estimates for heat equation in Besov spaces. Let $(P_t)_{t\geqslant0}$ be the Gaussian heat semigroup given by

$$P_tf=p_t*f\quad\text{where}\quad p_t(x)=(4\pi t)^{-\frac{d}{2}}\mathrm{e}^{-\frac{|x|^2}{4t}}.$$

For $\lambda\geqslant0$ and a time-dependent distribution $f_t(\cdot):\mathbb{R}_+\to\mathscr{S}'(\mathbb{R}^d)$, we define

$$u^\lambda_t(x):=\mathscr{I}^\lambda_t(f)(x):=\int_0^t\mathrm{e}^{-\lambda(t-s)}P_{t-s}f_s(x)\mathrm{d}s,\quad t>0.\tag{2.18}$$

In particular, in the distributional sense,

$$\partial_tu^\lambda=(\Delta-\lambda)u^\lambda+f.\tag{2.19}$$

We now show the following Schauder estimate, which shall play a basic role in the subcritical case.

Lemma 2.8. *Let $\alpha\in\mathbb{R}$, $1\leqslant q\leqslant q'\leqslant\infty$ and $1\leqslant p\leqslant p'\leqslant\infty$. Define*

$$\alpha':=\tfrac{2}{q}-\tfrac{2}{q'}+\tfrac{d}{p}-\tfrac{d}{p'}.$$

For any $T>0$, there is a constant $C=C(T,\alpha,d,p,q,p',q')>0$ such that for all $\lambda\geqslant0$,

$$\|\mathscr{I}^\lambda(f)\|_{\mathbb{L}^{q'}_T\mathbf{B}^{2+\alpha-\alpha'}_{p',q}}\lesssim_C\|f\|_{\mathbb{L}^q_T\mathbf{B}^{\alpha}_{p,q}},\tag{2.20}$$

and for $\theta\in(0,2+\frac{2}{q'}-\frac{2}{q})$,

$$\|\mathscr{I}^\lambda(f)\|_{\mathbb{L}^{q'}_T\mathbf{B}^{2+\alpha-\alpha'-\theta}_{p',1}}\lesssim_{C,\theta}(1+\lambda)^{-\frac{\theta}{2}}\|f\|_{\mathbb{L}^q_T\mathbf{B}^{\alpha}_{p,\infty}}.\tag{2.21}$$

Proof. Let $r\in[1,\infty]$ be defined by $1+\frac{1}{p'}=\frac{1}{r}+\frac{1}{p}$ and $T>0$. For any $j\geqslant-1$, by Young's inequality and Bernstein's inequality (see [1, Lemma 2.1, p52]), one sees that

$$\begin{aligned}\|\mathcal{R}_j\mathscr{I}^\lambda_t(f)\|_{p'}&\leqslant\int_0^t\mathrm{e}^{-\lambda(t-s)}\|\mathcal{R}_j(p_{t-s}*f_s)\|_{p'}\mathrm{d}s\\&\overset{(2.3)}{=}\int_0^t\mathrm{e}^{-\lambda(t-s)}\|\mathcal{R}_jp_{t-s}*\widetilde{\mathcal{R}}_jf_s\|_{p'}\mathrm{d}s\end{aligned}$$

$$\leqslant \int_0^t \mathrm{e}^{-\lambda(t-s)} \|\mathcal{R}_j p_{t-s}\|_r \|\widetilde{\mathcal{R}}_j f_s\|_p \mathrm{d}s$$
$$\lesssim 2^{(d-\frac{d}{r})j} \int_0^t \mathrm{e}^{-\lambda(t-s)} \|\mathcal{R}_j p_{t-s}\|_1 \|\widetilde{\mathcal{R}}_j f_s\|_p \mathrm{d}s.$$

Let $r' \in [1,\infty]$ be defined by $1+\frac{1}{q'} = \frac{1}{r'}+\frac{1}{q}$. By Young's inequality for time variable, we have

$$\left(\int_0^T \|\mathcal{R}_j \mathscr{I}^\lambda_t(f)\|^{q'}_{p'} \mathrm{d}t\right)^{\frac{1}{q'}} \lesssim 2^{(d-\frac{d}{r})j} \left(\int_0^T \mathrm{e}^{-r'\lambda s} \|\mathcal{R}_j p_s\|_1^{r'} \mathrm{d}s\right)^{\frac{1}{r'}} \left(\int_0^T \|\widetilde{\mathcal{R}}_j f_s\|_p^q \mathrm{d}s\right)^{\frac{1}{q}}.$$

It is well-known that for any $\theta \geqslant 0$, there is a constant $C = C(\theta, d) > 0$ such that (see [21])

$$\|\mathcal{R}_j p_s\|_1 \lesssim_C 1 \wedge (2^{2j}s)^{-\theta}, \quad j \geqslant -1,\ s \in (0,T].$$

If $r' < \infty$, then by the change of variable, we have

$$\left(\int_0^T \mathrm{e}^{-r'\lambda s} \|\mathcal{R}_j p_s\|_1^{r'} \mathrm{d}s\right)^{\frac{1}{r'}} \lesssim \left(\int_0^\infty \mathrm{e}^{-r'\lambda s}[1 \wedge (2^{2j}s)^{-2}] \mathrm{d}s\right)^{\frac{1}{r'}}$$
$$= \left(2^{-2j} \int_0^\infty \mathrm{e}^{-r' 2^{-2j}\lambda s}[1 \wedge s^{-2}] \mathrm{d}s\right)^{\frac{1}{r'}}$$
$$\leqslant \left(2^{-2j}(1 \wedge (r' 2^{-2j}\lambda)^{-1})\right)^{\frac{1}{r'}}.$$

If $r' = \infty$, then we directly have

$$\sup_{j\geqslant -1} \sup_{s \geqslant 0} \mathrm{e}^{-\lambda s} \|\mathcal{R}_j p_s\|_1 \lesssim 1.$$

Thus combining the above calculations, we obtain that for any $j \geqslant -1$ and $\lambda \geqslant 0$,

$$\|\mathcal{R}_j \mathscr{I}^\lambda_\cdot(f)\|_{\mathbb{L}^{q'}_T L^{p'}} = \left(\int_0^T \|\mathcal{R}_j \mathscr{I}^\lambda_t(f)\|^{q'}_{p'} \mathrm{d}t\right)^{\frac{1}{q'}}$$
$$\lesssim 2^{(d-\frac{d}{r}-\frac{2}{r'})j}(1 \wedge (2^{-2j}\lambda)^{-\frac{1}{r'}}) \left(\int_0^T \|\widetilde{\mathcal{R}}_j f_s\|_p^q \mathrm{d}s\right)^{\frac{1}{q}}$$
$$= 2^{(\alpha'-2)j}(1 \wedge (2^{-2j}\lambda)^{-\frac{1}{r'}}) \left(\int_0^T \|\widetilde{\mathcal{R}}_j f_s\|_p^q \mathrm{d}s\right)^{\frac{1}{q}},$$

where $\alpha' = d - \frac{d}{r} - \frac{2}{r'} = \frac{2}{q} - \frac{2}{q'} + \frac{d}{p} - \frac{d}{p'}$.

Now, for any $\gamma \in [1,q]$ and $\theta \geqslant 0$, by the definition of Besov norm and Minkowskii's inequality,

$$\|\mathscr{I}^\lambda(f)\|_{\mathbb{L}^{q'}_T \mathbf{B}^{2+\alpha-\alpha'-\theta}_{p',\gamma}} = \left\| \left(\sum_{j\geqslant -1} 2^{\gamma(2+\alpha-\alpha'-\theta)j} \|\mathcal{R}_j \mathscr{I}^\lambda_\cdot(f)\|^\gamma_{p'} \right)^{\frac{1}{\gamma}} \right\|_{\mathbb{L}^{q'}_T}$$
$$\leqslant \left(\sum_{j\geqslant -1} 2^{\gamma(2+\alpha-\alpha'-\theta)j} \|\mathcal{R}_j \mathscr{I}^\lambda_\cdot(f)\|^\gamma_{\mathbb{L}^{q'}_T L^{p'}} \right)^{\frac{1}{\gamma}}$$
$$\lesssim \left(\sum_{j\geqslant -1} 2^{\gamma(\alpha-\theta)j}(1 \wedge (2^{-2j}\lambda)^{-\frac{\gamma}{r'}}) \left(\int_0^T \|\widetilde{\mathcal{R}}_j f_s\|_p^q \mathrm{d}s\right)^{\frac{\gamma}{q}} \right)^{\frac{1}{\gamma}}$$
$$= \left(\sum_{j\geqslant -1} 2^{-\gamma\theta j}(1 \wedge (2^{-2j}\lambda)^{-\frac{\gamma}{r'}}) \left(\int_0^T 2^{q\alpha j} \|\widetilde{\mathcal{R}}_j f_s\|_p^q \mathrm{d}s\right)^{\frac{\gamma}{q}} \right)^{\frac{1}{\gamma}}.$$

In particular, for $\theta=0$, if we take $\gamma=q$, then we get (2.20). For $\theta>0$, if we take $\gamma=1$, then

$$\|\mathscr{I}^\lambda(f)\|_{\mathbb{L}^{q'}_T\mathbf{B}^{2+\alpha-\alpha'-\theta}_{p',1}}\lesssim\sum_{j\geqslant-1}2^{-\theta j}(1\wedge(2^{-2j}\lambda)^{-\frac{1}{r'}})\|f\|_{\mathbb{L}^q_T\mathbf{B}^\alpha_{p,\infty}}.$$

Note that for $\theta\in(0,\frac{2}{r'})=(0,2+\frac{2}{q'}-\frac{2}{q})$,

$$\sum_{j\geqslant-1}2^{-\theta j}(1\wedge(2^{-2j}\lambda)^{-\frac{1}{r'}})\lesssim_C 1\wedge\lambda^{-\frac{\theta}{2}}\lesssim_C(1+\lambda)^{-\frac{\theta}{2}}.$$

The proof is complete. □

Remark 2.9. *It is noted that (2.21) does not hold for $\theta=0$. Indeed, by (2.7) and Fourier's transform,*

$$\|\mathscr{I}^\lambda(f)\|_{\mathbb{L}^2_T\mathbf{B}^2_{2,2}}\asymp\|\mathscr{I}^\lambda(f)\|_{\mathbb{L}^2_T\mathbf{H}^2_2}\asymp\|f\|_{\mathbb{L}^2_T\mathbf{H}^0_2}\asymp\|f\|_{\mathbb{L}^2_T\mathbf{B}^0_{2,2}},$$

where $\asymp$ means that both sides are comparable up to a constant. Let $1\leqslant q\leqslant q'\leqslant\infty$ and $1\leqslant p\leqslant p'\leqslant\infty$ and $\theta\in(0,2-\frac{2}{q})$. By (2.20), (2.21) and $\mathbf{B}^\alpha_{p,q}\hookrightarrow\mathbf{B}^\alpha_{p,\infty}$, we immediately have

$$\|\mathscr{I}^\lambda(f)\|_{\mathbb{L}^{q'}_T\mathbf{B}^{2+\alpha-(\frac{2}{q}-\frac{2}{q'}+\frac{d}{p}-\frac{d}{p'})}_{p',q}}+(1+\lambda)^{\frac{\theta}{2}}\|\mathscr{I}^\lambda(f)\|_{\mathbb{L}^\infty_T\mathbf{B}^{2+\alpha-(\frac{2}{q}+\frac{d}{p}-\frac{d}{p'})-\theta}_{p',1}}\lesssim_C\|f\|_{\mathbb{L}^q_T\mathbf{B}^\alpha_{p,q}}.\tag{2.22}$$

3. A study of PDEs with distribution drifts

In this section, with $\lambda\geqslant0$, our objective is to investigate the solvability of the nonhomogeneous partial differential equation:

$$\partial_t u=\Delta u-\lambda u+b\cdot\nabla u+f,\quad u_0\equiv0,\tag{3.1}$$

where b and f are distributions. We will establish the well-posedness of PDE (3.1) in two cases: the subcritical case and the supercritical case, utilizing different methods. In the subcritical case, we employ the Duhamel formula, while in the supercritical case, we utilize the maximal principle and the energy method. Throughout this section we fix a terminal time $T>0$.

3.1. Subcritical case: $\frac{d}{p}+\frac{2}{q}<1+\alpha$.

In this section, we show the existence and uniqueness of weak solutions to PDE (3.1) by using Lemma 2.6 under the following subcritical conditions:

($\mathbf{H}^{\rm sub}$) Let $(\alpha_b,p_b,q_b)\in(-1,-\frac{1}{2}]\times[2,\infty]^2$ with $\frac{d}{p_b}+\frac{2}{q_b}<1+\alpha_b$. Suppose that

$$\kappa^b_1:=\|b\|_{\mathbb{L}^{q_b}_T\mathbf{B}^{\alpha_b}_{p_b,q_b}}<\infty\quad\text{and}\quad\kappa^b_2:=\|{\rm div} b\|_{\mathbb{L}^{q_b}_T\mathbf{B}^{-2-\alpha_b}_{p_b,q_b/(q_b-1)}}<\infty.$$

Remark 3.1. *Since $\frac{d}{p_b}+\frac{2}{q_b}<1+\alpha_b$ and $\alpha_b\in(-1,-\frac{1}{2}]$, it is easy to see that for any $q\geqslant\frac{q_b}{2}$,*

$$2+\alpha_b+\tfrac{d}{p_b}-\tfrac{2}{q_b}<2-\tfrac{4}{q_b}\leqslant2-\tfrac{2}{q}.\tag{3.2}$$

Remark 3.2. *Suppose that $b\in\mathbb{L}^{q_b}_T\mathbf{B}^{-1/2}_{p_b,1}$ with $\frac{d}{p_b}+\frac{2}{q_b}<\frac{1}{2}$. Then ($\mathbf{H}^{\rm sub}$) holds for $\alpha_b=-\frac{1}{2}$. Indeed, by (2.6) we have*

$$\kappa^b_1=\|b\|_{\mathbb{L}^{q_b}_T\mathbf{B}^{-1/2}_{p_b,q_b}}\lesssim\|b\|_{\mathbb{L}^{q_b}_T\mathbf{B}^{-1/2}_{p_b,1}},\quad\kappa^b_2=\|{\rm div} b\|_{\mathbb{L}^{q_b}_T\mathbf{B}^{-3/2}_{p_b,q_b/(q_b-1)}}\lesssim\|b\|_{\mathbb{L}^{q_b}_T\mathbf{B}^{-1/2}_{p_b,1}}.$$

If b is time-independent, then it was covered by the condition $b\in\mathbf{H}^{-1}_{p_b}$ with $p_b>2d$ in [48].

For convenience of notations, we introduce the following parameter set for later use:

$$\Theta:=(T,d,\alpha_b,p_b,q_b,\kappa^b_1,\kappa^b_2).\tag{3.3}$$

Now we show the following result.

Theorem 3.3. *Assume that* $(\mathbf{H}^{\rm sub})$ *holds. Let* $(\alpha,p,q)\in[-1,0]\times[1,\infty]^2$ *satisfy*

$$\tfrac{q_b}{2}\leqslant q\leqslant q_b,\quad p\leqslant p_b,\quad \alpha-\tfrac{d}{p}-\tfrac{2}{q}\geqslant \alpha_b-\tfrac{d}{p_b}-\tfrac{2}{q_b}.$$

For any $f\in\mathbb{L}^q_T\mathbf{B}^\alpha_{p,q}$, *there is a* $\lambda_0=\lambda_0(\Theta)>0$ *such that for any* $\lambda\geqslant\lambda_0$, *there is a unique* u^λ *with the regularity that for any* $\theta\in(0,2+\alpha_b+\frac{d}{p_b}-\frac{2}{q_b}]$,

$$\|u^\lambda\|_{\mathbb{L}^{q_b}_T\mathbf{B}^{2+\alpha_b}_{p_b,q_b}}+(1+\lambda)^{\frac{\theta}{2}}\|u^\lambda\|_{\mathbb{L}^\infty_T\mathbf{B}^{2+\alpha_b-2/q_b-\theta}_{p_b,1}}\leqslant C\|f\|_{\mathbb{L}^q_T\mathbf{B}^\alpha_{p,q}},\tag{3.4}$$

where $C=C(\Theta,\alpha,p,q)>0$, *which solves the following integral equation:*

$$u^\lambda_t=\mathscr{I}^\lambda_t(b\odot\nabla u^\lambda-\mathrm{div}b\preceq u^\lambda+f),$$

where $\mathscr{I}^\lambda$ *is defined by* (2.18).

Proof. We only show the a priori estimate (3.4). The existence follows by standard Picard's iteration. For simplicity of notations, we drop the superscript λ over u^λ.

By Schauder's estimate (2.22) and (2.16), we have for $\theta\in(0,2-\frac{2}{q_b})$,

$$\begin{aligned}&\|\mathscr{I}^\lambda(b\odot\nabla u)\|_{\mathbb{L}^{q_b}_T\mathbf{B}^{2+\alpha_b}_{p_b,q_b}}+(1+\lambda)^{\frac{\theta}{2}}\|\mathscr{I}^\lambda(b\odot\nabla u)\|_{\mathbb{L}^\infty_T\mathbf{B}^{2+\alpha_b-2/q_b-\theta}_{p_b,1}}\\&\qquad\lesssim_C\|b\odot\nabla u\|_{\mathbb{L}^{q_b}_T\mathbf{B}^{\alpha_b}_{p_b,q_b}}\lesssim_C\kappa^b_1\|u\|_{\mathbb{L}^\infty_T\mathbf{B}^{1+d/p_b}_{p_b,1}},\end{aligned}\tag{3.5}$$

where $C=C(\Theta,\theta)>0$. On the other hand, by (2.21), we have for any $\theta'\in(0,2-\frac{2}{q_b})$,

$$\|\mathscr{I}^\lambda(\mathrm{div}b\preceq u)\|_{\mathbb{L}^{q_b}_T\mathbf{B}^{2-d/p_b-2/q_b-\theta'}_{p_b,1}}\lesssim_C(1+\lambda)^{-\frac{\theta'}{2}}\|\mathrm{div}b\preceq u\|_{\mathbb{L}^{q_b/2}_T\mathbf{B}^0_{p_b/2,\infty}},\tag{3.6}$$

and for any $\theta''\in(0,2-\frac{4}{q_b})$,

$$\|\mathscr{I}^\lambda(\mathrm{div}b\preceq u)\|_{\mathbb{L}^\infty_T\mathbf{B}^{2-d/p_b-2/q_b-\theta''}_{p_b,1}}\lesssim_C(1+\lambda)^{-\frac{\theta''}{2}}\|\mathrm{div}b\preceq u\|_{\mathbb{L}^{q_b/2}_T\mathbf{B}^0_{p_b/2,\infty}}.\tag{3.7}$$

Since $\alpha_b\in(-1,-\frac12]$ and $\frac{d}{p_b}+\frac{2}{q_b}<1+\alpha_b$, we have

$$0<\vartheta_0:=-\alpha_b-\tfrac{d}{p_b}-\tfrac{2}{q_b}<2-\tfrac{2}{q_b}.$$

So, for fixed $\theta\in(0,2+\alpha_b+\frac{d}{p_b}-\frac{2}{q_b}]$ and $\vartheta\in(0,\vartheta_0)$, it holds that

$$\vartheta+\theta<2-\tfrac{4}{q_b}.$$

If we take $\theta'=\vartheta$ in (3.6) and $\theta''=\vartheta+\theta$ in (3.7), then by (2.17),

$$\begin{aligned}&\|\mathscr{I}^\lambda(\mathrm{div}b\preceq u)\|_{\mathbb{L}^{q_b}_T\mathbf{B}^{2+\alpha_b}_{p_b,1}}+(1+\lambda)^{\frac{\theta}{2}}\|\mathscr{I}^\lambda(\mathrm{div}b\preceq u)\|_{\mathbb{L}^\infty_T\mathbf{B}^{2+\alpha_b-2/q_b-\theta}_{p_b,1}}\\&\qquad\lesssim_C(1+\lambda)^{-\frac{\vartheta}{2}}\|\mathrm{div}b\preceq u\|_{\mathbb{L}^{q_b/2}_T\mathbf{B}^0_{p_b/2,\infty}}\lesssim_C(1+\lambda)^{-\frac{\vartheta}{2}}\kappa^b_2\|u\|_{\mathbb{L}^{q_b}_T\mathbf{B}^{2+\alpha_b}_{p_b,q_b}},\end{aligned}\tag{3.8}$$

where $C=C(\Theta,\theta,\vartheta)>0$. Moreover, by $\alpha-\frac{d}{p}-\frac{2}{q}\geqslant\alpha_b-\frac{d}{p_b}-\frac{2}{q_b}$ and (2.22), for $\theta\in(0,2-\frac{2}{q})$,

$$\|\mathscr{I}^\lambda(f)\|_{\mathbb{L}^{q_b}_T\mathbf{B}^{2+\alpha_b}_{p_b,q_b}}+(1+\lambda)^{\frac{\theta}{2}}\|\mathscr{I}^\lambda(f)\|_{\mathbb{L}^\infty_T\mathbf{B}^{2+\alpha_b-2/q_b-\theta}_{p_b,1}}\lesssim_C\|f\|_{\mathbb{L}^q_T\mathbf{B}^\alpha_{p,q}}.\tag{3.9}$$

Combining (3.5), (3.8), (3.9) and by (3.2), we obtain that for any $\theta\in(0,2+\alpha_b+\frac{d}{p_b}-\frac{2}{q_b}]$ and $\vartheta\in(0,\vartheta_0)$,

$$\begin{aligned}&\|u\|_{\mathbb{L}^{q_b}_T\mathbf{B}^{2+\alpha_b}_{p_b,q_b}}+(1+\lambda)^{\frac{\theta}{2}}\|u\|_{\mathbb{L}^\infty_T\mathbf{B}^{2+\alpha_b-2/q_b-\theta}_{p_b,1}}\\&\qquad\leqslant C_1\kappa^b_1\|u\|_{\mathbb{L}^\infty_T\mathbf{B}^{1+d/p_b}_{p_b,1}}+C_2(1+\lambda)^{-\frac{\vartheta}{2}}\kappa^b_2\|u\|_{\mathbb{L}^{q_b}_T\mathbf{B}^{2+\alpha_b}_{p_b,q_b}}+C_3\|f\|_{\mathbb{L}^q_T\mathbf{B}^\alpha_{p,q}},\end{aligned}\tag{3.10}$$

where $C_1=C_1(\Theta,\theta)>0$, $C_2=C_2(\Theta,\theta,\vartheta)>0$ and $C_3=C_3(\Theta,\theta,\alpha,p,q)>0$.

Since $\frac{2}{q_b}+\frac{d}{p_b}<1+\alpha_b$, one can choose $\theta_0=\theta_0(d,\alpha_b,p_b,q_b)$ being close to zero so that

$$\|u\|_{\mathbb{L}^\infty_T\mathbf{B}^{1+d/p_b}_{p_b,1}}\lesssim_C\|u\|_{\mathbb{L}^\infty_T\mathbf{B}^{2+\alpha_b-2/q_b-\theta_0}_{p_b,1}}.$$

Thus by (3.10) with $\theta=\vartheta=\theta_0$, one can take $\lambda_0=\lambda_0(\Theta)>0$ large enough so that for any $\lambda\geqslant\lambda_0$,

$$\|u\|_{\mathbb{L}^{q_b}_T\mathbf{B}^{2+\alpha_b}_{p_b,q_b}}+(1+\lambda)^{\frac{\theta_0}{2}}\|u\|_{\mathbb{L}^\infty_T\mathbf{B}^{2+\alpha_b-2/q_b-\theta_0}_{p_b,1}}\lesssim_C\|f\|_{\mathbb{L}^q_T\mathbf{B}^\alpha_{p,q}}.$$

The estimate (3.4) is then obtained by substituting this into (3.10). □

Remark 3.4. *By (2.19), it is easy to see that the unique solution in the above theorem also satisfies that for any $\varphi\in C^\infty_c(\mathbb{R}^d)$,*

$$\partial_t\langle u^\lambda,\varphi\rangle=\langle u^\lambda,\Delta\varphi\rangle-\lambda\langle u^\lambda,\varphi\rangle+\langle b\odot\nabla u^\lambda-\mathrm{div}b\preceq u^\lambda,\varphi\rangle+\langle f,\varphi\rangle.$$

In other words, u^λ is a weak solution of PDE (3.1).

3.2. **Supercritical case: $\frac{d}{p}+\frac{2}{q}<2+\alpha$.** We introduce the following supercritical index set

$$\mathscr{I}_d:=\Big\{(\alpha,p,q)\in[-1,0]\times[2,\infty]^2,\ \tfrac{d}{p}+\tfrac{2}{q}<2+\alpha\Big\},\tag{3.11}$$

and also the following energy space:

$$\mathscr{V}_T:=\left\{u:\|u\|_{\mathscr{V}_T}:=\sup_{t\in[0,T]}\|u(t)\|_2+\|\nabla u\|_{\mathbb{L}^2_T}<\infty\right\}.$$

We have the following simple lemma.

Lemma 3.5. *For any $(\alpha,p,q)\in\mathscr{I}_d$ and $T>0$, there is a constant $C=C(\alpha,p,q,d,T)>0$ such that for any $f\in\mathbb{L}^q_T\mathbf{H}^\alpha_p$ and $u,g\in\mathscr{V}_T$,*

$$\|\langle fu,g\rangle\|_{L^1_T}\lesssim_C\|f\|_{\mathbb{L}^q_T\mathbf{H}^\alpha_p}\|u\|_{\mathscr{V}_T}\|g\|_{\mathscr{V}_T}.$$

Proof. Let $(\alpha,p,q)\in\mathscr{I}_d$. Consider the case $p\in[2,\infty)$. Let $r\in(1,2]$ be defined by $\frac{1}{p}+\frac{1}{r}=1$. Let us choose $r_0\in(1,r]$ so that $-\alpha-\frac{d}{r}\leqslant1-\frac{d}{r_0}$. By Sobolev's embedding (2.9), we have

$$|\langle fu,g\rangle|\lesssim\|f\|_{\mathbf{H}^\alpha_p}\|ug\|_{\mathbf{H}^{-\alpha}_r}\lesssim\|f\|_{\mathbf{H}^\alpha_p}\|ug\|_{\mathbf{H}^1_{r_0}}.$$

Let $\frac{1}{r_0}=\frac{1}{r_1}+\frac{1}{2}$. By (2.5) and Hölder's inequality, we have

$$\|ug\|_{\mathbf{H}^1_{r_0}}\lesssim\|u\|_{r_1}\|g\|_2+\|u\|_{r_1}\|\nabla g\|_2+\|g\|_{r_1}\|\nabla u\|_2.$$

Hence, for $\frac{1}{q}+\frac{1}{q_1}+\frac{1}{2}=1$,

$$\|\langle fu,g\rangle\|_{L^1_T}\lesssim\|f\|_{\mathbb{L}^q_T\mathbf{H}^\alpha_p}\Big(\|u\|_{\mathbb{L}^{q_1}_TL^{r_1}}\|g\|_{\mathbb{L}^2_T\mathbf{H}^1_2}+\|u\|_{\mathbb{L}^2_T\mathbf{H}^1_2}\|g\|_{\mathbb{L}^{q_1}_TL^{r_1}}\Big).$$

Since $p,q\in[2,\infty)$ satisfies $\frac{d}{p}+\frac{2}{q}<2+\alpha$, one has $r_1,q_1\in[2,\infty)$ and $\frac{d}{r_1}+\frac{2}{q_1}>\frac{d}{2}$. By [49, Lemma 2.1], we conclude the desired estimate.

Next for $p=\infty$, by definition we have

$$|\langle fu,g\rangle|\leqslant\|f\|_{\mathbf{H}^\alpha_\infty}\|ug\|_{\mathbf{H}^{-\alpha}_1}\lesssim\|f\|_{\mathbf{H}^\alpha_\infty}\|ug\|_{\mathbf{H}^1_1},$$

and by Hölder's inequality,

$$\|ug\|_{\mathbf{H}^1_1}\lesssim\|u\|_2\|g\|_2+\|u\|_2\|\nabla g\|_2+\|g\|_2\|\nabla u\|_2.$$

Hence,

$$\|\langle fu,g\rangle\|_{L^1_T}\lesssim\|f\|_{\mathbb{L}^2_T\mathbf{H}^\alpha_\infty}\Big(\|u\|_{\mathbb{L}^\infty_TL^2}\|g\|_{\mathbb{L}^2_T\mathbf{H}^1_2}+\|u\|_{\mathbb{L}^2_T\mathbf{H}^1_2}\|g\|_{\mathbb{L}^\infty_TL^2}\Big).$$

The proof is complete. □

Next we make the following assumption on b.

($\mathbf{H}^{\rm sup}$) Let $(\alpha_b,p_b,q_b)\in\mathscr{I}_d$. Suppose that

$$\kappa^b_1:=\|b\|_{\mathbb{L}^{q_b}_T\mathbf{H}^{\alpha_b}_{p_b}}<\infty\quad\text{and}\quad\kappa^b_2:=\|\mathrm{div}b\|_{\mathbb{L}^2_TL^\infty}<\infty.$$

Recall the parameter set (3.3). Now we can show the following existence and uniqueness result.

Theorem 3.6. *Under* $(\mathbf{H}^{\rm sup})$*, for any* $(\alpha,p,q)\in\mathscr{I}_d$ *and* $f\in\mathbb{L}^q_T\mathbf{H}^\alpha_p$*, there is a weak solution* u *to PDE* (3.1) *in the sense that for any smooth function* $\varphi\in C^\infty_c(\mathbb{R}^d)$,

$$\partial_t\langle u,\varphi\rangle=\langle u,\Delta\varphi-b\cdot\nabla\varphi+{\rm div}b\,\varphi\rangle+\langle f,\varphi\rangle, \tag{3.12}$$

and for some $C=C(\Theta,\alpha,p,q)>0$,

$$\|u\|_{\mathbb{L}^\infty_T}+\|u\|_{\mathscr{V}_T}\lesssim_C\|f\|_{\mathbb{L}^q_T\mathbf{H}^\alpha_p}. \tag{3.13}$$

If in addition that $b\in\mathbb{L}^\infty_T\mathscr{B}+\mathbb{L}^2_TL^2$*, then the uniqueness holds in the class* (3.13).

Proof. **(Existence)** Let $b_n:=b*\phi_n$ and $f_n:=f*\phi_n$ be the mollifying approximation. Clearly,

$$\kappa^{b_n}_1\leqslant\kappa^b_1,\quad\kappa^{b_n}_2\leqslant\kappa^b_2.$$

Since $b_n,f_n\in\mathbb{L}^\infty_TC^\infty_b$, it is well-known that there is a classical solution $u_n\in\mathbb{L}^\infty_TC^\infty_b$ to the following PDE

$$\partial_tu_n=\Delta u_n+b_n\cdot\nabla u_n+f_n=\Delta u_n+{\rm div}(b_nu_n)-({\rm div}b_n)u_n+f_n,\quad u_n(0)=0.$$

By [49, Theorem 2.2], there is a constant $C=C(\Theta,\alpha,p,q)>0$ such that for all $n\in\mathbb{N}$,

$$\|u_n\|_{\mathbb{L}^\infty_T}+\|u_n\|_{\mathscr{V}_T}\lesssim_C\|f_n\|_{\mathbb{L}^q_T\mathbf{H}^\alpha_p}\lesssim_C\|f\|_{\mathbb{L}^q_T\mathbf{H}^\alpha_p}. \tag{3.14}$$

By (2.9) and (2.10), it is easy to see that

$$\begin{aligned}\|\partial_tu_n\|_{\mathbb{L}^2_T\mathbf{H}^{-3}_2}&\leqslant\|\Delta u_n+{\rm div}(b_nu_n)-({\rm div}b_n)u_n+f_n\|_{\mathbb{L}^2_T\mathbf{H}^{-3}_2}\\&\lesssim\|u_n\|_{\mathbb{L}^2_T\mathbf{H}^{-1}_2}+\|b_nu_n\|_{\mathbb{L}^2_T\mathbf{H}^{-2}_2}+\|({\rm div}b_n)u_n\|_{\mathbb{L}^2_T\mathbf{H}^{-3}_2}+\|f_n\|_{\mathbb{L}^2_T\mathbf{H}^{-3}_2}\leqslant C,\end{aligned}$$

where C is independent of n. This implies that

$$\|u_n(t)-u_n(s)\|_{\mathbf{H}^{-3}_2}\leqslant\|\partial_tu_n\|_{\mathbb{L}^2_T\mathbf{H}^{-3}_2}|\sqrt{|t-s|}\leqslant C\sqrt{|t-s|}.$$

By Aubin-Lions' lemma, there is a function $u\in\mathbb{L}^\infty_T\cap\mathscr{V}_T$ and subsequence $\{n_k\}$ such that for each t and $\varphi\in C^\infty_c(\mathbb{R}^d)$,

$$\lim_{k\to\infty}\langle u_{n_k}(t),\varphi\rangle=\langle u(t),\varphi\rangle,\quad u_{n_k}\stackrel{k\to\infty}{\to}u\text{ weakly in }\mathbb{L}^2_T\mathbf{H}^1_2, \tag{3.15}$$

and for each $R>0$,

$$\lim_{k\to\infty}\int_0^T\|u_{n_k}(s)-u(s)\|^2_{L^2(B_R)}{\rm d}s=0.$$

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

$$\|u(t)\|_\infty=\sup_{\varphi\in C^\infty_c,\|\varphi\|_1\leqslant1}\langle u(t),\varphi\rangle=\sup_{\varphi\in C^\infty_c,\|\varphi\|_1\leqslant1}\lim_{k\to\infty}\langle u_{n_k}(t),\varphi\rangle\leqslant\sup_n\|u_n(t)\|_\infty\leqslant C\|f\|_{\mathbb{L}^q_T\mathbf{H}^\alpha_p}.$$

Furthermore, for any $\varphi\in L^1(\mathbb{R}^d)$, we also have

$$\lim_{k\to\infty}\langle u_{n_k}(t),\varphi\rangle=\langle u(t),\varphi\rangle. \tag{3.16}$$

By taking weak limits, it is easy to see that u is a weak solution of PDE (3.1). Indeed, since $\frac{d}{p_b}+\frac{2}{q_b}<2+\alpha$, without loss of generality we may assume $q_b,p_b<\infty$. Otherwise, we may choose $q_b',p_b'<\infty$ so that $\frac{d}{p_b'}+\frac{2}{q_b'}<2+\alpha$. Noting that for any $\varphi\in C^\infty_c(\mathbb{R}^d)$,

$$\langle b_{n_k}u_{n_k},\nabla\varphi\rangle-\langle bu,\nabla\varphi\rangle=\langle(b_{n_k}-b)u_{n_k},\nabla\varphi\rangle+\langle u_{n_k}-u,b\cdot\nabla\varphi\rangle,$$

by Lemma 3.5 and $p_b,q_b\in[2,\infty)$, we have

$$\|\langle(b_{n_k}-b)u_{n_k},\nabla\varphi\rangle\|_{L^1_T}\lesssim_C\|(b_{n_k}-b)\chi\|_{\mathbb{L}^{q_b}_T\mathbf{H}^{\alpha_b}_{p_b}}\|u_{n_k}\|_{\mathscr{V}_T}\|\nabla\varphi\|_{\mathscr{V}_T}\stackrel{k\to\infty}{\to}0,$$

where $\chi\in C^\infty_c$ and $\chi=1$ on the support of φ, and due to $b\cdot\nabla\varphi\in\mathbb{L}^2_T\mathbf{H}^{-1}_2$, by (3.15),

$$\|\langle u_{n_k}-u,b\cdot\nabla\varphi\rangle\|_{L^1_T}\stackrel{k\to\infty}{\to}0.$$

Similarly, we have

$$\|\langle \mathrm{div} b_{n_k} u_{n_k} - \mathrm{div} b u, \varphi\rangle\|_{L^1_T} \overset{k\to\infty}{\to} 0.$$

(Uniqueness) Let u be any weak solution of PDE (3.1) with $f=0$ and satisfy

$$\|u\|_{\mathbb{L}^\infty_T} + \|u\|_{\mathscr{V}_T} < \infty.$$

Define $u_n := (u * \phi_n)\chi_n$, where $\chi_n = n^{-d}\chi(x/n)$ and $\chi \in C_c^\infty(\mathbb{R}^d)$ is a cuttoff function with $\chi(x) = 1$ on B_1. By the chain rule and the definition we have

$$\partial_t \langle u, u_n\rangle/2 = \langle \partial_t u, u_n\rangle = \langle u, \Delta u_n\rangle - \langle u, b\cdot\nabla u_n\rangle + \langle \mathrm{div} b u, u_n\rangle.$$

In particular,

$$\frac{\langle u(t), u_n(t)\rangle}{2} = -\int_0^t \langle \nabla u, \nabla u_n\rangle \mathrm{d}s - \int_0^t \Big[\langle u, b\cdot\nabla u_n\rangle - \langle \mathrm{div} b u, u_n\rangle\Big]\mathrm{d}s.$$

Suppose that $b = b_1 + b_2$, where $b_1 \in \mathbb{L}^\infty_T\mathscr{B}$ and $b_2 \in \mathbb{L}^2_T L^2$. Noting that

$$|\langle u, b_1\cdot\nabla u_n\rangle| \leqslant \|u\|_{\mathbf{H}^1_2}\|b_1\cdot\nabla u_n\|_{\mathbf{H}^{-1}_2} \leqslant \|u\|_{\mathbf{H}^1_2}\|b_1\|_{\mathscr{B}}\|u_n\|_{\mathbf{H}^1_2} \leqslant C\|b_1\|_{\mathscr{B}}\|u\|^2_{\mathbf{H}^1_2}, \tag{3.17}$$

and

$$|\langle u, b_2\cdot\nabla u_n\rangle| \leqslant \|u\|_\infty\|b_2\|_2\|\nabla u_n\|_2 \leqslant \|u\|_\infty\|b_2\|_2\|\nabla u\|_2,$$

by the dominated convergence theorem, we get

$$\frac{\|u(t)\|_2^2}{2} \leqslant -\int_0^t \|\nabla u\|_2^2\mathrm{d}s + \int_0^t \lim_{n\to\infty}\Big|\langle u, b\cdot\nabla u_n\rangle - \langle \mathrm{div} b u, u_n\rangle\Big|\mathrm{d}s.$$

By the definition of $\mathscr{B}$, it is easy to see that

$$\lim_{n\to\infty} b_1\cdot\nabla u_n \text{ exists in } \mathbf{H}^{-1}_2, \quad \lim_{n\to\infty} b_2\cdot\nabla u_n = b_2\cdot\nabla u \text{ in } L^1.$$

Hence, by $\mathrm{div} b \in L^\infty$ and commutating the order of limits (guaranteeing by the above two limits),

$$\begin{aligned}\lim_{n\to\infty}\langle u, b\cdot\nabla u_n\rangle = \lim_{n\to\infty}\lim_{m\to\infty}\langle u_m, b\cdot\nabla u_n\rangle &= \lim_{n\to\infty}\lim_{m\to\infty}\langle \mathrm{div} b\, u_m, u_n\rangle - \lim_{n\to\infty}\lim_{m\to\infty}\langle b\cdot\nabla u_m, u_n\rangle \\ &= \langle \mathrm{div} b\, u, u\rangle - \lim_{m\to\infty}\langle b\cdot\nabla u_m, u\rangle,\end{aligned}$$

which implies that

$$\lim_{n\to\infty}\langle u, b\cdot\nabla u_n\rangle = \langle \mathrm{div} b\, u, u\rangle/2.$$

Thus,

$$\frac{\|u(t)\|_2^2}{2} \leqslant -\int_0^t \|\nabla u\|_2^2\mathrm{d}s + \frac{1}{2}\int_0^t |\langle \mathrm{div} b\, u, u\rangle|\mathrm{d}s \leqslant \frac{1}{2}\int_0^t \|\mathrm{div} b\|_\infty\|u\|_2^2\mathrm{d}s.$$

Now by Gronwall's inequality, we obtain $u(t) = 0$. The uniqueness is proven. □

4. Subcritical case: Proof of Theorem 1.2

In this section, we establish the weak well-posedness of SDE (1.1) under the subcritical assumption $(\mathbf{H}^{\rm sub})$. To achieve this, we employ the classical Zvonkin transformation, ensuring both the existence and uniqueness of a weak solution.

Throughout this section we fix $T > 0$. To introduce a general notion of Krylov's estimate, we start by fixing a Banach space $\mathbb{B}$ that consists of distributions $f_t : [0,T] \to \mathscr{S}'(\mathbb{R}^d)$ so that $\mathbb{L}^1_T C_b \cap \mathbb{B}$ is dense in $\mathbb{B}$. Moreover, we make the following assumption:

- For any bounded measurable function $g : [0,T]\times\mathbb{R}^d \to \mathbb{R}$ with that the function $x \to g(t,x)$ is uniformly Lipschitz with respect to $t \in [0,T]$, the product gf is well-defined for all $f \in \mathbb{B}$, and there is a constant $C_g > 0$ such that for all $f \in \mathbb{B}$,

$$\|gf\|_{\mathbb{B}} \leqslant C_g\|f\|_{\mathbb{B}}. \tag{4.1}$$

It's worth noting that the space $\mathbb{B} := \mathbb{L}^q_T\mathbf{B}^\alpha_{p,q}$ with $\alpha \in [-1,0]$ and $p,q \in [1,\infty)$ satisfies the above assumption. Indeed, (4.1) follows by Lemma 2.5. Moreover, let $\phi_n(x) = n^d\phi(nx)$, where $\phi \in C^\infty_c(\mathbb{R}^d)$ is a smooth probability density function. For $f \in \mathbb{L}^q_T\mathbf{B}^\alpha_{p,q}$, if we define

$$f_n(t,x) = f(t,\cdot) * \phi_n(x),$$

then by definition, $f_n \in \mathbb{L}^q_T C^\infty_b$, and by the dominated convergence theorem,

$$\lim_{n\to\infty} \|f_n - f\|^q_{\mathbb{L}^q_T\mathbf{B}^\alpha_{p,q}} \leqslant \int_0^T \sum_{j\geqslant -1} 2^{\alpha jq} \lim_{n\to\infty} \|\phi_n * \mathcal{R}_j f(t) - \mathcal{R}_j f(t)\|^q_p \mathrm{d}t = 0.$$

It must be noticed that for $q = \infty$ or $p = \infty$, the above property does not hold since $C_b(\mathbb{R}^d)$ is not dense in L^∞. However, for $f \in \mathbb{L}^\infty_T\mathbf{B}^\alpha_{\infty,\infty}$, by (2.4), it holds that for any $\alpha' < \alpha$,

$$\|f_n - f\|_{\mathbb{L}^\infty_T\mathbf{B}^{\alpha'}_{\infty,\infty}} \leqslant C\|f\|_{\mathbb{L}^\infty_T\mathbf{B}^{\alpha'}_{\infty,\infty}} n^{\alpha-\alpha'}.$$

This is also the only reason that we require $p,q < \infty$ below.

Now we introduce the following notion used below.

Definition 4.1 (Krylov's estimate)**.** *Let $(X_t)_{t\in[0,T]}$ be a $\mathbb{R}^d$-valued stochastic process on a probability space $(\Omega,\mathscr{F},\mathbf{P})$. For $\theta \in (0,1)$, one says that $(X,\mathbb{B})$ satisfies Krylov's estimate with parameter θ if for any $f \in \mathbb{L}^1_T C_b \cap \mathbb{B}$, $m \in \mathbb{N}$, and any stopping times $\tau_0 \leqslant \tau_1 \leqslant T$ with $\tau_1 - \tau_0 \leqslant \delta$,*

$$\left\|\int_{\tau_0}^{\tau_1} f(s,X_s)\mathrm{d}s\right\|_{L^m(\Omega)} \leqslant C\delta^{\frac{1+\theta}{2}}\|f\|_{\mathbb{B}},$$

where the constant $C = C(m,\mathbb{B},d,T) > 0$.

Remark 4.2. *Here we require $\theta > 0$, that is crucial for Young's integral in (4.5) below.*

The following proposition is a direct conclusion of Krylov's estimate.

Proposition 4.3. *Suppose that $(X,\mathbb{B})$ satisfies Krylov's estimate with parameter $\theta \in (0,1)$. For any $f \in \mathbb{B}$, let $f_n \in \mathbb{L}^1_T C_b \cap \mathbb{B}$ be such that*

$$\lim_{n\to\infty} \|f_n - f\|_{\mathbb{B}} = 0.$$

Then the limit $\lim_{n\to\infty}\int_0^\cdot f_n(s,X_s)\mathrm{d}s$ exists in $L^p(\Omega, C([0,T]))$ for any $p \geqslant 1$ and the limit is denoted by $A^f_t := \lim_{n\to\infty}\int_0^t f_n(s,X_s)\mathrm{d}s$, which does not depend on the choice of approximation sequence $(f_n)_{n\in\mathbb{N}}$. Furthermore, for any $m \in \mathbb{N}$,

$$\sup_{t\in[0,T]} \|A^f_t\|_{L^m(\Omega)} + \sup_{s\neq t\in[0,T]} \frac{\|A^f_t - A^f_s\|_{L^m(\Omega)}}{|t-s|^{(1+\theta)/2}} \lesssim_C \|f\|_{\mathbb{B}}. \tag{4.2}$$

We now show a substitution formula for Young's integrals that will be used to show the uniqueness by Zvonkin's transformation.

Proposition 4.4. *Suppose that $(X,\mathbb{B})$ satisfies Krylov's estimate with parameter $\theta \in (0,1)$ and for any $m \in \mathbb{N}$,*

$$\|X_0\|_{L^m(\Omega)} + \|X_t - X_s\|_{L^m(\Omega)} \lesssim_C |t-s|^{1/2}, \quad \forall s,t \in [0,T]. \tag{4.3}$$

Let $g : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ be a bounded function satisfying

$$|g(t,x) - g(s,y)| \lesssim_C \sqrt{|t-s|} + |x-y|. \tag{4.4}$$

Then for any space-time function $f \in \mathbb{B}$, the integral $\int_0^t g(s,X_s)\mathrm{d}A^f_s$ is well-defined as Young's integral and

$$\int_0^t g(s,X_s)\mathrm{d}A^f_s = A^{f\cdot g}_t \quad a.s., \tag{4.5}$$

where A^f is defined in Proposition 4.3.

Proof. For a function $h : [0,T] \to \mathbb{R}$ and $\alpha \in (0,1)$, we write

$$[h]_\alpha := \sup_{s\neq t\in[0,T]} \frac{|h(t)-h(s)|}{|t-s|^\alpha}.$$

By (4.2), (4.3) and Kolmogorov's continuity criterion, for any $\alpha \in (\frac{1}{2}, \frac{1+\theta}{2})$, $\beta \in (0, \frac{1}{2})$ and $m \in \mathbb{N}$, there is a constant $C > 0$ such that for all $f \in \mathbb{B}$,

$$\big\|[A^f_\cdot]_\alpha\big\|_{L^m(\Omega)} \lesssim_C \|f\|_{\mathbb{B}}, \quad \big\|[X_\cdot]_\beta\big\|_{L^m(\Omega)} < \infty. \tag{4.6}$$

Below we fix $\alpha > \frac{1}{2}, \beta < \frac{1}{2}$ with $\alpha + \beta > 1$ and a sample point ω such that

$$[A^f_\cdot(\omega)]_\alpha < \infty, \quad [X_\cdot(\omega)]_\beta < \infty. \tag{4.7}$$

For any $0 \leqslant s < t \leqslant T$, we define

$$\Gamma_{s,t}(\omega) := g(s, X_s(\omega))(A^f_t(\omega) - A^f_s(\omega)).$$

For simplicity, we drop the dependence of ω. We note that for any $0 \leqslant s < u < t \leqslant T$,

$$\delta\Gamma_{s,u,t} := \Gamma_{s,t} - \Gamma_{s,u} - \Gamma_{u,t} = (g(s, X_s) - g(u, X_u))(A^f_t - A^f_u).$$

By (4.4) and (4.7), we have

$$|\delta\Gamma_{s,u,t}| \lesssim |u-s|^\beta |t-u|^\alpha [X_\cdot]_\beta [A^f_\cdot]_\alpha.$$

Since $\alpha + \beta > 1$, by the sewing lemma (cf. [17, Lemma 4.2]), we have

$$\Gamma_t := \int_0^t g(r, X_r)\mathrm{d}A^f_r := \lim_{|\pi|\to 0} \sum_{r,s\in\pi} \Gamma_{r,s} \quad \text{exists,}$$

where π is any partition of $[0,t]$, and

$$\left|\int_s^t (g(r, X_r) - g(s, X_s))\mathrm{d}A^f_r\right| = |\Gamma_t - \Gamma_s - \Gamma_{s,t}| \lesssim |t-s|^{\alpha+\beta}[X_\cdot]_\beta [A^f_\cdot]_\alpha.$$

In particular, taking $s = 0$, we get

$$\left|\int_0^t g(r, X_r)\mathrm{d}A^f_r\right| \leqslant \|g\|_\infty |A^f_t| + C[X_\cdot]_\beta [A^f_\cdot]_\alpha. \tag{4.8}$$

Now, let $\{f_n\}_{n=1}^\infty \subset \mathbb{L}^1_T C_b \cap \mathbb{B}$ be such that $\lim_{n\to\infty} \|f_n - f\|_{\mathbb{B}} = 0$. By definition, it is easy to see that

$$\int_0^t g(r, X_r)\mathrm{d}A^{f_n}_r = \int_0^t (gf_n)(r, X_r)\mathrm{d}r = A^{gf_n}_t \quad a.s. \tag{4.9}$$

By (4.8), (4.6) and (4.1), we have

$$\begin{aligned} &\mathbf{E}\left(\left|\int_0^t g(r, X_r)\mathrm{d}(A^{f_n}_r - A^f_r)\right|\right) + \mathbf{E}|A^{gf_n}_t - A^{gf}_t| \\ &= \mathbf{E}\left(\left|\int_0^t g(r, X_r)\mathrm{d}A^{f_n - f}_r\right|\right) + \mathbf{E}|A^{g(f_n - f)}_t| \\ &\lesssim \|f - f_n\|_{\mathbb{B}} + \|g(f - f_n)\|_{\mathbb{B}} \to 0 \quad \text{as } n \to \infty. \end{aligned}$$

Thus, there is a subsequence $\{n_k\}_{k=1}^\infty$ such that

$$\lim_{k\to\infty} \left|\int_0^t g(r, X_r)\mathrm{d}(A^{f_{n_k}}_r - A^f_r)\right| + |A^{gf_{n_k}}_t - A^{gf}_t| = 0 \quad a.s.$$

By taking $n = n_k$ and letting $k \to \infty$ in (4.9), we get (4.5). □

Below we always suppose that $(\mathbf{H}^{\rm sub})$ holds with $q_b, p_b < \infty$. Let

$$\theta_0 := 1 + \alpha_b - \tfrac{d}{p_b} - \tfrac{2}{q_b}, \quad b_n(t,x) := b(t,\cdot) * \phi_n(x).$$

By the property of convolution, we have

$$\begin{aligned} \kappa_1^{b_n} &= \|b_n\|_{\mathbb{L}^{q_b}_T \mathbf{B}^{\alpha_b}_{p_b,q_b}} \leqslant \|b\|_{\mathbb{L}^{q_b}_T \mathbf{B}^{\alpha_b}_{p_b,q_b}} = \kappa_1^b, \\ \kappa_2^{b_n} &= \|\mathrm{div} b_n\|_{\mathbb{L}^{q_b}_T \mathbf{B}^{-2-\alpha_b}_{p_b,q_b/(q_b-1)}} \leqslant \|\mathrm{div} b\|_{\mathbb{L}^{q_b}_T \mathbf{B}^{-2-\alpha_b}_{p_b,q_b/(q_b-1)}} = \kappa_2^b. \end{aligned} \tag{4.10}$$

We consider the following approximation SDE:

$$\mathrm{d}X^n_t = b_n(t, X^n_t)\mathrm{d}t + \sqrt{2}\mathrm{d}W_t, \quad X^n_0 \sim \mu \in \mathscr{P}(\mathbb{R}^d). \tag{4.11}$$

For simplicity of notations, we write

$$\mathscr{I}^{\alpha_b}_{p_b,q_b} := \Big\{(\alpha, p, q) \in [-1, 0] \times [1, \infty)^2 : \tfrac{q_b}{2} \leqslant q \leqslant q_b, \quad p \leqslant p_b, \quad \alpha - \tfrac{d}{p} - \tfrac{2}{q} \geqslant \alpha_b - \tfrac{d}{p_b} - \tfrac{2}{q_b}\Big\}.$$

Now we can show the following crucial Krylov estimate.

Lemma 4.5. *For any $(\alpha, p, q) \in \mathscr{I}^{\alpha_b}_{p_b,q_b}$, $(X^n, \mathbb{L}^q_T \mathbf{B}^\alpha_{p,q})$ satisfies Krylov's estimate with parameter θ_0 uniformly in n.*

Proof. Given $f \in C^\infty_c([0, T] \times \mathbb{R}^d)$, let $u_n \in \mathbb{L}^\infty_T C^\infty_b$ be the unique smooth solution to the following backward PDE

$$\partial_t u_n + \Delta u_n - \lambda u_n + b_n \cdot \nabla u_n + f = 0, \quad u_n(T) = 0.$$

For any $(\alpha, p, q) \in \mathscr{I}^{\alpha_b}_{p_b,q_b}$, by (3.4) and (4.10), there is a $\lambda_0 \geqslant 1$ so that for all $\theta \in (0, 2+\alpha_b+\frac{d}{p_b}-\frac{2}{q_b}]$ and $\lambda \geqslant \lambda_0$,

$$\sup_n \|u_n\|_{\mathbb{L}^\infty_T \mathbf{B}^{2+\alpha_b-2/q_b-\theta}_{p_b,1}} \lesssim (1+\lambda)^{-\frac{\theta}{2}} \|f\|_{\mathbb{L}^q_T \mathbf{B}^\alpha_{p,q}}. \tag{4.12}$$

By Itô's formula, we have

$$u_n(t, X^n_t) = u_n(0, x) + \int_0^t (\lambda u_n - f)(s, X^n_s)\mathrm{d}s + \sqrt{2}\int_0^t \nabla u_n(s, X^n_s)\mathrm{d}W_s.$$

In particular, for any stopping time $\tau_0 \leqslant \tau_1 \leqslant T$ with $\tau_1 - \tau_0 \leqslant \delta$,

$$\left|\int_{\tau_0}^{\tau_1} f(s, X^n_s)\mathrm{d}s\right| \leqslant 2\|u_n\|_\infty + \lambda\delta\|u_n\|_\infty + \sqrt{2}\left|\int_{\tau_0}^{\tau_1} \nabla u_n(s, X^n_s)\mathrm{d}W_s\right|.$$

By BDG's inequality, we have for any $m \geqslant 2$,

$$\begin{aligned} \left\|\int_{\tau_0}^{\tau_1} f(s, X^n_s)\mathrm{d}s\right\|_{L^m(\Omega)} &\lesssim (1+\lambda\delta)\|u_n\|_\infty + m\left\|\left(\int_{\tau_0}^{\tau_1} |\nabla u_n(s, X^n_s)|^2\mathrm{d}s\right)^{1/2}\right\|_{L^m(\Omega)} \\ &\lesssim (1+\lambda\delta)\|u_n\|_\infty + m\delta^{1/2}\|\nabla u_n\|_\infty. \end{aligned}$$

Here and below, the implicit constant is independent of m, n, δ, λ. Since

$$\theta_0 = 1 + \alpha_b - \tfrac{d}{p_b} - \tfrac{2}{q_b} \leqslant 2 + \alpha_b + \tfrac{d}{p_b} - \tfrac{2}{q_b},$$

by (4.12) and embedding (2.8), we have

$$\|u_n\|_\infty \lesssim \|u_n\|_{\mathbb{L}^\infty_T \mathbf{B}^{d/p_b}_{p_b,1}} = \|u_n\|_{\mathbb{L}^\infty_T \mathbf{B}^{2+\alpha_b-2/q_b-(1+\theta_0)}_{p_b,1}} \lesssim (1+\lambda)^{-\frac{1+\theta_0}{2}} \|f\|_{\mathbb{L}^q_T \mathbf{B}^\alpha_{p,q}},$$

and

$$\|\nabla u_n\|_\infty \lesssim \|u_n\|_{\mathbb{L}^\infty_T \mathbf{B}^{1+d/p_b}_{p_b,1}} = \|u_n\|_{\mathbb{L}^\infty_T \mathbf{B}^{2+\alpha_b-2/q_b-\theta_0}_{p_b,1}} \lesssim (1+\lambda)^{-\frac{\theta_0}{2}} \|f\|_{\mathbb{L}^q_T \mathbf{B}^\alpha_{p,q}}.$$

Thus

$$\left\|\int_{\tau_0}^{\tau_1} f(s, X^n_s)\mathrm{d}s\right\|_{L^m(\Omega)} \lesssim \left[(1+\lambda\delta)(1+\lambda)^{-\frac{1+\theta_0}{2}} + m\sqrt{\delta}(1+\lambda)^{-\frac{\theta_0}{2}}\right] \|f\|_{\mathbb{L}^q_T \mathbf{B}^\alpha_{p,q}}.$$

In particular, if we take $\lambda = \lambda_0 \vee \delta^{-1}$, then

$$\left\| \int_{\tau_0}^{\tau_1} f(s, X_s^n) \mathrm{d}s \right\|_{L^m(\Omega)} \lesssim m\delta^{\frac{1+\theta_0}{2}} \|f\|_{\mathbb{L}_T^q \mathbf{B}_{p,q}^\alpha}. \tag{4.13}$$

By a standard approximation, the above estimate still holds for any $f \in \mathbb{L}_T^1 C_b \cap \mathbb{L}_T^q \mathbf{B}_{p,q}^\alpha$. The proof is complete. □

Remark 4.6. *By Stirling's formula, estimate* (4.13) *also implies that for some $c_0 > 0$,*

$$\sup_n \mathbf{E} \exp\left\{ c_0 \left| \int_0^T f(s, X_s^n) \mathrm{d}s \right| / \|f\|_{\mathbb{L}_T^q \mathbf{B}_{p,q}^\alpha} \right\} < \infty.$$

By Theorem 3.3, there is a unique solution $u^\lambda : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ to the following $\mathbb{R}^d$-valued backward PDE system

$$\partial_t u^\lambda + \Delta u - \lambda u^\lambda + b \odot \nabla u^\lambda + \mathrm{div} b \preceq u^\lambda + b = 0, \quad u(T) = 0,$$

with the regularity that there is a $\lambda_0 > 0$ such that for all $\lambda \geqslant \lambda_0$ and $\theta \in (0, 2 + \alpha_b + \frac{d}{p_b} - \frac{2}{q_b}]$,

$$\|u^\lambda\|_{\mathbb{L}_T^\infty \mathbf{B}_{p_b,1}^{2+\alpha_b - 2/q_b - \theta}} \lesssim (1+\lambda)^{-\frac{\theta}{2}} \|b\|_{\mathbb{L}_T^{q_b} \mathbf{B}_{p_b,q_b}^{\alpha_b}}, \tag{4.14}$$

where the implicit constant does not depend on λ. In particular, by embedding (2.8), one can choose λ large enough so that

$$\sup_{t \in [0,T]} \|\nabla u^\lambda(t)\|_\infty \leqslant 1/2. \tag{4.15}$$

Let

$$\Phi(t,x) := x + u^\lambda(t,x).$$

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

$$x \mapsto \Phi(t,x) \text{ forms a } C^1\text{-diffeomorphism.}$$

In particular,

$$\|\nabla \Phi\|_\infty, \quad \|\nabla \Phi^{-1}\|_\infty \leqslant 4, \tag{4.16}$$

where Φ^{-1} stands for the inverse of $x \mapsto \Phi(t,x)$. Now we show the following crucial Zvonkin transformation.

Lemma 4.7 (Zvonkin's transformation)**.** *Let $(\mathfrak{F}, X, W)$ be a weak solution of SDE* (1.1) *in the sense of Definition 1.1. Suppose that for any $(\alpha, p, q) \in \mathscr{I}_{p_b,q_b}^{\alpha_b}$, $(X, \mathbb{L}_T^q \mathbf{B}_{p,q}^\alpha)$ satisfies Krylov's estimate with parameter θ_0 as above. Then $Y_t := \Phi(t, X_t)$ solves the following SDE:*

$$\mathrm{d}Y_t = \tilde{b}(t, Y_t)\mathrm{d}t + \tilde{\sigma}(t, Y_t)\mathrm{d}B_t, \tag{4.17}$$

where

$$\tilde{b}(t,y) := \lambda u(t, \Phi^{-1}(t,y)), \quad \tilde{\sigma}(t,y) := \sqrt{2}\nabla\Phi^{-1}(t, \Phi^{-1}(t,y)).$$

Moreover, $\tilde{b}$ and $\tilde{\sigma}$ are bounded measurable and for some $\gamma \in (0,1)$ and all $t \in [0,T]$, $y, y' \in \mathbb{R}^d$,

$$|\tilde{b}(t,y) - \tilde{b}(t,y')| \leqslant C|y - y'|, \quad |\tilde{\sigma}(t,y) - \tilde{\sigma}(t,y')| \leqslant C|y-y'|^\gamma, \tag{4.18}$$

and for some $c_0 > 0$,

$$|\tilde{\sigma}(t,y)\xi|^2 \geqslant c_0|\xi|^2, \quad \xi \in \mathbb{R}^d. \tag{4.19}$$

Proof. Let

$$u_n(t,x) := u(t,\cdot) * \phi_n(x),$$

where $\phi_n(x) = n^d \phi(nx)$ with $\phi \in C_c^\infty(\mathbb{R}^d)$ is a smooth probability density function. Then by Remark 3.4 with $\varphi = \phi_n(x - \cdot)$, we get

$$\partial_s u_n + \Delta u_n - \lambda u_n = -\phi_n * (b \odot \nabla u - \mathrm{div} b \preceq u + b) =: g_n, \tag{4.20}$$

and
$$\sup_{t\in[0,T]}\|\nabla u_n(t)\|_\infty\leqslant\sup_{t\in[0,T]}\|\nabla u(t)\|_\infty\leqslant 1/2.$$
Let
$$\Phi_n(t,x):=x+u_n(t,x).$$
Then
$$\|\nabla\Phi_n\|_\infty,\quad \|\nabla\Phi_n^{-1}\|_\infty\leqslant 4,$$
and by (4.20),
$$g_n\in\mathbb{L}^1_TL^\infty\Rightarrow\partial_t\Phi_n\in\mathbb{L}^1_TL^\infty.$$
Observe that by definition,
$$X_t=X_0+A^b_t+\sqrt{2}W_t,$$
where $A^b_t:=\lim_{n\to\infty}\int_0^t b_n(s,X_s)\mathrm{d}s$ by Proposition 4.3, satisfies
$$\sup_{t\in[0,T]}\|A^b_t\|_{L^m(\Omega)}+\sup_{s\neq t\in[0,T]}\frac{\|A^b_t-A^b_s\|_{L^m(\Omega)}}{|t-s|^{(1+\theta_0)/2}}\lesssim_C\|b\|_{\mathbb{L}^{q_b}_T\mathbf{B}^{\alpha_b}_{p_b,q_b}}.$$
Now by generalized Itô's formula (see [15]), (4.20) and (4.5), we have
$$\begin{aligned}Y^n_t:=\Phi_n(t,X_t)&=\Phi_n(0,x)+\int_0^t(\partial_su_n+\Delta u_n)(s,X_s)\mathrm{d}s\\&\quad+\int_0^t\mathrm{d}A^b_s\cdot\nabla\Phi_n(s,X_s)+\sqrt{2}\int_0^t\nabla\Phi_n(s,X_s)\mathrm{d}W_s\\&=\Phi_n(0,x)+\lambda\int_0^tu_n(s,X_s)\mathrm{d}s+A^{b\cdot\nabla\Phi_n+g_n}_t+\sqrt{2}\int_0^t\nabla\Phi_n(s,X_s)\mathrm{d}W_s.\end{aligned}\tag{4.21}$$
By (2.15), we make the following decomposition:
$$b\cdot\nabla\Phi_n+g_n=b+b\cdot\nabla u_n-\phi_n*(b\odot\nabla u-\mathrm{div}b\preceq u+b)=h^n_1+h^n_2,$$
where
$$h^n_1:=b-b*\phi_n+b\odot\nabla u_n-\phi_n*(b\odot\nabla u),\quad h^n_2:=\mathrm{div}b\preceq u_n-\phi_n*(\mathrm{div}b\preceq u).$$
In particular,
$$A^{b\cdot\nabla\Phi_n+g_n}_t=A^{h^n_1}_t+A^{h^n_2}_t.$$
By Krylov's estimate, we clearly have
$$\begin{aligned}\mathbf{E}|A^{h^n_1}_t|\lesssim\|h^n_1\|_{\mathbb{L}^{q_b}_T\mathbf{B}^{\alpha_b}_{p_b,q_b}}&\lesssim\|b-b*\phi_n\|_{\mathbb{L}^{q_b}_T\mathbf{B}^{\alpha_b}_{p_b,q_b}}+\|b\odot\nabla(u_n-u)\|_{\mathbb{L}^{q_b}_T\mathbf{B}^{\alpha_b}_{p_b,q_b}}\\&\quad+\|\phi_n*(b\odot\nabla u)-b\odot\nabla u\|_{\mathbb{L}^{q_b}_T\mathbf{B}^{\alpha_b}_{p_b,q_b}}.\end{aligned}$$
Note that by (2.16), (2.6), (2.4) and (4.14), there is an $\varepsilon>0$ so that for all $n\in\mathbb{N}$,
$$\|b\odot\nabla(u_n-u)\|_{\mathbb{L}^{q_b}_T\mathbf{B}^{\alpha_b}_{p_b,q_b}}\lesssim\kappa^b_1\|u_n-u\|_{\mathbb{L}^\infty_T\mathbf{B}^{1+d/p_b}_{p_b,1}}\lesssim n^{-\varepsilon}.$$
Thus, by the property of convolution, we have
$$\lim_{n\to\infty}\mathbf{E}|A^{h^n_1}_t|=0.$$
Moreover, since $\frac{d}{p_b}+\frac{2}{q_b}<1+\alpha_b$, one can choose ε with
$$0\leqslant-1-2\alpha_b<\varepsilon\leqslant-\big(\alpha_b+\tfrac{d}{p_b}+\tfrac{2}{q_b}\big)$$
so that $(-\varepsilon,p_b/2,q_b/2)\in\mathscr{I}^{\alpha_b}_{p_b,q_b}$. By Krylov's estimate again, we have
$$\begin{aligned}\mathbf{E}|A^{h^n_2}_t|\lesssim\|h^n_2\|_{\mathbb{L}^{q_b/2}_T\mathbf{B}^{-\varepsilon}_{p_b,q_b}}&\lesssim\|\phi_n*(\mathrm{div}b\preceq u)-\mathrm{div}b\preceq u\|_{\mathbb{L}^{q_b/2}_T\mathbf{B}^{-\varepsilon}_{p_b/2,q_b}}\\&\quad+\|\mathrm{div}b\preceq(u_n-u)\|_{\mathbb{L}^{q_b/2}_T\mathbf{B}^{-\varepsilon}_{p_b,q_b}}=:I_n+J_n.\end{aligned}$$

Note that by (2.17),

$$\lim_{n\to\infty} J_n \lesssim \lim_{n\to\infty} \|\mathrm{div} b \preceq (u_n - u)\|_{\mathbb{L}_T^{q_b/2}\mathbf{B}^0_{p_b,\infty}}$$
$$\lesssim \kappa_2^b \lim_{n\to\infty} \|u_n - u\|_{\mathbb{L}_T^{q_b}\mathbf{B}^{2+\alpha_b}_{p_b,q_b}} = 0.$$

Since $\|\mathrm{div} b \preceq u\|_{\mathbb{L}_T^{q_b/2}\mathbf{B}^{-\varepsilon}_{p_b/2,q_b}} \lesssim \|\mathrm{div} b \preceq u\|_{\mathbb{L}_T^{q_b/2}\mathbf{B}^{0}_{p_b/2,\infty}} < \infty$, by the property of convolution, we also have

$$\lim_{n\to\infty} I_n \to 0.$$

Combining the above limits, we obtain

$$\lim_{n\to\infty} \mathbf{E}|A_t^{b\cdot\nabla\Phi_n+g_n}| = 0.$$

By taking limits for both sides of (4.21), we conclude the proof of (4.17). As for (4.18) and (4.19), it follows by (4.15) and (4.16). □

Now we can show the following main result of this section.

Theorem 4.8. *Under* $(\mathbf{H}^{\mathrm{sub}})$ *with* $q_b, p_b < \infty$, *for any* $\mu \in \mathscr{P}(\mathbb{R}^d)$, *there is a weak solution* $(\mathfrak{F}, X, W)$ *to SDE* (1.1) *with initial distribution* μ, *and the uniqueness holds in the class that for any* $(\alpha, p, q) \in \mathscr{I}^{\alpha_b}_{p_b,q_b}$, $(X, \mathbb{L}_T^q\mathbf{B}^\alpha_{p,q})$ *satisfies Krylov's estimate with parameter* θ_0. *Moreover, for any* $t \in (0, T]$ *and* $X_0 = x$, X_t *has a density* $\rho_t(x, x')$, *which enjoys the following two sides Gaussian estimate: for all* $x, x' \in \mathbb{R}^d$,

$$C_0 t^{-d/2} \mathrm{e}^{-c_0|x-x'|^2/t} \leqslant \rho_t(x, x') \leqslant C_1 t^{-d/2} \mathrm{e}^{-c_1|x-x'|^2/t}, \tag{4.22}$$

where $C_0, C_1, c_0, c_1 > 0$ *only depend on the parameters* T, d, α_b, p_b, q_b *and* $\|b\|_{\mathbb{L}_T^{q_b}\mathbf{B}^{\alpha_b}_{p_b,q_b}}$.

Proof. **(Existence).** Consider the approximation SDE (4.11). By Lemma 4.5, the law of X^n in $\mathbb{C}_T$ is tight, and therefore, by Prohorov's theorem, it is relatively compact in $\mathscr{P}(\mathbb{C}_T)$. Then, by the weak convergence method, it is standard to derive the existence of a weak solution (see [47] and also the proof of Theorem 5.7 below).

(Uniqueness and heat kernel estimate). By Lemma 4.7, since the transformed SDE (4.17) has a bounded Lipschitz drift and bounded uniformly non-degenerate Hölder diffusion coefficients, it is well known that SDE (4.17) has a unique weak solution Y_t (see [38]), and for each $t \in (0, T]$ and starting point $Y_0 = y$, Y_t has a density $\tilde\rho_t(y, y')$, which enjoys two sides Gaussian estimate (for example, see [10]):

$$\tilde C_0 t^{-d/2} \mathrm{e}^{-\tilde c_0|y-y'|^2/t} \leqslant \tilde\rho_t(y, y') \leqslant \tilde C_1 t^{-d/2} \mathrm{e}^{-\tilde c_1|y-y'|^2/t}.$$

The desired estimate (4.22) follows by the change of variables (see [42]). □

Now we can give

Proof of Theorem 1.2. Since b is divergence free and by Lemma 2.4,

$$\mathbf{H}^{\alpha_b}_{p_b} \subset \cap_{\varepsilon>0, q_b\in[1,\infty]} \mathbf{B}^{\alpha_b-\varepsilon}_{p_b,q_b},$$

one sees that $(\mathbf{H}^{\mathrm{sub}})$ holds. If $q_b, p_b < \infty$, then the desired results follow by the above theorem. For case $p_b = \infty$, since $\frac{2}{q_b} < 1 + \alpha_b$, one can choose $\alpha_b' \in (-1, \alpha_b)$ so that $\frac{2}{q_b} < 1 + \alpha_b'$. In particular, $(\mathbf{H}^{\mathrm{sub}})$ still holds with $\kappa_1^b = \|b\|_{\mathbb{L}_T^{q_b}\mathbf{B}^{\alpha_b'}_{\infty,q_b}} \leqslant \|b\|_{\mathbb{L}_T^{q_b}\mathbf{B}^{\alpha_b}_{\infty,q_b}} < \infty$ and $\kappa_2^b = 0$. We emphasize that if b is not divergence-free, then $\kappa_2^b = \|\mathrm{div} b\|_{\mathbb{L}_T^{q_b}\mathbf{B}^{-2-\alpha_b'}_{\infty,q_b/(q_b-1)}}$ is possibly *infinite*. Thus by using (2.4), one can check that Lemma 4.7 still holds. The rest is the same as the proof of Theorem 4.8. □

5. Supercritical case: Proof of Theorem 1.4

Throughout this section we still fix $T > 0$ and always assume that supercritical assumption $(\mathbf{H}^{\rm sup})$ holds. In Subsection 5.1, we focus on establishing the existence of weak solutions. The crucial step in this regard is demonstrating the tightness of approximation solutions. Notably, in the supercritical case, due to the drift being a distribution, it's not possible to establish a Krylov estimate similar to that in Lemma 4.5 for any moments. Instead, we can only establish a Krylov estimate for the second-order moment, as demonstrated in Lemma 5.5, which is particularly useful for taking limits. However, this alone is not sufficient to establish tightness. To overcome this difficulty, we employ the stopping time technique and the strong Markov property of approximation SDEs. In Subsection 5.2, we discuss the uniqueness of generalized martingale problem and ultimately complete the proof of Theorem 1.4.

5.1. Existence of weak solutions.

We still consider the approximation SDE

$$X_t^n = X_0 + \int_0^t b_n(s, X_s^n)\mathrm{d}s + \sqrt{2}W_t, \tag{5.1}$$

where $b_n = b * \phi_n$. Recall the definition of $\mathscr{I}_d$ in (3.11) and the parameter set Θ in (3.3). We first show the following uniform Krylov estimate.

Lemma 5.1. *Under $(\mathbf{H}^{\rm sup})$, for any $(\alpha, p, q) \in \mathscr{I}_d$, there is a constant $C = C(\Theta, \alpha, p, q) > 0$ such that for all $t \in (0, T]$, stopping times $0 \leqslant \tau_0 \leqslant \tau_1 \leqslant t$ and $f \in C_c([0, t] \times \mathbb{R}^d)$,*

$$\sup_{n\in\mathbb{N}} \mathbf{E}\left(\int_{\tau_0}^{\tau_1} f(s, X_s^n)\mathrm{d}s\Big|\mathscr{F}_{\tau_0}\right) \lesssim_C \|f\|_{\mathbb{L}_t^q\mathbf{H}_p^\alpha}.$$

Proof. Without loss of generality, we assume $f \in C_c^\infty([0, T] \times \mathbb{R}^d)$. Fix $t \in [0, T]$. Let $u_n \in \mathbb{L}_T^\infty C_b^\infty(\mathbb{R}^d)$ be the smooth solution of the following backward PDE:

$$\partial_s u_n + \Delta u_n + b_n \cdot \nabla u_n = f, \quad u_n(t) = 0.$$

By (3.13), we have

$$\sup_n \left(\|u_n\|_{\mathbb{L}_t^\infty} + \|u_n\|_{\mathscr{V}_t}\right) \lesssim_C \|f\|_{\mathbb{L}_t^q\mathbf{H}_p^\alpha}. \tag{5.2}$$

Using Itô's formula for $(s, x) \to u_n(s, x)$, one sees that for any stopping times $0 \leqslant \tau_0 \leqslant \tau_1 \leqslant t$,

$$u_n(\tau_1, X_{\tau_1}^n) - u_n(\tau_0, X_{\tau_0}^n) = \int_{\tau_0}^{\tau_1} f(s, X_s^n)\mathrm{d}s + \sqrt{2}\int_{\tau_0}^{\tau_1} \nabla u_n(s, X_s^n)\mathrm{d}W_s.$$

Hence, by the optional stopping time theorem,

$$\mathbf{E}\left(\int_{\tau_0}^{\tau_1} f(s, X_s^n)\mathrm{d}s\Big|\mathscr{F}_{\tau_0}\right) = \mathbf{E}(u_n(\tau_1, X_{\tau_1}^n)|\mathscr{F}_{\tau_0}) - u_n(\tau_0, X_{\tau_0}^n). \tag{5.3}$$

By (5.2), we immediately obtain the result. □

Remark 5.2. *Note that if $\alpha = 0$, then we automatically have*

$$\mathbf{E}\left|\int_{\tau_0}^{\tau_1} f(s, X_s^n)\mathrm{d}s\right| \leqslant \mathbf{E}\left(\int_{\tau_0}^{\tau_1} |f(s, X_s^n)|\mathrm{d}s\right) \leqslant C\|f\|_{\mathbb{L}_t^q L^p}.$$

However, for $f \in \mathbb{L}_t^q\mathbf{H}_p^\alpha$ with $\alpha < 0$ and in the supercritical case, we can not show any absolute moment estimate since we do not have the maximum estimate of $\|\nabla u\|_\infty$ as done in the subcritical case. Thus we have to carefully treat the distribution-valued f.

We have the following tightness of the law $\mathbb{P}_n$ of approximation solution $X_\cdot^n$ in $\mathbb{C}_T$.

Lemma 5.3. *Under $(\mathbf{H}^{\rm sup})$, the family of probability measures $(\mathbb{P}_n)_{n\in\mathbb{N}}$ is tight.*

Proof. Let $\mathscr{T}_T$ be the set of all stopping times bounded by T. By Aldous' criterion of tightness, it suffices to prove that for some increasing function $\ell:(0,1)\to(0,\infty)$ with $\lim_{\delta\to0}\ell(\delta)=0$,

$$\sup_{\tau,\eta\in\mathscr{T}_T,\eta\leqslant\tau\leqslant\eta+\delta}\sup_n\mathbf{E}|X^n_\tau-X^n_\eta|\leqslant\ell(\delta),\quad\delta\in(0,1),$$

where $\ell(\delta)$ will be determined below. By the strong Markov property, we only need to show

$$\sup_{t\in[0,T]}\sup_{x_0\in\mathbb{R}^d}\sup_{\tau\leqslant\delta}\sup_n\mathbf{E}|X^{t,n}_\tau(x_0)-x_0|\leqslant\ell(\delta),$$

where $X^{t,n}_\cdot(x_0)$ stands for the solution of SDE (5.1) with starting point x_0 and drift

$$b^t_n(s,\cdot)=b_n(t+s,\cdot).$$

Define

$$h_\delta(x):=\sqrt{\ell(\delta)^2/9+|x-x_0|^2}.$$

By Itô's formula, we have

$$\mathbf{E}|X^{t,n}_\tau(x_0)-x_0|\leqslant\mathbf{E}h_\delta(X^{t,n}_\tau(x_0))=\ell(\delta)/3+\mathbf{E}\left(\int_0^\tau(\Delta+b_n(s)\cdot\nabla)h_\delta(X^{t,n}_s(x_0))\mathrm{d}s\right).$$

Note that for some C independent of δ and t,x_0,

$$|\nabla h_\delta(x)|=\frac{|x-x_0|}{\sqrt{\ell(\delta)^2/9+|x-x_0|^2}}\leqslant1,\quad|\nabla^2h_\delta|\leqslant2\ell(\delta)^{-1}.\tag{5.4}$$

Thus for $\tau\leqslant\delta$,

$$\mathbf{E}\left(\int_0^\tau\Delta h_\delta(X^{t,n}_s(x_0))\mathrm{d}s\right)\leqslant\delta\|\Delta h_\delta\|_\infty\leqslant2d\delta/\ell(\delta).$$

On the other hand, since $\frac{d}{p_b}+\frac{2}{q_b}<2+\alpha$, without loss of generality we assume $q_b<\infty$. Otherwise, we may choose $q_b'<\infty$ so that $\frac{d}{p_b}+\frac{2}{q_b'}<2+\alpha$. By Lemma 5.1 and (iv) of Lemma 2.4, we have

$$\mathbf{E}\left(\int_0^{\delta\wedge\tau}(b^t_n(s)\cdot\nabla h_\delta)(X^{t,n}_s(x_0))\mathrm{d}s\right)\leqslant C_1\|b^t_n\cdot\nabla h_\delta\|_{\mathbb{L}^{q_b}_\delta\mathbf{H}^{\alpha_b}_{p_b}}\leqslant C_2\|b^t_n\|_{\mathbb{L}^{q_b}_\delta\mathbf{H}^{\alpha_b}_{p_b}}\|\nabla h_\delta\|_{C^1_b},$$

where the constant $C_2>0$ only depends on Θ. By (5.4), we further have

$$\mathbf{E}\left(\int_0^{\delta\wedge\tau}(b^t_n(s)\cdot\nabla h_\delta)(X^{t,n}_s(x_0))\mathrm{d}s\right)\leqslant C_3\|b^t\|_{\mathbb{L}^{q_b}_\delta\mathbf{H}^{\alpha_b}_{p_b}}/\ell(\delta)\leqslant(C_3\|b^t\|_{\mathbb{L}^{q_b}_\delta\mathbf{H}^{\alpha_b}_{p_b}}\vee(6\delta))/\ell(\delta).$$

Finally if we take

$$\ell(\delta):=\sup_{t\in[0,T]}\sqrt{3[C_3\|b^t\|_{\mathbb{L}^{q_b}_\delta\mathbf{H}^{\alpha_b}_{p_b}}\vee(6d\delta)]},$$

then combining the above calculations, we obtain

$$\mathbf{E}|X^n_\tau(x_0)-x_0|\leqslant\ell(\delta)/3+2d\delta/\ell(\delta)+\ell(\delta)/3\leqslant\ell(\delta).$$

The proof is complete since $\lim_{\delta\to0}\sup_{t\in[0,T]}\int_t^{(t+\delta)\wedge T}|g(s)|\mathrm{d}s=0$ for $g\in L^1([0,T])$. □

By the above lemma and Prokhorov's theorem, $(\mathbb{P}_n)_{n\in\mathbb{N}}$ is relatively compact in $\mathscr{P}(\mathbb{C}_T)$. Next, we want to show that any accumulation point of the sequence $(\mathbb{P}_n)_{n\in\mathbb{N}}$ is indeed a weak solution of SDE (1.1). However, to take the limit, we need a different type of Krylov's estimate, which, in turn, requires that the initial distribution has an L^2-distribution.

For approximation SDE (5.1), it is well-known that for any $t\in(0,T]$, X^n_t admits a smooth density $\rho_n(t,x)$ that satisfies the Fokker-Planck equation

$$\partial_t\rho_n=\Delta\rho_n-\mathrm{div}(b_n\rho_n).\tag{5.5}$$

We need the following simple lemma.

Lemma 5.4. *Suppose $\rho_0 \in L^2$ and let $\chi \in C_c^\infty(\mathbb{R}^d)$ with $\chi = 1$ on B_1 and $\chi = 0$ on B_2^c. We have*
$$\|\rho_n\|_{\mathscr{V}_T} \leqslant (\kappa_2^b + 1)\mathrm{e}^{\kappa_2^b}\|\rho_0\|_2, \tag{5.6}$$
and
$$\lim_{R\to\infty}\sup_n \|\rho_n(1-\chi(\cdot/R))\|_{\mathscr{V}_T} = 0. \tag{5.7}$$

Proof. First of all, since $b_n \in \mathbb{L}_T^2 C_b^\infty$, by the maximum principle, we have
$$\rho_n \in \mathbb{L}_T^\infty L^\infty.$$
Let $\varphi \in C_c^\infty(\mathbb{R}^d)$. By multiplying both sides of (5.5) by φ, we get
$$\partial_t \rho_n\varphi = \Delta(\rho_n\varphi) - \mathrm{div}(b_n\rho_n\varphi) + g_n^\varphi, \tag{5.8}$$
where
$$g_n^\varphi := b_n\rho_n\cdot\nabla\varphi - 2\nabla\rho_n\cdot\nabla\varphi - \rho_n\Delta\varphi.$$
Multiplying both sides of (5.8) by $\rho_n\varphi$ and integrating on $\mathbb{R}^d$, we have
$$\frac{1}{2}\partial_t\|\rho_n\varphi\|_2^2 + \|\nabla(\rho_n\varphi)\|_2^2 = \langle b_n\rho_n\varphi_n, \nabla(\rho_n\varphi)\rangle + \langle g_n^\varphi, \rho_n\varphi\rangle.$$
Noting that
$$\langle b_n\rho_n\varphi_n, \nabla(\rho_n\varphi)\rangle = -\frac{1}{2}\langle \mathrm{div} b_n, (\rho_n\varphi)^2\rangle \leqslant \frac{1}{2}\|\mathrm{div} b_n\|_\infty\|\rho_n\varphi\|_2^2 \leqslant \frac{1}{2}\|\mathrm{div} b\|_\infty\|\rho_n\varphi\|_2^2$$
and
$$\langle g_n^\varphi, \rho_n\varphi\rangle = \langle b_n, \varphi\rho_n^2\nabla\varphi\rangle + \langle\rho_n^2, |\nabla\varphi|^2\rangle,$$
we have
$$\begin{aligned}\|\rho_n(t)\varphi\|_2^2 + 2\int_0^t\|\nabla(\rho_n\varphi)\|_2^2\mathrm{d}s &\leqslant \|\rho_0\varphi\|_2^2 + \int_0^t\|\mathrm{div} b\|_\infty\|\rho_n\varphi\|_2^2\mathrm{d}s\\ &\quad + \int_0^t\Big(\langle b_n, \varphi\rho_n^2\nabla\varphi\rangle + \langle\rho_n^2, |\nabla\varphi|^2\rangle\Big)\mathrm{d}s.\end{aligned}$$
By Gronwall's inequality, we derive
$$\|\rho_n\varphi\|_{\mathscr{V}_T} \leqslant (\kappa_2^b + 1)\mathrm{e}^{\kappa_2^b}\Big(\|\rho_0\varphi\|_2 + \|\langle b_n, \varphi\rho_n^2\nabla\varphi\rangle\|_{L_T^1} + \|\langle\rho_n^2, |\nabla\varphi|^2\rangle\|_{L_T^1}\Big). \tag{5.9}$$
Let $\chi_R(x) := \chi(x/R)$. Replacing φ in (5.9) by χ_R and letting $R\to\infty$, by the dominated convergence theorem, it is easy to derive
$$\|\rho_n\|_{\mathscr{V}_T} \leqslant (\kappa_2^b + 1)\mathrm{e}^{\kappa_2^b}\|\rho_0\|_2.$$
On the other hand, replacing φ in (5.9) by $\bar\chi_R := 1 - \chi_R$, we get
$$\|\rho_n\bar\chi_R\|_{\mathscr{V}_T} \leqslant (\kappa_2^b + 1)\mathrm{e}^{\kappa_2^b}\Big(\|\rho_0\bar\chi_R\|_2 + \|\langle b_n, \bar\chi_R\rho_n^2\nabla\bar\chi_R\rangle\|_{L_T^1} + \|\langle\rho_n^2, |\nabla\bar\chi_R|^2\rangle\|_{L_T^1}\Big).$$
By Lemma 3.5 and the definition of $\mathscr{V}_T$, we have
$$\begin{aligned}\|\langle b_n, \bar\chi_R\rho_n^2\nabla\bar\chi_R\rangle\|_{L_T^1} &\lesssim \|b_n\|_{\mathbb{L}_T^{q_b}\mathbf{H}_{q_b}^{\alpha_b}}\|\rho_n\|_{\mathscr{V}_T}\|\bar\chi_R\nabla\bar\chi_R\rho_n\|_{\mathscr{V}_T}\\ &\lesssim \|b\|_{\mathbb{L}_T^{q_b}\mathbf{H}_{q_b}^{\alpha_b}}\|\rho_n\|_{\mathscr{V}_T}^2/R \lesssim \|\rho_0\|_2^2/R\end{aligned}$$
and
$$\|\langle\rho_n^2, |\nabla\bar\chi_R|^2\rangle\|_{L_T^1} \lesssim \|\rho_n\|_{\mathbb{L}_T^2}^2/R \lesssim \|\rho_0\|_2^2/R.$$
Hence,
$$\sup_n\|\rho_n\bar\chi_R\|_{\mathscr{V}_T} \lesssim \|\rho_0\bar\chi_R\|_2 + \|\rho_0\|_2^2/R,$$
which converges to zero as $R\to\infty$. The proof is complete. □

We have the following Krylov estimate.

Lemma 5.5. *Suppose $\rho_0 \in L^2$. Under $(\mathbf{H}^{\rm sup})$, for any $(\alpha, p, q) \in \mathscr{I}_d$, there is a constant $C = C(\Theta, \alpha, p, q) > 0$ such that for all $t \in (0, T]$ and $f \in C_c([0,t] \times \mathbb{R}^d)$,*

$$\mathbf{E}\left|\int_0^t f(s, X^n_s) \mathrm{d}s\right|^2 \lesssim_C \|f\|^2_{\mathbb{L}^q_t \mathbf{H}^\alpha_p} \|\rho_0\|_2. \tag{5.10}$$

Proof. Without loss of generality, we assume $f \in C^\infty_c([0,T] \times \mathbb{R}^d)$. Fix $t \in [0,T]$. Let u solve the following backward PDE

$$\partial_s u_n + \Delta u_n + b_n \cdot \nabla u_n + f = 0, \quad u_n(t) = 0.$$

By (5.3) with $\tau_0 = s < t$ and $\tau_1 = t$, we have

$$\mathbf{E}\left(\int_s^t f(r, X^n_r) \mathrm{d}r \Big| \mathscr{F}_s\right) = u_n(s, X^n_s).$$

Therefore,

$$\begin{aligned}\mathbf{E}\left|\int_0^t f(r, X^n_r) \mathrm{d}r\right|^2 &= 2\mathbf{E}\left[\int_0^t f(s, X^n_s)\left(\int_s^t f(r, X^n_r)\mathrm{d}r\right)\mathrm{d}s\right] \\ &= 2\mathbf{E}\left[\int_0^t f(s, X^n_s)\mathbf{E}\left(\int_s^t f(r, X^n_r)\mathrm{d}r\Big|\mathscr{F}_s\right)\mathrm{d}s\right] \\ &= 2\mathbf{E}\left[\int_0^t f(s, X^n_s) u_n(s, X^n_s)\mathrm{d}s\right] \\ &= 2\int_0^t \langle f(s)u_n(s), \rho_n(s)\rangle \mathrm{d}s \leqslant 2\|\langle f u_n, \rho_n\rangle\|_{L^1_t}. \end{aligned} \tag{5.11}$$

By Lemma 3.5, we get

$$\mathbf{E}\left|\int_0^t f(s, X^n_s)\mathrm{d}s\right|^2 \lesssim \|f\|_{\mathbb{L}^q_t \mathbf{H}^\alpha_p} \|\rho_n\|_{\mathscr{V}_t} \|u_n\|_{\mathscr{V}_t}.$$

The desired estimate follows by (5.6) and (5.2). □

Remark 5.6. *One can not show the following stronger estimate*

$$\mathbf{E}\left(\sup_{t\in[0,T]}\left|\int_0^t f(s, X^n_s)\mathrm{d}s\right|^2\right) \lesssim_C \|f\|^2_{\mathbb{L}^q_T \mathbf{H}^\alpha_p} \|\rho_0\|_2$$

since we are using the estimate (5.6) of the density and f is a distribution.

Now we can show the following existence result of weak solutions.

Theorem 5.7. *Suppose that $\rho_0 \in L^2$. Under $(\mathbf{H}^{\rm sup})$, there is a weak solution to SDE (1.1) such that for all $t \in (0, T]$ and $f \in C_c([0,t] \times \mathbb{R}^d)$,*

$$\mathbf{E}\left|\int_0^t f(s, X_s)\mathrm{d}s\right|^2 \lesssim_C \|f\|^2_{\mathbb{L}^q_t \mathbf{H}^\alpha_p} \|\rho_0\|_2, \tag{5.12}$$

where $C = C(\Theta, \alpha, p, q) > 0$. Moreover, for each $t \in (0, T]$, X_t admits a density $\rho_t \in L^2$ with $\|\rho\|_{\mathscr{V}_T} \leqslant (\kappa^b_2 + 1)\mathrm{e}^{\kappa^b_2}\|\rho_0\|_2$ and in the distributional sense

$$\partial_t \rho = \Delta\rho - \mathrm{div}(b\rho). \tag{5.13}$$

Proof. By Lemma 5.3 and Prokhorov's criterion, there is a subsequence n_k and a probability measure $\mathbb{P} \in \mathscr{P}(\mathbb{C}_T)$ so that $\mathbb{P}_{n_k}$ weakly converges to $\mathbb{P}$ as $k \to \infty$. Without loss of generality, we assume that $\mathbb{P}_n$ weakly converges to $\mathbb{P}$ as $n \to \infty$. By Skorokhod's representation theorem, if necessary, changing the probability space, we may assume that there is a continuous process $(X_t)_{t\in[0,T]}$ such that

$$X^n_\cdot \to X_\cdot \quad \text{a.s. in } \mathbb{C}_T.$$

Estimate (5.12) follows by taking limits for (5.10). To show that X is a weak solution of SDE (1.1), it suffices to show that

$$\lim_{n\to\infty}\mathbf{E}\left|\int_0^t b_n(r,X^n_r)\mathrm{d}r-\int_0^t b_n(r,X_r)\mathrm{d}r\right|^2=0.$$

Noticing that for fixed $m\in\mathbb{N}$, by the dominated convergence theorem,

$$\lim_{n\to\infty}\mathbf{E}\left|\int_0^t b_m(r,X^n_r)\mathrm{d}r-\int_0^t b_m(r,X_r)\mathrm{d}r\right|^2=0,$$

once we have shown

$$\lim_{m,m'\to\infty}\sup_n\mathbf{E}\left|\int_0^t (b_m-b_{m'})(r,X^n_r)\mathrm{d}r\right|^2=0, \tag{5.14}$$

then it follows that

$$\lim_{n\to\infty}\int_0^t b_n(r,X_r)\mathrm{d}r=:A^b_t \text{ in } L^2.$$

Let $\chi\in C^\infty_c(\mathbb{R}^d)$ with $\chi=1$ on B_1 and $\chi=0$ on B^c_2. For $R>0$, define cutoff function

$$\chi_R(x):=\chi(x/R).$$

Since $\frac{d}{p_b}+\frac{2}{q_b}<2+\alpha$, one can choose finite $\bar q_b,\bar p_b\in[2,\infty)$ with $\bar q_b\leqslant q_b,\bar p_b\leqslant p_b$ and such that $\frac{d}{\bar p_b}+\frac{2}{\bar q_b}<2+\alpha$. Then for fixed $R>0$, by Lemma 5.5, we have

$$\lim_{m,m'\to\infty}\sup_n\mathbf{E}\left|\int_0^t ((b_m-b_{m'})\chi_R)(r,X^n_r)\mathrm{d}r\right|^2\lesssim\lim_{m,m'\to\infty}\|(b_m-b_{m'})\chi_R\|^2_{\mathbb{L}^{\bar q_b}_t\mathbf{H}^{\alpha_b}_{\bar p_b}}=0.$$

Thus for proving (5.14), it remains to show that for $\bar\chi_R(x):=1-\chi_R(x)$,

$$\lim_{R\to\infty}\sup_{m,m'}\sup_n\mathbf{E}\left|\int_0^t ((b_m-b_{m'})\bar\chi_R)(r,X^n_r)\mathrm{d}r\right|^2=0. \tag{5.15}$$

Let $u^{m,m'}_{n,R}=u$ solve the following backward PDE:

$$\partial_s u+\Delta u+b_n\cdot\nabla u_n+(b_m-b_{m'})\bar\chi_R=0,\quad u(t)=0.$$

By (5.11) and Lemma 3.5, we have

$$\begin{aligned}\mathbf{E}\left|\int_0^t ((b_m-b_{m'})\bar\chi_R)(r,X^n_r)\mathrm{d}r\right|^2&\leqslant 2\|\langle(b_m-b_{m'})\bar\chi_R u^{m,m'}_{n,R},\rho_n\rangle\|_{L^1_t}\\&\lesssim\|b_m-b_{m'}\|_{\mathbb{L}^{q_b}_t\mathbf{H}^{\alpha_b}_{p_b}}\|\bar\chi_R\rho_n\|_{\mathscr{V}_t}\|u^{m,m'}_{n,R}\|_{\mathscr{V}_t}\lesssim\|b\|^2_{\mathbb{L}^{q_b}_t\mathbf{H}^{\alpha_b}_{p_b}}\|\bar\chi_R\rho_n\|_{\mathscr{V}_t},\end{aligned}$$

which implies (5.15) by (5.7). Finally, for (5.13), as in the proof of existence in Theorem 3.6, it follows by (5.5) and (5.6). $\square$

5.2. **Generalized martingale problems.** Let $(\alpha,p,q)\in\mathscr{I}_d$ (see (3.11) for a definition). For any $f\in\mathbb{L}^q_T\mathbf{H}^\alpha_p$, by Theorem 3.6, there exists a weak solution to the following backward PDE

$$\partial_t u+\Delta u+b\cdot\nabla u=f,\quad u(T)\equiv 0 \tag{5.16}$$

with

$$\|u\|_{\mathbb{L}^\infty_T\cap\mathscr{V}_T}\leqslant C\|f\|_{\mathbb{L}^q_T\mathbf{H}^\alpha_p}.$$

Note that the weak solution may be not unique if we do not assume $b\in\mathbb{L}^\infty_T\mathscr{B}+\mathbb{L}^2_T\mathbb{L}$. However, if we consider the following regularized PDE

$$\partial_t u_n+\Delta u_n+b_n\cdot\nabla u_n=f,\quad u_n(T)\equiv 0. \tag{5.17}$$

Then for each n, there is a unique solution u_n denoted by $\mathcal{S}^n_f$ so that

$$\|\mathcal{S}^n_f\|_{\mathbb{L}^\infty_T\cap\mathscr{V}_T}\leqslant C\|f\|_{\mathbb{L}^q_T\mathbf{H}^\alpha_p},$$

where C is independent of n. In other words, $f \mapsto \mathcal{S}^n_f$ is a bounded linear operator from $\mathbb{L}^q_T\mathbf{H}^\alpha_p$ to $\mathbb{L}^\infty_T \cap \mathscr{V}_T$. By Banach-Steinhaus theorem, there is a subsequence n_k and a bounded linear operator $\mathcal{S} : \mathbb{L}^q_T\mathbf{H}^\alpha_p \to \mathbb{L}^\infty_T \cap \mathscr{V}_T$ so that for each $f \in \mathbb{L}^q_T\mathbf{H}^\alpha_p$ and $t \in [0, T]$, $\varphi \in L^1(\mathbb{R}^d)$ (see the proof of Theorem 3.6),

$$\lim_{k\to\infty} \langle \mathcal{S}^{n_k}_f(t), \varphi\rangle = \langle \mathcal{S}_f(t), \varphi\rangle, \tag{5.18}$$

and for each $R > 0$,

$$\lim_{k\to\infty} \int_0^T \|\mathcal{S}^{n_k}_f(s) - \mathcal{S}_f(s)\|^2_{L^2(B_R)}\mathrm{d}s = 0. \tag{5.19}$$

In the following we fix such an operator $\mathcal{S}$ and construct a unique generalized martingale solution associated with $\mathcal{S}$. Of course, its definition depends on the choice of subsequence n_k as well as the mollifiers ϕ_{n_k}. We emphasize that when $b \in \mathbb{L}^\infty_T\mathcal{B} + \mathbb{L}^2_T\mathbb{L}$, $\mathcal{S}_f$ is nothing, but the unique solution of (5.16), which is independent of the choice of n_k and the mollifiers.

Definition 5.8. *Let $\mu \in \mathscr{P}(\mathbb{R}^d)$. We call a probability measure $\mathbb{P} \in \mathscr{P}(\mathbb{C}_T)$ a generalized martingale solution of SDE (1.1) starting from μ and associated with the operator $\mathcal{S}$, if $\mathbb{P} \circ (w_0)^{-1} = \mu$ and for any $f \in C^\infty_c([0, T] \times \mathbb{R}^d)$,*

$$M^f_t := \mathcal{S}_f(t, w_t) - \mathcal{S}_f(0, w_0) - \int_0^t f(r, w_r)\mathrm{d}r, \quad w_\cdot \in \mathbb{C}_T,$$

is an almost surely martingale under $\mathbb{P}$ with respect to the natural filtration $\mathscr{B}_s$ in the sense that there is a Lebesgue null set $\mathscr{N} \subset (0, T)$ such that for all $0 \leqslant s < t \leqslant T$ with $s, t \notin \mathscr{N}$,

$$\mathbb{E}(M^f_t|\mathscr{B}_s) = M^f_s.$$

Remark 5.9. *Since, for each $t \in [0, T]$, $\mathcal{S}_f(t, \cdot) \in L^\infty$, in order to make sense of M^f_t, we have assumed that the marginal distribution $\mathbb{P} \circ w_t^{-1}$ has a density. However, it's essential to note that although $t \mapsto M^f_t$ is defined pointwise, we lack information about the path properties of $t \mapsto M^f_t$. For instance, we have no knowledge of its cádlág property since we only have $\mathcal{S}_f \in \mathbb{L}^\infty_T$.*

Now we show the following main result.

Theorem 5.10. *Assume* $(\mathbf{H}^{\sup})$ *holds. If the initial distribution μ has an L^2-density w.r.t. the Lebesgue measure, then there is a unique generalized martingale solution to SDE (1.1) associated with the operator $\mathcal{S}$ in the sense of Definition 5.8.*

Proof. **(Existence)** Let $\mathbb{P}$ be any accumulation point of $(\mathbb{P}_n)_{n\in\mathbb{N}}$. We want to show that $\mathbb{P}$ is a martingale solution in the sense of Definition 5.8. Up to taking a subsequence, without loss of generality we assume that

$$\mathbb{P}_n \to \mathbb{P} \quad \text{as } n \to \infty \text{ in } \mathscr{P}(\mathbb{C}_T) \text{ weakly.}$$

For $f \in C^\infty_c([0, T] \times \mathbb{R}^d)$, let $u_n = \mathcal{S}^n_f$ solve the following backward PDE:

$$\partial_t u_n + \Delta u_n + b_n \cdot \nabla u_n = f, \quad u_n(T) = 0.$$

By Itô's formula, we have

$$u_n(t, X^n_t) = u_n(0, X^n_0) + \int_0^t f(s, X^n_s)\mathrm{d}s + \sqrt{2}\int_0^t \nabla u_n(s, X^n_s)\mathrm{d}W_s. \tag{5.20}$$

Define

$$M^n_t := u_n(t, w_t) - u_n(0, w_0) - \int_0^t f(r, w_r)\mathrm{d}r.$$

Since $\mathbb{P}_n = \mathbf{P} \circ (X^n_\cdot)^{-1}$, by (5.20) one sees that M^n_t is a martingale under $\mathbb{P}_n$ with respect to $\mathscr{B}_s$. More precisely, for any $s < t$ and bounded continuous $\mathscr{B}_s$-measurable functional G_s,

$$\mathbb{E}^{\mathbb{P}_n}((M^n_t - M^n_s)G_s) = 0,$$

equivalently,

$$\mathbb{E}^{\mathbb{P}_n}\Big(((u_n(t,w_t)-u_n(s,w_s))G_s\Big)=\mathbb{E}^{\mathbb{P}_n}\left(G_s\int_s^t f(r,w_r)\mathrm{d}r\right).$$

Clearly, we have

$$\lim_{n\to\infty}\mathbb{E}^{\mathbb{P}_n}\left(G_s\int_s^t f(r,w_r)\mathrm{d}r\right)=\mathbb{E}^{\mathbb{P}}\left(G_s\int_s^t f(r,w_r)\mathrm{d}r\right). \tag{5.21}$$

By (5.19), there is a Lebesgue null set $\mathscr{N}\subset[0,T]$ depending on f such that for all $t\notin\mathscr{N}$ and $R>0$,

$$\lim_{k\to\infty}\|\mathcal{S}_f^{n_k}(t)-\mathcal{S}_f(t)\|_{L^2(B_R)}=0. \tag{5.22}$$

We want to show that for $s,t\notin\mathscr{N}$ with $s<t$,

$$\lim_{n\to\infty}\mathbb{E}^{\mathbb{P}_n}\Big((u_n(t,w_t)-u_n(s,w_s))G_s\Big)=\mathbb{E}^{\mathbb{P}}\Big((\mathcal{S}_f(t,w_t)-\mathcal{S}_f(s,w_s))G_s\Big). \tag{5.23}$$

For fixed $m\in\mathbb{N}$, since $x\mapsto u_m(t,x)$ is bounded continuous, we have

$$\lim_{n\to\infty}\mathbb{E}^{\mathbb{P}_n}\Big((u_m(t,w_t)G_s\Big)=\mathbb{E}^{\mathbb{P}}\Big(u_m(t,w_t)G_s\Big).$$

On the other hand, noting that for $R>0$,

$$\begin{aligned}\mathbb{E}^{\mathbb{P}_n}\big|(u_m(t,w_t)-\mathcal{S}_f(t,w_t))\mathbb{1}_{B_R}(w_t)\big|&=\int_{B_R}\big|u_m(t,x)-\mathcal{S}_f(t,x)\big|\rho_n(t,x)\mathrm{d}x\\&\leqslant\|u_m(t,\cdot)-\mathcal{S}_f(t,\cdot)\|_{L^2(B_R)}\|\rho_n(t)\|_{L^2(B_R)},\end{aligned}$$

by (5.22) and (5.9) we have

$$\lim_{m\to\infty}\sup_n\mathbb{E}^{\mathbb{P}_n}\big|(u_m(t,w_t)-\mathcal{S}_f(t,w_t))\mathbb{1}_{B_R}(w_t)\big|=0$$

and

$$\lim_{m\to\infty}\mathbb{E}^{\mathbb{P}}\big|(u_m(t,w_t)-\mathcal{S}_f(t,w_t))\mathbb{1}_{B_R}(w_t)\big|=0.$$

Moreover, we also have

$$\mathbb{E}^{\mathbb{P}_n}\big|(u_m(t,w_t)-\mathcal{S}_f(t,w_t))\mathbb{1}_{B^c_R}(w_t)\big|=\int_{B^c_R}\big|u_m(t,x)-\mathcal{S}_f(t,x)\big|\rho_n(t,x)\mathrm{d}x\leqslant\|u_m(t)\|_\infty\mu^n_t(B^c_R),$$

where $\mu^n_t=\mathbb{P}^n\circ w_t^{-1}$. By the tightness of $(\mu^n_t)_{n\in\mathbb{N}}$, we have

$$\lim_{R\to\infty}\sup_{m,n}\mathbb{E}^{\mathbb{P}_n}\big|(u_m(t,w_t)-\mathcal{S}_f(t,w_t))\mathbb{1}_{B^c_R}(w_t)\big|\leqslant C\lim_{R\to\infty}\sup_n\mu^n_t(B^c_R)=0.$$

Thus we get (5.23). Hence, for all $s,t\notin\mathscr{N}$ with $s<t$ and any bounded continuous $\mathscr{B}_s$-measurable G_s,

$$\mathbb{E}^{\mathbb{P}}\Big((\mathcal{S}_f(t,w_t)-\mathcal{S}_f(s,w_s))G_s\Big)=\mathbb{E}^{\mathbb{P}}\left(G_s\int_s^t f(r,w_r)\mathrm{d}r\right),$$

which gives

$$\mathbb{E}(M^f_t|\mathscr{B}_s)=M^f_s.$$

In particular, by noticing that $u_n(T)=\mathcal{S}_f(T)=0$, (5.23) holds for $t=T$ and $s\notin\mathscr{N}$. Now, we show (5.23) holds for $s=0$ and $t\notin\mathscr{N}$ as well. We only need to show that for any bounded continuous function $g:\mathbb{R}^d\to\mathbb{R}$,

$$\lim_{n\to\infty}\mathbb{E}^{\mathbb{P}_n}u_n(0,w_0)g(w_0)=\mathbb{E}^{\mathbb{P}}\mathcal{S}_f(0,w_0)g(w_0).$$

We note that $\mathbb{P}_n\circ(w_0)^{-1}=\mathbb{P}\circ(w_0)^{-1}=\mu_0$ and by (5.18),

$$\lim_{n\to\infty}\langle u_n(0),\rho_0 g\rangle=\langle\mathcal{S}_f(0),\rho_0 g\rangle.$$

Therefore, we have $\mathbb{E}(M^f_t|\mathscr{F}_0)=0$. The existence of a generalized martingale solution is proven.

(Uniqueness) Let $\mathbb{P}_1, \mathbb{P}_2$ be two generalized martingale solutions with initial distribution $\mu_0(\mathrm{d}x) = \rho_0(x)\mathrm{d}x$ in the sense of Definition 5.8. By the definition, for any $f \in C_c([0,T]\times\mathbb{R}^d)$, there is a Lebesgue full measure set $\mathscr{A}_f \subset [0,T]$ containing 0 and T such that for any $s,t \in \mathscr{A}_f$ and bounded $\mathscr{B}_s$-measurable G_s,

$$\mathbb{E}^{\mathbb{P}_i}((M_t^f - M_s^f)G_s) = 0, \quad i = 1,2. \tag{5.24}$$

Since $C_c([0,T]\times\mathbb{R}^d)$ is separable, we can choose $\mathscr{A} = \mathscr{A}_f$ independent of f.

Firstly, taking $(s,t) = (0,T)$, $G \equiv 1$ in (5.24), we have

$$\mathbb{E}^{\mathbb{P}_1}\left(\int_0^T f(t,w_t)\mathrm{d}t\right) = \langle \mathcal{S}_f(0), \rho_0\rangle = \mathbb{E}^{\mathbb{P}_2}\left(\int_0^T f(t,w_t)\mathrm{d}t\right).$$

which implies that $\mathbb{P}_1$ and $\mathbb{P}_2$ have the same one dimensional time marginal distribution. In fact, taking $f(t,x) = h(t)g(x)$ in the above equality, where $h \in C([0,T])$ and $g \in C_c(\mathbb{R}^d)$, by Fubini's theorem, one finds that for Lebesgue almost all $t \in (0,T)$ and any $g \in C_c(\mathbb{R}^d)$,

$$\mathbb{E}^{\mathbb{P}_1} g(w_t) = \mathbb{E}^{\mathbb{P}_2} g(w_t).$$

Since $t \mapsto w_t$ is continuous, we further have for any $t \in [0,T]$ and $g \in C_c(\mathbb{R}^d)$,

$$\mathbb{E}^{\mathbb{P}_1} g(w_t) = \mathbb{E}^{\mathbb{P}_2} g(w_t) \Rightarrow \mathbb{P}_1 \circ w_t^{-1} = \mathbb{P}_2 \circ w_t^{-1}.$$

Below we show that $\mathbb{P}_1$ and $\mathbb{P}_2$ have the same finite dimensional time marginal distribution by induction. We assume that for some $n \in \mathbb{N}$, $\mathbb{P}_1$ and $\mathbb{P}_2$ have the same n-dimensional time marginal distribution, that is, for any $0 \leqslant t_1 < t_2 < \cdots < t_n \leqslant T$,

$$\mathbb{P}_1 \circ (w_{t_1}, \cdots, w_{t_n})^{-1} = \mathbb{P}_2 \circ (w_{t_1}, \cdots, w_{t_n})^{-1}.$$

Let $t_0 = 0$ and $0 \leqslant t_1 < t_2 < \cdots < t_{n-1} \leqslant T$ be fixed. Let $g \in C_c(\mathbb{R}^{nd})$ and $t_n \in (t_{n-1}, T) \cap \mathscr{A}$. By (5.24) with $(s,t) = (t_n, T)$ and $G_s = g(w_{t_1}\cdots, w_{t_n})$, we have

$$\mathbb{E}^{\mathbb{P}_i}\left(\int_{t_n}^T f(s,w_s)\mathrm{d}s\, g(w_{t_1}\cdots,w_{t_n})\right) = \mathbb{E}^{\mathbb{P}_i}\Big(\mathcal{S}_f(t_n, w_{t_n}) g(w_{t_1}\cdots, w_{t_n})\Big), \quad i = 1,2,$$

which implies by the induction hypothesis that

$$\mathbb{E}^{\mathbb{P}_1}\left(\int_{t_n}^T f(s,w_s)\mathrm{d}s\, g(w_{t_1}\cdots,w_{t_n})\right) = \mathbb{E}^{\mathbb{P}_2}\left(\int_{t_n}^T f(s,w_s)\mathrm{d}s\, g(w_{t_1}\cdots,w_{t_n})\right).$$

Therefore, for any $t_n \in (t_{n-1}, T) \cap \mathscr{A}$ and Lebesgue almost all $t_{n+1} \in (t_n, T]$,

$$\mathbb{E}^{\mathbb{P}_1}\left(g_1(w_{t_{n+1}}) g(w_{t_1}\cdots,w_{t_n})\right) = \mathbb{E}^{\mathbb{P}_2}\left(g_1(w_{t_{n+1}}) g(w_{t_1}\cdots,w_{t_n})\right).$$

By the continuity of $t \mapsto w_t$, we get for all $t_n, t_{n+1} \in (t_{n-1}, T]$ with $t_n < t_{n+1}$,

$$\mathbb{P}_1 \circ (w_{t_1}, ..., w_{t_n}, w_{t_{n+1}})^{-1} = \mathbb{P}_2 \circ (w_{t_1}, ..., w_{t_n}, w_{t_{n+1}})^{-1}.$$

Namely, $\mathbb{P}_1 = \mathbb{P}_2$. The proof is complete. □

Now we can give

Proof of Theorem 1.4. Let n_k be any subsequence and n_k' be the subsubsequence of n_k used in the definition of $\mathcal{S}_f$. By Theorems 5.10 and 5.7, any accumulation point of $\mathbb{P}_{n_k'}$ is a generalized martingale solution, and meanwhile, a weak solution. By the uniqueness of generalized martingale solutions, for any initial value μ, without further selecting a subsequence, $\mathbb{P}_{n_k'}$ weakly converges to the weak solution of SDE (1.1). Finally, the Markov property follows by the uniqueness of generalized martingale solutions (see [12]). If $b = b_1 + b_2$, where $b_1 \in \mathbb{L}_T^\infty\mathcal{B}$ and $b_2 \in \mathbb{L}_T^2 L^2$, then PDE (5.16) has a unique weak solution. Thus, no necessary to select a subsequence, $\mathcal{S}_f$ is well-defined, and as above, the whole sequence $\mathbb{P}_n$ weakly converges to the weak solution of SDE (1.1). □

6. Applications to diffusion in random environment

In this section, we introduce certain Gaussian noise b so that our results can be applied to them. In particular, we construct the diffusion process in random environment using this noise.

6.1. **Vector-valued Gaussian fields.** In this section we introduce the notion of vector-valued Gaussian field. Fix $m\in\mathbb{N}$ and let $\mu(\mathrm{d}\xi):=\{\mu_{ij}(\mathrm{d}\xi)\}_{i,j=1}^m$ be a matrix-valued signed Radon measure on $\mathbb{R}^d$. We assume that

(A) For each Borel measurable set $A\subset\mathbb{R}^d$, $(\mu_{ij}(A))_{i,j=1}^m$ is a symmetric semi-positive definite Hermit matrix, that is, for any complex numbers $(a_i)_{i=1}^m$,

$$\mu_{ij}(A)=\mu_{ji}(A),\quad \sum_{i,j=1}^m a_i\bar{a}_j\mu_{ij}(A)\geqslant 0. \tag{6.1}$$

Moreover, $\mu(\mathrm{d}\xi)=\mu(-\mathrm{d}\xi)$ and for some $\ell\in\mathbb{R}$,

$$\sup_{i,j=1,\cdots,m}\int_{\mathbb{R}^d}(1+|\xi|)^\ell|\mu_{ij}|_{\rm var}(\mathrm{d}\xi)<\infty, \tag{6.2}$$

where $|\cdot|_{\rm var}$ stands for the total variation measure of a signed measure.

For two $\mathbb{R}^m$-valued Schwartz functions $f,g\in\mathscr{S}(\mathbb{R}^d;\mathbb{R}^m)$ and $k=0,1,2,\cdots$, we define

$$\langle f,g\rangle_{\mu;k}:=\sum_{i,j=1}^m\int_{\mathbb{R}^d}(1+|\xi|^{2k})\widehat{f}_i(\xi)\overline{\widehat{g_j}(\xi)}\mu_{ij}(\mathrm{d}\xi).$$

Since $\xi\to\mu_{ij}(\mathrm{d}\xi)$ is symmetric, one sees that $\langle f,g\rangle_{\mu;k}=\langle g,f\rangle_{\mu;k}\in\mathbb{R}$ is an inner product on $\mathscr{S}(\mathbb{R}^d;\mathbb{R}^m)$, and by (6.1), for any $f\in\mathscr{S}(\mathbb{R}^d;\mathbb{R}^m)$,

$$\|f\|_{\mu;k}:=\left(\sum_{i,j=1}^m\int_{\mathbb{R}^d}(1+|\xi|^{2k})\widehat{f}_i(\xi)\overline{\widehat{f}_j(\xi)}\mu_{ij}(\mathrm{d}\xi)\right)^{1/2}\geqslant 0.$$

Hence, $\|\cdot\|_{\mu;k}$ is a seminorm on linear $\mathscr{S}(\mathbb{R}^d;\mathbb{R}^m)$. Clearly,

$$\|\cdot\|_\mu:=\|\cdot\|_{\mu;0}\leqslant\|\cdot\|_{\mu;k}\leqslant\|\cdot\|_{\mu;k+1}.$$

It must be noticed that $\|f\|_\mu=0$ does not imply $f=0$. Let E_0 be the null space of $\|\cdot\|_\mu$, i.e.,

$$E_0:=\{f\in\mathscr{S}(\mathbb{R}^d;\mathbb{R}^m):\|f\|_\mu=0\}.$$

We define $\mathbb{H}_k$ being the completion of the quotient space $\mathscr{S}(\mathbb{R}^d;\mathbb{R}^m)/E_0$ with respect to the seminorm $\|\cdot\|_{\mu;k}$. Thus we obtain a Hilbert space and

$$\mathbb{H}_{k+1}\subset\mathbb{H}_k\subset\mathbb{H}_0.$$

Remark 6.1. *Assume $\ell\geqslant 0$ in (6.2), then the Dirac measure $\delta_0^{\otimes m}$ belongs to $\mathbb{H}_0$. Indeed, for $f_n(x)=n^df(nx)$, where $f\in\mathscr{S}(\mathbb{R}^d;\mathbb{R}^m)$ satisfies $\int_{\mathbb{R}^d}f_i(x)\mathrm{d}x=1$ for each $i=1,\cdots,m$, we have*

$$\|f_n-f_{n'}\|_\mu^2=\sum_{i,j=1}^m\int_{\mathbb{R}^d}(\widehat{f}_i(\xi/n)-\widehat{f}_i(\xi/n'))\overline{(\widehat{f}_j(\xi/n)-\widehat{f}_j(\xi/n'))}\mu_{ij}(\mathrm{d}\xi).$$

Since $\ell\geqslant 0$ in (6.2) and $\widehat{f}_i(0)=1$, by the dominated convergence theorem, we get

$$\lim_{n\to\infty}\|f_n-f_{n'}\|_\mu=0.$$

In general, the element in $\mathbb{H}_0$ may be not a Schwartz distribution.

For any $x \in \mathbb{R}^d$, $k \in \mathbb{N}$ and $f \in \mathscr{S}(\mathbb{R}^d;\mathbb{R}^m)$, define

$$\tau_x f(\cdot) := f(\cdot - x) \in \mathscr{S}(\mathbb{R}^d;\mathbb{R}^m), \quad \nabla^k f(\cdot) := \partial_x^k f(\cdot) \in \mathscr{S}(\mathbb{R}^d;\mathbb{R}^m \otimes (\mathbb{R}^d)^{\otimes k}).$$

Noting that

$$\widehat{\tau_x f}(\xi) = \mathrm{e}^{-\mathrm{i}x\cdot\xi}\hat{f}(-\xi), \quad \widehat{\nabla^k f}(\xi) = \mathrm{i}^k \xi^{\otimes k}\hat{f}(\xi),$$

we clearly have

$$\|\tau_x f\|_{\mathbb{H}_0} \leqslant \|f\|_{\mathbb{H}_0}, \quad \|\nabla^k f\|_{\mathbb{H}_0} \leqslant \|f\|_{\mathbb{H}_k}.$$

In particular, we can extend $\tau_x : \mathbb{H}_0 \to \mathbb{H}_0$ and $\nabla^k : \mathbb{H}_k \to \mathbb{H}_0$ being bounded linear operators.

We have the following simple lemma.

Lemma 6.2. *For any $f \in \mathbb{H}_0$, the function $x \mapsto \tau_x f \in \mathbb{H}_0$ is continuous, and for any $f \in \mathbb{H}_k$, $x \mapsto \tau_x f \in \mathbb{H}_0$ is k-order Fréchet differentiable and*

$$\nabla_x^k \tau_x f = \tau_x \nabla^k f \quad \text{in } \mathbb{H}_0. \tag{6.3}$$

Proof. By the boundedness $\|\tau_x f\|_{\mathbb{H}_0} \leqslant \|f\|_{\mathbb{H}_0}$, we may assume $f \in \mathscr{S}(\mathbb{R}^d;\mathbb{R}^m)$. Thus, by definition,

$$\|\tau_x f - \tau_y f\|_\mu^2 = \sum_{i,j=1}^m \int_{\mathbb{R}^d} |\mathrm{e}^{-\mathrm{i}\xi\cdot x} - \mathrm{e}^{-\mathrm{i}\xi\cdot y}|^2 \widehat{f}_i(\xi)\overline{\widehat{f}_j(\xi)}\mu_{ij}(\mathrm{d}\xi),$$

which implies by the dominated convergence theorem,

$$\lim_{x\to y} \|\tau_x f - \tau_y f\|_\mu^2 = 0.$$

For (6.3), it suffices to show it for $k = 1$. For $h, x \in \mathbb{R}^d$ with $h \neq 0$, by definition we have

$$\left\|\frac{\tau_{x+h} f - \tau_x f}{h} - \tau_x \nabla f\right\|_\mu^2 = \sum_{i,j=1}^m \int_{\mathbb{R}^d} \left|\frac{\mathrm{e}^{-\mathrm{i}\xi\cdot h} - 1}{h} + \mathrm{i}\xi\right|^2 \widehat{f}_i(\xi)\overline{\widehat{f}_j(\xi)}\mu_{ij}(\mathrm{d}\xi).$$

Noting that

$$\sum_{i,j=1}^m \int_{\mathbb{R}^d} \sup_h \left|\frac{\mathrm{e}^{-\mathrm{i}\xi\cdot h} - 1}{h} + \mathrm{i}\xi\right|^2 |\widehat{f}_i(\xi)\overline{\widehat{f}_j(\xi)}\mu_{ij}(\mathrm{d}\xi) \leqslant \sum_{i,j=1}^m \int_{\mathbb{R}^d} |\xi|^2 \widehat{f}_i(\xi)\overline{\widehat{f}_j(\xi)}\mu_{ij}(\mathrm{d}\xi) < \infty,$$

by the dominant convergence theorem, we have

$$\lim_{h\to 0} \left\|\frac{\tau_{x+h} f - \tau_x f}{h} - \tau_x \nabla f\right\|_\mu = 0.$$

The proof is complete. □

Now we introduce the following notion of vector-valued Gaussian random field (see [39]).

Definition 6.3. *Let U be the real valued Gaussian random field on $\mathbb{H}_0$, that is, U is a continuous linear operator from $\mathbb{H}_0$ to $L^2(\Omega, \mathbf{P})$, and for each $f \in \mathbb{H}_0$, $U(f)$ is a real-valued Gaussian random variable with mean zero and variance $\|f\|_{\mathbb{H}_0}^2$. In particular,*

$$\mathbf{E}\big(U(f)U(g)\big) = \langle f, g\rangle_{\mathbb{H}_0}. \tag{6.4}$$

We shall call U an $\mathbb{R}^m$-valued Gaussian noise over $\mathbb{R}^d$ with matrix-valued spectral measure μ.

We call U an $\mathbb{R}^m$-valued Gaussian noise. The reason is that for given real valued Schwartz function $h \in \mathscr{S}(\mathbb{R}^d;\mathbb{R})$ and $i = 1, \cdots, m$, if we define

$$f_i := (0, \cdots, 0, h, 0, \cdots, 0) \in \mathscr{S}(\mathbb{R}^d;\mathbb{R}^m), \quad U_i(h) := U(f_i), \tag{6.5}$$

then it is easy to see that

$$U = (U_1, \cdots, U_m).$$

Moreover, for any $f, g \in \mathscr{S}(\mathbb{R}^d;\mathbb{R})$, we have

$$\mathbf{E}(U_i(f)U_j(g)) = \int_{\mathbb{R}^d} \hat{f}(\xi)\overline{\hat{g}(\xi)}\mu_{ij}(\mathrm{d}\xi),$$

which gives the covariance between different components. In particular, for $f = (f_1, \cdots, f_m) \in \mathscr{S}(\mathbb{R}^d;\mathbb{R}^m)$, we shall write

$$U(f) = \langle U, f\rangle = U_1(f_1) + \cdots + U_m(f_m).$$

On the other hand, in the above definition, the Gaussian field U is defined as a bounded linear operator from $\mathbb{H}_0$ to $L^2(\Omega, \mathbf{P})$. A natural question arises: can we view U as an m-dimensional random distribution of the spatial variable $x \in \mathbb{R}^d$? We shall provide an affirmative answer and show its Besov regularity in the variable x below.

To achieve this, we first introduce the convolution of U with $\varphi \in \mathbb{H}_0$. For a given $\varphi \in \mathbb{H}_0$, we define

$$(\varphi * U)(x) := U_\varphi(x) := U(\tau_x\varphi) \in L^2(\Omega, \mathbf{P}).$$

Theorem 6.4. *For any $\varphi \in \mathbb{H}_0$, the function $x \mapsto U_\varphi(x)$ is continuous in $L^2(\Omega, \mathbf{P})$, and for any $k \in \mathbb{N}$ and $\varphi \in \mathbb{H}_k$, $x \mapsto U_\varphi(x)$ is k-order Fréchet differentiable, and for each $x \in \mathbb{R}^d$,*

$$\nabla^k_x U_\varphi(x) = \langle U, \tau_x \nabla^k \varphi\rangle, \quad \mathbf{P}-a.s. \tag{6.6}$$

Moreover, for any $p \geqslant 2$, there is a constant $C = C(d, p) > 0$ such that for all $\varphi \in \mathbb{H}_k$,

$$\sup_{x\in\mathbb{R}^d} \|\nabla^k U_\varphi(x)\|_{L^p(\Omega)} \leqslant C\|\varphi\|_{\mu;k}. \tag{6.7}$$

Proof. Note that

$$\mathbf{E}|U_\varphi(x) - U_\varphi(y)|^2 = \mathbf{E}|U(\tau_x\varphi - \tau_y\varphi)|^2 = \|\tau_x\varphi - \tau_y\varphi\|^2_{\mathbb{H}_0}.$$

The continuity of $x \mapsto U_\varphi(x)$ now follows by Lemma 6.2. Similarly, by (6.3), we have (6.6). Moreover, by the the hypercontractivity of Gaussian random variables, we have for any $p \geqslant 2$,

$$\begin{aligned}\|\nabla^k U_\varphi(x)\|_{L^p(\Omega)} \lesssim \|U_{\nabla^k\varphi}(x)\|_{L^2(\Omega)} &= \left(\sum_{i,j=1}^m \int_{\mathbb{R}^d} \widehat{\nabla^k\varphi_i}(\xi)\overline{\widehat{\nabla^k\varphi_j}(\xi)}\mu_{ij}(\mathrm{d}\xi)\right)^{1/2} \\ &= \left(\sum_{i,j=1}^m \int_{\mathbb{R}^d} |\xi|^{2k}\widehat{\varphi}_i(\xi)\overline{\widehat{\varphi_j}(\xi)}\mu_{ij}(\mathrm{d}\xi)\right)^{1/2} \leqslant \|\varphi\|_{\mu;k}.\end{aligned}$$

This completes the proof. $\square$

Now for $f \in \mathscr{S}(\mathbb{R}^d;\mathbb{R})$, we define

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

Lemma 6.5. *For any $\varphi \in \mathbb{H}_0$, U_φ is an $\mathbb{R}$-valued Gaussian noise with spectral measure*

$$\mu_\varphi(\mathrm{d}\xi) = \sum_{i,j=1}^m \hat{\varphi}_i(\xi)\hat{\varphi}_j(-\xi)\mu_{ij}(\mathrm{d}\xi).$$

Proof. For any $f, g \in \mathscr{S}(\mathbb{R}^d;\mathbb{R})$, by definition and Fubini's theorem, we have

$$\begin{aligned}\mathbf{E}\big(U_\varphi(f)U_\varphi(g)\big) &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathbf{E}\big(U_\varphi(x)U_\varphi(y)\big)f(x)g(y)\mathrm{d}x\mathrm{d}y \\ &= \sum_{i,j=1}^m \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\left[\int_{\mathbb{R}^d} \widehat{\tau_x\varphi_i}(\xi)\overline{\widehat{\tau_y\varphi_j}(\xi)}\mu_{ij}(\mathrm{d}\xi)\right]f(x)g(y)\mathrm{d}x\mathrm{d}y \\ &= \sum_{i,j=1}^m \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\left[\int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}(\xi\cdot x-\xi\cdot y)}\widehat{\varphi}_i(\xi)\overline{\widehat{\varphi_j}(\xi)}\mu_{ij}(\mathrm{d}\xi)\right]f(x)g(y)\mathrm{d}x\mathrm{d}y \\ &= \sum_{i,j=1}^m \int_{\mathbb{R}^d} \hat{f}(-\xi)\hat{g}(\xi)\widehat{\varphi}_i(\xi)\widehat{\varphi_j}(-\xi)\mu_{ij}(\mathrm{d}\xi) = \langle f, g\rangle_{\mu_\varphi}.\end{aligned}$$

The result now follows. □

Below, for a Banach space $\mathbb{B}$, let $\mathbf{B}^s_p(\mathbb{B}) := \mathbf{B}^s_{p,\infty}(\mathbb{B})$ be the $\mathbb{B}$-valued Besov space, and for any $\rho_\kappa(x) = (1+|x|)^{-\kappa}$, where $\kappa \in \mathbb{R}$, let $\mathbf{B}^s_p(\rho_\kappa; \mathbb{B})$ be the weighted Besov space, which is defined by (see [22])

$$\|f\|_{\mathbf{B}^s_p(\rho_\kappa;\mathbb{B})} := \|f\rho_k\|_{\mathbf{B}^s_p(\mathbb{B})}.$$

We have the following regularity result.

Theorem 6.6. *Under* **(A)**, *for any $p \in [1,\infty)$, $\kappa > d/p$ and $\beta < \ell/2$, it holds that*

$$U \in \mathbf{B}^{\ell/2}_\infty(L^p(\Omega;\mathbb{R}^m)) \subset \mathbf{B}^{\ell/2}_p(\rho_\kappa; L^p(\Omega;\mathbb{R}^m)) \subset L^p(\Omega; \mathbf{B}^\beta_p(\rho_\kappa;\mathbb{R}^m)). \tag{6.8}$$

Proof. First of all, recalling the definition of $(\phi_j)_{j\geqslant -1}$ in (2.1), one sees that for any $\gamma \in \mathbb{R}$,

$$|\phi_j(\xi)| \lesssim_C 1 \wedge \big(2^{\gamma j}(1+|\xi|)^{-\gamma}\big), \quad j \geqslant -1, \quad \xi \in \mathbb{R}^d. \tag{6.9}$$

Indeed, for $j=-1$, since $\mathrm{supp}\phi_{-1} \subset \mathbf{B}_1$, it is obvious. For $j \geqslant 0$, noting that

$$K_j := \mathrm{supp}\phi_j \subset \{\xi : 2^{j-1} \leqslant |\xi| \leqslant 2^{j+1}\},$$

we have for any $\gamma \in \mathbb{R}$,

$$|\phi_j(\xi)| \leqslant \mathbf{1}_{K_j}(\xi) \lesssim 2^{j\gamma}/(1+|\xi|)^\gamma.$$

For each $i = 1,\cdots,m$, under (6.2), by definition (6.5) and the hypercontractivity of Gaussian random variables, we have for any $p \in [1,\infty)$ and $j = 0,1,\cdots$,,

$$\begin{aligned}\|\mathcal{R}_j U_i(x)\|_{L^p(\Omega)} &\lesssim \|\mathcal{R}_j U_i(x)\|_{L^2(\Omega)} = \|U_i(\tau_x\check{\phi}_j)\|_{L^2(\Omega)} = \left(\int_{\mathbb{R}^d} |\phi_j(\xi)|^2 \mu_{ii}(\mathrm{d}\xi)\right)^{1/2} \\ &\lesssim 2^{-\ell j/2}\left(\int_{\mathbb{R}^d}(1+|\xi|)^\ell \mu_{ii}(\mathrm{d}\xi)\right)^{1/2} \lesssim 2^{-\ell j/2},\end{aligned} \tag{6.10}$$

and for $\kappa > d/p$,

$$\|\rho_\kappa \mathcal{R}_j U\|^p_{L^p(\mathbb{R}^d;L^p(\Omega))} = \int_{\mathbb{R}^d}(1+|x|)^{-p\kappa}\|\mathcal{R}_j U(x)\|^p_{L^p(\Omega)}\mathrm{d}x \lesssim 2^{-p\ell j/2}\int_{\mathbb{R}^d}(1+|x|)^{-p\kappa}\mathrm{d}x.$$

Hence, by [22, Theorem 2.7],

$$U \in \mathbf{B}^{\ell/2}_\infty(L^p(\Omega;\mathbb{R}^m)) \subset \mathbf{B}^{\ell/2}_p(\rho_\kappa; L^p(\Omega;\mathbb{R}^m)).$$

Finally, for $\beta < \ell/2$, by Fubini's theorem, we have

$$\Big\|\sup_{j\geqslant 0} 2^{\beta j}\|\rho_\kappa \mathcal{R}_j U\|_p\Big\|_{L^p(\Omega)} \leqslant \sum_{j\geqslant 0} 2^{\beta j}\|\rho_\kappa\mathcal{R}_j U\|_{L^p(\mathbb{R}^d\times\Omega)} \leqslant \sum_{j\geqslant 0} 2^{(\beta-\ell/2)j}\|U\|_{\mathbf{B}^{\ell/2}_p(\rho_\kappa;L^p(\Omega;\mathbb{R}^m))}.$$

The proof is completed . □

6.2. **Examples: Divergence-free Gaussian field.** In this section we present concrete examples to illustrate our results. The following example is standard.

Example 6.7. Let $m = 1$. For $\gamma \in [0,d)$, let

$$\mu(\mathrm{d}\xi) := |\xi|^{-\gamma}\mathrm{d}\xi.$$

Obviously, (6.2) holds. In this case, for $\gamma > 0$, it is well known that for some $c_{d,\gamma} > 0$ (see [37, p117, Lemma 2]),

$$\hat{\mu}(x) = c_{d,\gamma}|x|^{\gamma-d}.$$

In particular, for any $f, g \in \mathscr{S}(\mathbb{R}^d)$,

$$\mathbf{E}\big(U(f)U(g)\big) = \int_{\mathbb{R}^d}\hat{f}(\xi)\hat{g}(-\xi)\mu(\mathrm{d}\xi) = c_{d,\gamma}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\frac{f(x)g(y)}{|x-y|^{d-\gamma}}\mathrm{d}x\mathrm{d}y.$$

For $\gamma = 0$, we have

$$\hat{\mu}(x) = \delta_0(\mathrm{d}x)$$

and

$$\mathbf{E}\big(U(f)U(g)\big) = \int_{\mathbb{R}^d} f(x)g(x)\mathrm{d}x.$$

Namely, X is a space white noise on $\mathbb{R}^d$. As in (6.10) and by the change of variables, we have

$$\|\mathcal{R}_j U(x)\|_{L^p(\Omega)} \lesssim \left(\int_{\mathbb{R}^d} |\phi_j(\xi)|^2 \mu(\mathrm{d}\xi)\right)^{1/2} = 2^{j(d-\gamma)/2} \left(\int_{\mathbb{R}^d} |\phi_0(\xi)|^2 \frac{\mathrm{d}\xi}{|\xi|^\gamma}\right)^{1/2}.$$

Hence, for any $p \in [1, \infty)$, $\kappa > d/p$ and $\beta < (\gamma - d)/2$,

$$U \in \mathbf{B}^{(\gamma-d)/2}_\infty(L^p(\Omega)) \subset \mathbf{B}^{(\gamma-d)/2}_p(\rho_\kappa; L^p(\Omega)) \subset L^p(\Omega, \mathbf{B}^\beta_p(\rho_\kappa)). \tag{6.11}$$

In particular, bigger γ means that Gaussian field U has better regularity.

Example 6.8 (Divergence free vector-valued Gaussian field)**.** For given $\gamma \in [0, d)$, define

$$\mu^{(\gamma)}(\mathrm{d}\xi) := |\xi|^{-\gamma}\Big(\mathbb{I}_{d\times d} - \frac{\xi \otimes \xi}{|\xi|^2}\Big)\mathrm{d}\xi. \tag{6.12}$$

It is easy to see that (6.2) holds for $\ell < \gamma - d$, and for $a := (a_i)_{i=1}^d \subset \mathbb{C}$ and $\xi \in \mathbb{R}^d$,

$$a^T\Big(\mathbb{I}_{d\times d} - \frac{\xi \otimes \xi}{|\xi|^2}\Big)\bar{a} = |a|^2 - \sum_{i,j}(a_i\xi_i\bar{a}_j\xi_j)/|\xi|^2 = |a|^2 - \Big|\sum_i (a_i\xi_i)\Big|^2/|\xi|^2 \geqslant 0.$$

Thus **(A)** holds. Let $b^{(\gamma)}$ be the $\mathbb{R}^d$-valued Gaussian random field with matrix-valued spectral measure $\mu^{(\gamma)}(\mathrm{d}x)$. Then by (6.8), for any $p \in [1, \infty)$, $\kappa > d/p$ and $\beta < \frac{\gamma-d}{2}$,

$$b^{(\gamma)} \in L^p(\Omega; \mathbf{B}^\beta_p(\rho_\kappa; \mathbb{R}^m)).$$

In particular, there is a $\mathbf{P}$-null set $N \subset \Omega$ so that for all $\omega \notin N$,

$$b^{(\gamma)}(\omega, \cdot) \in \mathbf{B}^\beta_p(\rho_\kappa; \mathbb{R}^m),$$

and

$$\mathrm{div} b^{(\gamma)}(\omega, \cdot) = 0.$$

In fact, for any $f \in \mathscr{S}(\mathbb{R}^d)$,

$$\begin{aligned}\mathbf{E}|\langle \mathrm{div} b^{(\gamma)}, f\rangle|^2 = \mathbf{E}|\langle b^{(\gamma)}, \nabla f\rangle|^2 &= \int_{\mathbb{R}^d} \widehat{\nabla f}(\xi)\Big(\mathbb{I}_{d\times d} - \frac{\xi\otimes\xi}{|\xi|^2}\Big)\big[\overline{\widehat{\nabla f}(\xi)}\big]^*|\xi|^{-\gamma}\mathrm{d}\xi \\ &= \int_{\mathbb{R}^d} |\hat{f}(\xi)|^2 \xi\Big(\mathbb{I}_{d\times d} - \frac{\xi\otimes\xi}{|\xi|^2}\Big)\xi^*|\xi|^{-\gamma}\mathrm{d}\xi = 0,\end{aligned} \tag{6.13}$$

where $*$ stands for the transpose of a row vector. In particular, one can apply Theorem 1.2 to solve the following SDE with $\gamma \in (d-2, d)$:

$$\mathrm{d}X_t = b^{(\gamma)}(\omega, X_t)\mathrm{d}t + \sqrt{2}\mathrm{d}W_t,$$

where X_t can be thought of the diffusion in random environment (see [13]).

Remark 6.9. *Consider $d = 3$. Let $\mathbf{U}^{(\gamma)} := (U_1, U_2, U_3)$ be the 3-dimensional Gaussian field with spectral measure matrix $\mu(\mathrm{d}\xi) = |\xi|^{-\gamma}\mathbb{I}_{3\times 3}\mathrm{d}\xi$, where $\gamma \in (1, 3)$. Let $b^{(\gamma)}$ be the $\mathbb{R}^3$-valued Gaussian field in the above example. It is easy to see that (see [13])*

$$b^{(\gamma)} = \nabla \times (-\Delta)^{-\frac{1}{2}}\mathbf{U}^{(\gamma)}, \quad \nabla\times := \begin{pmatrix} 0 & -\partial_3 & \partial_2 \\ \partial_3 & 0 & -\partial_1 \\ -\partial_2 & \partial_1 & 0 \end{pmatrix}.$$

Remark 6.10. *Consider $d=2$. Let U be the two-dimensional spatial white noise with spectral measure $\mu(\mathrm{d}\xi)=\mathrm{d}\xi$ (see Example 6.7). Let $\mathscr{S}_0(\mathbb{R}^2)$ be the class of Schwartz function $f\in\mathscr{S}(\mathbb{R}^2)$ with $0\notin \mathrm{supp}(\hat f)$. Let $b^{(0)}$ be the $\mathbb{R}^2$-valued Gaussian field in Example 6.8. Then it is easy to see that for any $f,g\in\mathscr{S}_0(\mathbb{R}^2)$,*

$$\mathbf{E}\big((\nabla^\perp(-\Delta)^{-\frac12}U)(f)(\nabla^\perp(-\Delta)^{-\frac12})U(g)\big)=\mathbf{E}(b^{(0)}(f)b^{(0)}(g)),$$

where $\nabla^\perp:=(\partial_{x_2},-\partial_{x_1})$. In other words,

$$b^{(0)}=\nabla^\perp(-\Delta)^{-\frac12}U \text{ on } \mathscr{S}_0(\mathbb{R}^2).$$

However, $\mathscr{S}_0(\mathbb{R}^2)$ is not a deterministic class since for $f\in\mathscr{S}_0(\mathbb{R}^2)$, $\int_{\mathbb{R}^2}f(x)\mathrm{d}x=0$. So we can not claim $b^{(0)}=\nabla^\perp(-\Delta)^{-\frac12}U$ as distributions.

Example 6.11. Let b_ε be the Gaussian field with matrix-valued spectral measure (see Remark 1.6)

$$\mu_\varepsilon(\mathrm{d}\xi):=\mathbf{1}_{|\xi|<1/\varepsilon}\Big(\mathbb{I}_{d\times d}-\frac{\xi\otimes\xi}{|\xi|^2}\Big)\mathrm{d}\xi.$$

By Remark 6.1 and Theorem 6.6, for $p\in[1,\infty)$ and $\kappa>d/p$, we have

$$b_\varepsilon(x):=b_\varepsilon(\tau_x\delta_0)\in\cap_{\beta>0}L^p(\Omega;\mathbf{B}^\beta_p(\rho_\kappa;\mathbb{R}^m)).$$

Let $0<\varepsilon<\varepsilon'\leqslant 1/2$. Noting that

$$(\mathbb{I}-\Delta)^{-\frac{\beta}{2}}(b_\varepsilon-b_{\varepsilon'})(x)=(b_\varepsilon-b_{\varepsilon'})(\tau_x(\mathbb{I}-\Delta)^{-\frac{\beta}{2}}\delta_0),$$

for $\beta>1$ and $p\geqslant 2$, we have for all $x\in\mathbb{R}^d$,

$$\begin{aligned}\|(\mathbb{I}-\Delta)^{-\frac{\beta}{2}}(b_\varepsilon-b_{\varepsilon'})(x)\|_{L^p(\Omega)}&\lesssim\|(\mathbb{I}-\Delta)^{-\frac{\beta}{2}}(b_\varepsilon-b_{\varepsilon'})(x)\|_{L^2(\Omega)}\\&=\left(\int_{\mathbb{R}^2}\mathbf{1}_{1/\varepsilon'|\leqslant\xi|<1/\varepsilon}(1+|\xi|^2)^{-\frac{\beta}{2}}\mathrm{d}\xi\right)^{\frac12}\lesssim|\varepsilon-\varepsilon'|^{\frac{\beta-1}{2}}.\end{aligned}$$

Therefore, as in Theorem 6.6, for any $p\in[1,\infty)$, $\kappa>d/p$ and $\beta>1$, there exists a vector field $b\in L^p(\Omega,\mathbf{H}^{-\beta}_p(\rho_\kappa))$ with $\mathrm{div}b=0$ and such that

$$\lim_{\varepsilon\to0}\mathbf{E}\|\rho_\kappa(b_\varepsilon-b)\|^p_{\mathbf{H}^{-\beta}_p}=0.$$

On the other hand, fix $p\in[2,\infty)$ and $\kappa>d/p$. It is easy to see that

$$\sup_{x\in\mathbb{R}^d}\|(\mathbb{I}-\Delta)^{-\frac12}b_\varepsilon(x)\|_{L^p(\Omega)}\leqslant C\sqrt{\ln(1/\varepsilon)},$$

and so,

$$\sup_{\varepsilon\in(0,1/2)}\mathbf{E}\left(\frac{\|\rho_\kappa b_\varepsilon\|_{\mathbf{H}^{-1}_p}}{\sqrt{\ln(1/\varepsilon)}}\right)^p=0.$$

By a tightness argument and Skorokhod's representation theorem, there is a subsequence $\varepsilon_n\to0$, a probability space $(\widetilde\Omega,\widetilde{\mathscr{F}},\widetilde{\mathbb{P}})$ and a family of Gaussian process $\{\widetilde b_n\}_{n=1}^\infty$ thereon such that

$$\widetilde{\mathbb{P}}\circ\widetilde b_n^{-1}=\mathbf{P}\circ\left(\frac{b_{\varepsilon_n}}{\sqrt{\ln(1/\varepsilon_n)}}\right)^{-1}\quad\text{in }\mathbf{H}^{-1}_p(\rho_\kappa),$$

and, there is a $\widetilde{\mathbb{P}}$-null set N such that for all $\widetilde\omega\notin N$,

$$\sup_{n\in\mathbb{N}}\|\widetilde b_n(\widetilde\omega,\cdot)\|_{\mathbf{H}^{-1}_p(\rho_\kappa)}<\infty. \tag{6.14}$$

6.3. **Stochastic heat equations.** Let $\gamma\in(-\infty,d)$ and $\eta^{(\gamma)}$ be an $\mathbb{R}^d$-valued time-white Gaussian noise with matrix-valued spectral measure $\mu^{(\gamma)}(\mathrm{d}\xi)\mathrm{d}t$, where $\mu^{(\gamma)}$ is given by (6.12). More precisely, $\eta^{(\gamma)}$ is a Gaussian field over $\mathbb{R}_+\times\mathbb{R}^2$ with covariance given by that for any $f,g\in L^2([0,\infty);\mathscr{S}(\mathbb{R}^d;\mathbb{R}^d))$,

$$\mathbf{E}\big(\eta^{(\gamma)}(f)\eta^{(\gamma)}(g)\big)=\sum_{i,j=1}^d\int_0^\infty\!\!\int_{\mathbb{R}^2}\hat f_i(t,\xi)\hat g_j(t,-\xi)\mu^{(\gamma)}_{ij}(\mathrm{d}\xi)\mathrm{d}t.$$

Now, we consider the following $\mathbb{R}^d$-valued stochastic heat equation (abbreviated as SHE) on $\mathbb{R}^d$:

$$\partial_t u^{(\gamma)}=\Delta u^{(\gamma)}+\eta^{(\gamma)},\quad u^{(\gamma)}_0=0. \tag{6.15}$$

Let

$$p_t(x)=(4\pi t)^{-1}\exp\{-|x|^2/(4t)\}.$$

By Duhamael's formula, the solution is determined by the following formula:

$$u^{(\gamma)}(t,x):=\int_0^t\int_{\mathbb{R}^d}p_{t-s}(x-y)\eta^{(\gamma)}(\mathrm{d}y,\mathrm{d}s):=\eta^{(\gamma)}(\mathbf{1}_{(0,t)}\tau_x p_{t-\cdot}).$$

Noting that for any $j\geqslant 0$ and $i=1,\cdots,m$,

$$\mathcal{R}_j u^{(\gamma)}_i(t,x)=\eta^{(\gamma)}_i(\mathbf{1}_{(0,t)}\tau_x\mathcal{R}_j p_{t-\cdot}),$$

as in Theorem 6.6, by (2.1) and the change of variables, we have

$$\begin{aligned}\|\mathcal{R}_j u^{(\gamma)}_i(t,x)\|^2_{L^p(\Omega)}\lesssim\|\mathcal{R}_j u^{(\gamma)}_i(t,x)\|^2_{L^2(\Omega)}&=\int_0^t\!\!\int_{\mathbb{R}^d}|\tau_x\mathcal{R}_j p_{t-s}(\xi)|^2\mu^{(\gamma)}_{ii}(\mathrm{d}\xi)\mathrm{d}s\\&=\int_0^t\!\!\int_{\mathbb{R}^d}|\phi_j(\xi)|^2\mathrm{e}^{-2(t-s)|\xi|^2}|\xi|^{-\gamma}(1-|\xi_i|^2/|\xi|^2)\mathrm{d}\xi\mathrm{d}s\\&\leqslant 2^{j(d-\gamma)}\int_0^t\!\!\int_{\mathbb{R}^d}|\phi_0(\xi)|^2\mathrm{e}^{-2^{2j+1}s|\xi|^2}|\xi|^{-\gamma}\mathrm{d}\xi\mathrm{d}s\\&\lesssim 2^{j(d-\gamma-2)}\int_{\mathbb{R}^d}|\phi_0(\xi)|^2|\xi|^{-2-\gamma}\mathrm{d}\xi,\end{aligned}$$

and also,

$$\|\mathcal{R}_0 u^{(\gamma)}_{-1}(t,x)\|^2_{L^p(\Omega)}\lesssim\int_{\mathbb{R}^d}|\phi_{-1}(\xi)|^2|\xi|^{-\gamma}\mathrm{d}\xi<\infty.$$

Hence, for $\gamma<d-2$,

$$u^{(\gamma)}\in L^\infty(\mathbb{R}_+;\mathbf{B}^{1+\frac{\gamma-d}{2}}_\infty(L^p(\Omega;\mathbb{R}^d))).$$

Moreover, as (6.13), we also have

$$\mathrm{div}u^{(\gamma)}=0.$$

In particular, we can use Theorem 1.2 to *time-dependent* vector field $b^{(\gamma)}(t,x)=u^{(\gamma)}(t,x)$ with $\gamma\in(d-4,d-2)$.

Acknowledgement: We would like to thank Rongchan Zhu and Xiangchan Zhu for their invaluable contributions, particularly for their insights into the background of the critical cases and the uniqueness of generalized martingale solutions. This work is partially supported by NNSFC grants of China (Nos. 12131019), and the German Research Foundation (DFG) through the Collaborative Research Centre(CRC) 1283 "Taming uncertainty and profiting from randomness and low regularity in analysis, stochastics and their applications".

References

[1] H. Bahouri, J.-Y. Chemin and R. Danchin. *Fourier Analysis and Nonlinear Partial Differential Equations.* Grundlehren Math. Wiss., 343, Springer, Heidelberg, 2011. xvi+523 pp.

[2] R.F. Bass R.F. and Z.Q. Chen. Stochastic differential equations for Dirichlet processes. *Probability Theory and Related Fields.* 121(3) (2001), pp.422-446.

[3] R.F. Bass R.F. and Z.Q. Chen. Brownian motion with singular drift. *The Annals of Probability.* 31(2) (2003), 791-817.

[4] L. Beck, F. Flandoli, M. Gubinelli and M. Maurelli. Stochastic ODEs and stochastic linear PDEs with critical drift: regularity, duality and uniqueness *Electron. J. Probab.* **24** (2019), Paper No. 136, 72 pp.

[5] J. Bergh, J. Löfström.*Interpolation spaces.* Grundlehren Math. Wiss., 223, Springer, Heidelberg, 1976.

[6] O. Butkovsky, K. Lê, L. Mytnik. Stochastic equations with singular drift driven by fractional Brownian motion. http://arxiv.org/abs/2302.11937v1.

[7] G. Cannizzaro and K. Chouk. Multidimensional SDEs with singular drift and universal construction of the polymer measure with white noise potential. *Ann. Probab.* **46** (2018), 1710–1763.

[8] G. Cannizzaro, L. Haunschmid-Sibitz, F. Toninelli, $\sqrt{\log t}$-superdiffusivity for a Brownian particle in the curl of the 2D GFF. *Ann. Probab.* 50 (2022), no. 6, 2475–2498.

[9] G. Chatzigeorgiou, P. Morfe, F. Otto and L. Wang. The Gaussian free-field as a stream function: asymptotics of effective diffusivity in infra-red cut-off. arXiv:2212.14244 (2022).

[10] Z.-Q. Chen, E. Hu, L. Xie and X. Zhang. Heat kernels for non-symmetric diffusion operators with jumps. *J. Differential Equations* 263 (2017) 6576-6634.

[11] F. Delarue and R. Diel. Rough paths and 1d SDE with a time dependent distributional drift: application to polymers. *Probab. Theory Related Fields* **165** (2016), 1–63.

[12] N. Ethier and G. Kurtz. *Markov Processes: Characterization and Convergence.* Wiley Ser. Probab. Math. Statist. Probab. Math. Statist. John Wiley & Sons, Inc., New York, 1986. x+534 pp.

[13] A. Fannjiang. Normal and anomalous diffusions in random flows. Mathematics of Multiscale Materials, Edited by Kenneth M. Golden Geoffrey R. Gritnmett Richard D. James Graeme W. Milton Pabitra N. Sen. The IMA Volumes in Mathematics and its Applications Volume 99, 1998, 81-100.

[14] F. Flandoli, E. Issoglio and F. Russo. Multidimensional stochastic differential equations with distributional drift. *Trans. Am. Math. Soc.* **369** (2017), 1665-1688.

[15] H. Föllmer. Calcul d'Itô sans probabilités, *Séminaire de Probabilités XV*, 143-150, Springer, Berlin (1981).

[16] J. Fontbona and M. Martinez. Paths Clustering and an Existence Result for Stochastic Vortex Systems. *Journal of Statistical Physics*, Vol. 128, No. 3, August 2007

[17] P. K. Friz and M. Hairer. *A course on rough paths. With an introduction to regularity structures.* Universitext, Springer, Cham, 2014. xiv+251 pp.

[18] L. Galeati and M. Gerencsér. Solution theory of fractional SDEs in complete subcritical regimes. https://arxiv.org/pdf/2207.03475.pdf.

[19] L. Gräfner, A. Djurdjevac, X. Ji and N. Perkowski. Scaling-supercritical singular SDEs with almost divergence-free distributional drift. *In preparation* (2023).

[20] L. Gräfner and N. Perkowski. Energy solutions and generators of singular SPDEs. *Lecture note.* (2023). Link

[21] Z. Hao, M. Wu and X. Zhang. Schauder estimates for nonlocal kinetic equations and applications. *J. Math. Pures Appl. (9)* **140** (2020), 139–184.

[22] Z. Hao, X. Zhang, R. Zhu and X. Zhu. Singular kinetic equations and applications. arXiv:2108.05042. To appear in *Annals of Probability* (2023+).

[23] P.-E. Jabin and Z. Wang. Quantitative estimates of propagation of chaos for stochastic systems with $\mathbf{W}^{-1,\infty}$ kernels. *Invent. math.* (2018) 214:523–591.

[24] D. Kinzebulatov and K.R. Madou. Strong solutions of SDEs with singular (form-bounded) drift via Roeckner-Zhao approach. arXiv:2306.04825 (2023).

[25] H. Kremp and N. Perkowski. Multidimensional SDE with distributional drift and Lévy noise. *Bernoulli* **28** (2022), 1757–1783.

[26] N. V. Krylov. On time inhomogeneous stochastic Itô equations with drift in L^{d+1}. *Ukrainian Math. J.*, 72 (2021), no. 9, 1420–1444.

[27] N. V. Krylov. On diffusion processes with drift in L^d. *Probab. Theory Related Fields* 179 (2021), no. 1-2, 165–199.

[28] N. V. Krylov. On strong solutions of Itô's equations with $\sigma \in W^1_d$ and $b \in L_d$. *Ann. Probab.* **49** (2021), 3142–3167.

[29] N.V. Krylov and M. Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields* **131** (2005), 154–196.

[30] S.V. Lototsky and A. Shah. Gaussian fields and stochastic heat equations. *Differential Integral Equations* **33** (2020), 527–554.

[31] H. Osada. A stochastic differential equation arising from the vortex problem. *Proc. Jpn. Acad. Ser. A Math. Sci.* 61, 333–336 (1986)
[32] H. Osada. Diffusion processes with generators of generalized divergence form. *J. Math. Kyoto Univ.* **27** (1987), 597–619.
[33] M. Röckner and G. Zhao. SDEs with critical time dependent drifts: strong solutions. arXiv:2103.05803 (2021).
[34] M. Röckner and G. Zhao. SDEs with critical time dependent drifts: weak solutions. *Bernoulli* **29** (2023), 757–784.
[35] G. Seregin, L. Silvestre, V. Šverák, A. Zlatoš. On divergence-free drifts. *J. Differential Equations* **252** (2012), 505–540.
[36] S. Sheffield. Gaussian free fields for mathematicians. *Probab. Theory Relat. Fields* **139** (2007), 521–541.
[37] E.M. Stein. *Singular Integrals and Differentiability Properties of Functions*. Springer-Verlag, Berlin, 1971.
[38] D. W. Stroock and S.S. Varadhan. *Multidimensional diffusion processes.* Springer-Verlag, Berlin, 1979.
[39] J. Svante, *Gaussian Hilbert Spaces.* Cambridge University Press, 1997.
[40] S. Takanobu. On the existence and uniqueness of SDE describing an n-particle system interacting via a singular potential. *Proc. Japan Acad. Ser. A Math. Sci.* 61:287–290 (1985).
[41] A.J. Veretennikov. On the strong solutions of stochastic differential equations. *Theory Probab. Appl.* **24** (1979), 354–366.
[42] L. Xie and X. Zhang. Ergodicity of stochastic differential equations with jumps and singular coefficients. *Annales de l'Institut Henri Poincaré - Probabilités et Statistiques*. 2020, Vol. 56, No. 1, 175-229.
[43] X. Zhang. Strong solutions of SDEs with singular drift and Sobolev diffusion coefficients. *Stochastic Process. Appl.*, **115/11** pp. 1805–1818(2005).
[44] X. Zhang. Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients. *Electron. J. Probab.* **16** (2011), 1096-1116.
[45] X. Zhang. Stochastic differential equations with Sobolev diffusion and singular drift. *Annals of Applied Probability* **26** (2016), 2697-2732.
[46] X. Zhang. Maximum principle for non-uniformly parabolic equations and applications. arXiv:2012.05026. To appear in *Annali della Scuola Normale Superiore di Pisa, Classe di Scienze*, (2023+).
[47] X. Zhang and G. Zhao: Singular Brownian diffusion processes. *Commun. Math. Stat.* **6** (2018), 533-581.
[48] X. Zhang and G. Zhao: Heat kernel and ergodicity of SDEs with distributional drifts. arXiv:1710.10537 (2017).
[49] X. Zhang and G. Zhao: Stochastic Lagrangian Path for Leray's Solutions of 3D Navier-Stokes Equations. *Commun. Math. Phys.* **381** (2021), 491–525.

Zimo Hao: Fakultät für Mathematik, Universität Bielefeld, 33615, Bielefeld, Germany, Email: zhao@math.uni-bielefeld.de

Xicheng Zhang: School of Mathematics and Statistics, Beijing Institute of Technology, Beijing 100081, China, Email: XichengZhang@gmail.com