

SINGULAR KINETIC EQUATIONS AND APPLICATIONS

BY ZIMO HAO^{1,4,a}, XICHENG ZHANG^{2,b}, RONGCHAN ZHU^{2,c}, XIANGCHAN ZHU^{3,d}¹*Fakultät für Mathematik, Universität Bielefeld, D-33501 Bielefeld, Germany*, ^azhao@math.uni-bielefeld.de²*Department of Mathematics, Beijing Institute of Technology, Beijing 100081, China*, ^bxichengzhang@gmail.com;
^czhurongchan@126.com³*Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China*,
^dzhuxiangchan@126.com⁴*School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei 430072, China*

In this paper we study singular kinetic equations on \mathbb{R}^{2d} by the paracontrolled distribution method introduced in [GIP15]. We first develop paracontrolled calculus in the kinetic setting, and use it to establish the global well-posedness for the linear singular kinetic equations under the assumptions that the products of singular terms are well-defined. We also demonstrate how the required products can be defined in the case that singular term is a Gaussian random field by probabilistic calculation. Interestingly, although the terms in the zeroth Wiener chaos of regularization approximation are not zero, they converge in suitable weighted Besov spaces and no renormalization is required. As applications the global well-posedness for a nonlinear kinetic equation with singular coefficients is obtained by the entropy method. Moreover, we also solve the martingale problem for nonlinear kinetic distribution dependent stochastic differential equations with singular drifts.

CONTENTS

1	Introduction	2
1.1	Main results	5
1.2	Sketch of proofs and structure of the paper	7
1.3	Further relevant literature	8
1.4	Notations and conventions	9
2	Preliminaries	9
2.1	Weighted anisotropic Besov spaces	10
2.1.1	Definition	10
2.1.2	Embeddings and interpolation	12
2.1.3	Characterization	14
2.2	Paraproduct calculus	16
3	Kinetic semigroups and commutator estimates	17
3.1	Kinetic semigroup estimates	18
3.2	Kinetic Hölder spaces and characterization	21
3.3	Schauder's estimates	25
3.4	Commutator estimates	27
3.5	Renormalized pairs	31
4	Linear kinetic equations with distribution drifts	33
4.1	Paracontrolled solutions	34

MSC2020 subject classifications: Primary 60H15; secondary 35R60.*Keywords and phrases:* Paracontrolled calculus, Kinetic equations, Weighted anisotropic Besov spaces, Second order mean-field SDE, distributional drift.

4.2	Well-posedness for (4.1)	40
5	Well-posedness of singular mean field equations	43
6	Nonlinear martingale problem with singular drifts	52
7	Existence of renormalized pairs in probabilistic sense	59
7.1	Statement of main result	59
7.2	Examples for (\mathbf{A}^β)	61
7.3	Proof of Theorem 7.2	63
A	Characterizations for $\mathbf{B}_{p,q}^{s,a}(\rho)$	69
B	Proof of Lemma 7.4	77
	Funding	80
	References	80

1. Introduction. In this paper we are concerned with the following nonlinear kinetic equation with singular drifts in \mathbb{R}^{2d} :

$$(1.1) \quad \partial_t u = \Delta_v u - v \cdot \nabla_x u - b \cdot \nabla_v u - K * \langle u \rangle \cdot \nabla_v u, \quad u(0) = u_0,$$

where $u : \mathbb{R}_+ \times \mathbb{R}^{2d} \rightarrow \mathbb{R}$ is a function of time variable t , position x and velocity v , $\langle u \rangle(t, x) := \int_{\mathbb{R}^d} u(t, x, v) dv$ stands for the mass, $K : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is a kernel function and

$$K * \langle u \rangle(t, x) := \int_{\mathbb{R}^d} K(x - y) \langle u \rangle(t, y) dy,$$

and for some $\alpha \in (\frac{1}{2}, \frac{2}{3})$ and $T > 0$,

$$(1.2) \quad b = (b_1, \dots, b_d) \in (L_T^\infty \mathbf{C}_a^{-\alpha}(\rho))^d,$$

is a Gaussian random field and the example of b which we have in mind is white noise in v and colored in x . Here ρ is a polynomial weight and $\mathbf{C}_a^{-\alpha}(\rho)$ stands for the weighted anisotropic Hölder space introduced in Subsection 2.1. The aim of this paper is to establish the well-posedness for the above singular SPDE and the associated distributional dependent SDEs (see (1.8) below) under suitable assumptions. In Subsection 1.1 we state the main results under suitable analytic assumptions, which could be verified by probabilistic assumptions on the covariance of b in Section 7.

The kinetic equation was originally introduced by Landau in 1936 to study the plasma phenomenon in physics, which is a nonlinear PDE with square and nonlocal second order term (see [Lan36], [AV04] and references therein). As model equations, we consider the following two linear kinetic equations

$$(1.3) \quad \mathcal{L}u := (\partial_t - v \cdot \nabla_x - \Delta_v)u = f,$$

$$\mathcal{L}^*u = (\partial_t + v \cdot \nabla_x - \Delta_v)u = f.$$

where \mathcal{L} is also called Kolmogorov operator since in [Kol34], he first wrote down the fundamental solution of \mathcal{L} (see (3.2) below). These two equations have the following relation:

$$\tau \mathcal{L}u = \mathcal{L}^*(\tau u), \quad \tau u(t, x, v) = u(t, x, -v)$$

and transform τ influences nothing in our formulation.

Now we consider the following scaling transform: for $\lambda > 0$ and $a, b, c > 0$, let

$$u_\lambda(t, x, v) := \lambda^a u(\lambda^b t, \lambda^c x, \lambda v), \quad f_\lambda(t, x, v) := f(\lambda^b t, \lambda^c x, \lambda v).$$

It is easy to check that

$$(1.4) \quad \mathcal{L}u_\lambda = f_\lambda \iff a = -2, b = 2, c = 3.$$

Next we consider the improvement of the regularities in x and v for (1.3). Suppose that for some $\alpha \in (0, 1)$ and $\beta, \gamma > 0$, there is a constant $C > 0$ such that for all $\lambda > 0$,

$$(1.5) \quad [u\lambda]_{\mathbf{C}_x^{\alpha+\gamma}} \lesssim_C [f\lambda]_{\mathbf{C}_x^\alpha}, \quad [u\lambda]_{\mathbf{C}_v^{\alpha+\beta}} \lesssim_C [f\lambda]_{\mathbf{C}_v^\alpha},$$

where for any $\gamma > 0$,

$$[g]_{\mathbf{C}_x^\gamma} := \sup_{h \in \mathbb{R}^d} \|\delta_{x;h}^{([\gamma]+1)} g\|_\infty / |h|^\gamma$$

with $\delta_{x;h}^{(1)} g(x, v) := g(x+h, v) - g(x, v)$ and $\delta_{x;h}^{(M+1)} = \delta_{x;h}^{(1)} \delta_{x;h}^{(M)}$, similarly for $[g]_{\mathbf{C}_v^\beta}$. Note that

$$[u\lambda]_{\mathbf{C}_x^{\alpha+\gamma}} = \lambda^{3(\alpha+\gamma)-2} [u]_{\mathbf{C}_x^{\alpha+\gamma}}, \quad [f\lambda]_{\mathbf{C}_x^\alpha} = \lambda^{3\alpha} [f]_{\mathbf{C}_x^\alpha},$$

and

$$[u\lambda]_{\mathbf{C}_v^{\alpha+\beta}} = \lambda^{\alpha+\beta-2} [u]_{\mathbf{C}_v^{\alpha+\beta}}, \quad [f\lambda]_{\mathbf{C}_v^\alpha} = \lambda^\alpha [f]_{\mathbf{C}_v^\alpha}.$$

Under scaling invariant (1.5), we must have

$$\gamma = 2/3, \quad \beta = 2.$$

In other words, the gains of the regularities for kinetic equation (1.3) in x and v are $\frac{2}{3}$ and 2, respectively. Thus the following Schauder estimate is expected: for any $\alpha, \beta > 0$, there is a constant $C = C(\alpha, \beta, d) > 0$ such that

$$(1.6) \quad \|u\|_{L_T^\infty \mathbf{C}_x^{\alpha+2/3}} + \|u\|_{L_T^\infty \mathbf{C}_v^{\beta+2}} \lesssim_C \|f\|_{L_T^\infty \mathbf{C}_x^\alpha} + \|f\|_{L_T^\infty \mathbf{C}_v^\beta},$$

where \mathbf{C}_x^α and \mathbf{C}_v^β stand for the Hölder spaces in directions x and v , respectively. Due to different scaling and regularity between x and v variables, we study (1.1) in the anisotropic Hölder space (see Subsection 2.1 for definition).

When $\alpha = \beta/3 > 0$, Schauder's estimate (1.6) has been studied extensively in [Lo05], [Pr09] (see [HWZ20], [IS21] for nonlocal version), and the maximal L^p -regularity estimates were obtained in [Bo02] (see also [CZ18], [HMP19] and [ZZ21] for stochastic version). We mention that the structure of Lie group was introduced to define the kinetic Hölder spaces for the Schauder estimates in [IS21] (see also earlier work [Po04]). In the current work, we introduce the kinetic Hölder space, which is equivalent to the one introduced in [IS21], without using the notion of Lie group.

One motivation for studying kinetic equation (1.1) with distribution valued coefficient b is to develop solution theory for degenerate singular SPDEs. When $\alpha > \frac{1}{2}$, due to the singularity of the coefficients b in (1.2), the best regularity of the solution to (1.1) is in $L_T^\infty \mathbf{C}_a^{2-\alpha}$, which makes the linear term $b \cdot \nabla_v u$ not well defined in the classical sense. Such kind of problems also arise in the understanding of singular SPDEs, such as famous KPZ equations [KPZ86], which have been intensely studied recently. Hairer in [Hai14] developed the regularity structure theory to give a meaning to a large class of singular SPDEs. Parallel to that, a paracontrolled distribution method was proposed by Gubinelli, Imkeller and Perkowski [GIP15], which is also a powerful tool for studying singular SPDEs. The key idea of these theories is to use the structure of solutions to give a meaning to the terms which are not classically defined. These terms are well-defined with the help of probabilistic calculation and renormalization for the "enhanced noise", i.e. the noise and the higher order terms appearing in the decomposition of the equations. Based on these ideas the solution theories for quasilinear parabolic singular SPDEs, Schrödinger and wave equations driven by singular noise have

been developed in [OW19, OSSW18, GHa19, OSSW21] and [DW18, GKO18, GKO18a] (see also the references therein). In this paper we aim to develop paracontrolled distribution calculus for the degenerate kinetic SPDEs with singular coefficients.

Going back to kinetic equation (1.1), it is natural to work on the whole space since the velocity v physically takes values in the whole space, where the coefficients b , which come from the noise and the renormalized terms, stay in the weighted Besov spaces. This prevents us from using a fixed point argument in the same space. To the best of our knowledge, there are two methods to solve this problem. One is to use a clever construction of exponential weight depending on time variable proposed in [HL18]. The other one is to use localization trick developed in [ZZZ22]. In this paper we follow the localization method in [ZZZ22] to solve this problem. We deduce a priori estimates for (1.1) and by a compactness argument obtain the existence of solutions. The localization argument also implies uniqueness. We refer to Section 1.2 for more details on the idea of the proof. Compared to the local solutions for singular SPDEs mentioned above, a priori estimates and the global well-posedness for different parabolic singular SPDEs have been obtained, see [MW17, MW17a, GH19] for the dynamical Φ_d^4 -model and [PR19, ZZZ22] for the KPZ equations and singular HJB equations.

Another motivation is that equation (1.1) can be viewed as the mean field limit of empirical measures for a second order interacting particle system in random environment. More precisely, consider the following N -interacting particle system in \mathbb{R}^d , where each particle obeys the Newtonian second law perturbed by time Gaussian noise \dot{B}_t^i and environment noise W :

$$\ddot{X}_t^{N,i} = W(X_t^{N,i}, \dot{X}_t^{N,i}) + \frac{1}{N} \sum_{j \neq i} K(X_t^{N,i} - X_t^{N,j}) + \sqrt{2} \dot{B}_t^i, \quad i = 1, \dots, N,$$

where $(B_t^i)_{i \in \mathbb{N}}$ is a sequence of d -dimensional independent standard Brownian motions on a stochastic basis $(\Omega, \mathcal{F}, \mathbf{P}; (\mathcal{F}_t)_{t \geq 0})$, $K : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is the interaction kernel, and $W : \mathbb{R}^{2d} \rightarrow \mathbb{R}^d$ is a vector-valued distribution and stands for the environmental noise, which acts on all particles. We will see in Section 7 that our condition on W allows for spatial white noise in v direction for $d = 1$, which may be derived from average of a sequence of i.i.d random variables (see e.g. [PR21, Remark 2.2]). The factor $\frac{1}{N}$ in front of the interacting force K is called mean-field scaling which keeps the total mass of order 1. If we introduce a new velocity variable $V_t^{N,i} := \dot{X}_t^{N,i}$ and let $Z_t^{N,i} := (X_t^{N,i}, V_t^{N,i})$, then the above second order SDE can be written as the familiar form:

$$(1.7) \quad \begin{cases} dX_t^{N,i} = V_t^{N,i} dt, & i = 1, 2, \dots, N, \\ dV_t^{N,i} = \left[W(Z_t^{N,i}) + \frac{1}{N} \sum_{j \neq i} K(X_t^{N,i} - X_t^{N,j}) \right] dt + \sqrt{2} dB_t^i. \end{cases}$$

On the other hand, for each $i \in \mathbb{N}$, consider the following kinetic distributional dependent SDE (abbreviated as DDSDE)

$$(1.8) \quad d\bar{X}_t^i = \bar{V}_t^i dt, \quad d\bar{V}_t^i = W(\bar{Z}_t^i) dt + (K * \mu_{\bar{X}_t^i})(\bar{X}_t^i) dt + \sqrt{2} dB_t^i,$$

where $\bar{Z}_t^i := (\bar{X}_t^i, \bar{V}_t^i)$ and for a probability measure μ in \mathbb{R}^d ,

$$K * \mu(x) := \int_{\mathbb{R}^d} K(x - y) \mu(dy).$$

When W and K are globally Lipschitz, it is well-known that there are unique solutions to (1.7) and (1.8), and the following propagation of chaos holds (see [Szn91, Theorem 1.4]): Suppose $Z_0^{N,i} = \bar{Z}_0^i$ and $\{\bar{Z}_0^i\}$ are i.i.d. random variables. Then for each $i \in \mathbb{N}$ and $T > 0$,

$$(1.9) \quad \sup_N \sqrt{N} \mathbf{E} \left(\sup_{t \in [0, T]} |Z_t^{N,i} - \bar{Z}_t^i| \right) < \infty.$$

Note that $(\bar{Z}^i)_{i \in \mathbb{N}}$ are i.i.d. random processes. Let $\mu = (\mu(t))_{t \geq 0}$ be the distribution of $(\bar{Z}^i)_{i \in \mathbb{N}}$. By Itô's formula, one sees that $\mu(t)$ solves the following non-linear Fokker-Planck equation: for any $\phi \in C_b^2(\mathbb{R}^{2d})$,

$$(1.10) \quad \partial_t \langle \mu, \phi \rangle = \langle \mu, \Delta_v \phi + v \cdot \nabla_x \phi + (W + K * \langle \mu \rangle) \cdot \nabla_v \phi \rangle.$$

With a little confusion of notation with $\langle \mu \rangle$, we also write

$$\langle \mu, \phi \rangle := \int_{\mathbb{R}^{2d}} \phi(z) \mu(dz).$$

Now, let $u_N(t) := \frac{1}{N} \sum_{i=1}^N \delta_{Z_t^{N,i}}$ be the empirical distribution measure. By (1.7) and Itô's formula again, one finds that for any $\phi \in C_b^2(\mathbb{R}^{2d})$,

$$(1.11) \quad d\langle u_N, \phi \rangle = \langle u_N, \Delta_v \phi + v \cdot \nabla_x \phi + (W + K * \langle u_N \rangle) \cdot \nabla_v \phi \rangle dt + \frac{\sqrt{2}}{N} \sum_{i=1}^N \nabla_v \phi(Z_t^{N,i}) dB_t^i.$$

In particular, each term in (1.11) converges to the corresponding one in (1.10) in suitable sense. For examples, by Itô's isometry, we have

$$\mathbf{E} \left| \frac{1}{N} \sum_{i=1}^N \int_0^t \nabla_v \phi(Z_s^{N,i}) dB_s^i \right|^2 = \frac{1}{N^2} \sum_{i=1}^N \mathbf{E} \int_0^t |\nabla_v \phi(Z_s^{N,i})|^2 ds \leq \frac{t \|\nabla_v \phi\|_\infty}{N} \rightarrow 0.$$

Note that if $W, K \in C_b^\infty$, then $\mu(t)$ has a smooth density $u(t, z)$ so that

$$(1.12) \quad \partial_t u = \Delta_v u - v \cdot \nabla_x u - \operatorname{div}_v((W + K * \langle u \rangle)u).$$

We also mention that when W depends on the random environment ω , the empirical measure u_N also converges to the solution to equation (1.12), which corresponds to the conditional law of \bar{Z} w.r.t. W with $\mu_{\bar{X}}$ in (1.8) also given by conditional law. This means that the conditional propagation of chaos holds (see [CF16] for more details). In particular, if $\operatorname{div}_v W \equiv 0$, then the above equation reduces to the form of (1.1). In physics this assumption is natural which is satisfied if the force only depends on the position. We refer to Section 1.3 for more background, more references in this direction. In the following we regard (1.12) and (1.1) as random PDEs, i.e. we fix a path of W , and solve the SPDE path by path.

1.1. Main results. The main goal of this article is to give a meaning to the kinetic equation (1.1) and establish the global well-posedness of (1.1) under (1.2). As mentioned above, since $b \cdot \nabla_v u$ does not make sense, we need to use paracontrolled method and perform renormalizations by probabilistic calculations to give a rigorous meaning to $b \cdot \nabla_v u$.

First of all, we consider the following linear PDE with distributions b, f :

$$(1.13) \quad \partial_t u = \Delta_v u + v \cdot \nabla_x u + b \cdot \nabla_v u + f, \quad u(0) = u_0.$$

To state our main results, we introduce some parameters and notations. Let $\vartheta := \frac{9}{2-3\alpha}$ for some $\alpha \in (1/2, 2/3)$. For given $\kappa_0 < 0$, $\kappa_1 \in (0, \frac{1}{2\vartheta+2}]$, $\kappa_2 \in \mathbb{R}$ and $\kappa_3 := (2\vartheta+1)\kappa_1 + \kappa_2$, in the statement of our main results below, we shall use the following weight functions:

$$\rho_i(x, v) := (1 + |x|^{1/3} + |v|)^{-\kappa_i}, \quad i = 0, 1, 2, 3.$$

Let $\mathbb{B}_T^\alpha(\rho_1, \rho_2)$ be the space of renormalized pairs and $\mathbb{B}_T^\alpha(\rho_1)$ the space of renormalized vector fields introduced in Definition 3.15. Formally $(b, f) \in \mathbb{B}_T^\alpha(\rho_1, \rho_2)$ and $b \in \mathbb{B}_T^\alpha(\rho_1)$ mean $b, f \in L_T^\infty(C_a^{-\alpha}(\rho_1))$ and for $\mathcal{J} = \mathcal{L}^{-1}$, $b \circ \nabla_v \mathcal{J} f \in L_T^\infty C_a^{1-2\alpha}(\rho_1 \rho_2)$, $b \circ \nabla_v \mathcal{J} b \in$

$L_T^\infty \mathbf{C}_a^{1-2\alpha}(\rho_1^2)$, are well-defined respectively, which in general could be realized by a probabilistic calculation. Here \circ is the paraproduct introduced in Subsection 2.2. The example we have in mind is a Gaussian random forcing and our assumption allow, for example, when $d = 1$, b to be white in v variable and colored in x variable. Compared to the heat semigroup, the interesting point is that the terms in the zeroth Wiener chaos are not zero and converge in the corresponding weighted Besov space. In fact, the terms in the zeroth Wiener chaos subtracting formally divergence terms which by symmetry are zero will converge. Hence no renormalization appears in the smooth approximation of equation (1.1).

The following result provides the well-posedness of the linear singular PDE (1.13).

THEOREM 1.1. *Suppose that $(b, f) \in \mathbb{B}_T^\alpha(\rho_1, \rho_2)$ and $b \in \mathbb{B}_T^\alpha(\rho_1)$. For any $T > 0$ and $\varphi \in \mathbf{C}_a^\gamma(\rho_2/\rho_1)$, where $\gamma > 1 + \alpha$, there is a unique paracontrolled solution $u \in \mathbb{S}_{T,a}^{2-\alpha}(\rho_3)$ to PDE (1.13) in the sense of Definition 4.1, where $\mathbb{S}_{T,a}^{2-\alpha}(\rho_3)$ is the kinetic Hölder space introduced in Definition 3.5.*

In Section 4 we prove this result. Along the way of proving Theorem 1.1, we develop paracontrolled calculus in the kinetic setting and prove a commutator estimate for the kinetic semigroup. We refer to Section 1.2 for more details on this point. The complete version of Theorem 1.1 is given in Theorem 4.7.

Next we consider the nonlinear kinetic Fokker-Planck equation (1.12).

THEOREM 1.2. *Let $T > 0$. Suppose that $W \in \mathbb{B}_T^\alpha(\rho_1)$ with $\operatorname{div}_v W = 0$ and $K \in \mathbf{C}^{\beta/3}(\mathbb{R}^d)$ with $\beta > \alpha - 1$. For $\gamma > 1 + \alpha$, and any probability density function u_0 with $u_0 \in L^1(\rho_0) \cap \mathbf{C}_a^\gamma$ there exists at least a probability density paracontrolled solution $u \in \mathbb{S}_{T,a}^{2-\alpha}(\rho_3)$ to equation (1.12).*

If in addition that K is bounded, then for any initial data $u_0 \in L^1(\rho_0) \cap \mathbf{C}_a^\gamma$ with finite entropy $\int u_0 \ln u_0 < \infty$, the solution is unique.

The complete version of Theorem 1.2 is given in Theorem 5.4. By $\operatorname{div}_v W \equiv 0$ we can write (1.12) in the non-divergence form

$$\partial_t u = \Delta_v u - v \cdot \nabla_x u - (W + K * \langle u \rangle) \cdot \nabla_v u$$

and Theorem 1.1 can be applied. As mentioned above the solution to (1.12) can be viewed as a probability density. Hence in this paper we concentrate on such kind of solutions. Formally from the equation we see the integral of solution is a constant. Also if the initial value is nonnegative, then a maximum principle implies the solution is always nonnegative. As usual, the key point to prove this theorem is to establish the a priori estimates (5.7) and (5.8) in Section 5 about entropy. Compared with the previous work in [JW16], our assumptions are more flexible. We refer to Section 1.2 for details on the idea of the proof.

Finally, as an application we also obtain the well-posedness for the associated nonlinear martingale problem of (1.8).

THEOREM 1.3. *Let $T > 0$. Suppose that $W \in \mathbb{B}_T^\alpha(\rho_1)$ and $K \in \mathbf{C}^{\beta/3}(\mathbb{R}^d)$ with $\beta > \alpha - 1$. For any initial probability distribution ν with finite moment $\int |z|_a^\delta \nu(dz) < \infty$, where $\delta > \frac{(4\vartheta+4)\kappa_1}{2-\alpha}$, there exists a martingale solution to nonlinear SDE (1.8) starting from ν . Moreover, if K is bounded measurable, then the solution is unique.*

The complete version of Theorem 1.3 is given in Theorem 6.3. Our martingale problem is considered in the sense of Either and Kurtz [EK86, Section 4.3, p173], which is a general

notion. The usual martingale problem is that for all functions u in the domain of generator $\mathcal{L}^\mu := \Delta_v + v \cdot \nabla_x + (b + K * \mu_t) \cdot \nabla_v$, the process $u(t, X_t, V_t) - u(0, x, v) - \int_0^t (\partial_t + \mathcal{L}^\mu)u(s, X_s, V_s)ds$ with $\mu_t = \text{Law}(X_t)$ is a martingale. However, due to singularity of b , smooth functions might be not in the domain of \mathcal{L}^μ . We can find such u by solving the backward Kolmogorov equation. We refer to Section 1.2 for more details on this point. This type of martingale problem has been treated in [DD16, CC18, KP22] for linear non-degenerated singular SDEs. To the best of our knowledge, this is the first well-posedness result for singular degenerate nonlinear SDEs. As an application of Theorem 1.3, one may derive the convergence of the marginal laws of the interacting particle system (1.7) to the nonlinear SDE (1.8) by using a general result from [La23] about the propagation of chaos.

1.2. Sketch of proofs and structure of the paper. In Section 2, we recall some facts about the anisotropic weighted Besov spaces and the associated paracontrolled calculus. In particular, a quite useful characterization of anisotropic weighted Besov spaces is stated in Theorem 2.7, whose proof is given in Appendix A.

For the kinetic semigroup, we introduce a new weighted kinetic Hölder space associated with the transport term $v \cdot \nabla_x$ (see Definition 3.5). On this space, Schauder's estimate for the kinetic semigroup is established (see Lemma 3.11). The key point to use the paracontrolled calculus for the kinetic equation (1.1) is a commutator estimate for the kinetic semigroup which we establish in Subsection 3.4. Note that it seems impossible to show a commutator estimate in the form $[\mathcal{L}_v, f \prec]g$ as in [GIP15] for $\mathcal{L}_v := \Delta_v + v \cdot \nabla_x$, since the loss of regularity from \mathcal{L}_v and the gain of regularity from the kinetic semigroup do not match i.e. the kinetic operator loses 1 regularity in x direction while the Schauder estimate for the kinetic semigroup only gains $2/3$ regularity in x direction. Moreover, the commutator for the kinetic semigroup under the action of block operator \mathcal{R}_j^α is not like the heat semigroup and there is an extra transport term left, which leads to a commutator estimate in the kinetic Hölder space introduced in Definition 3.5 (see Lemma 3.13). We refer the readers to the argument at the beginning of Section 3.4 for more details on this point. In Subsection 3.5, we give the notion of renormalized pairs as in [ZZZ22] as mentioned in Subsection 1.1.

Sections 4 and 5 are devoted to the well-posedness of equations (1.13) and (1.12). We first use paracontrolled calculus in the kinetic setting, characterization of the weighted Hölder space and localization trick developed in [ZZZ22] to derive uniform bounds in a polynomial growth weighted Besov space for the solutions to the linear equation (1.13). The new point is that we prove a localization result for paracontrolled solution (see Proposition 4.4). This localization property allows us to establish a priori estimate (4.31) for *any* paracontrolled solution of (1.13), which automatically yields the uniqueness. Note that the proof of the uniqueness in [ZZZ22] is to adopt the exponential weight technique developed in [HL18]. For the nonlinear equation mentioned above, we concentrate on probability density solutions. In this case, to prove existence of solutions and the convergence of the nonlinear term in (1.1), we need to show the convergence of the approximation solutions in L^1 -space, which follows from a moment estimate for some SDEs by a probabilistic method. Usually one obtains such kind of moment estimates for distributional drift SDE by using the Zvonkin transform to kill the singular drift term (see e.g. [ZZ18]). However, the required C^1 -diffeomorphism in Zvonkin transform cannot be constructed since in x -direction the regularity cannot be C^1 . In Section 5 we use Theorem 4.7 to deduce a Krylov type estimate, which can be used to control the distribution drift (see Lemma 5.8). The uniqueness proof follows from a priori entropy estimate and L^1 -estimate. To deal with the distributional drift term, we use linear approximations and Theorem 4.7.

In Section 6 we consider the martingale problem associated with (1.8) and establish the well-posedness. As mentioned in Subsection 1.1, we solve this martingale problem by analyzing the Kolomogorov backward equation. Since this is a nonlinear martingale problem, the corresponding Kolomogorov equation should be nonlinear. However, it is not known a-priori that the law of the solutions to (1.8) is absolutely continuous w.r.t. Lebesgue measure. As a result, we consider the linear equation for fixed μ and we can apply Theorem 1.1 directly. More precisely, we consider the following equation for fixed $\mu : [0, T] \rightarrow \mathcal{P}(\mathbb{R}^d)$:

$$(1.14) \quad \partial_t u + \mathcal{L}^\mu u = f, \quad u(T) = u^T,$$

for a sufficiently large class of functions f and u^T , and therefore we replace the martingale problem with the requirement that the process $u(t, X_t, V_t) - u(0, x, v) - \int_0^t f(s, X_s, V_s) ds$ with $\mu_t = \text{Law}(X_t)$ is a martingale. For the existence of a martingale solution, we use the standard tightness argument. Moreover, to obtain the convergence, we prove the continuity of the nonlinear term (see Lemma 6.5). For the uniqueness of martingale solutions, we first show the uniqueness of the solutions to the linear equations (i.e. $K \equiv 0$), and then use Girsanov's transformation and Gronwall's inequality.

Section 7 is concerned with the probabilistic analysis connected to the construction of the stochastic objects needed in the sequel. More precisely, we consider a class of stationary Gaussian distributions X of class $\mathbf{C}_a^{-\alpha}(\rho_\kappa)$. This class includes one dimensional spatial white noise in v direction and colored in x direction; any covariance operator $|\partial_x|^{-\lambda}$ with $\lambda > 5/9$ when $d = 1$ is admissible. For such X we construct the generalized products $\nabla_v \mathcal{J} X \circ X$ as probabilistic limits of smooth approximations.

1.3. Further relevant literature. The study of mean field limit and propagation of chaos for interacting particle system originated from McKean [McK67], see for instance the classical reference [Szn91]. As mentioned above, DDSDE which is also called McKean-Vlasov equation is closely related to mean field limit. To the best of our knowledge, Vlasov [Vl68] first proposed McKean-Vlasov's equations, which arise in many applications, such as multi-agent systems (see [BRTV98, BT97]), filtering (see [CX10]) and so on. Recently the research on the mean field limit for the 1st order system, with singular interaction kernels has experienced immense improvements including those results focusing on the vortex model [Osa86, FHM14] and more general singular kernels as in [JW18] and Serfaty [Ser20]. When $W \equiv 0$ and $K(x) \in L^\infty(\mathbb{R}^d)$, Jabin and Wang [JW16] studied the well-posedness of PDE (1.10) and propagation of chaos. In the pioneering work by Funaki [Fu84] the martingale problem for a non-linear PDE is clearly formulated. After that global well-posedness of DDSDE has been studied a lot in the literature (see [MV20] [Wa18] [RZ21] and references therein). Recently, there are also a lot of work concerning the mean field limit and the limiting DDSDE for interacting particle systems influenced by a common environmental noise as (1.7) (see e.g. [CF16, R20, HSS21] and reference therein). However, so far as we know, most work concentrate on the first order system, which is related to a parabolic SPDE, and the related common noise W is trace-class type noise, i.e. the noise W is function valued w.r.t. spatial variable.

In many applications such as control problems and Coulomb potential from physics, the coefficients for the related DDSDE are very singular. Hence, studying the nonlinear kinetic equation and DDSDE with singular coefficients counts for much. In the present paper, we can obtain global well-posedness for these nonlinear equations with singular environmental noise W , which so far as we know, has not been obtained in the literature.

The study of SDEs with distributional drifts has also attracted much interest in recent years (see [DD16, ZZ18, CC18, KP22] etc.). Such singular diffusions arise as models for

stochastic processes in random media. When $d = 1$, based on the rough path method, Delarue and Dielthe [DD16] studied the SDE with rough drift. In [CC18], based on the theory of paracontrolled calculus, Cannizzaro and Chouk proved the well-posedness for the martingale problem with singular drift in higher dimensions (see also [KP22] when Brownian motion is replaced by α -stable processes). For the second order system (1.8), to the best of our knowledge, there is no such kind of result. Finally, we also mention that when $K \equiv 0$, the strong and weak well-posedness of SDE (1.8) with Hölder drift W was studied in [Ch17] [WZ16] and [Zh18].

1.4. Notations and conventions. Throughout this paper, we use C or c with or without subscripts to denote an unrelated constant, whose value may change in different places. We also use $:=$ as a way of definition. By $A \lesssim_C B$ and $A \asymp_C B$ or simply $A \lesssim B$ and $A \asymp B$, we mean that for some unimportant constant $C \geq 1$,

$$A \leq CB, \quad C^{-1}B \leq A \leq CB.$$

For convenience, we collect some commonly used notations and definitions below.

$\mathbf{B}_{p,q}^{s,a}(\rho)$: weighted Besov space (Def. 2.3)	$\mathbf{B}_{p,q}^{s,a} := \mathbf{B}_{p,q}^{s,a}(1)$
$\mathbf{C}_a^s(\rho) := \mathbf{B}_{\infty,\infty}^{s,a}(\rho)$, $\mathbb{C}_{T,a}^s(\rho) := L^\infty([0, T]; \mathbf{C}_a^s(\rho))$	$\mathbf{C}_a^s := \mathbf{C}_a^s(1)$
$\mathbb{S}_{T,a}^\alpha(\rho)$: Kinetic Hölder space (3.19)	$\mathbb{S}_{T,a}^\alpha := \mathbb{S}_{T,a}^\alpha(1)$
$\mathbb{B}_T^\alpha(\rho)$: Space of renormalized pair (Def. 3.15)	$\mathbb{B}_T^\alpha := \mathbb{B}_T^\alpha(1)$
$f \prec g, f \succ g, f \circ g$: Paraproduct (Sec. 2.2)	$f \succcurlyeq g := f \succ g + f \circ g$
$\text{com}(f, g, h) := (f \prec g) \circ h - f(g \circ h)$ (Sec. 2.2)	$\mathcal{L}_\lambda := \partial_t - v \cdot \nabla_x - \Delta_v + \lambda$
$\Gamma_t z := (x + tv, v)$, $\Gamma_t f(z) := f(\Gamma_t z)$	$\mathcal{I}_\lambda := \mathcal{L}_\lambda^{-1}$
$P_t f = \Gamma_t p_t * \Gamma_t f = \Gamma_t(p_t * f)$: Kinetic semigroup	$B_r^a := \{x : x _a \leq r\}$
$\varrho(x, v) := ((1 + x ^2)^{1/3} + 1 + v ^2)^{-1/2}$	$\mathcal{P}_w := \{\varrho^\kappa, \kappa \in \mathbb{R}\}$
Commutator: $[\mathcal{A}_1, \mathcal{A}_2]f := \mathcal{A}_1(\mathcal{A}_2 f) - \mathcal{A}_2(\mathcal{A}_1 f)$	$\mathbb{N}_0 := \mathbb{N} \cup \{0\}$
$\delta_h f(x) := f(x + h) - f(x)$	$\delta_h^{(k)} := \delta_h \delta_h^{(k-1)}$

2. Preliminaries. In this section we introduce the basic notations and recall various preliminary results concerning weighted anisotropic Besov spaces (see [Di96], [Tri06]). Since the precise results that we need are difficult to locate in the literature, and for the readers' convenience, we give details of the proof in Subsection 2.1. In Subsection 2.2 we present paraproduct calculus on the anisotropic Besov spaces which follows in the same way as the classical argument.

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

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

Fix $n \in \mathbb{N}$. Let $m = (m_1, \dots, m_n) \in \mathbb{N}^n$ with $m_1 + \dots + m_n = N$ and $a = (a_1, \dots, a_n) \in [1, \infty)^n$ be also fixed. We introduce the following distance in \mathbb{R}^N by

$$|x - y|_a := \sum_{i=1}^n |x_i - y_i|^{1/a_i}, \quad x_i, y_i \in \mathbb{R}^{m_i},$$

where $|\cdot|$ denotes the Euclidean norm in \mathbb{R}^{m_i} . For $x = (x_1, \dots, x_n)$, $t > 0$ and $s \in \mathbb{R}$, we denote

$$(2.1) \quad t^{sa}x := (t^{sa_1}x_1, \dots, t^{sa_n}x_n) \in \mathbb{R}^N, \quad B_t^a := \left\{x \in \mathbb{R}^N : |x|_a \leq t\right\}.$$

Clearly we have

$$(2.2) \quad |t^a x|_a = t|x|_a, \quad t \geq 0.$$

2.1. Weighted anisotropic Besov spaces. In this subsection we first give the definition of weighted Besov spaces and then deduce several important properties of these spaces. Next we study how they relate to each other via continuous embeddings and interpolations. Besides, a characterization for the weighted Besov spaces is given for later use. Finally we give an interpolation result.

2.1.1. Definition. To introduce the anisotropic Besov space, we first give the weighted version of dyadic partition of unity. Let ϕ_{-1}^a be a symmetric nonnegative C^∞ -function on \mathbb{R}^N with

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

For $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^{m_1} \times \dots \times \mathbb{R}^{m_n}$ and $j \geq 0$, we define

$$(2.3) \quad \phi_j^a(\xi) := \phi_{-1}^a(2^{-(j+1)}\xi) - \phi_{-1}^a(2^{-aj}\xi).$$

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

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

DEFINITION 2.1. For given $j \geq -1$, the block operator \mathcal{R}_j^a is defined on $\mathcal{S}'(\mathbb{R}^N)$ by

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

with the convention $\mathcal{R}_j^a \equiv 0$ for $j \leq -2$. In particular, for $j \geq 0$,

$$(2.4) \quad \mathcal{R}_j^a f(x) = 2^{a \cdot m_j} \int_{\mathbb{R}^N} \check{\phi}_0^a(2^{aj}y) f(x-y) dy,$$

where $a \cdot m = a_1 m_1 + \dots + a_n m_n$.

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

$$(2.5) \quad \mathcal{R}_j^a = \mathcal{R}_j^a \tilde{\mathcal{R}}_j^a, \quad \text{where } \tilde{\mathcal{R}}_j^a := \mathcal{R}_{j-1}^a + \mathcal{R}_j^a + \mathcal{R}_{j+1}^a,$$

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

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

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

$$(2.6) \quad \widetilde{\mathcal{R}}_j^a f(x) = 2^{a \cdot mj} \int_{\mathbb{R}^N} \widetilde{\phi}_0^a(2^{aj}y) f(x-y) dy, \quad j \geq 1,$$

where

$$\widetilde{\phi}_0^a(\xi) := 2^{a \cdot m} \phi_0(2^a \xi) + \phi_0(\xi) + 2^{-a \cdot m} \phi_0(2^{-a} \xi).$$

The cut-off low frequency operator S_k is defined by

$$(2.7) \quad S_k f := \sum_{j=-1}^{k-1} \mathcal{R}_j^a f \rightarrow f, \quad k \rightarrow \infty.$$

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

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

The Bony decomposition of fg is formally given by (cf. [BCD11])

$$(2.8) \quad fg = f \prec g + f \circ g + g \prec f.$$

The key point of Bony's decomposition is

$$(2.9) \quad \mathcal{R}_j^a (S_{k-1} f \mathcal{R}_k^a g) = 0 \quad \text{for } |k-j| > 3.$$

Indeed, by Fourier's transform, we have

$$(\mathcal{R}_j^a (S_{k-1} f \mathcal{R}_k^a g))^\wedge = \phi_j^a \cdot \sum_{i=-1}^{k-2} (\phi_i^a \hat{f}) * (\phi_k^a \hat{g}).$$

Since the support of $\sum_{i=-1}^{k-2} (\phi_i^a \hat{f}) * (\phi_k^a \hat{g})$ is contained in $B_{2^{k+1}}^a \setminus B_{2^k/6}^a$, we have

$$\phi_j^a \cdot \left(\sum_{i=-1}^{k-2} (\phi_i^a \hat{f}) * (\phi_k^a \hat{g}) \right) = 0, \quad |k-j| > 3,$$

which in turn implies (2.9).

Now we recall the following definition about the admissible weights from [Tri06].

DEFINITION 2.2. A C^∞ -smooth function $\rho : \mathbb{R}^N \rightarrow (0, \infty)$ is called an admissible weight if for each $j \in \mathbb{N}$, there is a constant $C_j > 0$ such that

$$(2.10) \quad |\nabla^j \rho(x)| \leq C_j \rho(x), \quad \forall x \in \mathbb{R}^N,$$

and for some $C, \kappa > 0$,

$$(2.11) \quad \rho(x) \leq C \rho(y) (1 + |x - y|_a^\kappa), \quad \forall x, y \in \mathbb{R}^N.$$

The set of all the admissible weights is denoted by \mathcal{W} .

For $\rho \in \mathcal{W}$ and $p \in [1, \infty]$, we define

$$\|f\|_{L^p(\rho)} := \|\rho f\|_p := \left(\int_{\mathbb{R}^N} |\rho(x) f(x)|^p dx \right)^{1/p}.$$

Let ρ_1, ρ_2, ρ_3 be three weight functions. Suppose that for some $C_1 > 0$,

$$\rho_1(x) \leq C_1 \rho_2(y) \rho_3(x - y), \quad \forall x, y \in \mathbb{R}^N.$$

By the classical Young inequality, we have the following weighted version

$$(2.12) \quad \|f * g\|_{L^q(\rho_1)} \leq C_1 C_2 \|f\|_{L^r(\rho_2)} \|g\|_{L^p(\rho_3)},$$

where $r, p, q \in [1, \infty]$ satisfy $1/q + 1 = 1/p + 1/r$ and $C_2 = C_2(r, p, q) > 0$.

Now we introduce the following weighted anisotropic Besov spaces (see [Di96]).

DEFINITION 2.3. Let $\rho \in \mathcal{W}$, $p, q \in [1, \infty]$ and $s \in \mathbb{R}$. The weighted anisotropic Besov space $\mathbf{B}_{p,q}^{s,a}(\rho)$ is defined by

$$\mathbf{B}_{p,q}^{s,a}(\rho) := \left\{ f \in \mathcal{S}'(\mathbb{R}^N) : \|f\|_{\mathbf{B}_{p,q}^{s,a}(\rho)} := \left(\sum_{j \geq -1} 2^{sjq} \|\mathcal{R}_j^a f\|_{L^p(\rho)}^q \right)^{1/q} < \infty \right\}.$$

For simplicity of notation, we write

$$\mathbf{C}_a^s(\rho) := \mathbf{B}_{\infty,\infty}^{s,a}(\rho), \quad \mathbf{C}_a^s := \mathbf{C}_a^s(1), \quad \mathbf{B}_{p,q}^{s,a} := \mathbf{B}_{p,q}^{s,a}(1),$$

and when $a = (1, 1, \dots, 1)$ we shall drop the index a in above notations.

2.1.2. Embeddings and interpolation. We first prove the following inequality of Bernstein's type.

LEMMA 2.4. Let $\rho \in \mathcal{W}$ be an admissible weight.

- (i) For any $k \in \mathbb{N}_0$, $1 \leq p \leq q \leq \infty$ and $i = 1, 2, \dots, n$, there is a constant $C = C(\rho, m, p, q, a, k, i) > 0$ such that for all $j \geq -1$,

$$(2.13) \quad \|\nabla_{x_i}^k \mathcal{R}_j^a f\|_{L^q(\rho)} \lesssim_C 2^{j(a_i k + a \cdot m(\frac{1}{p} - \frac{1}{q}))} \|\mathcal{R}_j^a f\|_{L^p(\rho)},$$

where $\nabla_{x_i}^k$ denotes the k -order gradient with respect to x_i , and

$$(2.14) \quad \|\mathcal{R}_j^a f\|_{L^p(\rho)} \lesssim_C \|f\|_{L^p(\rho)}.$$

- (ii) For any $s \in \mathbb{R}$ and $p \in [1, \infty]$, there is a constant $C = C(\rho, m, p, a) > 0$ such that for all $j \geq -1$,

$$(2.15) \quad \|J_s \mathcal{R}_j^a f\|_{L^p(\rho)} \asymp_C 2^{sj} \|\mathcal{R}_j^a f\|_{L^p(\rho)},$$

where $\widehat{J_s f}(\xi) := \left(\sum_{i=1}^n (1 + |\xi_i|^2)^{1/(2a_i)} \right)^s \widehat{f}(\xi)$.

PROOF. We only prove (2.13) and (2.15) for $j \geq 1$. For $j = -1, 0$, they follow directly from definition and $\phi_{-1}^a, \phi_0^a \in \mathcal{S}(\mathbb{R}^N)$.

- (i) By (2.5), (2.11) and (2.12), we have

$$\begin{aligned} \|\nabla_{x_i}^k \mathcal{R}_j^a f\|_{L^q(\rho)} &= \|\nabla_{x_i}^k \widetilde{\mathcal{R}}_j^a \mathcal{R}_j^a f\|_{L^q(\rho)} = \|(\nabla_{x_i}^k \widetilde{\phi}_j^a) * \mathcal{R}_j^a f\|_{L^q(\rho)} \\ &\lesssim \|(1 + |\cdot|^\kappa)_a \nabla_{x_i}^k \widetilde{\phi}_j^a\|_{L^r} \|\mathcal{R}_j^a f\|_{L^p(\rho)}, \end{aligned}$$

where $1/p + 1/r = 1 + 1/q$, κ is from (2.11) and

$$\widetilde{\phi}_j^a := \phi_{j-1}^a + \phi_j^a + \phi_{j+1}^a.$$

Since $\kappa \geq 0$, by (2.6) we have

$$\begin{aligned} \|(1 + |\cdot|_a^\kappa) \nabla_{x_i}^k \widetilde{\phi}_j^a\|_{L^r} &\leq 2^{a_i k j} 2^{(a \cdot m) j (1 - \frac{1}{r})} \left(\int_{\mathbb{R}^N} |\nabla_{x_i}^k \widetilde{\phi}_0^a(x)|^r (1 + |2^{-aj} x|_a^\kappa)^r dx \right)^{1/r} \\ &\leq 2^{j(a_i k + a \cdot m(\frac{1}{p} - \frac{1}{q}))} \left(\int_{\mathbb{R}^N} |\nabla_{x_i}^k \widetilde{\phi}_0^a(x)|^r (1 + |x|_a^\kappa)^r dx \right)^{1/r}. \end{aligned}$$

Thus we get (2.13). For (2.14), it is similar.

(ii) By (2.5), (2.11) and (2.12), we similarly have

$$\|J_s \mathcal{R}_j^a f\|_{L^p(\rho)} \lesssim \|(1 + |\cdot|_a^\kappa) J_s \widetilde{\phi}_j^a\|_{L^1} \|\mathcal{R}_j^a f\|_{L^p(\rho)}.$$

Note that by definition and the change of variable,

$$(J_s \widetilde{\phi}_j^a)^\wedge(\xi) = \left(\sum_{i=1}^n (1 + |\xi_i|^2)^{1/(2a_i)} \right)^s \widetilde{\phi}_j^a(\xi) = 2^{sj} F_{s,j}(2^{-aj} \xi),$$

where

$$F_{s,j}(\xi) := \left(\sum_{i=1}^n (2^{-2a_i j} + |\xi_i|^2)^{1/(2a_i)} \right)^s \widetilde{\phi}_0^a(\xi).$$

Since $\text{supp}(\phi_0^a) \subset B_2^a \setminus B_{1/2}^a$, we have for any $k \in \mathbb{N}_0$ and $i = 1, \dots, n$,

$$\sup_{j \geq 1} \int_{\mathbb{R}^N} |\nabla_{\xi_i}^k F_{s,j}(\xi)| d\xi < \infty,$$

which in turn implies that

$$\sup_{j \geq 1} \sup_{x \in \mathbb{R}^N} (1 + |x_i|^k) |\check{F}_{s,j}(x)| < \infty.$$

Hence,

$$\begin{aligned} \|J_s \widetilde{\phi}_j^a (1 + |\cdot|_a^\kappa)\|_{L^1} &= 2^{sj} 2^{a \cdot m j} \int_{\mathbb{R}^N} |\check{F}_{s,j}(2^{aj} x)| (1 + |x|_a^\kappa) dx \\ &= 2^{sj} \int_{\mathbb{R}^N} |\check{F}_{s,j}(x)| (1 + |2^{-aj} x|_a^\kappa) dx \\ &\leq 2^{sj} \int_{\mathbb{R}^N} |\check{F}_{s,j}(x)| (1 + |x|_a^\kappa) dx \lesssim 2^{sj}. \end{aligned}$$

Thus, for $j \geq 1$,

$$\|J_s \mathcal{R}_j^a f\|_{L^p(\rho)} \lesssim_C 2^{sj} \|\mathcal{R}_j^a f\|_{L^p(\rho)}.$$

Since $J_s J_{-s} = \text{Id}$, we also have another side inequality. \square

REMARK 2.5. By definition and (2.15), one sees that for any $p, q \in [1, \infty]$ and $s, s' \in \mathbb{R}$, J_s is an isomorphism between $\mathbf{B}_{p,q}^{s'+s,a}$ and $\mathbf{B}_{p,q}^{s',a}$, i.e.,

$$(2.16) \quad J_s \mathbf{B}_{p,q}^{s'+s,a} = \mathbf{B}_{p,q}^{s',a}.$$

As an easy consequence of Bernstein's inequality, we have the following embedding theorem of weighted anisotropic Besov spaces.

THEOREM 2.6. *Let $\rho \in \mathcal{W}$, $s_1, s_2 \in \mathbb{R}$, $1 \leq r \leq p \leq \infty$ be such that*

$$s_2 = s_1 + (a \cdot m)\left(\frac{1}{r} - \frac{1}{p}\right).$$

For any $q \in [1, \infty]$, there is a constant $C = C(\rho, m, a, p, q, r, s_1, s_2) > 0$ such that

$$(2.17) \quad \|f\|_{\mathbf{B}_{p,q}^{s_1,a}(\rho)} \leq C \|f\|_{\mathbf{B}_{r,q}^{s_2,a}(\rho)}.$$

Moreover, for any $1 \leq q_1 \leq q_2 \leq \infty$ and $\rho_2 \leq \rho_1$,

$$(2.18) \quad \|f\|_{\mathbf{B}_{p,q_2}^{s,a}(\rho_2)} \leq \|f\|_{\mathbf{B}_{p,q_1}^{s,a}(\rho_1)},$$

and for $\theta \in [0, 1]$ and $p_1, p_2 \in [1, \infty]$, $\rho_1, \rho_2 \in \mathcal{W}$, $s, s_1, s_2 \in \mathbb{R}$ with

$$\frac{\theta}{p_1} + \frac{1-\theta}{p_2} = \frac{1}{p}, \quad \theta s_1 + (1-\theta)s_2 = s,$$

the following interpolation inequality holds,

$$(2.19) \quad \|f\|_{\mathbf{B}_{p,q}^{s,a}(\rho_1^\theta \rho_2^{1-\theta})} \leq \|f\|_{\mathbf{B}_{p_1,q}^{s_1,a}(\rho_1)}^\theta \|f\|_{\mathbf{B}_{p_2,q}^{s_2,a}(\rho_2)}^{1-\theta}.$$

PROOF. (2.17) is straightforward by Lemma 2.4 with $k = 0$. (2.18) and (2.19) are direct consequences of the definition and Hölder's inequality. \square

2.1.3. *Characterization.* Now we give a characterization of $\mathbf{B}_{p,q}^{s,a}(\rho)$. To this end, we introduce the following notations. For $f : \mathbb{R}^N \rightarrow \mathbb{R}$ and $h \in \mathbb{R}^N$, the first order difference operator is defined by

$$\delta_h f(x) := f(x+h) - f(x),$$

and for $M \in \mathbb{N}$, the M -order difference operator is defined recursively by

$$\delta_h^{(M+1)} f(x) := \delta_h \delta_h^{(M)} f(x).$$

By induction, it is easy to see that

$$(2.20) \quad \delta_h^{(M)} f(x) = \sum_{k=0}^M (-1)^{M-k} \binom{M}{k} f(x+kh), \quad h \in \mathbb{R}^N,$$

where $\binom{M}{k}$ is the binomial coefficient. The following characterization of $\mathbf{B}_{p,q}^{s,a}(\rho)$ is probably well-known to experts. Since we cannot find them in the literature, for the readers' convenience, we provide detailed proofs in Appendix A.

THEOREM 2.7. *Let $\rho \in \mathcal{W}$. For any $s \in (0, \infty)$ and $p, q \in [1, \infty]$, there exists a constant $C = C(\rho, a, m, p, q, s) \geq 1$ such that for all $f \in \mathbf{B}_{p,q}^{s,a}(\rho)$,*

$$(2.21) \quad \|f\|_{\mathbf{B}_{p,q}^{s,a}(\rho)} \asymp_C \|f\|_{\tilde{\mathbf{B}}_{p,q}^{s,a}(\rho)} \asymp_C \|\rho f\|_{\tilde{\mathbf{B}}_{p,q}^{s,a}},$$

where

$$\|f\|_{\tilde{\mathbf{B}}_{p,q}^{s,a}(\rho)} := \left(\int_{|h|_a \leq 1} \left(\frac{\|\delta_h^{([s]+1)} f\|_{L^p(\rho)}}{|h|_a^s} \right)^q \frac{dh}{|h|_a^{a \cdot m}} \right)^{1/q} + \|f\|_{L^p(\rho)},$$

where $[s]$ denotes the integer part of s . Moreover, for any $s \in \mathbb{R}$ and $p, q \in [1, \infty]$,

$$(2.22) \quad \|f\|_{\mathbf{B}_{p,q}^{s,a}(\rho)} \asymp_C \|\rho f\|_{\mathbf{B}_{p,q}^{s,a}}.$$

REMARK 2.8. For $\rho \equiv 1$, the characterization of (2.21) is proven in [ZZ21, Lemma 2.8]. In particular, for $s > 0$, since $\mathbf{C}_a^s(\rho) = \mathbf{B}_{\infty,\infty}^{s,a}(\rho)$, by (2.21) we have

$$(2.23) \quad \|f\|_{\mathbf{C}_a^s(\rho)} \prec_C \|f\|_{\mathbf{C}_{x_1}^{s/a_1}(\rho)} + \cdots + \|f\|_{\mathbf{C}_{x_n}^{s/a_n}(\rho)},$$

where for $i = 1, \dots, n$,

$$\|f\|_{\mathbf{C}_{x_i}^s(\rho)} := \|f\|_{L^\infty(\rho)} + \sup_{|h_i| \leq 1} \frac{\|\delta_{h_i}^{([s]+1)} f\|_{L^\infty(\rho)}}{|h_i|^{s/a_i}},$$

and

$$\delta_{h_i} f(x) := f(\cdots, x_{i-1}, x_i + h_i, x_{i+1}, \cdots) - f(\cdots, x_{i-1}, x_i, x_{i+1}, \cdots).$$

As a corollary, we have the following result.

COROLLARY 2.9. Let $\rho \in \mathcal{W}$. For any $\alpha \in \mathbb{R}$, $s > 0$ and $p, q \in [1, \infty]$, there is a constant $C = C(\rho, a, m, p, q, \alpha, s) > 0$ such that for all $f \in \mathbf{B}_{p,q}^{\alpha+s,a}(\rho)$ and $h \in \mathbb{R}^N$,

$$(2.24) \quad \|\delta_h^{([s]+1)} f\|_{\mathbf{B}_{p,q}^{\alpha,a}(\rho)} \lesssim_C |h|_a^s (1 + |h|_a^\kappa) \|f\|_{\mathbf{B}_{p,q}^{\alpha+s,a}(\rho)},$$

where $\kappa \geq 0$ is from (2.11).

PROOF. By (2.21), for $|h|_a \leq 1$, we have

$$\|\delta_h^{([s]+1)} f\|_{L^p(\rho)} \lesssim |h|_a^s \|f\|_{\mathbf{B}_{p,\infty}^{s,a}(\rho)}.$$

For $|h|_a \geq 1$, by (2.20) and (2.11) we have

$$\|\delta_h^{([s]+1)} f\|_{L^p(\rho)} \lesssim (1 + |h|_a^\kappa) \|f\|_{L^p(\rho)} \lesssim (1 + |h|_a^\kappa) \|f\|_{\mathbf{B}_{p,\infty}^{s,a}(\rho)}.$$

Therefore,

$$(2.25) \quad \|\delta_h^{([s]+1)} f\|_{L^p(\rho)} \lesssim |h|_a^s (1 + |h|_a^\kappa) \|f\|_{\mathbf{B}_{p,\infty}^{s,a}(\rho)}.$$

Noting that

$$\|\mathcal{R}_j^a f\|_{\mathbf{B}_{p,\infty}^{s,a}(\rho)} = \sup_{k \geq -1} 2^{ks} \|\mathcal{R}_k^a \mathcal{R}_j^a f\|_{L^p(\rho)} \lesssim 2^{js} \|\mathcal{R}_j^a f\|_{L^p(\rho)},$$

by (2.25), we have

$$\begin{aligned} \|\delta_h^{([s]+1)} f\|_{\mathbf{B}_{p,q}^{\alpha,a}(\rho)}^q &= \sum_{j \geq -1} 2^{\alpha j q} \|\delta_h^{([s]+1)} \mathcal{R}_j^a f\|_{L^p(\rho)}^q \\ &\lesssim |h|_a^{qs} (1 + |h|_a^\kappa)^q \sum_{j \geq -1} 2^{\alpha j q} \|\mathcal{R}_j^a f\|_{\mathbf{B}_{p,\infty}^{s,a}(\rho)}^q \\ &\lesssim |h|_a^{qs} (1 + |h|_a^\kappa)^q \sum_{j \geq -1} 2^{(\alpha+s)jq} \|\mathcal{R}_j^a f\|_{L^p(\rho)}^q \\ &= |h|_a^{qs} (1 + |h|_a^\kappa)^q \|f\|_{\mathbf{B}_{p,q}^{\alpha+s,a}(\rho)}^q. \end{aligned}$$

The proof is complete. □

We finish this subsection with an interpolation lemma for later use.

LEMMA 2.10. *Let $\{T_j\}_{j=-1}^\infty$ be a family of linear operators from $\mathcal{S}'(\mathbb{R}^N)$ to some Banach space \mathbb{X} . Assume that for some $\beta_0 < \beta_1$ and any $j \geq -1$, there are constants $C_{ij} > 0, i = 0, 1$ such that*

$$\|T_j f\|_{\mathbb{X}} \leq C_{ij} 2^{-j\beta_i} \|f\|_{\mathbf{C}_a^{\beta_i}}, \quad i = 0, 1.$$

Then for any $\beta \in (\beta_0, \beta_1)$, there is a constant $C = C(a, m, \beta, \beta_0, \beta_1) > 0$ such that

$$\|T_j f\|_{\mathbb{X}} \leq C(C_{0j} + C_{1j}) 2^{-j\beta} \|f\|_{\mathbf{C}_a^\beta}, \quad j \geq -1.$$

PROOF. Since for any $k \geq -1$,

$$\|\mathcal{R}_k^a f\|_{\mathbf{C}_a^{\beta_i}} \lesssim 2^{(\beta_i - \beta)k} \|f\|_{\mathbf{C}_a^\beta}, \quad i = 0, 1,$$

we have by the assumptions,

$$\begin{aligned} \|T_j f\|_{\mathbb{X}} &\leq \sum_{k \geq -1} \|T_j \mathcal{R}_k^a f\|_{\mathbb{X}} \leq C_{0j} 2^{-j\beta_0} \sum_{k > j} \|\mathcal{R}_k^a f\|_{\mathbf{C}_a^{\beta_0}} + C_{1j} 2^{-j\beta_1} \sum_{k \leq j} \|\mathcal{R}_k^a f\|_{\mathbf{C}_a^{\beta_1}} \\ &\lesssim \left(C_{0j} 2^{-j\beta_0} \sum_{k > j} 2^{(\beta_0 - \beta)k} + C_{1j} 2^{-j\beta_1} \sum_{k \leq j} 2^{(\beta_1 - \beta)k} \right) \|f\|_{\mathbf{C}_a^\beta} \\ &\lesssim (C_{0j} + C_{1j}) 2^{-j\beta} \|f\|_{\mathbf{C}_a^\beta}. \end{aligned}$$

The proof is complete. \square

2.2. Paraproduct calculus. In this subsection we recall some basic ingredients in the paracontrolled calculus developed by Bony [Bon81] and [GIP15]. The first important fact is that the product fg of two distributions $f \in \mathbf{C}_a^\alpha$ and $g \in \mathbf{C}_a^\beta$ is well defined if and only if $\alpha + \beta > 0$ as given in the following lemma.

LEMMA 2.11. *Let $\rho_1, \rho_2 \in \mathcal{W}$. We have for any $\beta \in \mathbb{R}$,*

$$(2.26) \quad \|f \prec g\|_{\mathbf{C}_a^\beta(\rho_1 \rho_2)} \lesssim_C \|f\|_{L^\infty(\rho_1)} \|g\|_{\mathbf{C}_a^\beta(\rho_2)},$$

and for any $\alpha < 0$ and $\beta \in \mathbb{R}$,

$$(2.27) \quad \|f \prec g\|_{\mathbf{C}_a^{\alpha+\beta}(\rho_1 \rho_2)} \lesssim_C \|f\|_{\mathbf{C}_a^\alpha(\rho_1)} \|g\|_{\mathbf{C}_a^\beta(\rho_2)}.$$

Moreover, for any $\alpha, \beta \in \mathbb{R}$ with $\alpha + \beta > 0$,

$$(2.28) \quad \|f \circ g\|_{\mathbf{C}_a^{\alpha+\beta}(\rho_1 \rho_2)} \lesssim_C \|f\|_{\mathbf{C}_a^\alpha(\rho_1)} \|g\|_{\mathbf{C}_a^\beta(\rho_2)}.$$

In particular, for any $\alpha, \beta \in \mathbb{R}$ with $\alpha + \beta > 0$,

$$(2.29) \quad \|f \cdot g\|_{\mathbf{C}_a^{\alpha \wedge \beta}(\rho_1 \rho_2)} \lesssim_C \|f\|_{\mathbf{C}_a^\alpha(\rho_1)} \|g\|_{\mathbf{C}_a^\beta(\rho_2)}.$$

PROOF. Totally the same as [GIP15, Lemma 2.1] and [GH19, Lemma 2.14]. \square

For two abstract operators $\mathcal{A}_1, \mathcal{A}_2$ acting on functions, we shall use the following notation to denote the commutator between \mathcal{A}_1 and \mathcal{A}_2 :

$$[\mathcal{A}_1, \mathcal{A}_2]f := \mathcal{A}_1(\mathcal{A}_2 f) - \mathcal{A}_2(\mathcal{A}_1 f).$$

We have the following simple commutator estimate (see [GIP15, Lemma 2.2]).

LEMMA 2.12. *For any $\rho_1, \rho_2 \in \mathcal{W}$ and $\alpha \in (0, 1)$, there is a constant $C = C(\rho_1, \rho_2, \alpha, a, m) > 0$ such that for all $j \geq -1$,*

$$(2.30) \quad \|[\mathcal{R}_j^a, f]g\|_{L^\infty(\rho_1 \rho_2)} \lesssim_C 2^{-\alpha j} \|f\|_{\mathbf{C}_a^\alpha(\rho_1)} \|g\|_{L^\infty(\rho_2)}.$$

The following lemma is a weighted anisotropic version of Lemma 2.4 in [GIP15].

LEMMA 2.13. *Let $\rho_1, \rho_2, \rho_3 \in \mathcal{W}$. For any $\alpha \in (0, 1)$ and $\beta, \gamma \in \mathbb{R}$ with $\alpha + \beta + \gamma > 0$ and $\beta + \gamma < 0$, there exists a trilinear bounded operator com on $\mathbf{C}_a^\alpha(\rho_1) \times \mathbf{C}_a^\beta(\rho_2) \times \mathbf{C}_a^\gamma(\rho_3)$ such that*

$$(2.31) \quad \|\text{com}(f, g, h)\|_{\mathbf{C}_a^{\alpha+\beta+\gamma}(\rho_1 \rho_2 \rho_3)} \lesssim_C \|f\|_{\mathbf{C}_a^\alpha(\rho_1)} \|g\|_{\mathbf{C}_a^\beta(\rho_2)} \|h\|_{\mathbf{C}_a^\gamma(\rho_3)},$$

where

$$\text{com}(f, g, h) = (f \prec g) \circ h - f(g \circ h).$$

In addition, if $\beta < 0$ and $\alpha + \beta > 0$, then $[h \circ, f]g$ can be extended to be a bounded linear operator on $\mathbf{C}_a^\alpha(\rho_1) \times \mathbf{C}_a^\beta(\rho_2) \times \mathbf{C}_a^\gamma(\rho_3)$ with

$$(2.32) \quad \|[h \circ, f]g\|_{\mathbf{C}_a^{\alpha+\beta+\gamma}(\rho_1 \rho_2 \rho_3)} \lesssim_C \|f\|_{\mathbf{C}_a^\alpha(\rho_1)} \|g\|_{\mathbf{C}_a^\beta(\rho_2)} \|h\|_{\mathbf{C}_a^\gamma(\rho_3)}.$$

PROOF. By Lemmas 2.11 and 2.12, estimate of (2.31) is completely the same as in [GIP15]. For (2.32), note that

$$[h \circ, f]g = h \circ (gf) - f(h \circ g) = h \circ (f \succ g) + \text{com}(f, g, h).$$

By Lemma 2.11, we have

$$\begin{aligned} \|h \circ (f \succ g)\|_{\mathbf{C}_a^{\alpha+\beta+\gamma}(\rho_1 \rho_2 \rho_3)} &\lesssim \|h\|_{\mathbf{C}_a^\gamma(\rho_3)} \|f \succ g\|_{\mathbf{C}_a^{\alpha+\beta}(\rho_1 \rho_2)} \\ &\lesssim \|h\|_{\mathbf{C}_a^\gamma(\rho_3)} \|f\|_{\mathbf{C}_a^\alpha(\rho_1)} \|g\|_{\mathbf{C}_a^\beta(\rho_2)}, \end{aligned}$$

which together with (2.31) yields (2.32). \square

3. Kinetic semigroups and commutator estimates. In this section we introduce basic estimates about the kinetic semigroup. Compared to the heat semigroup, due to the presence of the transport term, the kinetic semigroup does not commute with block operator \mathcal{R}_j^a (see (3.13) below), which brings some new features. In Subsection 3.2, we introduce a kinetic Hölder space which admits velocity transport in time direction, as well as a localization characterization for weighted Hölder space proved in [ZZZ22], which will be used to obtain the well-posedness of linear equation (1.13). In Subsection 3.3, we establish the Schauder estimate in kinetic Hölder spaces. In Subsection 3.4, we prove a commutator estimate for the kinetic semigroup which is essential to apply the paracontrolled calculus for the kinetic equations. Finally, in Subsection 3.5 we introduce the renormalized pairs used in the definition of paracontrolled solutions.

In the remainder of this paper, we consider the following case of the weighted anisotropic Besov spaces:

$$N = 2d, \quad d \in \mathbb{N}, \quad n = 2, \quad m_1 = m_2 = d, \quad a = (3, 1).$$

For $t > 0$, let P_t be the kinetic semigroup defined by

$$(3.1) \quad P_t f(z) := \Gamma_t p_t * \Gamma_t f(z) = \Gamma_t(p_t * f)(z), \quad z = (x, v) \in \mathbb{R}^{2d},$$

where for $t \in \mathbb{R}$,

$$\Gamma_t f(z) := f(\Gamma_t z), \quad \Gamma_t z := (x + tv, v),$$

and

$$(3.2) \quad p_t(z) = p_t(x, v) = \left(\frac{4\pi t^4}{3} \right)^{-d/2} \exp \left(- \frac{3|x|^2 + |3x - 2tv|^2}{4t^3} \right)$$

is the density of the following process

$$Z_t := (X_t, V_t) = \left(\sqrt{2} \int_0^t B_s ds, \sqrt{2} B_t \right),$$

where B_t is a d -dimensional standard Brownian motion. The reason of choosing multi-scale parameter $a = (3, 1)$ is the following scaling property (see also (1.4)):

$$(X_{\lambda t}, V_{\lambda t}) \stackrel{(d)}{=} (\lambda^{\frac{3}{2}} X_t, \lambda^{\frac{1}{2}} V_t), \quad \lambda > 0.$$

Note that

$$(3.3) \quad \Gamma_t \Gamma_s = \Gamma_{t+s}, \quad p_t(z) = t^{-a \cdot m/2} p_1(t^{-a/2} z),$$

and for $\varphi \in C_b^\infty(\mathbb{R}^{2d})$,

$$\partial_t P_t \varphi = (\Delta_v + v \cdot \nabla_x) P_t \varphi.$$

Notation: Let \mathcal{P}_w be the set of all polynomial weights with the form:

$$(3.4) \quad \rho(z) = \varrho(z)^\kappa, \quad \kappa \in \mathbb{R},$$

where for $z = (x, v)$,

$$(3.5) \quad \varrho(x, v) := ((1 + |x|^2)^{1/3} + 1 + |v|^2)^{-1/2} \asymp (1 + |z|_a)^{-1}.$$

Clearly, for some $C_0 = C_0(\kappa, d) > 0$,

$$(3.6) \quad \rho(z) \leq C_0 \rho(\bar{z}) (1 + |z - \bar{z}|_a^{|\kappa|}),$$

and for any $j \in \mathbb{N}$ and some $C_j = C_j(\kappa, d) > 0$,

$$(3.7) \quad |\nabla_v^j \rho(z)| \leq C_j \rho(z) \varrho^j(z), \quad |\nabla_x^j \rho(z)| \leq C_j \rho(z) \varrho^{2j}(z),$$

and for any $T > 0$, there is a constant $C_T = C(T, \kappa, d) > 0$ such that

$$(3.8) \quad C_T^{-1} \rho(z) \leq \Gamma_t \rho(z) \leq C_T \rho(z), \quad z \in \mathbb{R}^{2d}, \quad t \in [0, T].$$

Moreover, for $\rho_1, \rho_2 \in \mathcal{P}_w$, we have

$$\rho_1 / \rho_2, \rho_1 \rho_2, \rho_1 \vee \rho_2, \rho_1 \wedge \rho_2 \in \mathcal{P}_w.$$

3.1. Kinetic semigroup estimates. In this subsection, we establish the basic properties of the kinetic semigroup. To this end, we first recall the estimate for the heat kernel of Kolmogorov operator $\Delta_v + v \cdot \nabla_x$ under the action of block operator \mathcal{R}_j^a and a crucial decomposition (3.13) from [HWZ20].

LEMMA 3.1. *For any $\alpha, \beta, \gamma \geq 0$ and $T > 0$, there is a $C = C(T, d, \alpha, \beta, \gamma) > 0$ such that for all $j \geq -1$ and $t \in (0, T]$,*

$$(3.9) \quad \int_{\mathbb{R}^{2d}} |x|^\beta |v|^\gamma |\mathcal{R}_j^a \Gamma_t p_t(x, v)| dx dv \lesssim_C 2^{-(3\beta+\gamma)j} ((t^{1/2} 2^j)^{-\alpha} \wedge 1).$$

In particular, for any $\rho \in \mathcal{P}_w$, $T > 0$ and $\alpha \geq 0$, there is a constant $C = C(T, d, \alpha, \rho) > 0$ such that for all $j \geq -1$ and $t \in (0, T]$,

$$(3.10) \quad \|\mathcal{R}_j^a \Gamma_t p_t\|_{L^1(\rho)} \lesssim_C (t^{1/2} 2^j)^{-\alpha} \wedge 1.$$

PROOF. When $j \geq 0$, by [HWZ20, Lemma 5.1 (5.9)], we have for any $n \in \mathbb{N}_0$,

$$\begin{aligned} \mathcal{J}_j(t) &:= \int_{\mathbb{R}^{2d}} |x|^\beta |v|^\gamma |\mathcal{R}_j^a \Gamma_t p_t(x, v)| dx dv \\ &\lesssim \left(\hbar^{3n} + \hbar^n \right) \left(2^{-(3\beta+\gamma)j} + t^{\frac{3\beta+\gamma}{2}} \right) \\ &= 2^{-(3\beta+\gamma)j} \left(\hbar^{3n} + \hbar^n \right) \left(1 + \hbar^{-(3\beta+\gamma)} \right), \end{aligned}$$

where $\hbar := t^{-\frac{1}{2}} 2^{-j}$. Since $n \in \mathbb{N}_0$ is arbitrary, we clearly have for any $\alpha \geq 0$,

$$\mathcal{J}_j(t) \lesssim_C 2^{-(3\beta+\gamma)j} \hbar^\alpha = 2^{-(3\beta+\gamma)j} (t^{1/2} 2^j)^{-\alpha}.$$

When $j = -1$, we have

$$\mathcal{J}_{-1}(t) \leq C \leq C T^{\alpha/2} t^{-\alpha/2}.$$

Thus we get (3.9) by the arbitrariness of $\alpha \geq 0$. Estimate (3.10) follows directly by (3.9). \square

We recall the following important observation from [HWZ20, Lemma 6.7].

LEMMA 3.2. *For $t \geq 0$ and $j \in \mathbb{N}_0$, define*

$$\Theta_j^t := \left\{ \ell \geq -1 : 2^\ell \leq 2^4(2^j + t2^{3j}), 2^j \leq 2^4(2^\ell + t2^{3\ell}) \right\}.$$

(i) *For any $\ell \notin \Theta_j^t$, it holds that*

$$(3.11) \quad \mathcal{R}_j^a \Gamma_t \mathcal{R}_\ell^a = 0.$$

(ii) *For any $0 \neq \beta \in \mathbb{R}$, there is a constant $C = C(\beta) > 0$ such that*

$$(3.12) \quad \sum_{\ell \in \Theta_j^t} 2^{\beta\ell} \lesssim_C 2^{j\beta} (1 + t2^{2j})^{|\beta|}, \quad j \in \mathbb{N}_0, t \geq 0.$$

REMARK 3.3. By (3.1), one sees that

$$\mathcal{R}_j^a P_t f = (\Gamma_t p_t) * (\mathcal{R}_j^a \Gamma_t f) = \sum_{\ell \geq -1} (\Gamma_t p_t) * (\mathcal{R}_j^a \Gamma_t \mathcal{R}_\ell^a f).$$

In view of (3.11), we have the following decomposition of the kinetic semigroup:

$$(3.13) \quad \mathcal{R}_j^a P_t f = \sum_{\ell \in \Theta_j^t} (\Gamma_t p_t) * (\mathcal{R}_j^a \Gamma_t \mathcal{R}_\ell^a f) = \sum_{\ell \in \Theta_j^t} \mathcal{R}_j^a P_t \mathcal{R}_\ell^a f, \quad j \in \mathbb{N}_0.$$

By (3.13), we can derive the following results that quantify the regularizing and continuity properties of the kinetic semigroup P_t .

LEMMA 3.4. (i) *For any $\rho \in \mathcal{P}_w$, $\alpha \geq 0$, $\beta \in \mathbb{R}$ and $T > 0$, there is a constant $C = C(\rho, T, d, \alpha, \beta) > 0$ such that for all $j \geq -1$, $t \in (0, T]$ and $f \in \mathbf{C}_a^\beta(\rho)$,*

$$(3.14) \quad \|\mathcal{R}_j^a P_t f\|_{L^\infty(\rho)} \lesssim_C 2^{-j\beta} (1 \wedge (t^{\frac{1}{2}} 2^j)^{-\alpha}) \|f\|_{\mathbf{C}_a^\beta(\rho)}.$$

In particular, for any $\alpha \geq 0$,

$$(3.15) \quad \|P_t f\|_{\mathbf{C}_a^{\alpha+\beta}(\rho)} \lesssim_C t^{-\alpha/2} \|f\|_{\mathbf{C}_a^\beta(\rho)}.$$

(ii) For any $\rho \in \mathcal{P}_w$, $k \in \mathbb{N}_0$, $\beta < k$ and $T > 0$, there is a constant $C = C(T, k, \rho, \beta) > 0$ such that for all $t \in (0, T]$ and $f \in \mathbf{C}_a^\beta(\rho)$,

$$(3.16) \quad \|\nabla_v^k P_t f\|_{L^\infty(\rho)} \lesssim_C t^{(\beta-k)/2} \|f\|_{\mathbf{C}_a^\beta(\rho)}.$$

(iii) For any $\rho \in \mathcal{P}_w$, $T > 0$ and $\beta \in (0, 2)$, there is a constant $C = C(\rho, d, \beta, T) > 0$ such that for all $t \in [0, T]$ and $f \in \mathbf{C}_a^\beta(\rho)$,

$$(3.17) \quad \|P_t f - \Gamma_t f\|_{L^\infty(\rho)} \lesssim_C t^{\beta/2} \|f\|_{\mathbf{C}_a^\beta(\rho)}.$$

PROOF. (i) By the interpolation lemma 2.10, we only show (3.14) for $\beta \neq 0$. Let ρ be as in (3.4). For $j \in \mathbb{N}_0$, by (3.13), (3.6) and (2.12), we have

$$(3.18) \quad \begin{aligned} \|\mathcal{R}_j^a P_t f\|_{L^\infty(\rho)} &\leq \sum_{\ell \in \Theta_j^t} \|\mathcal{R}_j^a \Gamma_t p_t * \Gamma_t \mathcal{R}_\ell^a f\|_{L^\infty(\rho)} \\ &\lesssim \|(1 + |\cdot|_a^{|\kappa|}) \mathcal{R}_j^a \Gamma_t p_t\|_{L^1} \sum_{\ell \in \Theta_j^t} \|\Gamma_t \mathcal{R}_\ell^a f\|_{L^\infty(\rho)}. \end{aligned}$$

Moreover, by (3.8), we have

$$\begin{aligned} \sum_{\ell \in \Theta_j^t} \|\Gamma_t \mathcal{R}_\ell^a f\|_{L^\infty(\rho)} &\lesssim \sum_{\ell \in \Theta_j^t} \|\Gamma_t(\rho \mathcal{R}_\ell^a f)\|_{L^\infty} = \sum_{\ell \in \Theta_j^t} \|\rho \mathcal{R}_\ell^a f\|_{L^\infty} \\ &\leq \sum_{\ell \in \Theta_j^t} 2^{-\ell\beta} \|f\|_{\mathbf{C}_a^\beta(\rho)} \stackrel{(3.12)}{\lesssim} 2^{-\beta j} (1 + (t4^j)^{|\beta|}) \|f\|_{\mathbf{C}_a^\beta(\rho)}. \end{aligned}$$

Therefore, by (3.10) and (3.18), for any $l \geq 0$,

$$\|\mathcal{R}_j^a P_t f\|_{L^\infty(\rho)} \lesssim ((t^{1/2} 2^j)^{-2l} \wedge 1) 2^{-\beta j} (1 + (t4^j)^{|\beta|}) \|f\|_{\mathbf{C}_a^\beta(\rho)},$$

which implies (3.14) for $j \in \mathbb{N}_0$ by taking $l = \frac{\alpha}{2} + |\beta|$ and $l = \frac{\alpha}{2}$, respectively. For $j = -1$, it is obvious. Moreover, (3.15) follows directly by (3.14).

(ii) For (3.16), by Lemma 2.4, we have

$$\begin{aligned} \|\nabla_v^k P_t f\|_{L^\infty(\rho)} &\leq \sum_{j=-1}^{\infty} 2^{kj} \|\mathcal{R}_j^a P_t f\|_{L^\infty(\rho)} \leq \sum_{j=-1}^{\infty} 2^{(k-\beta)j} (1 \wedge (t^{\frac{1}{2}} 2^j)^{-2k}) \|f\|_{\mathbf{C}_a^\beta(\rho)} \\ &\lesssim \|f\|_{\mathbf{C}_a^\beta(\rho)} \int_{-\infty}^{\infty} 2^{(k-\beta)s} (1 \wedge (t^{\frac{1}{2}} 2^s)^{-2k}) ds \\ &= \|f\|_{\mathbf{C}_a^\beta(\rho)} t^{(\beta-k)/2} \int_0^{\infty} s^{k-\beta} (1 \wedge s^{-2k}) \frac{\ln 2 ds}{s}, \end{aligned}$$

which gives (3.16).

(iii) Note that

$$P_t f - \Gamma_t f \stackrel{(3.1)}{=} \Gamma_t(p_t * f - f),$$

and by $p_t(z) = p_t(-z)$,

$$p_t * f(z) - f(z) = \frac{1}{2} \int_{\mathbb{R}^{2d}} p_t(\bar{z}) (\delta_{\bar{z}} f(z) + \delta_{-\bar{z}} f(z)) d\bar{z}.$$

By (3.8), (2.25) and (3.6), we have

$$\begin{aligned} \|P_t f - \Gamma_t f\|_{L^\infty(\rho)} &\lesssim \|p_t * f - f\|_{L^\infty(\rho)} \leq \frac{1}{2} \int_{\mathbb{R}^{2d}} p_t(\bar{z}) \|\delta_{\bar{z}} f + \delta_{-\bar{z}} f\|_{L^\infty(\rho)} d\bar{z} \\ &\lesssim \left(\int_{\mathbb{R}^{2d}} p_t(\bar{z}) |\bar{z}|_a^\beta (1 + |\bar{z}|_a^\kappa) d\bar{z} \right) \|f\|_{C_a^\beta(\rho)}, \end{aligned}$$

where $\kappa > 0$ is from (3.6) and we have used that for $\beta \in [1, 2)$,

$$\delta_{\bar{z}} f + \delta_{-\bar{z}} f = \delta_{\bar{z}}^{(2)} f(\cdot - \bar{z}).$$

Thus we obtain (3.17) by (3.3). \square

3.2. Kinetic Hölder spaces and characterization. In this subsection we first give the definition of weighted kinetic Hölder space which is different from the classical one in time direction. Then we give a characterization of the space. By using this we could transform the results for a PDE in un-weighted Besov spaces to a PDE in weighted Besov spaces. See Section 4 for more details. Finally we give an estimate for the derivative of a function in v direction in kinetic Hölder spaces.

For $T > 0$, $\alpha \in \mathbb{R}$ and $\rho \in \mathcal{P}_w$, let $C_{T,a}^\alpha(\rho)$ be the space of all space-time distributions with finite norm

$$\|f\|_{C_{T,a}^\alpha(\rho)} := \sup_{0 \leq t \leq T} \|f(t)\|_{C_a^\alpha(\rho)} < \infty.$$

We introduce the following weighted kinetic Hölder space.

DEFINITION 3.5 (Kinetic Hölder space). Let $\rho \in \mathcal{P}_w$, $\alpha \in (0, 2)$ and $T > 0$. Define

$$(3.19) \quad \mathbb{S}_{T,a}^\alpha(\rho) := \left\{ f : \|f\|_{\mathbb{S}_{T,a}^\alpha(\rho)} := \|f\|_{C_{T,a}^\alpha(\rho)} + \|f\|_{C_{T,\Gamma}^{\alpha/2} L^\infty(\rho)} < \infty \right\},$$

where for $\beta \in (0, 1)$,

$$\|f\|_{C_{T,\Gamma}^\beta L^\infty(\rho)} := \sup_{0 \leq t \leq T} \|f(t)\|_{L^\infty(\rho)} + \sup_{s \neq t \in [0, T]} \frac{\|f(t) - \Gamma_{t-s} f(s)\|_{L^\infty(\rho)}}{|t - s|^\beta}.$$

For $\rho = 1$, we simply write

$$\mathbb{S}_{T,a}^\alpha := \mathbb{S}_{T,a}^\alpha(1), \quad C_{T,\Gamma}^\beta L^\infty := C_{T,\Gamma}^\beta L^\infty(1).$$

REMARK 3.6. (i) In the above definition, the appearance of Γ_t reflects the transport role of $v \cdot \nabla_x$ (see also (3.17) for the same reason). It is noticed that this definition is essentially equivalent to the one introduced in [IS21] by using the language of group.

(ii) Lemma 3.11 below stated the Schauder estimate in kinetic Hölder space. If f is independent of time variable, we can check that the Schauder estimate holds in the classical Hölder space.

Next we show a localization characterization for $\mathbb{S}_{T,a}^\alpha(\rho)$ which will be used in Section 4 to deduce global estimate. Let χ be a nonnegative smooth function with

$$(3.20) \quad \chi(z) = 1, \quad |z|_a \leq 1/8, \quad \chi(z) = 0, \quad |z|_a > 1/4,$$

and for $r > 0$ and $z_0 \in \mathbb{R}^{2d}$,

$$(3.21) \quad \chi_r^{z_0}(z) := \chi\left(\frac{z - z_0}{r^a}\right), \quad \phi_r^{z_0}(z) := \chi_{r(1+|z_0|_a)}^{z_0}(z),$$

where we have used the notation (2.1). The following characterization of weighted Hölder spaces is due to [ZZZZ22, Lemma 3.8].

LEMMA 3.7. *Let $\alpha > 0$ and $r \in (0, 1]$. For any $\rho \in \mathcal{P}_w$, there is a constant $C = C(r, \alpha, d, \rho) > 0$ such that*

$$(3.22) \quad \|\phi_r^{z_0}\|_{\mathbf{C}_a^\alpha(\rho)} \lesssim_C \rho(z_0), \quad z_0 \in \mathbb{R}^{2d},$$

and for any $j \in \mathbb{N}$ and ϱ given in (3.5),

$$(3.23) \quad \|\nabla_v^j \phi_r^{z_0}\|_{\mathbf{C}_a^\alpha(\rho)} + \|v \cdot \nabla_x \phi_r^{z_0}\|_{\mathbf{C}_a^\alpha(\rho)} \lesssim_C (\varrho\rho)(z_0), \quad z_0 \in \mathbb{R}^{2d}.$$

Moreover, for any $\rho_1, \rho_2 \in \mathcal{P}_w$, there is a constant $C = C(r, \alpha, d, \rho_1, \rho_2) > 0$ such that

$$(3.24) \quad \|f\|_{\mathbf{C}_a^\alpha(\rho_1\rho_2)} \asymp_C \sup_{z_0 \in \mathbb{R}^{2d}} (\rho_1(z_0) \|\phi_r^{z_0} f\|_{\mathbf{C}_a^\alpha(\rho_2)})$$

and

$$(3.25) \quad \|f\|_{L^\infty(\rho_1\rho_2)} \asymp_C \sup_{z_0 \in \mathbb{R}^{2d}} (\rho_1(z_0) \|\phi_r^{z_0} f\|_{L^\infty(\rho_2)}).$$

PROOF. We only prove (3.23) and (3.24). (3.22) and (3.25) follow similarly. For $j, k \in \mathbb{N}_0$, by the chain rule and definitions (3.21) and (2.1),

$$\nabla_v^j \nabla_x^k \phi_r^{z_0}(z) = [r(1 + |z_0|_a)]^{-j-3k} (\nabla_v^j \nabla_x^k \chi) \left(\frac{z - z_0}{[r(1 + |z_0|_a)]^a} \right).$$

Note that

$$\mathcal{D} := \text{supp}(\nabla_v^j \nabla_x^k \chi) \subset \left\{ z : |z - z_0|_a \leq \frac{1}{4} r(1 + |z_0|_a) \right\}.$$

Since $r < 1$, it is easy to see that for $z = (x, v) \in \mathcal{D}$,

$$|v| \leq |z - z_0|_a + |z_0|_a \leq \frac{1}{4} + \frac{5}{4} |z_0|_a, \quad 1 + |z|_a \asymp 1 + |z_0|_a.$$

Hence,

$$\|v \cdot \nabla_x \phi_r^{z_0}\|_{L^\infty(\rho)} = \sup_{z \in \mathbb{R}^{2d}} (|v \cdot \nabla_x \phi_r^{z_0}(z)| \rho(z)) \lesssim (\varrho\rho)(z_0),$$

and by the chain rule, for any $m \in \mathbb{N}$,

$$\|\nabla^m (v \cdot \nabla_x \phi_r^{z_0})\|_{L^\infty(\rho)} \lesssim (\varrho\rho)(z_0).$$

Thus by (2.21) we have

$$\|v \cdot \nabla_x \phi_r^z\|_{\mathbf{C}_a^\alpha(\rho)} \lesssim \|\nabla^{[\alpha]+1} (v \cdot \nabla_x \phi_r^z)\|_{L^\infty(\rho)} + \|v \cdot \nabla_x \phi_r^z\|_{L^\infty(\rho)} \lesssim (\varrho\rho)(z_0).$$

Similarly,

$$\|\nabla_v^j \phi_r^z\|_{\mathbf{C}_a^\alpha(\rho)} \lesssim (\varrho\rho)(z_0).$$

Next we prove (3.24). By definition, it is easy to see that there exists δ such that for h with $|h|_a < \delta$

$$\delta_h^{([\alpha]+1)} f(z_0) = \delta_h^{([\alpha]+1)} (\phi_r^{z_0} f)(z_0).$$

Therefore, by Theorem 2.7

$$\begin{aligned} \|f\|_{\mathbf{C}_a^\alpha(\rho_1\rho_2)} &\lesssim \sup_{z_0} (\rho_1\rho_2)(z_0) \left[\sup_{|h|_a \leq \delta} \frac{\delta_h^{([\alpha]+1)} (\phi_r^{z_0} f)(z_0)}{|h|_a^\alpha} + (\phi_r^{z_0} f)(z_0) \right] \\ &\lesssim \sup_{z_0} \|\phi_r^{z_0} f\|_{\mathbf{C}_a^\alpha(\rho_1\rho_2)} \lesssim \sup_{z_0} \|\phi_{2r}^{z_0}\|_{\mathbf{C}_a^\alpha(\rho_1)} \|\phi_r^{z_0} f\|_{\mathbf{C}_a^\alpha(\rho_2)} \\ &\stackrel{(3.22)}{\lesssim} \sup_{z_0} \rho_1(z_0) \|\phi_r^{z_0} f\|_{\mathbf{C}_a^\alpha(\rho_2)}. \end{aligned}$$

Here we used $\phi_r^{z_0} = \phi_{2r}^{z_0} \phi_r^{z_0}$ in the third inequality. On the other hand,

$$\sup_{z_0} \rho_1(z_0) \|\phi_r^{z_0} f\|_{\mathbf{C}_a^\alpha(\rho_2)} \lesssim \sup_{z_0} \rho_1(z_0) \|\phi_r^{z_0}\|_{\mathbf{C}_a^\alpha(\rho_1^{-1})} \|f\|_{\mathbf{C}_a^\alpha(\rho_1 \rho_2)} \stackrel{(3.22)}{\lesssim} \|f\|_{\mathbf{C}_a^\alpha(\rho_1 \rho_2)}.$$

Thus (3.24) follows. The proof is complete. \square

To give the characterization of the kinetic Hölder spaces, we need the following elementary lemma.

LEMMA 3.8. *For any $z_0 \in \mathbb{R}^{2d}$ and $|t| \leq r^3 < 1$, it holds that*

$$(3.26) \quad \phi_r^{z_0} \Gamma_t \phi_{8r}^{z_0} = \phi_r^{z_0},$$

and for $j = 0, 1$, there is a constant $C = C(r, d) > 0$ such that

$$(3.27) \quad \|\Gamma_t \nabla_v^j \phi_r^{z_0} - \nabla_v^j \phi_r^{z_0}\|_{L^\infty} \lesssim_C |t| / (1 + |z_0|_a^{2+j}).$$

PROOF. For $|t| \leq r^3 < 1$, by Young's inequality, we have

$$|tv|^{1/3} \leq r \left(\frac{2}{3} + \frac{|v|}{3} \right).$$

Equality (3.26) follows by

$$\text{supp}(\phi_r^{z_0}) \subset B_{r(1+|z_0|_a)/4}^a(z_0) \subset \Gamma_t B_{r(1+|z_0|_a)}^a(z_0)$$

and $\phi_{8r}^{z_0} \equiv 1$ on $B_{r(1+|z_0|_a)}^a(z_0)$. For (3.27), note that for $z = (x, v)$,

$$\begin{aligned} |\Gamma_t \nabla_v^j \phi_r^{z_0}(z) - \nabla_v^j \phi_r^{z_0}(z)| &\leq \sup_{s \in [0, t]} |t| |v| |\nabla_x \nabla_v^j \phi_r^{z_0}(\Gamma_s z)| \cdot \mathbf{1}_{\{|\Gamma_s z - z_0|_a \leq r(1+|z_0|_a)/4\}} \\ &\leq |t| |z|_a \frac{\|\nabla_x \nabla_v^j \chi\|_{L^\infty}}{(r(1+|z_0|_a))^{3+j}} \mathbf{1}_{\{|z - z_0|_a \leq r(1+|z_0|_a)\}} \\ &\lesssim |t| / (1 + |z_0|_a^{2+j}). \end{aligned}$$

The proof is complete. \square

Now we are in a position to give the following characterization of $\mathbb{S}_{T,a}^\alpha(\rho)$.

LEMMA 3.9. *For any $\alpha \in (0, 2)$, $r \in (0, 1/8)$, $\rho \in \mathcal{P}_w$ and $T > 0$, there is a constant $C = C(T, r, \alpha, d, \rho) > 0$ such that*

$$(3.28) \quad \|f\|_{\mathbb{S}_{T,a}^\alpha(\rho)} \asymp_C \sup_{z_0} \left(\rho(z_0) \|\phi_r^{z_0} f\|_{\mathbb{S}_{T,a}^\alpha} \right).$$

PROOF. By (3.24), we only need to prove that for any $\alpha \in (0, 1)$,

$$\|f\|_{\mathbf{C}_{T,\Gamma}^\alpha L^\infty(\rho)} \asymp \sup_z \left(\rho(z) \|\phi_r^z f\|_{\mathbf{C}_{T,\Gamma}^\alpha L^\infty} \right).$$

By definition and (3.8), (3.25), it suffices to show

$$\begin{aligned} (3.29) \quad & \sup_{0 \leq t \leq T} \|f(t)\|_{L^\infty(\rho)} + \sup_{0 < |t-s| \leq r^3} \sup_z \frac{\rho(z) \|\phi_r^z f(t) - \phi_r^z \Gamma_{t-s} f(s)\|_{L^\infty}}{|t-s|^\alpha} \\ & \asymp \sup_{0 \leq t \leq T} \|f(t)\|_{L^\infty(\rho)} + \sup_{0 < |t-s| \leq r^3} \sup_z \frac{\rho(z) \|\phi_r^z f(t) - \Gamma_{t-s}(\phi_r^z f)(s)\|_{L^\infty}}{|t-s|^\alpha}. \end{aligned}$$

Since $\phi_r^z = \phi_r^z \phi_{8r}^z$, it follows from (3.26) that

$$\phi_r^z \Gamma_t f - \Gamma_t(\phi_r^z f) = \phi_r^z \Gamma_t(\phi_{8r}^z f) - \Gamma_t(\phi_r^z \phi_{8r}^z f), \quad \forall t \in [0, r^3].$$

Thus by the fact $\Gamma_t(fg) = \Gamma_t f \Gamma_t g$ and (3.27), one sees that

$$\begin{aligned} & \sup_{0 < |t-s| \leq r^3} \sup_z \frac{\rho(z) \|\phi_r^z \Gamma_{t-s} f(s) - \Gamma_{t-s}(\phi_r^z f)(s)\|_{L^\infty}}{|t-s|^\alpha} \\ &= \sup_{0 < |t-s| \leq r^3} \sup_z \frac{\rho(z) \|\phi_r^z \Gamma_{t-s}(\phi_{8r}^z f(s)) - \Gamma_{t-s}(\phi_r^z \phi_{8r}^z f)(s)\|_{L^\infty}}{|t-s|^\alpha} \\ &\lesssim \sup_{0 < |t-s| \leq r^3} \sup_z \frac{\rho(z) \|\Gamma_{t-s}(\phi_{8r}^z f(s))\|_{L^\infty} \|\phi_r^z - \Gamma_{t-s} \phi_r^z\|_{L^\infty}}{|t-s|^\alpha} \\ &\lesssim \sup_{0 < |t-s| \leq r^3} \sup_z \frac{\rho(z) \|\phi_{8r}^z f(s)\|_{L^\infty} |t-s|}{|t-s|^\alpha} \lesssim \sup_{0 \leq s \leq T} \|f(s)\|_{L^\infty(\rho)}. \end{aligned}$$

The proof is complete. \square

Next we give a result regarding derivative of a function in kinetic Hölder spaces.

LEMMA 3.10. *For any $\alpha \in (1, 2)$, $T > 0$ and $\rho \in \mathcal{P}_w$, there is a constant $C = C(\rho, T, \alpha, d) > 0$ such that for all $f \in \mathbb{S}_{T,a}^\alpha(\rho)$,*

$$(3.30) \quad \|\nabla_v f\|_{\mathbb{S}_{T,a}^{\alpha-1}(\rho)} \lesssim_C \|f\|_{\mathbb{S}_{T,a}^\alpha(\rho)}.$$

PROOF. First of all, we prove (3.30) for $\rho = 1$. Fix $\alpha \in (1, 2)$, $s, t \in [0, T]$ and $z \in \mathbb{R}^{2d}$. By definition, one sees that for all $\bar{z} = (0, \bar{v}) \in \mathbb{R}^{2d}$,

$$\begin{aligned} \mathcal{I}_1 &:= |f(t, \bar{z} + z) - f(s, \Gamma_{t-s} z) - \bar{v} \cdot (\nabla_v f)(s, \Gamma_{t-s} z)| \\ &\leq |f(t, \bar{z} + z) - f(s, \Gamma_{t-s}(\bar{z} + z))| + |f(s, \Gamma_{t-s}(\bar{z} + z)) - f(s, \bar{z} + \Gamma_{t-s} z)| \\ &\quad + |f(s, \bar{z} + \Gamma_{t-s} z) - f(s, \Gamma_{t-s} z) - \bar{v} \cdot (\nabla_v f)(s, \Gamma_{t-s} z)| \\ &\lesssim |t-s|^{\frac{\alpha}{2}} \|f\|_{\mathbf{C}_{T,\Gamma}^{\alpha/2} L^\infty} + |(t-s)\bar{v}|^{\frac{\alpha}{3}} \|f(s)\|_{\mathbf{C}_x^{\alpha/3}} + |\bar{v}|^\alpha \|\nabla_v f(s)\|_{\mathbf{C}_v^{\alpha-1}} \\ &\lesssim (|t-s|^{\frac{\alpha}{2}} + |\bar{v}|^\alpha) \|f\|_{\mathbb{S}_{T,a}^\alpha}, \end{aligned}$$

where we used $|(t-s)\bar{v}|^{\frac{\alpha}{3}} \leq \frac{3}{2}|t-s|^{\frac{\alpha}{2}} + 3|\bar{v}|^\alpha$ in the last step. By exchanging $s \leftrightarrow t$ and in place of z by $\Gamma_{t-s} z$, we also have

$$\mathcal{I}_2 := |f(s, \bar{z} + \Gamma_{t-s} z) - f(t, z) - \bar{v} \cdot (\nabla_v f)(t, z)| \lesssim (|t-s|^{\frac{\alpha}{2}} + |\bar{v}|^\alpha) \|f\|_{\mathbb{S}_{T,a}^\alpha}.$$

Let ω be the unit vector in \mathbb{R}^d so that

$$\mathcal{I} := |(\nabla_v f)(t, z) - (\nabla_v f)(s, \Gamma_{t-s} z)| = \omega \cdot [(\nabla_v f)(t, z) - (\nabla_v f)(s, \Gamma_{t-s} z)].$$

Let $\bar{v} = (t-s)^{\frac{1}{2}} \omega$ and $\bar{z} = (0, \bar{v})$. Then

$$\begin{aligned} (t-s)^{\frac{1}{2}} \mathcal{I} &\leq \mathcal{I}_1 + \mathcal{I}_2 + |f(t, \bar{z} + z) - f(s, \bar{z} + \Gamma_{t-s} z)| + |f(t, z) - f(s, \Gamma_{t-s} z)| \\ &\leq \mathcal{I}_1 + \mathcal{I}_2 + |(t-s)\bar{v}|^{\frac{\alpha}{3}} \|f(s, \cdot)\|_{\mathbf{C}_x^\alpha} + 2\|f(t) - \Gamma_{t-s} f(s)\|_{L^\infty} \\ &\lesssim (t-s)^{\frac{\alpha}{2}} \|f\|_{\mathbb{S}_{T,a}^\alpha}. \end{aligned}$$

Hence,

$$(3.31) \quad \mathcal{I} = |(\nabla_v f)(t, z) - (\nabla_v f)(s, \Gamma_{t-s} z)| \lesssim (t-s)^{\frac{\alpha-1}{2}} \|f\|_{\mathbb{S}_{T,a}^\alpha}.$$

Moreover, by Bernstein's inequality in Lemma 2.4, it is clear that

$$\|\nabla_v f\|_{\mathbb{C}_{T,a}^{\alpha-1}} \lesssim \|f\|_{\mathbb{C}_{T,a}^\alpha},$$

which together with (3.31) implies (3.30) for $\rho = 1$.

Next, for $\beta \in (0, 2)$, note that by (3.27) and the definition of kinetic Hölder space,

$$\|\nabla_v \phi_r^z g\|_{\mathbb{C}_{T,\Gamma}^{\beta/2} L^\infty} \lesssim (1 + |z|_a)^{-1} \|g\|_{\mathbb{C}_{T,\Gamma}^{\beta/2} L^\infty}$$

and

$$\|\nabla_v \phi_r^z g\|_{\mathbb{C}_a^\beta} \lesssim \|\nabla_v \phi_r^z\|_{\mathbb{C}_a^\beta} \|g\|_{\mathbb{C}_a^\beta} \lesssim (1 + |z|_a)^{-1} \|g\|_{\mathbb{C}_a^\beta}.$$

Hence,

$$(3.32) \quad \|\nabla_v \phi_r^z g\|_{\mathbb{S}_{T,a}^\beta} \lesssim (1 + |z|_a)^{-1} \|g\|_{\mathbb{S}_{T,a}^\beta}.$$

Now, for any $r \in (0, 1/16)$, by Lemma 3.9 and (3.32) we have

$$\begin{aligned} \|\nabla_v f\|_{\mathbb{S}_{T,a}^{\alpha-1}(\rho)} &\asymp \sup_z \rho(z) \|\phi_r^z \nabla_v f\|_{\mathbb{S}_{T,a}^{\alpha-1}} \\ &\lesssim \sup_z \rho(z) \|\nabla_v(\phi_r^z f)\|_{\mathbb{S}_{T,a}^{\alpha-1}} + \sup_z \rho(z) \|\nabla_v \phi_r^z f\|_{\mathbb{S}_{T,a}^{\alpha-1}} \\ &\lesssim \sup_z \rho(z) \|\phi_r^z f\|_{\mathbb{S}_{T,a}^\alpha} + \sup_z \rho(z) \|\nabla_v \phi_r^z(\phi_{2r}^z f)\|_{\mathbb{S}_{T,a}^{\alpha-1}} \\ &\lesssim \|f\|_{\mathbb{S}_{T,a}^\alpha(\rho)} + \sup_z \rho(z) \|\phi_{2r}^z f\|_{\mathbb{S}_{T,a}^{\alpha-1}} \lesssim \|f\|_{\mathbb{S}_{T,a}^\alpha(\rho)}. \end{aligned}$$

Here in the second inequality we used (3.30) for $\rho = 1$ we proved above. The proof is complete. \square

3.3. Schauder's estimates. For given $\lambda \geq 0$ and $f \in L_{\text{loc}}^1(\mathbb{R}_+; \mathcal{S}'(\mathbb{R}^{2d}))$, we consider the following model kinetic equation:

$$\mathcal{L}_\lambda u := (\partial_t - \Delta_v + \lambda - v \cdot \nabla_x) u = f, \quad u(0) = 0.$$

By Duhamel's formula, the unique solution of the above equation is given by

$$(3.33) \quad u(t, \cdot) = \int_0^t e^{-\lambda(t-s)} P_{t-s} f(s, \cdot) ds := \mathcal{J}_\lambda f(t, \cdot).$$

In other words, \mathcal{J}_λ is the inverse of \mathcal{L}_λ . For $q \in [1, \infty]$, $T > 0$ and a Banach space \mathbb{B} , we write

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

Now we can show the following Schauder estimate.

LEMMA 3.11. (*Schauder estimates*) *Let $\rho \in \mathcal{P}_w$, $\beta \in (0, 2)$ and $\theta \in (\beta, 2]$. For any $q \in [\frac{2}{2-\theta}, \infty]$ and $T > 0$, there is a constant $C = C(d, \beta, \theta, q, T) > 0$ such that for all $\lambda \geq 0$ and $f \in L_T^q \mathbb{C}_a^{-\beta}(\rho)$,*

$$(3.34) \quad \|\mathcal{J}_\lambda f\|_{\mathbb{S}_{T,a}^{\theta-\beta}(\rho)} \lesssim_C (\lambda \vee 1)^{\frac{\theta}{2} + \frac{1}{q} - 1} \|f\|_{L_T^q \mathbb{C}_a^{-\beta}(\rho)}.$$

PROOF. Let $q \in [\frac{2}{2-\theta}, \infty]$ and $\frac{1}{p} + \frac{1}{q} + \frac{\theta}{2} = 1$. By (3.14) and Hölder's inequality, we have for $\beta \in \mathbb{R}$,

$$\begin{aligned} \|\mathcal{R}_j^a \mathcal{J}_\lambda f(t)\|_{L^\infty(\rho)} &\lesssim \int_0^t e^{-\lambda(t-s)} 2^{j\beta} (1 \wedge ((t-s)4^j)^{-2}) \|f(s)\|_{\mathbf{C}_a^{-\beta}(\rho)} ds \\ &\lesssim 2^{j\beta} \left(\int_0^t e^{-\lambda ps} ds \right)^{\frac{1}{p}} \left(\int_0^t (1 \wedge (s4^j)^{-\frac{4}{\theta}}) ds \right)^{\frac{\theta}{2}} \|f\|_{L_T^q \mathbf{C}_a^{-\beta}(\rho)} \\ &\lesssim 2^{j(\beta-\theta)} (\lambda \vee 1)^{-\frac{1}{p}} \left(\int_0^\infty (1 \wedge s^{-\frac{4}{\theta}}) ds \right)^{\frac{\theta}{2}} \|f\|_{L_T^q \mathbf{C}_a^{-\beta}(\rho)}. \end{aligned}$$

This implies that for $\beta \in \mathbb{R}$,

$$(3.35) \quad \|\mathcal{J}_\lambda f\|_{\mathbf{C}_{T,a}^{\theta-\beta}(\rho)} \lesssim (\lambda \vee 1)^{\frac{\theta}{2} + \frac{1}{q} - 1} \|f\|_{L_T^q \mathbf{C}_a^{-\beta}(\rho)}.$$

On the other hand, let $u := \mathcal{J}_\lambda f$. For any $0 \leq t_1 < t_2 \leq T$, we have

$$\begin{aligned} u(t_2) - \Gamma_{t_2-t_1} u(t_1) &= \int_0^{t_1} \left(e^{-\lambda(t_2-s)} - e^{-\lambda(t_1-s)} \right) P_{t_2-s} f(s) ds \\ &\quad + \left(P_{t_2-t_1} - \Gamma_{t_2-t_1} \right) \mathcal{J}_\lambda f(t_1) + \int_{t_1}^{t_2} e^{-\lambda(t_2-s)} P_{t_2-s} f(s) ds \\ &:= I_1 + I_2 + I_3. \end{aligned}$$

Let

$$q' := q/(q-1).$$

For I_1 , by (3.16) and Hölder's inequality, we have for $\beta > 0$,

$$\begin{aligned} \|I_1\|_{L^\infty(\rho)} &\leq |e^{-\lambda(t_2-t_1)} - 1| \int_0^{t_1} e^{-\lambda(t_1-s)} \|P_{t_2-s} f(s)\|_{L^\infty(\rho)} ds \\ &\lesssim [\lambda(t_2-t_1) \wedge 1] \int_0^{t_1} e^{-\lambda(t_1-s)} (t_2-s)^{-\frac{\beta}{2}} \|f(s)\|_{\mathbf{C}_a^{-\beta}(\rho)} ds \\ &\leq [\lambda(t_2-t_1)]^{\frac{\theta}{2}} (t_2-t_1)^{-\frac{\beta}{2}} \left(\int_0^{t_1} e^{-\lambda q' s} ds \right)^{\frac{1}{q'}} \|f\|_{L_T^q \mathbf{C}_a^{-\beta}(\rho)} \\ &\lesssim (t_2-t_1)^{\frac{\theta-\beta}{2}} (\lambda \vee 1)^{\frac{\theta}{2} + \frac{1}{q} - 1} \|f\|_{L_T^q \mathbf{C}_a^{-\beta}(\rho)}. \end{aligned}$$

For I_2 , by (3.17) and (3.35), we have for $\beta \in (\theta-2, \theta)$,

$$\begin{aligned} \|I_2\|_{L^\infty(\rho)} &\leq (t_2-t_1)^{\frac{\theta-\beta}{2}} \|\mathcal{J}_\lambda f\|_{\mathbf{C}_{T,a}^{\theta-\beta}(\rho)} \\ &\lesssim (t_2-t_1)^{\frac{\theta-\beta}{2}} (\lambda \vee 1)^{\frac{\theta}{2} + \frac{1}{q} - 1} \|f\|_{L_T^q \mathbf{C}_a^{-\beta}(\rho)}. \end{aligned}$$

For I_3 , by (3.16), Hölder's inequality and the change of variable, we have for $\beta \in (0, \theta)$,

$$\begin{aligned} \|I_3\|_{L^\infty(\rho)} &\lesssim \int_{t_1}^{t_2} e^{-\lambda(t_2-s)} (t_2-s)^{-\frac{\beta}{2}} \|f(s)\|_{\mathbf{C}_a^{-\beta}(\rho)} ds \\ &\lesssim \left(\int_0^{t_2-t_1} e^{-q' \lambda s} s^{-\frac{q'\beta}{2}} ds \right)^{\frac{1}{q'}} \|f\|_{L_T^q \mathbf{C}_a^{-\beta}(\rho)} \\ &\lesssim \left((t_2-t_1)^{\frac{1}{q'} - \frac{\beta}{2}} \wedge \lambda^{\frac{\beta}{2} - \frac{1}{q'}} \right) \|f\|_{L_T^q \mathbf{C}_a^{-\beta}(\rho)} \\ &\lesssim (t_2-t_1)^{\frac{\theta-\beta}{2}} \lambda^{\frac{\theta}{2} + \frac{1}{q} - 1} \|f\|_{L_T^q \mathbf{C}_a^{-\beta}(\rho)}, \end{aligned}$$

where in the third inequality we have used interpolation inequality $a^\gamma \wedge b^{-\gamma} \leq a^\delta b^{\delta-\gamma}$ for all $a, b > 0$ and $0 < \delta \leq \gamma$. Combining the above calculations, we obtain

$$\|\mathcal{J}_\lambda f\|_{C_T^{(\theta-\beta)/2} L^\infty(\rho)} \lesssim (\lambda \vee 1)^{\frac{\theta}{2} + \frac{1}{q} - 1} \|f\|_{L_T^q \mathbf{C}_a^{-\beta}(\rho)}.$$

The proof is complete. \square

3.4. Commutator estimates. In this subsection we prove important commutator estimates about the kinetic semigroups, which are essential for applying paracontrolled calculus to the kinetic equations. Compared with the case of the classical heat semigroups (see [GIP15]), the kinetic semigroup is not a Fourier multiplier and there is a Γ_t in the commutator as stated in the left hand side of (3.36) below, which leads to a commutator for \mathcal{J}_λ in the kinetic Hölder space (see Lemma 3.13 below). In the estimate of the main term $I_j^{(0)}$ in the proof, we find that the commutator for the operator $\Gamma_t p_t^*$ gains regularity from f , while the commutator for Γ_t cannot have this property. In particular, the decomposition (3.13) plays a crucial role in the following proof.

LEMMA 3.12. *Let $\rho_1, \rho_2 \in \mathcal{P}_w$. For any $\alpha \in (0, 1)$, $\beta \in \mathbb{R}$, $\delta \geq 0$ and $T > 0$, there is a constant $C = C(\rho_1, \rho_2, \alpha, \beta, \delta, T, d) > 0$ such that for all $f \in \mathbf{C}_a^\alpha(\rho_1)$, $g \in \mathbf{C}_a^\beta(\rho_2)$ and $t \in (0, T]$, $j \geq -1$,*

$$(3.36) \quad \begin{aligned} & \|\mathcal{R}_j^a P_t(f \prec g) - \mathcal{R}_j^a(\Gamma_t f \prec P_t g)\|_{L^\infty(\rho_1 \rho_2)} \\ & \lesssim_C t^{-\frac{\delta}{2}} 2^{-(\alpha+\beta+\delta)j} \|f\|_{\mathbf{C}_a^\alpha(\rho_1)} \|g\|_{\mathbf{C}_a^\beta(\rho_2)}. \end{aligned}$$

PROOF. Without loss of generality, we only prove (3.36) for $j \geq 3$ and $\beta \neq 0, -\alpha$. For $\beta = 0$ or $-\alpha$, it follows by the interpolation Lemma 2.10. First of all, by (2.5), (3.13) and the definition of \prec , we have

$$\mathcal{R}_j^a P_t(f \prec g) = \sum_{\ell \sim j} \sum_{i \in \Theta_\ell^t} \sum_{k \geq -1} \mathcal{R}_j^a \mathcal{R}_\ell^a P_t \mathcal{R}_i^a (S_{k-1} f \mathcal{R}_k^a g),$$

where

$$\ell \sim j \iff |\ell - j| \leq 3.$$

Noting that by (2.9),

$$\mathcal{R}_i^a (S_{k-1} f \mathcal{R}_k^a g) = 0 \text{ for } i \in \Theta_\ell^t \text{ and } k \notin \Theta_\ell^t \pm 3,$$

where

$$\Theta_\ell^t \pm 3 := \{k \geq 0 : |k - i| \leq 3, i \in \Theta_\ell^t\},$$

we further have

$$\begin{aligned} \mathcal{R}_j^a P_t(f \prec g) &= \sum_{\ell \sim j} \sum_{i \in \Theta_\ell^t} \sum_{k \in \Theta_\ell^t \pm 3} \mathcal{R}_j^a \mathcal{R}_\ell^a P_t \mathcal{R}_i^a (S_{k-1} f \mathcal{R}_k^a g) \\ &\stackrel{(3.13)}{=} \sum_{\ell \sim j} \sum_{k \in \Theta_\ell^t \pm 3} \mathcal{R}_j^a \mathcal{R}_\ell^a P_t (S_{k-1} f \mathcal{R}_k^a g). \end{aligned}$$

Similarly, by (2.9) we also have

$$\mathcal{R}_j^a(\Gamma_t f \prec P_t g) = \sum_{\ell \sim j} \mathcal{R}_j^a (S_{\ell-1} \Gamma_t f \cdot \mathcal{R}_\ell^a P_t g) =: I_j^{(1)} + I_j^{(2)},$$

where

$$I_j^{(1)} := \sum_{\ell \sim j} \mathcal{R}_j^a(\Gamma_t S_{\ell-1} f \cdot \mathcal{R}_\ell^a P_t g),$$

$$I_j^{(2)} := \sum_{\ell \sim j} \mathcal{R}_j^a([S_{\ell-1}, \Gamma_t] f \cdot \mathcal{R}_\ell^a P_t g).$$

For $I_j^{(1)}$, by (3.13) again, we can write

$$\begin{aligned} I_j^{(1)} &= \sum_{\ell \sim j} \sum_{k \in \Theta_\ell^t \pm 3} \mathcal{R}_j^a(\Gamma_t S_{\ell-1} f \cdot \mathcal{R}_\ell^a P_t \mathcal{R}_k^a g) \\ &= \sum_{\ell \sim j} \sum_{k \in \Theta_\ell^t \pm 3} \mathcal{R}_j^a(\Gamma_t (S_{\ell-1} - S_{k-1}) f \cdot \mathcal{R}_\ell^a P_t \mathcal{R}_k^a g) \\ &\quad + \sum_{\ell \sim j} \sum_{k \in \Theta_\ell^t \pm 3} \mathcal{R}_j^a(\Gamma_t S_{k-1} f \cdot \mathcal{R}_\ell^a P_t \mathcal{R}_k^a g) =: I_j^{(11)} + I_j^{(12)}. \end{aligned}$$

Combining the above calculations, we obtain

$$\mathcal{R}_j^a P_t(f \prec g) - \mathcal{R}_j^a(\Gamma_t f \prec P_t g) = I_j^{(0)} - I_j^{(11)} - I_j^{(2)},$$

where

$$I_j^{(0)} := \sum_{\ell \sim j} \sum_{k \in \Theta_\ell^t \pm 3} \mathcal{R}_j^a \left(\mathcal{R}_\ell^a P_t(S_{k-1} f \mathcal{R}_k^a g) - \Gamma_t S_{k-1} f \cdot \mathcal{R}_\ell^a P_t \mathcal{R}_k^a g \right).$$

For $I_j^{(0)}$, let $F_k := S_{k-1} f$ and $G_k := \mathcal{R}_k^a g$. Note that

$$\begin{aligned} J_{k\ell} &:= \mathcal{R}_\ell^a P_t(S_{k-1} f \mathcal{R}_k^a g) - \Gamma_t S_{k-1} f \cdot \mathcal{R}_\ell^a P_t \mathcal{R}_k^a g \\ &\stackrel{(3.1)}{=} (\mathcal{R}_\ell^a \Gamma_t p_t) * \Gamma_t(F_k G_k) - \Gamma_t F_k (\mathcal{R}_\ell^a \Gamma_t p_t * \Gamma_t G_k) \\ &= \int_{\mathbb{R}^{2d}} \mathcal{R}_\ell^a \Gamma_t p_t(\bar{z}) (\Gamma_t F_k(z - \bar{z}) - \Gamma_t F_k(z)) \Gamma_t G_k(z - \bar{z}) d\bar{z}. \end{aligned}$$

By (2.25), (3.8) and (3.9), there exists $\delta_0 > 0$ such that for any $m \geq 0$,

$$\begin{aligned} \|J_{k\ell}\|_{L^\infty(\rho_1 \rho_2)} &\lesssim \left(\int_{\mathbb{R}^{2d}} |\mathcal{R}_\ell^a \Gamma_t p_t(\bar{z})| |\Gamma_t \bar{z}|_a^\alpha (1 + |\bar{z}|_a^{\delta_0}) d\bar{z} \right) \|F_k\|_{\mathbf{C}_a^\alpha(\rho_1)} \|G_k\|_{L^\infty(\rho_2)} \\ &\lesssim 2^{-\alpha\ell} (1 \wedge (t4^\ell)^{-m}) \|f\|_{\mathbf{C}_a^\alpha(\rho_1)} 2^{-\beta k} \|g\|_{\mathbf{C}_a^\beta(\rho_2)}. \end{aligned}$$

Hence, by (3.12), for $\beta \neq 0$,

$$\begin{aligned} \|I_j^{(0)}\|_{L^\infty(\rho_1 \rho_2)} &\lesssim \sum_{\ell \sim j} \sum_{k \in \Theta_\ell^t \pm 3} 2^{-\alpha\ell - \beta k} (1 \wedge (t4^\ell)^{-|\beta| - \frac{\delta}{2}}) \|f\|_{\mathbf{C}_a^\alpha(\rho_1)} \|g\|_{\mathbf{C}_a^\beta(\rho_2)} \\ &\lesssim 2^{-j(\alpha+\beta)} (1 + (t4^j)^{|\beta|}) (1 \wedge (t4^j)^{-|\beta| - \frac{\delta}{2}}) \|f\|_{\mathbf{C}_a^\alpha(\rho_1)} \|g\|_{\mathbf{C}_a^\beta(\rho_2)} \\ &\lesssim 2^{-j(\alpha+\beta)} (t4^j)^{-\frac{\delta}{2}} \|f\|_{\mathbf{C}_a^\alpha(\rho_1)} \|g\|_{\mathbf{C}_a^\beta(\rho_2)}. \end{aligned}$$

For $I_j^{(11)}$, by (3.14) we have for any $m \geq 0$,

$$\|I_j^{(11)}\|_{L^\infty(\rho_1 \rho_2)} \lesssim \sum_{\ell \sim j} \sum_{k \in \Theta_\ell^t \pm 3} \|\Gamma_t (S_{\ell-1} - S_{k-1}) f \cdot \mathcal{R}_\ell^a P_t \mathcal{R}_k^a g\|_{L^\infty(\rho_1 \rho_2)}$$

$$\begin{aligned}
&\lesssim \sum_{\ell \sim j} \sum_{k \in \Theta_\ell^t \pm 3} \|(S_{\ell-1} - S_{k-1})f\|_{L^\infty(\rho_1)} \|\mathcal{R}_\ell^a P_t \mathcal{R}_k^a g\|_{L^\infty(\rho_2)} \\
&\lesssim \sum_{\ell \sim j} \sum_{k \in \Theta_\ell^t \pm 3} \sum_{i=k \wedge \ell}^{k \vee \ell} \|\mathcal{R}_i^a f\|_{L^\infty(\rho_1)} (1 \wedge (t4^\ell)^{-m}) \|\mathcal{R}_k^a g\|_{\mathbf{C}_a^0(\rho_2)} \\
&\lesssim \sum_{\ell \sim j} \sum_{k \in \Theta_\ell^t \pm 3} \left(\sum_{i=k \wedge \ell-1}^{k \vee \ell-2} 2^{-i\alpha} \right) (1 \wedge (t4^j)^{-m}) \|f\|_{\mathbf{C}_a^\alpha(\rho_1)} \|\mathcal{R}_k^a g\|_{L^\infty(\rho_2)} \\
&\lesssim \sum_{\ell \sim j} \sum_{k \in \Theta_\ell^t \pm 3} 2^{-(k \wedge \ell)\alpha} 2^{-k\beta} (1 \wedge (t4^j)^{-m}) \|f\|_{\mathbf{C}_a^\alpha(\rho_1)} \|g\|_{\mathbf{C}_a^\beta(\rho_2)} \\
&\lesssim 2^{-(\alpha+\beta)j} (1+t4^j)^{\alpha+|\beta|} (1 \wedge (t4^j)^{-m}) \|f\|_{\mathbf{C}_a^\alpha(\rho_1)} \|g\|_{\mathbf{C}_a^\beta(\rho_2)},
\end{aligned}$$

where in the last step we have used $2^{-(k \wedge \ell)\alpha} \leq 2^{-k\alpha} + 2^{-\ell\alpha}$, (3.12) and $\beta \neq 0, -\alpha$. Taking $m = \alpha + |\beta| + \delta/2$, we get

$$\|I_j^{(11)}\|_{L^\infty(\rho_1 \rho_2)} \lesssim 2^{-(\alpha+\beta)j} (t4^j)^{-\delta/2} \|f\|_{\mathbf{C}_a^\alpha(\rho_1)} \|g\|_{\mathbf{C}_a^\beta(\rho_2)}.$$

For $I_j^{(2)}$, noting that

$$[S_{\ell-1}, \Gamma_t]f = \sum_{i=-1}^{\ell-2} [\mathcal{R}_i^a, \Gamma_t]f = - \sum_{i=\ell-1}^{\infty} [\mathcal{R}_i^a, \Gamma_t]f,$$

and by the definition of \mathcal{R}_i^a ,

$$[\mathcal{R}_i^a, \Gamma_t]f(z) = \int_{\mathbb{R}^{2d}} \check{\phi}_i^a(\bar{z}) (f(\Gamma_t(z - \bar{z})) - f(\Gamma_t z - \bar{z})) d\bar{z},$$

we have for some $\delta_0 > 0$ and all i ,

$$\begin{aligned}
&\|[\mathcal{R}_i^a, \Gamma_t]f\|_{L^\infty(\rho_1)} \\
&\leq \sup_{z \in \mathbb{R}^{2d}} \rho_1(z) \int_{\mathbb{R}^{2d}} |\check{\phi}_i^a(\bar{z})| |f(\Gamma_t(z - \bar{z})) - f(\Gamma_t z - \bar{z})| d\bar{z} \\
&\stackrel{(3.6)}{\lesssim} \sup_{z \in \mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} |\check{\phi}_i^a(\bar{z})| (1 + |\bar{z}|_a)^{\delta_0} |f(\Gamma_t(z - \bar{z})) - f(\Gamma_t z - \bar{z})| \rho_1(z - \bar{z}) d\bar{z} \\
&\stackrel{(3.8)}{\lesssim} \sup_{z \in \mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} |\check{\phi}_i^a(\bar{z})| (1 + |\bar{z}|_a)^{\delta_0} |f(\Gamma_t(z - \bar{z})) - f(\Gamma_t z - \bar{z})| \rho_1(\Gamma_t(z - \bar{z})) d\bar{z} \\
&\stackrel{(2.25)}{\lesssim} \int_{\mathbb{R}^{2d}} |\check{\phi}_i^a(\bar{z})| (1 + |\bar{z}|_a)^{\delta_0} (t|\bar{v}|)^{\frac{\alpha}{3}} (1 + t|\bar{v}|)^{\delta_0} \|f\|_{\mathbf{C}_x^{\alpha/3}(\rho_1)} d\bar{z} \\
&\lesssim (t2^{-i})^{\frac{\alpha}{3}} \|f\|_{\mathbf{C}_a^\alpha(\rho_1)},
\end{aligned}$$

Moreover,

$$\|[S_{\ell-1}, \Gamma_t]f\|_{L^\infty(\rho_1)} \lesssim \sum_{i=\ell-1}^{\infty} (t2^{-i})^{\frac{\alpha}{3}} \|f\|_{\mathbf{C}_a^\alpha(\rho_1)} \lesssim t^{\frac{\alpha}{3}} 2^{-\frac{\alpha\ell}{3}} \|f\|_{\mathbf{C}_a^\alpha(\rho_1)}.$$

Hence, by (3.14),

$$\|I_j^{(2)}\|_{L^\infty(\rho_1 \rho_2)} \lesssim \sum_{\ell \sim j} \|[S_{\ell-1}, \Gamma_t]f\|_{L^\infty(\rho_1)} \|\mathcal{R}_\ell^a P_t g\|_{L^\infty(\rho_2)}$$

$$\begin{aligned}
&\lesssim \sum_{\ell \sim j} t^{\frac{\alpha}{3}} 2^{-\frac{\alpha \ell}{3}} \|f\|_{\mathbf{C}_a^\alpha(\rho_1)} 2^{-\beta \ell} (t 4^\ell)^{-\frac{\alpha}{3} - \frac{\delta}{2}} \|g\|_{\mathbf{C}_a^\beta(\rho_2)} \\
&\lesssim t^{-\delta/2} 2^{-(\alpha+\beta+\delta)j} \|f\|_{\mathbf{C}_a^\alpha(\rho_1)} \|g\|_{\mathbf{C}_a^\beta(\rho_2)}.
\end{aligned}$$

The proof is complete. \square

Using this lemma we can show the following crucial commutator estimate.

LEMMA 3.13. *Let $\rho_1, \rho_2 \in \mathcal{P}_w$ and $\alpha \in (0, 1)$, $\beta \in \mathbb{R}$. For $k = 0, 1$, $T > 0$ and $\theta \in [0, 2]$, there is a constant $C > 0$ such that for all $\lambda \geq 0$,*

$$(3.37) \quad \|[\nabla_v^k \mathcal{J}_\lambda, f \prec]g\|_{\mathbf{C}_{T,a}^{\alpha+\beta+\theta-k}(\rho_1 \rho_2)} \lesssim_C (\lambda \vee 1)^{\frac{\theta-2}{2}} \|f\|_{\mathbf{S}_{T,a}^\alpha(\rho_1)} \|g\|_{\mathbf{C}_{T,a}^\beta(\rho_2)}.$$

PROOF. For $k = 0$, by definition (3.33) of \mathcal{J}_λ , we can write

$$\begin{aligned}
[\mathcal{J}_\lambda, f \prec]g(t) &= \int_0^t e^{-\lambda(t-s)} \left(P_{t-s}(f(s) \prec g(s)) - f(t) \prec P_{t-s}g(s) \right) ds \\
&= \int_0^t e^{-\lambda(t-s)} \left(P_{t-s}(f(s) \prec g(s)) - \Gamma_{t-s}f(s) \prec P_{t-s}g(s) \right) ds \\
&\quad + \int_0^t e^{-\lambda(t-s)} (\Gamma_{t-s}f(s) - f(t)) \prec P_{t-s}g(s) ds \\
&=: I_1(t) + I_2(t).
\end{aligned}$$

For $I_1(t)$, by Lemma 3.12 with $\delta = 0, 4$ we have

$$\|\mathcal{R}_j^a I_1(t)\|_{L^\infty(\rho_1 \rho_2)} \lesssim 2^{-(\alpha+\beta)j} \int_0^t e^{-\lambda s} ((4^j s)^{-2} \wedge 1) ds \|f\|_{\mathbf{C}_{T,a}^\alpha(\rho_1)} \|g\|_{\mathbf{C}_{T,a}^\beta(\rho_2)}.$$

Note that by Hölder's inequality,

$$\begin{aligned}
\int_0^t e^{-\lambda s} ((4^j s)^{-2} \wedge 1) ds &\leq \left(\int_0^t e^{-\frac{2\lambda s}{2-\theta}} ds \right)^{\frac{2-\theta}{2}} \left(\int_0^t ((4^j s)^{-2} \wedge 1)^{\frac{2}{\theta}} ds \right)^{\frac{\theta}{2}} \\
(3.38) \quad &\lesssim (\lambda \vee 1)^{\frac{\theta-2}{2}} 2^{-\theta j}.
\end{aligned}$$

Thus,

$$\|\mathcal{R}_j^a I_1(t)\|_{L^\infty(\rho_1 \rho_2)} \lesssim (\lambda \vee 1)^{\frac{\theta-2}{2}} 2^{-(\alpha+\beta+\theta)j} \|f\|_{\mathbf{C}_{T,a}^\alpha(\rho_1)} \|g\|_{\mathbf{C}_{T,a}^\beta(\rho_2)}.$$

For $I_2(t)$, for any $\gamma > 0$, note that by (2.26),

$$\begin{aligned}
&\|\mathcal{R}_j^a ((\Gamma_{t-s}f(s) - f(t)) \prec P_{t-s}g(s))\|_{L^\infty(\rho_1 \rho_2)} \\
&\lesssim 2^{-(\gamma+\beta)j} \|\Gamma_{t-s}f(s) - f(t)\|_{L^\infty(\rho_1)} \|P_{t-s}g(s)\|_{\mathbf{C}_a^{\gamma+\beta}(\rho_2)} \\
&\lesssim 2^{-(\gamma+\beta)j} (t-s)^{\frac{\alpha-\gamma}{2}} \|f\|_{\mathbf{S}_{T,a}^\alpha(\rho_1)} \|g(s)\|_{\mathbf{C}_a^\beta(\rho_2)},
\end{aligned}$$

which implies by (3.38) again,

$$\begin{aligned}
\|\mathcal{R}_j^a I_2(t)\|_{L^\infty(\rho_1 \rho_2)} &\lesssim 2^{-(\alpha+\beta)j} \int_0^t e^{-\lambda s} (1 \wedge (s 4^j)^{-2}) ds \|f\|_{\mathbf{S}_{T,a}^\alpha(\rho_1)} \|g\|_{\mathbf{C}_{T,a}^\beta(\rho_2)} \\
&\lesssim 2^{-(\alpha+\beta+\theta)j} (\lambda \vee 1)^{\frac{\theta-2}{2}} \|f\|_{\mathbf{S}_{T,a}^\alpha(\rho_1)} \|g\|_{\mathbf{C}_{T,a}^\beta(\rho_2)}.
\end{aligned}$$

Thus we obtain (3.37) for $k = 0$. Note that

$$[\nabla_v \mathcal{J}_\lambda, f \prec]g = \nabla_v [\mathcal{J}_\lambda, f \prec]g + \nabla_v f(t) \prec \mathcal{J}_\lambda g.$$

Estimate (3.37) for $k = 1$ follows by what we have proved and Lemma 2.11 and (3.35). Thus we complete the proof. \square

The following commutator estimate is straightforward by Lemma 3.13, Lemmas 2.11 and Lemma 2.13. Since we will use it many times later, we write it as a lemma.

LEMMA 3.14. *Let $\rho_1, \rho_2, \rho_3 \in \mathcal{P}_w$. For any $\alpha \in (1, 2)$, $\gamma \in \mathbb{R}$ and $\beta < 0$ with $\alpha + \beta > 1$, $\alpha + \beta + \gamma > 0$ and $1 + \beta + \gamma < 0$, we have*

$$\| [b \circ \nabla_v \mathcal{J}_\lambda, \phi] f \|_{\mathbb{C}_{T,a}^{\alpha+\beta+\gamma}(\rho_1 \rho_2 \rho_3)} \lesssim \| \phi \|_{\mathbb{S}_{T,a}^{\alpha-1}(\rho_1)} \| f \|_{\mathbb{C}_{T,a}^\beta(\rho_2)} \| b \|_{\mathbb{C}_{T,a}^\gamma(\rho_3)}.$$

PROOF. Note that

$$\begin{aligned} [b \circ \nabla_v \mathcal{J}_\lambda, \phi] f &= b \circ \nabla_v \mathcal{J}_\lambda (\phi \succ f) + b \circ \nabla_v \mathcal{J}_\lambda (\phi \prec f) - \phi (b \circ \nabla_v \mathcal{J}_\lambda f) \\ &= b \circ \nabla_v \mathcal{J}_\lambda (\phi \succ f) + b \circ [\nabla_v \mathcal{J}_\lambda, \phi \prec] f + \text{com}(\phi, \nabla_v \mathcal{J}_\lambda f, b). \end{aligned}$$

By (2.28), (3.34) and (2.27), we have

$$\begin{aligned} \| b \circ \nabla_v \mathcal{J}_\lambda (\phi \succ f) \|_{\mathbb{C}_{T,a}^{\alpha+\beta+\gamma}(\rho_1 \rho_2 \rho_3)} &\lesssim \| \nabla_v \mathcal{J}_\lambda (\phi \succ f) \|_{\mathbb{C}_{T,a}^{\alpha+\beta}(\rho_1 \rho_2)} \| b \|_{\mathbb{C}_{T,a}^\gamma(\rho_3)} \\ &\lesssim \| \phi \succ f \|_{\mathbb{C}_{T,a}^{\alpha+\beta-1}(\rho_1 \rho_2)} \| b \|_{\mathbb{C}_{T,a}^\gamma(\rho_3)} \\ &\lesssim \| \phi \|_{\mathbb{C}_{T,a}^{\alpha-1}(\rho_1)} \| f \|_{\mathbb{C}_{T,a}^\beta(\rho_2)} \| b \|_{\mathbb{C}_{T,a}^\gamma(\rho_3)}. \end{aligned}$$

By (2.28), (2.29) and (3.37), we have

$$\begin{aligned} \| b \circ [\nabla_v \mathcal{J}_\lambda, \phi \prec] f \|_{\mathbb{C}_{T,a}^{\alpha+\beta+\gamma}(\rho_1 \rho_2 \rho_3)} &\lesssim \| [\nabla_v \mathcal{J}_\lambda, \phi \prec] f \|_{\mathbb{C}_{T,a}^{\alpha+\beta}(\rho_1 \rho_2)} \| b \|_{\mathbb{C}_{T,a}^\gamma(\rho_3)} \\ &\lesssim \| \phi \|_{\mathbb{S}_T^{\alpha-1}(\rho_1)} \| f \|_{\mathbb{C}_{T,a}^\beta(\rho_2)} \| b \|_{\mathbb{C}_{T,a}^\gamma(\rho_3)}. \end{aligned}$$

By (2.31), (3.34) and (2.29), we have

$$\begin{aligned} \| \text{com}(\phi, \nabla_v \mathcal{J}_\lambda f, b) \|_{\mathbb{C}_{T,a}^{\alpha+\beta+\gamma}(\rho_1 \rho_2 \rho_3)} &\lesssim \| \phi \|_{\mathbb{C}_{T,a}^{\alpha-1}(\rho_3)} \| \nabla_v \mathcal{J}_\lambda f \|_{\mathbb{C}_{T,a}^{1+\beta}(\rho_2)} \| b \|_{\mathbb{C}_{T,a}^\gamma(\rho_1)} \\ &\lesssim \| \phi \|_{\mathbb{C}_{T,a}^{\alpha-1}(\rho_3)} \| f \|_{\mathbb{C}_{T,a}^\beta(\rho_2)} \| b \|_{\mathbb{C}_{T,a}^\gamma(\rho_1)}. \end{aligned}$$

Combining the above calculations, we obtain the desired estimate. \square

3.5. Renormalized pairs. In this subsection we introduce the renormalized pairs. Fix $\alpha \in (\frac{1}{2}, \frac{2}{3})$ and $\rho_1, \rho_2 \in \mathcal{P}_w$. For $T > 0$, let $b = (b_1, \dots, b_d)$ and f be $d + 1$ -distributions in $\mathbb{C}_{T,a}^{-\alpha}(\rho_1)$ and $\mathbb{C}_{T,a}^{-\alpha}(\rho_2)$ respectively. We introduce the following important quantity for later use: for $q \in [1, \infty]$,

$$(3.39) \quad \mathbb{A}_{T,q}^{b,f}(\rho_1, \rho_2) := \sup_{\lambda \geq 0} \| b \circ \nabla_v \mathcal{J}_\lambda f \|_{L_T^q \mathbb{C}_a^{1-2\alpha}(\rho_1 \rho_2)} + (\| b \|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_1)} + 1) \| f \|_{L_T^q \mathbb{C}_a^{-\alpha}(\rho_2)}.$$

By (2.28), $b(t) \circ \nabla_v \mathcal{J}_\lambda f(t)$ is not well-defined for $\alpha > \frac{1}{2}$ since by Schauder's estimate, we only have (see Lemma 3.11)

$$\nabla_v \mathcal{J}_\lambda f \in \mathbb{C}_{T,a}^{1-\alpha}(\rho_2).$$

However, in the probabilistic sense, it is possible to give a meaning to $b \circ \nabla_v \mathcal{J}_\lambda f$ when b, f are some Gaussian noises (see Section 7 for general probabilistic assumptions and examples for Gaussian noises to satisfy the requirement in Definition 3.15 below). This motivates us to introduce the following notion.

DEFINITION 3.15. We call the above $(b, f) \in \mathbb{C}_{T,a}^{-\alpha}(\rho_1) \times \mathbb{C}_{T,a}^{-\alpha}(\rho_2)$ a renormalized pair if there exists a sequence of $(b_n, f_n) \in L_T^\infty C_b^\infty \times L_T^\infty C_b^\infty$ with

$$(3.40) \quad \sup_{n \in \mathbb{N}} \mathbb{A}_{T,\infty}^{b_n, f_n}(\rho_1, \rho_2) < \infty$$

and such that

$$(3.41) \quad \lim_{n \rightarrow \infty} (\|b_n - b\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_1)} + \|f_n - f\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_2)}) = 0,$$

and for each $\lambda \geq 0$, there exists a distribution $b \circ \nabla_v \mathcal{J}_\lambda f \in \mathbb{C}_{T,a}^{1-2\alpha}(\rho_1 \rho_2)$ such that

$$(3.42) \quad \lim_{n \rightarrow \infty} \|b_n \circ \nabla_v \mathcal{J}_\lambda f_n - b \circ \nabla_v \mathcal{J}_\lambda f\|_{\mathbb{C}_{T,a}^{1-2\alpha}(\rho_1 \rho_2)} = 0.$$

The set of all the above renormalized pair is denoted by $\mathbb{B}_T^\alpha(\rho_1, \rho_2)$. If for each $i = 1, \dots, d$, $(b, b_i) \in \mathbb{B}_T^\alpha(\rho_1, \rho_1)$, we simply say $b \in \mathbb{B}_T^\alpha(\rho_1)$, a renormalized vector field.

A renormalized pair $(b, f) \in \mathbb{B}_T^\alpha(\rho_1, \rho_2)$ is always associated with certain approximation sequence $(b_n, f_n)_{n \in \mathbb{N}}$. The key point is of course the convergence in (3.42), which in general does not imply that for $(b, f), (b', f) \in \mathbb{B}_T^\alpha(\rho_1, \rho_2)$,

$$(b + b', f) \in \mathbb{B}_T^\alpha(\rho_1, \rho_2).$$

In other words, $\mathbb{B}_T^\alpha(\rho_1, \rho_2)$ is not a linear space. But we have the following easy lemma.

LEMMA 3.16. For $(b, f) \in \mathbb{B}_T^\alpha(\rho_1, \rho_2)$ and $b' \in \mathbb{C}_{T,a}^\beta(\rho_1)$ with $\beta > \alpha - 1$, we have

$$(b + b', f) \in \mathbb{B}_T^\alpha(\rho_1, \rho_2).$$

PROOF. Let $(b_n, f_n)_{n \in \mathbb{N}}$ be the approximation sequence in the definition of $(b, f) \in \mathbb{B}_T^\alpha(\rho_1, \rho_2)$. Let φ_n be any mollifiers in \mathbb{R}^{2d} and define $b'_n(t, \cdot) := b'(t, \cdot) * \varphi_n(\cdot)$. By definition, it is easy to see that

$$\sup_{n \in \mathbb{N}} \mathbb{A}_{T,\infty}^{b_n + b'_n, f_n}(\rho_1, \rho_2) \leq \sup_{n \in \mathbb{N}} (\mathbb{A}_{T,\infty}^{b_n, f_n}(\rho_1, \rho_2) + \mathbb{A}_{T,\infty}^{b'_n, f_n}(\rho_1, \rho_2)) < \infty.$$

For any $\gamma \in (\alpha - 1, \beta)$, by (2.24) we clearly have

$$\lim_{n \rightarrow \infty} \|b'_n - b'\|_{\mathbb{C}_{T,a}^\gamma(\rho_1)} = 0,$$

and by (2.28) and (3.41),

$$\begin{aligned} & \lim_{n \rightarrow \infty} \|b'_n \circ \nabla_v \mathcal{J}_\lambda f_n - b' \circ \nabla_v \mathcal{J}_\lambda f\|_{\mathbb{C}_{T,a}^0(\rho_1 \rho_2)} \\ & \leq \lim_{n \rightarrow \infty} \|b'_n\|_{\mathbb{C}_{T,a}^\gamma(\rho_1)} \|\nabla_v \mathcal{J}_\lambda (f_n - f)\|_{\mathbb{C}_{T,a}^{1-\alpha}(\rho_2)} \\ & \quad + \lim_{n \rightarrow \infty} \|b'_n - b'\|_{\mathbb{C}_{T,a}^\gamma(\rho_1)} \|\nabla_v \mathcal{J}_\lambda f\|_{\mathbb{C}_{T,a}^{1-\alpha}(\rho_2)} = 0. \end{aligned}$$

The proof is complete. \square

To eliminate the parameter λ in (3.39), we give the following lemma, the proof of which follows from [ZZZZ22, Lemma 2.16].

LEMMA 3.17. Let $\mathcal{J}_s^t(f) = \int_s^t P_{t-r} f(r) dr$. For any $t > 0$, we have

$$(3.43) \quad \sup_{\lambda \geq 0} \|b(t) \circ \nabla_v \mathcal{J}_\lambda f(t)\|_{\mathbb{C}_a^{1-2\alpha}(\rho)} \leq 2 \sup_{s \in [0, t]} \|b(t) \circ \nabla_v \mathcal{J}_s^t(f)\|_{\mathbb{C}_a^{1-2\alpha}(\rho)}.$$

The following localized property about the operation \circ is useful.

LEMMA 3.18. *Let $T > 0$, $\rho_1, \rho_2, \rho_3, \rho_4 \in \mathcal{P}_w$, $\alpha \in (\frac{1}{2}, \frac{2}{3})$ and $\gamma \in (\alpha, 1)$. Suppose*

$$(b, f) \in \mathbb{B}_T^\alpha(\rho_1, \rho_2), \quad \psi \in \mathbb{C}_{T,a}^\gamma(\rho_3), \quad \phi \in \mathbb{S}_{T,a}^\gamma(\rho_4).$$

Then $(b\psi, f\phi) \in \mathbb{B}_T^\alpha(\rho_1\rho_3, \rho_2\rho_4)$ with approximation sequence $(b_n\psi, f_n\phi)$, and there is a $C > 0$ depending only on $T, \gamma, \alpha, d, \rho_i$ such that for all $\lambda \geq 0$,

$$(3.44) \quad \begin{aligned} & \| (b\psi) \circ \nabla_v \mathcal{J}_\lambda(f\phi) - \psi\phi(b \circ \nabla_v \mathcal{J}_\lambda f) \|_{\mathbb{C}_{T,a}^{1+\gamma-2\alpha}(\rho_1\rho_2\rho_3\rho_4)} \\ & \lesssim_C \|b\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_1)} \|f\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_2)} \|\psi\|_{\mathbb{C}_{T,a}^\gamma(\rho_3)} \|\phi\|_{\mathbb{S}_{T,a}^\gamma(\rho_4)}. \end{aligned}$$

PROOF. By approximation, we only prove the a priori estimate (3.44). Note that

$$\begin{aligned} I &:= (b\psi) \circ \nabla_v \mathcal{J}_\lambda(f\phi) - \psi\phi(b \circ \nabla_v \mathcal{J}_\lambda f) \\ &= [(b\psi) \circ \nabla_v \mathcal{J}_\lambda, \phi]f + \phi[\nabla_v \mathcal{J}_\lambda f \circ, \psi]b =: I_1 + I_2. \end{aligned}$$

For I_1 , since $1 + \gamma - 2\alpha > 0$ and $1 - 2\alpha < 0$, $\gamma + 1 - \alpha > 1$ by Lemma 3.14, we have

$$\begin{aligned} \|I_1\|_{\mathbb{C}_{T,a}^{1+\gamma-2\alpha}(\rho_1\rho_2\rho_3\rho_4)} & \lesssim \|\phi\|_{\mathbb{S}_{T,a}^\gamma(\rho_4)} \|f\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_2)} \|b\psi\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_1\rho_3)} \\ & \lesssim \|\phi\|_{\mathbb{S}_{T,a}^\gamma(\rho_4)} \|f\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_2)} \|b\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_1)} \|\psi\|_{\mathbb{C}_{T,a}^\gamma(\rho_3)}. \end{aligned}$$

For I_2 , similarly by (2.32), we have

$$\begin{aligned} \|I_2\|_{\mathbb{C}_{T,a}^{1+\gamma-2\alpha}(\rho_1\rho_2\rho_3\rho_4)} & \lesssim \|\phi\|_{\mathbb{C}_{T,a}^\gamma(\rho_4)} \|[\nabla_v \mathcal{J}_\lambda f \circ, \psi]b\|_{\mathbb{C}_{T,a}^{1+\gamma-2\alpha}(\rho_1\rho_2\rho_3)} \\ & \lesssim \|\phi\|_{\mathbb{C}_{T,a}^\gamma(\rho_4)} \|\nabla_v \mathcal{J}_\lambda f\|_{\mathbb{C}_{T,a}^{1-\alpha}(\rho_2)} \|b\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_1)} \|\psi\|_{\mathbb{C}_{T,a}^\gamma(\rho_3)}. \end{aligned}$$

Thus we obtain (3.44) by (3.35). \square

4. Linear kinetic equations with distribution drifts. Now that the necessary facts about the kinetic semigroup and weighted Besov spaces are established, the next two sections are devoted to the actual construction of the solution to the stochastic kinetic equation. The aim of this section is to show the well-posedness of the following linear singular kinetic equation: for $\lambda \geq 0$,

$$(4.1) \quad \mathcal{L}_\lambda u := (\partial_t - \Delta_v - v \cdot \nabla_x + \lambda)u = b \cdot \nabla_v u + f, \quad u(0) = \varphi,$$

where $b = (b_1, \dots, b_d)$ and f satisfy that for some $\alpha \in (\frac{1}{2}, \frac{2}{3})$ and $\rho_1, \rho_2 \in \mathcal{P}_w$,

$$(4.2) \quad \begin{cases} (b, f) \in \mathbb{B}_T^\alpha(\rho_1, \rho_2), \quad b \in \mathbb{B}_T^\alpha(\rho_1), \quad T > 0, \\ \text{have the same approximation sequence } (b_n, f_n). \end{cases}$$

For simplicity of notations, we shall write

$$\ell_T^b(\rho_1) := \sum_{i=1}^d \mathbb{A}_{T,\infty}^{b,b_i}(\rho_1, \rho_1) + 1.$$

We also write

$$\ell_T^b = \ell_T^b(1) \quad \mathbb{A}_{T,q}^{b,b} = \sum_{i=1}^d \mathbb{A}_{T,q}^{b,b_i}(1, 1).$$

In Subsection 4.1, we first introduce the notion of paracontrolled solutions, and then establish a localization property for paracontrolled solutions. Such a localization is natural for

classical solutions by the chain rule. However, for paracontrolled solutions, it is quite involved since the renormalized pair is defined in the approximation level. In Subsection 4.2, following the same argument as in [ZZZ22, Section 3] and using estimate and commutators for the kinetic semigroup, we show the well-posedness for PDE (4.2) in weighted anisotropic Hölder spaces. We emphasize that unlike using the exponential weight technique in [ZZZ22, Section 3], the uniqueness is a direct consequence of the a priori estimate (4.31) below.

4.1. *Paracontrolled solutions.* To introduce the paracontrolled solution of PDE (4.1), we make the following paracontrolled ansatz as in [GIP15]:

$$(4.3) \quad u = P_t \varphi + u^\sharp + \nabla_v u \prec \mathcal{J}_\lambda b + \mathcal{J}_\lambda f,$$

where u^\sharp solves the following equation

$$(4.4) \quad u^\sharp = \mathcal{J}_\lambda (\nabla_v u \succ b + b \circ \nabla_v u) + [\mathcal{J}_\lambda, \nabla_v u \prec] b.$$

Note that $b \circ \nabla_v u$ is not well-defined in the classical sense. We give its definition by paracontrolled ansatz and renormalized pair as follows: By (4.3), we can write

$$(4.5) \quad \begin{aligned} b \circ \nabla_v u &= b \circ \nabla_v u^\sharp + b \circ \nabla_v (\nabla_v u \prec \mathcal{J}_\lambda b) + b \circ \nabla_v \mathcal{J}_\lambda f + b \circ \nabla_v P_t \varphi \\ &= b \circ \nabla_v u^\sharp + b \circ (\nabla_v^2 u \prec \mathcal{J}_\lambda b) + (b \circ \nabla_v \mathcal{J}_\lambda b) \cdot \nabla_v u \\ &\quad + \text{com}(\nabla_v u, \nabla_v \mathcal{J}_\lambda b, b) + b \circ \nabla_v \mathcal{J}_\lambda f + b \circ \nabla_v P_t \varphi. \end{aligned}$$

This motivates us to introduce the following definition.

DEFINITION 4.1. Let $T > 0$, $\rho_1 \in \mathcal{P}_w$ be a bounded weight and $\rho_2, \rho_3 \in \mathcal{P}_w$ be any weights. Under (4.2) and $\varphi \in \mathbf{C}_a^{1+\alpha+\varepsilon}(\rho_2/\rho_1)$ for some $\varepsilon > 0$, we call $u \in \mathbb{S}_T^{2-\alpha}(\rho_3)$ a paracontrolled solution of PDE (4.1) corresponding to (b, f) if for some $\rho_4 \in \mathcal{P}_w$,

$$(4.6) \quad u - P_t \varphi - \nabla_v u \prec \mathcal{J}_\lambda b - \mathcal{J}_\lambda f =: u^\sharp \in \mathbb{C}_{T,a}^{3-2\alpha}(\rho_4)$$

satisfies (4.4) with $b \circ \nabla_v u$ given by (4.5) which is well-defined by (4.8) below.

REMARK 4.2. In the above definition, if we consider $\bar{u} = u - P_t \varphi$, then the initial value is reduced to zero. In this case, the nonhomogeneous f shall be replaced by

$$\bar{f} = f + b \cdot \nabla_v P_t \varphi \in \mathbb{C}_{T,a}^{-\alpha}(\rho_2).$$

By Lemma 3.18 with $\psi = 1, \phi = \nabla_v P_t \varphi$ and $\rho_3 = 1, \rho_4 = \rho_2/\rho_1$,

$$\|b \circ \nabla_v \mathcal{J}_\lambda (b \cdot \nabla_v P_t \varphi)\|_{\mathbb{C}_{T,a}^{1-2\alpha}(\rho_1 \rho_2)} \lesssim \|\varphi\|_{\mathbf{C}_a^{1+\alpha+\varepsilon}(\rho_2/\rho_1)} \ell_T^b(\rho_1).$$

Thus, we still have

$$(b, \bar{f}) \in \mathbb{B}_T^\alpha(\rho_1, \rho_2),$$

and $(\bar{u}, \bar{u}^\sharp)$ is a paracontrolled solution of (4.1) with $f = \bar{f}$ and $\bar{u}(0) = 0$, where

$$\bar{u}^\sharp = u^\sharp + \nabla_v P_t \varphi \prec \mathcal{J}_\lambda b - \mathcal{J}_\lambda (b \cdot \nabla_v P_t \varphi).$$

In the following, for simplicity, we may and shall assume $\varphi \equiv 0$ by this procedure.

We have the following a priori estimate about the regularity of u^\sharp .

THEOREM 4.3. *Let $u \in \mathbb{S}_{T,a}^{2-\alpha}(\rho_3)$ be a paracontrolled solution to (4.1) in the sense of Definition 4.1 with $\varphi \equiv 0$. For any $\varepsilon > \frac{2\alpha-1}{2-3\alpha}$ and $\rho_4 := \rho_1^{1+\varepsilon}((\rho_1\rho_3) \wedge \rho_2)$, there is a constant $C = C(T, \varepsilon, \alpha, d, \rho_i, \ell_T^b(\rho_1)) > 0$ such that for all $\lambda \geq 0$,*

$$(4.7) \quad \|u^\sharp\|_{\mathbb{C}_{T,a}^{3-2\alpha}(\rho_4)} \lesssim_C \|u\|_{\mathbb{S}_{T,a}^{2-\alpha}(\rho_3)} + \mathbb{A}_{T,\infty}^{b,f}(\rho_1, \rho_2).$$

PROOF. First of all, we show that for any $\gamma, \beta \in (\alpha, 2-2\alpha]$ and $\rho_5 \leq (\rho_1\rho_3) \wedge \rho_2$,

$$(4.8) \quad \|b \circ \nabla_v u\|_{\mathbb{C}_{T,a}^{1-2\alpha}(\rho_1\rho_5)} \lesssim \ell_T^b(\rho_1) \left(\|u\|_{\mathbb{C}_{T,a}^{\alpha+\gamma}(\rho_3)} + \|u^\sharp\|_{\mathbb{C}_{T,a}^{\beta+1}(\rho_5)} \right) + \mathbb{A}_{T,\infty}^{b,f}(\rho_1, \rho_2).$$

To prove this, it suffices to estimate each term in (4.5).

- Since $\beta > \alpha$, by (2.28), we have

$$\|b \circ \nabla_v u^\sharp\|_{L^\infty(\rho_1\rho_5)} \lesssim \|b\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_1)} \|\nabla_v u^\sharp\|_{\mathbb{C}_{T,a}^\beta(\rho_5)} \leq \ell_T^b(\rho_1) \|u^\sharp\|_{\mathbb{C}_{T,a}^{\beta+1}(\rho_5)}.$$

- Since $\gamma > \alpha$ and $\gamma + \alpha - 2 < 0$, by (2.27), (2.28), we have

$$\begin{aligned} \|b \circ (\nabla_v^2 u \prec \mathcal{J}_\lambda b)\|_{\mathbb{C}_{T,a}^{1-2\alpha}(\rho_1^2\rho_3)} &\lesssim \|b\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_1)} \|\nabla_v^2 u \prec \mathcal{J}_\lambda b\|_{\mathbb{C}_{T,a}^\gamma(\rho_1\rho_3)} \\ &\lesssim \|b\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_1)} \|\nabla_v^2 u\|_{\mathbb{C}_{T,a}^{\gamma+\alpha-2}(\rho_3)} \|\mathcal{J}_\lambda b\|_{\mathbb{C}_{T,a}^{2-\alpha}(\rho_1)} \\ &\lesssim \|b\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_1)}^2 \|u\|_{\mathbb{C}_{T,a}^{\gamma+\alpha}(\rho_3)} \lesssim \ell_T^b(\rho_1) \|u\|_{\mathbb{C}_{T,a}^{\alpha+\gamma}(\rho_3)}. \end{aligned}$$

- Since $\gamma > \alpha$, by (2.29) we have

$$\begin{aligned} \|\nabla_v u(b \circ \nabla_v \mathcal{J}_\lambda b)\|_{\mathbb{C}_{T,a}^{1-2\alpha}(\rho_1^2\rho_3)} &\lesssim \|\nabla_v u\|_{\mathbb{C}_{T,a}^{\gamma+\alpha-1}(\rho_3)} \|b \circ \nabla_v \mathcal{J}_\lambda b\|_{\mathbb{C}_{T,a}^{1-2\alpha}(\rho_1^2)} \\ &\lesssim \ell_T^b(\rho_1) \|u\|_{\mathbb{C}_{T,a}^{\alpha+\gamma}(\rho_3)}. \end{aligned}$$

- Since $\gamma > \alpha$, by (2.31), we have

$$\begin{aligned} \|\text{com}(\nabla_v u, \nabla_v \mathcal{J}_\lambda b, b)\|_{\mathbb{C}_{T,a}^{1-2\alpha}(\rho_1^2\rho_3)} &\lesssim \|b\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_1)} \|\nabla_v u\|_{\mathbb{C}_{T,a}^{\gamma+\alpha-1}(\rho_3)} \|\nabla_v \mathcal{J}_\lambda b\|_{\mathbb{C}_{T,a}^{1-\alpha}(\rho_1)} \\ &\lesssim \|b\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_1)}^2 \|u\|_{\mathbb{C}_{T,a}^{\gamma+\alpha}(\rho_3)} \lesssim \ell_T^b(\rho_1) \|u\|_{\mathbb{C}_{T,a}^{\alpha+\gamma}(\rho_3)}. \end{aligned}$$

Combining the above estimates and by $\rho_1\rho_5 \leq \rho_1^2\rho_3$, we get (4.8).

On the other hand, by (2.27), we have

$$\|\nabla_v u \succ b\|_{\mathbb{C}_{T,a}^{1-2\alpha}(\rho_1\rho_3)} \lesssim \|u\|_{\mathbb{C}_{T,a}^{2-\alpha}(\rho_3)} \|b\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_1)},$$

and by (3.37) with $(k, \theta) = (0, 2)$ and (3.30),

$$\|[\mathcal{J}_\lambda, \nabla_v u \prec] b\|_{\mathbb{C}_{T,a}^{3-2\alpha}(\rho_1\rho_3)} \lesssim \|\nabla_v u\|_{\mathbb{S}_T^{1-\alpha}(\rho_3)} \|b\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_1)} \lesssim \|u\|_{\mathbb{S}_T^{2-\alpha}(\rho_3)} \|b\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_1)}.$$

Thus, by (4.4), (4.8) and Schauder's estimate (3.35), thanks to $\rho_5 \leq \rho_3$, we obtain for $\beta \in (\alpha, 2-2\alpha)$,

$$(4.9) \quad \|u^\sharp\|_{\mathbb{C}_{T,a}^{3-2\alpha}(\rho_1\rho_5)} \lesssim \|u\|_{\mathbb{S}_T^{2-\alpha}(\rho_3)} + \|u^\sharp\|_{\mathbb{C}_{T,a}^{\beta+1}(\rho_5)} + \mathbb{A}_{T,\infty}^{b,f}(\rho_1, \rho_2).$$

For $\varepsilon > \frac{2\alpha-1}{2-3\alpha}$, one can choose β close to α so that

$$\theta := \frac{\varepsilon}{1+\varepsilon} = \frac{\alpha+\beta-1}{1-\alpha}.$$

Let

$$\rho_4 := \rho_1^{1+\varepsilon}((\rho_1\rho_3) \wedge \rho_2), \quad \rho_5 := \rho_4^\theta((\rho_1\rho_3) \wedge \rho_2)^{1-\theta}.$$

Noting that $\rho_1 \rho_5 = \rho_4$, by (2.19) and Young's inequality, we have for any $\delta > 0$,

$$\begin{aligned} \|u^\sharp\|_{\mathbb{C}_{T,a}^{\beta+1}(\rho_5)} &\lesssim \|u^\sharp\|_{\mathbb{C}_{T,a}^{3-2\alpha}(\rho_4)}^\theta \|u^\sharp\|_{\mathbb{C}_{T,a}^{2-\alpha}((\rho_1 \rho_3) \wedge \rho_2)}^{1-\theta} \\ &\leq \delta \|u^\sharp\|_{\mathbb{C}_{T,a}^{3-2\alpha}(\rho_4)} + C_\delta \|u^\sharp\|_{\mathbb{C}_{T,a}^{2-\alpha}((\rho_1 \rho_3) \wedge \rho_2)}. \end{aligned}$$

Substituting this into (4.9) and by $\rho_4 = \rho_1 \rho_5$ and letting δ small enough, we get

$$(4.10) \quad \|u^\sharp\|_{\mathbb{C}_{T,a}^{3-2\alpha}(\rho_4)} \lesssim \|u\|_{\mathbb{S}_T^{2-\alpha}(\rho_3)} + \|u^\sharp\|_{\mathbb{C}_{T,a}^{2-\alpha}((\rho_1 \rho_3) \wedge \rho_2)} + \mathbb{A}_{T,\infty}^{b,f}(\rho_1, \rho_2).$$

On the other hand, by (4.6), (2.27) and (3.34) we have

$$\begin{aligned} \|u^\sharp\|_{\mathbb{C}_{T,a}^{2-\alpha}((\rho_1 \rho_3) \wedge \rho_2)} &\lesssim \|u\|_{\mathbb{C}_{T,a}^{2-\alpha}(\rho_3)} + \|\nabla_v u \prec \mathcal{J}_\lambda b\|_{\mathbb{C}_{T,a}^{2-\alpha}(\rho_1 \rho_3)} + \|\mathcal{J}_\lambda f\|_{\mathbb{C}_{T,a}^{2-\alpha}(\rho_2)} \\ &\lesssim \|u\|_{\mathbb{C}_{T,a}^{2-\alpha}(\rho_3)} + \|\nabla_v u\|_{L_T^\infty(\rho_3)} \|\mathcal{J}_\lambda b\|_{\mathbb{C}_{T,a}^{2-\alpha}(\rho_1)} + \|f\|_{\mathbb{C}_{T,a}^{2-\alpha}(\rho_2)} \\ (4.11) \quad &\lesssim \sqrt{\ell_T^b(\rho_1)} \|u\|_{\mathbb{C}_{T,a}^{2-\alpha}(\rho_3)} + \|f\|_{\mathbb{C}_{T,a}^{2-\alpha}(\rho_2)}. \end{aligned}$$

Substituting this into (4.10), we complete the proof. \square

For the uniqueness part we need the following localization result about the paracontrolled solutions.

PROPOSITION 4.4. Let u be a paracontrolled solution to PDE (4.1) with $\varphi = 0$. Let $\phi, \psi \in C_c^\infty(\mathbb{R}^{2d})$ with $\psi \equiv 1$ on the support of ϕ . Then $\bar{u} := u\phi \in \mathbb{S}_{T,a}^{2-\alpha}$ is also a paracontrolled solution to PDE (4.1) corresponding to $(\bar{b}, g) \in \mathbb{B}_T^\alpha$, where

$$\bar{b} := b\psi, \quad g := \phi f - u\Delta_v \phi - 2\nabla_v \phi \cdot \nabla_v u - (v \cdot \nabla_x \phi)u - (b \cdot \nabla_v \phi)u.$$

PROOF. Without loss of generality we assume that $\lambda = 0$. First of all, we claim $(\bar{b}, g) \in \mathbb{B}_T^\alpha$. Since $\nabla_v \phi u \in \mathbb{S}_{T,a}^{2-\alpha}$, by Lemma 3.18, $(\bar{b}, \phi f - (b \cdot \nabla_v \phi)u) \in \mathbb{B}_T^\alpha$. Note that

$$b' := -u\Delta_v \phi - 2\nabla_v \phi \cdot \nabla_v u - (v \cdot \nabla_x \phi)u \in \mathbb{S}_{T,a}^{1-\alpha} \subset \mathbb{C}_{T,a}^{1-\alpha}.$$

By Lemma 3.16, we have $(\bar{b}, g) = (\bar{b}, \phi f - (b \cdot \nabla_v \phi)u + b') \in \mathbb{B}_T^\alpha$.

By definition, one needs to show that

$$(4.12) \quad \bar{u} - \nabla_v \bar{u} \prec \mathcal{J} \bar{b} - \mathcal{J} g =: \bar{u}^\sharp \in \mathbb{C}_{T,a}^{3-2\alpha}$$

satisfies

$$(4.13) \quad \bar{u}^\sharp = \mathcal{J}(\nabla_v \bar{u} \succ \bar{b} + \bar{b} \circ \nabla_v \bar{u}) + [\mathcal{J}, \nabla_v \bar{u}] \bar{b},$$

with

$$\begin{aligned} (4.14) \quad \bar{b} \circ \nabla_v \bar{u} &:= \bar{b} \circ \nabla_v \bar{u}^\sharp + \bar{b} \circ (\nabla_v^2 \bar{u} \prec \mathcal{J} \bar{b}) + (\bar{b} \circ \nabla_v \mathcal{J} \bar{b}) \cdot \nabla_v \bar{u} \\ &\quad + \text{com}(\nabla_v \bar{u}, \nabla_v \mathcal{J} \bar{b}, \bar{b}) + \bar{b} \circ \nabla_v \mathcal{J} g. \end{aligned}$$

Since u is a paracontrolled solution, by definition we have

$$(4.15) \quad u = \mathcal{J}(b \star \nabla_v u + f),$$

where

$$(4.16) \quad b \star \nabla_v u := \nabla_v u \succ b + b \circ \nabla_v u + \nabla_v u \prec b.$$

Let $(b_n, f_n) \in L_T^\infty C_b^\infty$ be as in (4.2). We introduce an approximation of u by

$$(4.17) \quad u_n := u^\sharp + \nabla_v u \prec \mathcal{J} b_n + \mathcal{J} f_n, \quad \bar{b}_n := b_n \psi, \quad \bar{u}_n := u_n \phi,$$

and

$$(4.18) \quad \bar{b} \odot \nabla_v \bar{u} := (b \star \nabla_v u) \phi + (b \cdot \nabla_v \phi) u - \nabla_v \bar{u} \prec \bar{b} - \nabla_v \bar{u} \succ \bar{b}.$$

In the classical case, it is easy to see $\bar{b} \odot \nabla_v \bar{u} = \bar{b} \circ \nabla_v \bar{u}$. In the paracontrolled case this is not obvious and we introduce $\bar{b} \odot \nabla_v \bar{u}$ which can be easily checked as limit of $\bar{b}_n \circ \nabla_v \bar{u}_n$ (see step (ii) below). Moreover, it is not hard to prove that \bar{u}^\sharp satisfies (4.13) with $\bar{b} \circ \nabla_v \bar{u}$ replaced by $\bar{b} \odot \nabla_v \bar{u}$ (see step (iii) below). Finally we use approximations to prove $\bar{b} \odot \nabla_v \bar{u} = \bar{b} \circ \nabla_v \bar{u}$ (see step (iv) below). Our proof is divided into the following four steps:

(i) We show that u_n is a suitable approximation of u and for some $\rho \in \mathcal{P}_w$,

$$(4.19) \quad \lim_{n \rightarrow \infty} \|b_n \cdot \nabla_v u_n - b \star \nabla_v u\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho)} = 0.$$

(ii) We prove $\bar{b} \odot \nabla_v \bar{u} \in \mathbb{C}_{T,a}^{1-2\alpha}$ and

$$(4.20) \quad \lim_{n \rightarrow \infty} \|\bar{b}_n \circ \nabla_v \bar{u}_n - \bar{b} \odot \nabla_v \bar{u}\|_{\mathbb{C}_{T,a}^{-\alpha}} = 0.$$

(iii) We show that for \bar{u}^\sharp being defined by (4.12) satisfies the following,

$$(4.21) \quad \mathbb{C}_{T,a}^{3-2\alpha} \ni \bar{u}^\sharp = \mathcal{J}(\nabla_v \bar{u} \succ \bar{b} + \bar{b} \odot \nabla_v \bar{u}) + [\mathcal{J}, \nabla_v \bar{u} \prec] \bar{b}.$$

(iv) With $\bar{b} \odot \nabla_v \bar{u}$ being defined by (4.14), we prove

$$(4.22) \quad \bar{b} \odot \nabla_v \bar{u} = \bar{b} \circ \nabla_v \bar{u}.$$

Proof of (i): First of all, by (4.6), (4.17), (2.26) and (3.34), we have

$$\|u_n - u\|_{\mathbb{C}_{T,a}^{2-\alpha}((\rho_1 \rho_3) \wedge \rho_2)} \lesssim \|\nabla_v u\|_{\mathbb{L}_T^\infty(\rho_3)} \|b_n - b\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_1)} + \|f_n - f\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_2)},$$

which implies by (3.41) that

$$(4.23) \quad \lim_{n \rightarrow \infty} \|u_n - u\|_{\mathbb{C}_{T,a}^{2-\alpha}((\rho_1 \rho_3) \wedge \rho_2)} = 0.$$

Next, by (2.27), (4.23) and (3.41), we also have for some $\rho \in \mathcal{P}_w$,

$$(4.24) \quad \lim_{n \rightarrow \infty} \|b_n \prec \nabla_v u_n - b \prec \nabla_v u\|_{\mathbb{C}_{T,a}^{1-2\alpha}(\rho)} = 0,$$

and by (2.26), (4.23) and (3.41),

$$(4.25) \quad \lim_{n \rightarrow \infty} \|b_n \succ \nabla_v u_n - b \succ \nabla_v u\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho)} = 0.$$

Moreover, note that by (4.17),

$$\begin{aligned} b_n \circ \nabla_v u_n &= b_n \circ \nabla_v u^\sharp + b_n \circ (\nabla_v^2 u \prec \mathcal{J} b_n) + (b_n \circ \nabla_v \mathcal{J} b_n) \cdot \nabla_v u \\ &\quad + \text{com}(\nabla_v u, \nabla_v \mathcal{J} b_n, b_n) + b_n \circ \nabla_v \mathcal{J} f_n. \end{aligned}$$

By (3.41), (3.42) and Lemmas 2.11 and 2.13, it is easy to see that each term of the above RHS converges to the one in (4.5) in $\mathbb{C}_{T,a}^{1-2\alpha}(\rho)$ for some $\rho \in \mathcal{P}_w$. Thus,

$$(4.26) \quad \lim_{n \rightarrow \infty} \|b_n \circ \nabla_v u_n - b \circ \nabla_v u\|_{\mathbb{C}_{T,a}^{1-2\alpha}(\rho)} = 0.$$

Since $-\alpha < 1 - 2\alpha$, combining (4.24), (4.25) and (4.26), we obtain (4.19).

Proof of (ii): In this step we first use the chain rule for approximations and then take the limit. Since $\psi \phi = \phi$, by the chain rule we have

$$\bar{b}_n \cdot \nabla_v \bar{u}_n = (b_n \psi) \cdot \nabla_v (u_n \phi) = (b_n \cdot \nabla_v u_n) \phi + (b_n \cdot \nabla_v \phi) u_n.$$

Hence, by Bony's decomposition,

$$\bar{b}_n \circ \nabla_v \bar{u}_n = (b_n \cdot \nabla_v u_n) \phi + (b_n \cdot \nabla_v \phi) u_n - \nabla \bar{u}_n \prec \bar{b}_n - \nabla \bar{u}_n \succ \bar{b}_n.$$

Since $\phi, \psi \in C_c^\infty(\mathbb{R}^{2d})$, by (3.41) and (4.23), we have

$$\lim_{n \rightarrow \infty} \|(b_n \cdot \nabla_v \phi) u_n - (b \cdot \nabla_v \phi) u\|_{\mathbb{C}_{T,a}^{1-2\alpha}} = 0,$$

and by Lemma 2.11,

$$\lim_{n \rightarrow \infty} \|\bar{b}_n \prec \nabla_v \bar{u}_n - \bar{b} \prec \nabla_v \bar{u}\|_{\mathbb{C}_{T,a}^{1-2\alpha}} = 0$$

$$\lim_{n \rightarrow \infty} \|\bar{b}_n \succ \nabla_v \bar{u}_n - \bar{b} \succ \nabla_v \bar{u}\|_{\mathbb{C}_{T,a}^{-\alpha}} = 0,$$

which together with (4.19) and (4.18) yields (4.20). On the other hand, we use regularity of $\bar{b}_n \circ \nabla_v \bar{u}_n$ to improve the regularity. Note that

$$\begin{aligned} \bar{b}_n \circ \nabla_v \bar{u}_n &= (b_n \psi) \circ (\nabla_v u_n \phi) + (b_n \psi) \circ (\nabla_v \phi u_n) \\ &= [(\nabla_v u_n \phi) \circ, \psi] b_n + \psi [b_n \circ, \phi] \nabla_v u_n \\ &\quad + \psi \phi (b_n \circ \nabla_v u_n) + (b_n \psi) \circ (\nabla_v \phi u_n). \end{aligned}$$

Moreover, by (4.20), (2.32) and (4.26), one sees that

$$(4.27) \quad \|\bar{b} \odot \nabla_v \bar{u}\|_{\mathbb{C}_{T,a}^{1-2\alpha}} \leq \sup_n \|\bar{b}_n \circ \nabla_v \bar{u}_n\|_{\mathbb{C}_{T,a}^{1-2\alpha}} < \infty.$$

Proof of (iii): By the chain rule, we have in the distributional sense

$$\mathcal{L} \bar{u} = \mathcal{L}(u \phi) = \phi \mathcal{L} u - u \Delta_v \phi - 2 \nabla_v \phi \cdot \nabla_v u - (v \cdot \nabla_x \phi) u.$$

Taking the inverse $\mathcal{L}^{-1} = \mathcal{J}$, and by (4.15) and definition (4.18), we get

$$\begin{aligned} \bar{u} &= \mathcal{J}((b \star \nabla_v u + f) \phi - u \Delta_v \phi - 2 \nabla_v \phi \cdot \nabla_v u - (v \cdot \nabla_x \phi) u) \\ &= \mathcal{J}(\bar{b} \odot \nabla_v \bar{u} + \nabla_v \bar{u} \prec \bar{b} + \nabla_v \bar{u} \succ \bar{b} + g), \end{aligned}$$

which, combining with definition (4.12), yields (4.21). Moreover, since by (2.27) and (4.27),

$$\nabla_v \bar{u} \succ \bar{b} + \bar{b} \odot \nabla_v \bar{u} \in \mathbb{C}_{T,a}^{1-2\alpha},$$

by (3.34) and (3.37), we clearly have

$$(4.28) \quad \bar{u}^\sharp \in \mathbb{C}_{T,a}^{3-2\alpha}.$$

Proof of (iv): To show (4.22), we first find a suitable approximation for $\bar{b} \circ \nabla_v \bar{u}$. Let

$$g_n := f_n \phi - u \Delta_v \phi - 2 \nabla_v \phi \cdot \nabla_v u - (b_n \cdot \nabla_v \phi) u - (v \cdot \nabla_x \phi) u.$$

By Lemmas 3.16 and 3.18, one sees that $(b_n \psi, g_n)$ is the approximation sequence of (\bar{b}, g) and $g_n \rightarrow g$ in $\mathbb{C}_{T,a}^{-\alpha}$. Noting that

$$\begin{aligned} \bar{b}_n \circ \nabla_v (\bar{u}^\sharp + \nabla_v \bar{u} \prec \mathcal{J} \bar{b}_n + \mathcal{J} g_n) &= \bar{b}_n \circ \nabla_v \bar{u}^\sharp + \bar{b}_n \circ (\nabla_v^2 \bar{u} \prec \mathcal{J} \bar{b}_n) + (\bar{b}_n \circ \nabla_v \mathcal{J} \bar{b}_n) \cdot \nabla_v \bar{u} \\ &\quad + \text{com}(\nabla_v \bar{u}, \nabla_v \mathcal{J} \bar{b}_n, \bar{b}_n) + \bar{b}_n \circ \nabla_v \mathcal{J} g_n, \end{aligned}$$

by (4.28), (3.41), (3.42), Lemmas 2.11, 2.13 and some tedious calculations, we have

$$(4.29) \quad \lim_{n \rightarrow \infty} \bar{b}_n \circ \nabla_v (\bar{u}^\sharp + \nabla_v \bar{u} \prec \mathcal{J} \bar{b}_n + \mathcal{J} g_n) = \bar{b} \circ \nabla_v \bar{u} \text{ in } \mathbb{C}_{T,a}^{1-2\alpha}.$$

Here we use the decomposition in Lemma 3.18 to deduce the convergence of $\bar{b}_n \circ \nabla_v \mathcal{J} \bar{b}_n$ to $\bar{b} \circ \nabla_v \mathcal{J} \bar{b}$. Hence, by (4.20) and (4.29), it remains to prove that in suitable space,

$$(4.30) \quad \lim_{n \rightarrow \infty} \bar{b}_n \circ \nabla_v (\bar{u}_n - \bar{u}^\sharp - \nabla_v \bar{u} \prec \mathcal{J} \bar{b}_n - \mathcal{J} g_n) =: \lim_{n \rightarrow \infty} \Lambda_n = 0.$$

Note that by (4.12),

$$\bar{u}^\sharp = \bar{u} - \nabla_v \bar{u} \prec \mathcal{J} \bar{b} - \mathcal{J} g = \phi \left(u^\sharp + (\nabla_v u \prec \mathcal{J} b) + \mathcal{J} f \right) - \nabla_v \bar{u} \prec \mathcal{J} \bar{b} - \mathcal{J} g,$$

which together with (4.17) yields

$$\Lambda_n = \bar{b}_n \circ \nabla_v ((\nabla_v u \prec \mathcal{J} B_n) \phi - \nabla_v \bar{u} \prec \mathcal{J} (B_n \psi) + \phi \mathcal{J} F_n - \mathcal{J} G_n),$$

where

$$B_n := b_n - b, \quad F_n := f_n - f, \quad G_n := g_n - g.$$

By commutator estimates (see Lemmas 2.11 and 2.13) and (3.41), (3.42), it is easy to see that

$$\lim_{n \rightarrow \infty} \left(\bar{b}_n \circ \nabla_v ((\nabla_v u \prec \mathcal{J} B_n) \phi) - \phi \nabla_v u (\bar{b}_n \circ \nabla_v \mathcal{J} B_n) \right) = 0$$

and

$$\lim_{n \rightarrow \infty} \left(\bar{b}_n \circ \nabla_v (\nabla_v \bar{u} \prec \mathcal{J} (B_n \psi)) - \psi \nabla_v \bar{u} (\bar{b}_n \circ \nabla_v \mathcal{J} B_n) \right) = 0.$$

Moreover, noting that

$$\phi \mathcal{J} F_n - \mathcal{J} G_n = -[\mathcal{J}, \phi] F_n + \mathcal{J} (B_n \cdot \nabla_v \phi u),$$

by Lemma 3.14 and Lemma 3.13, we also have

$$\lim_{n \rightarrow \infty} \left(\bar{b}_n \circ \nabla_v (\phi \mathcal{J} F_n - \mathcal{J} G_n) - (\nabla_v \phi u) (\bar{b}_n \circ \nabla_v \mathcal{J} B_n) \right) = 0.$$

Finally, since $\psi \nabla_v \bar{u} = \nabla_v (\phi u)$, we have

$$(\psi \nabla_v \bar{u} - \phi \nabla_v u - \nabla_v \phi u) (\bar{b}_n \circ \nabla_v \mathcal{J} B_n) \equiv 0,$$

which together with the above three limits yields (4.30). The proof is complete. \square

REMARK 4.5. The above result clearly holds for classical solutions by the chain rule. However, for the paracontrolled solution we cannot directly apply the chain rule since the paracontrolled solution is in the renormalized sense, i.e., $b \cdot \nabla_v \mathcal{J} b$ and $b \cdot \nabla_v \mathcal{J} f$ are understood in the approximation sense. Therefore, we have to first construct suitable smooth approximations for the solution so that we can use the chain rule. In the last step, an obvious difficulty is that although

$$\lim_{n \rightarrow \infty} \|b_n - b\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho)} = 0, \quad \lim_{n \rightarrow \infty} \|b_n \circ \nabla_v \mathcal{J} b_n - b \circ \nabla_v \mathcal{J} b\|_{\mathbb{C}_{T,a}^{1-2\alpha}(\rho)} = 0,$$

it does not imply that

$$\lim_{n \rightarrow \infty} b_n \circ \nabla_v \mathcal{J} (b_n - b) = 0 \text{ in any space.}$$

4.2. *Well-posedness for (4.1).* The aim of this section is to show the well-posedness of PDE (4.1). To this end we first establish the following Schauder estimate with the coefficients in unweighted kinetic Hölder spaces. The proofs are essentially the same as in [ZZZ22, Section 3.2] by using Lemmas 3.10, 3.11, 3.13 and Theorem 4.3. In the following we omit the details of the proof and only recall the idea of proof. The proof is divided into two steps: first we prove a Schauder estimate depending polynomially on the coefficients for λ large enough. Then by a classical maximum principle we extend the result to all $\lambda \geq 0$. In the case of (4.1) the maximal principle is easy when $b, f \in L_T^\infty \mathbf{C}_b^\infty$, as the fundamental solution exists in this case (see [DM10]). The only difference is that we do not introduce the notion \ll . Instead by similar argument as in the proof of Theorem 4.7 below we obtain the following results.

THEOREM 4.6. *Let $T > 0$ and $\varphi = 0$. For any $(b, f) \in \mathbb{B}_T^\alpha$, there is a unique paracontrolled solution u to PDE (4.1) in the sense of Definition 4.1. Moreover, there are $q > 1$ large enough only depending on α and $c_1, c_2 > 0$ such that*

$$\|u\|_{\mathbb{L}_T^\infty} \leq c_1 (\ell_T^b)^{\frac{5}{2-3\alpha}} \mathbb{A}_{T,q}^{b,f}, \quad \|u\|_{\mathbb{S}_{T,a}^{2-\alpha}} \leq c_2 (\ell_T^b)^{\frac{9}{2-3\alpha}} \mathbb{A}_{T,\infty}^{b,f}.$$

By using the characterization of $\mathbb{S}_{T,a}^\alpha(\rho)$ established in Lemma 3.9 and Theorem 4.6 we obtain the following well-posedness of PDE (4.1) in weighted kinetic Hölder spaces.

THEOREM 4.7. *Let $\alpha \in (\frac{1}{2}, \frac{2}{3})$ and $\vartheta := \frac{9}{2-3\alpha}$. Let $\kappa_1 > 0$ and $\kappa_2 \in \mathbb{R}$ with*

$$(2\vartheta + 2)\kappa_1 \leq 1, \quad \kappa_3 := (2\vartheta + 1)\kappa_1 + \kappa_2.$$

With notations in (3.5), let

$$\rho_i := \varrho^{\kappa_i} \in \mathcal{P}_w, \quad i = 1, 2, 3.$$

Under (4.2), for any $T > 0$ and $\varphi \in \mathbf{C}_a^\gamma(\rho_2/\rho_1)$, where $\gamma > 1 + \alpha$, there is a unique paracontrolled solution $u \in \mathbb{S}_{T,a}^{2-\alpha}(\rho_3)$ to PDE (4.1) in the sense of Definition 4.1 so that

$$(4.31) \quad \|u\|_{\mathbb{S}_{T,a}^{2-\alpha}(\rho_3)} \lesssim_C \|\varphi\|_{\mathbf{C}_a^\gamma(\rho_2/\rho_1)} + \mathbb{A}_{T,\infty}^{b,f}(\rho_1, \rho_2),$$

where $C = C(T, d, \alpha, \kappa_i, \ell_T^b(\rho_1)) > 0$. Moreover, let $(b_n, f_n) \in L_T^\infty \mathbf{C}_b^\infty \times L_T^\infty \mathbf{C}_b^\infty$ be the approximation in Definition 3.15, and $\varphi_n \in \mathbf{C}_b^\infty$ with

$$\sup_n \|\varphi_n\|_{\mathbf{C}_a^\gamma(\rho_2/\rho_1)} < \infty,$$

and φ_n converges to φ in \mathbb{R}^{2d} locally uniformly. Let u_n be the classical solution of PDE (4.1) corresponding to (b_n, f_n) and φ_n . Then for any $\beta > \alpha$ and $\rho_4 \in \mathcal{P}_w$ with $\lim_{z \rightarrow \infty} (\rho_4/\rho_3)(z) = 0$, we have

$$(4.32) \quad \lim_{n \rightarrow \infty} \|u_n - u\|_{\mathbb{S}_{T,a}^{2-\beta}(\rho_4)} = 0.$$

PROOF. We mainly concentrate on showing the a priori estimate (4.31) for any paracontrolled solution u of PDE (4.1). Without loss of generality we may assume $\lambda = 0$ and $\varphi = 0$ (see Remark 4.2). We fix $0 < r < \frac{1}{16}$. Note that $\phi_{2r}^z = 1$ on the support of ϕ_r^z . For each $z \in \mathbb{R}^d$, by Proposition 4.4, $u_z := u\phi_r^z$ is a paracontrolled solution to the following PDE:

$$\partial_t u_z = \Delta_v u_z + v \cdot \nabla_x u_z + b_z \cdot \nabla_v u_z + g_z, \quad u_z(0) = 0,$$

where $b_z := b\phi_{2r}^z$ and

$$g_z := f\phi_r^z - 2\nabla_v u \cdot \nabla_v \phi_r^z - (\Delta_v \phi_r^z + v \cdot \nabla_x \phi_r^z)u - b \cdot \nabla_v \phi_r^z u.$$

By Theorem 4.6, there are $q > 1$ large enough and two constants $c_1, c_2 > 0$ such that for all $z \in \mathbb{R}^d$,

$$(4.33) \quad \|u_z\|_{\mathbb{S}_T^{2-\alpha}} \leq c_1(\ell_T^{b_z})^\vartheta \mathbb{A}_{T,\infty}^{b_z, g_z}, \quad \|u_z\|_{\mathbb{L}_T^\infty} \leq c_2(\ell_T^{b_z})^\vartheta \mathbb{A}_{T,q}^{b_z, g_z}.$$

Below, for simplicity of notations, we drop the time variable. By the definition of g_z , Lemma 2.11, (3.22) and (3.23), we have

$$(4.34) \quad \begin{aligned} \|g_z\|_{\mathbf{C}_a^{-\alpha}} &\leq \|f\phi_r^z\|_{\mathbf{C}_a^{-\alpha}} + 2\|\nabla_v u \cdot \nabla_v \phi_r^z\|_{\mathbf{C}_a^{-\alpha}} + \|b \cdot \nabla_v \phi_r^z u\|_{\mathbf{C}_a^{-\alpha}} \\ &\quad + \|u(\Delta_v \phi_r^z + v \cdot \nabla_x \phi_r^z)\|_{L^\infty} \\ &\lesssim \|f\|_{\mathbf{C}_a^{-\alpha}(\rho_2)} \|\phi_r^z\|_{\mathbf{C}_a^1(\rho_2^{-1})} + \|\nabla_v u\|_{\mathbf{C}_a^{-\alpha}(\rho_3)} \|\nabla_v \phi_r^z\|_{\mathbf{C}_a^1(\rho_3^{-1})} \\ &\quad + \|b\|_{\mathbf{C}_a^{-\alpha}(\rho_1)} \|u\|_{\mathbf{C}_a^1(\rho_3)} \|\nabla_v \phi_r^z\|_{\mathbf{C}_a^1(\rho_1 \rho_3)^{-1}} \\ &\quad + \|u\|_{L^\infty(\rho_3)} \|\Delta_v \phi_r^z + v \cdot \nabla_x \phi_r^z\|_{L^\infty(\rho_3^{-1})} \\ &\lesssim \rho_2^{-1}(z) \|f\|_{\mathbf{C}_a^{-\alpha}(\rho_2)} + (\varrho \rho_1^{-1} \rho_3^{-1})(z) \|u\|_{\mathbf{C}_a^1(\rho_3)}. \end{aligned}$$

Hence,

$$(4.35) \quad \|g_z\|_{L_T^q \mathbf{C}_a^{-\alpha}} \lesssim \rho_2^{-1}(z) \|f\|_{L_T^q \mathbf{C}_a^{-\alpha}(\rho_2)} + (\varrho \rho_1^{-1} \rho_3^{-1})(z) \|u\|_{L_T^q \mathbf{C}_a^1(\rho_3)}.$$

Moreover, we have for any $\tilde{\lambda} \geq 0$,

$$\begin{aligned} \|(b_z \circ \nabla_v \mathcal{J}_{\tilde{\lambda}} g_z)\|_{\mathbf{C}_a^{1-2\alpha}} &\leq \|b_z \circ \nabla_v \mathcal{J}_{\tilde{\lambda}}(f\phi_r^z)\|_{\mathbf{C}_a^{1-2\alpha}} + \|b_z \circ \nabla_v \mathcal{J}_{\tilde{\lambda}}(b \cdot \nabla_v \phi_r^z u)\|_{\mathbf{C}_a^{1-2\alpha}} \\ &\quad + \|b_z \circ \nabla_v \mathcal{J}_{\tilde{\lambda}}(u(\Delta_v \phi_r^z + v \cdot \nabla_x \phi_r^z) + 2\nabla_v u \cdot \nabla_v \phi_r^z)\|_{L^\infty} \\ &=: I_1^z + I_2^z + I_3^z. \end{aligned}$$

For I_1^z , by (3.44) with $\rho_3 = \rho_1^{-1}$, $\rho_4 = \rho_2^{-1}$ and $\phi = \phi_r^z$, we have

$$I_1^z \lesssim \|\phi_{2r}^z\|_{\mathbf{C}_a^1(\rho_1^{-1})} \|\phi_r^z\|_{\mathbf{C}_a^1(\rho_2^{-1})} \mathbb{A}_{t,\infty}^{b,f}(\rho_1, \rho_2) \lesssim (\rho_1^{-1} \rho_2^{-1})(z) \mathbb{A}_{t,\infty}^{b,f}(\rho_1, \rho_2).$$

For I_2^z , by (3.44) with $\rho_3 = \rho_1^{-2}$, $\rho_4 \equiv 1$, and $\psi = \nabla \phi_r^z u$, we have

$$I_2^z \lesssim \|\phi_{2r}^z\|_{\mathbf{C}_a^1(\rho_1^{-2})} \|\nabla_v \phi_r^z u\|_{\mathbb{S}_{t,a}^1} \mathbb{A}_{t,\infty}^{b,b}(\rho_1, \rho_1) \lesssim (\varrho \rho_1^{-2} \rho_3^{-1})(z) \|u\|_{\mathbb{S}_{t,a}^1(\rho_3)},$$

where by (3.32) and (3.28), we have

$$\|\nabla_v \phi_r^z u\|_{\mathbb{S}_{t,a}^1} \lesssim \varrho(z) \|\phi_{2r}^z u\|_{\mathbb{S}_{t,a}^1} \lesssim (\varrho \rho_3^{-1})(z) \|u\|_{\mathbb{S}_{t,a}^1(\rho_3)}.$$

For I_3^z , as in (4.34), by (2.28), Lemma 3.11 and (3.23), we have

$$\begin{aligned} I_3^z &\lesssim \|b_z\|_{\mathbf{C}_a^{-\alpha}} \|\nabla_v \mathcal{J}_{\tilde{\lambda}}(u(\Delta_v \phi_r^z + v \cdot \nabla_x \phi_r^z) + 2\nabla_v u \cdot \nabla_v \phi_r^z)\|_{\mathbf{C}_a^1} \\ &\lesssim \rho_1^{-1}(z) \|b\|_{\mathbf{C}_a^{-\alpha}(\rho_1)} \|u(\Delta_v \phi_r^z + v \cdot \nabla_x \phi_r^z) + 2\nabla_v u \cdot \nabla_v \phi_r^z\|_{\mathbb{C}_{t,a}^0} \\ &\lesssim (\varrho \rho_1^{-1} \rho_3^{-1})(z) \|u\|_{\mathbb{C}_{t,a}^1(\rho_3)} \lesssim (\varrho \rho_1^{-2} \rho_3^{-1})(z) \|u\|_{\mathbb{C}_{t,a}^1(\rho_3)}, \end{aligned}$$

where in the second step we used

$$(4.36) \quad \|b_z\|_{\mathbf{C}_a^{-\alpha}} \lesssim \|b\|_{\mathbf{C}_a^{-\alpha}(\rho_1)} \|\phi_r^z\|_{\mathbf{C}_a^1(\rho_1^{-1})} \lesssim \rho_1^{-1}(z) \|b\|_{\mathbf{C}_a^{-\alpha}(\rho_1)},$$

and in the last step we used ρ_1 is bounded. Combining the above calculations, we get for any $t \in [0, T]$,

$$\|(b_z \circ \nabla_v \mathcal{J}_{\tilde{\lambda}} g_z)(t)\|_{\mathbf{C}_a^{1-2\alpha}} \leq (\rho_1^{-1} \rho_2^{-1})(z) \mathbb{A}_{t,\infty}^{b,f}(\rho_1, \rho_2) + (\varrho \rho_1^{-2} \rho_3^{-1})(z) \|u\|_{\mathbb{S}_{t,a}^1(\rho_3)}.$$

Now by the definition of $\mathbb{A}_{T,q}^{b_z, g_z}$, (4.35), (4.36) and the calculations above, we get

$$(4.37) \quad \begin{aligned} \mathbb{A}_{T,q}^{b_z, g_z} &= \sup_{\tilde{\lambda}} \|b_z \circ \nabla_v \mathcal{J}_{\tilde{\lambda}} g_z\|_{L_T^q \mathbf{C}_a^{1-2\alpha}} + (\|b_z\|_{\mathbf{C}_{T,a}^{-\alpha}} + 1) \|g_z\|_{L_T^q \mathbf{C}_a^{-\alpha}} \\ &\lesssim (\rho_1^{-1} \rho_2^{-1})(z) \mathbb{A}_{T,\infty}^{b,f}(\rho_1, \rho_2) + (\varrho \rho_1^{-2} \rho_3^{-1})(z) \left(\int_0^T \|u\|_{\mathbb{S}_{t,a}^1(\rho_3)}^q dt \right)^{1/q}. \end{aligned}$$

On the other hand, by (3.22) and (3.44) with $\rho_3 = \rho_1^{-1}$, $\rho_4 = \rho_1^{-1}$ and $\phi = \psi = \phi_r^z$, we have

$$\|b_z \circ \nabla \mathcal{J}_{\tilde{\lambda}} b_z\|_{\mathbf{C}_{T,a}^{1-2\alpha}} \lesssim \rho_1^{-2}(z) (\|b \circ \nabla \mathcal{J}_{\tilde{\lambda}} b\|_{\mathbf{C}_{T,a}^{1-2\alpha}(\rho_1^2)} + \|b\|_{\mathbf{C}_{T,a}^{-\alpha}(\rho_1)}^2).$$

Hence, by (4.36)

$$\ell_T^{b_z} = \mathbb{A}_{T,\infty}^{b_z, b_z}(1, 1) + 1 \lesssim \rho_1^{-2}(z) \ell_T^b(\rho_1),$$

Then, by (4.33) and (4.37) with $q = \infty$, we have

$$\begin{aligned} \|u_z\|_{\mathbb{S}_{T,a}^{2-\alpha}} &\lesssim \rho_1^{-2\vartheta}(z) \left[(\rho_1^{-1} \rho_2^{-1})(z) \mathbb{A}_{T,\infty}^{b,f}(\rho_1, \rho_2) + (\varrho \rho_1^{-2} \rho_3^{-1})(z) \|u\|_{\mathbb{S}_{T,a}^1(\rho_3)} \right] \\ &= (\rho_1^{-1-2\vartheta} \rho_2^{-1})(z) \mathbb{A}_{T,\infty}^{b,f}(\rho_1, \rho_2) + (\varrho \rho_1^{-2-2\vartheta} \rho_3^{-1})(z) \|u\|_{\mathbb{S}_{T,a}^1(\rho_3)}, \end{aligned}$$

and

$$\|u_z\|_{\mathbb{L}_T^\infty} \lesssim (\rho_1^{-1-2\vartheta} \rho_2^{-1})(z) \mathbb{A}_{T,\infty}^{b,f}(\rho_1, \rho_2) + (\varrho \rho_1^{-2-2\vartheta} \rho_3^{-1})(z) \left(\int_0^T \|u\|_{\mathbb{S}_{t,a}^1(\rho_3)}^q dt \right)^{1/q}.$$

From these two estimates, and noting that

$$\rho_3 = \rho_1^{1+2\vartheta} \rho_2, \quad \varrho \rho_1^{-2-2\vartheta} \leq 1,$$

by Lemmas 3.9 and 3.7, we get

$$(4.38) \quad \|u\|_{\mathbb{S}_{T,a}^{2-\alpha}(\rho_3)} \lesssim \mathbb{A}_{T,\infty}^{b,f}(\rho_1, \rho_2) + \|u\|_{\mathbb{S}_{T,a}^1(\rho_3)}$$

and

$$(4.39) \quad \|u\|_{\mathbb{L}_T^\infty(\rho_3)} \lesssim \mathbb{A}_{T,\infty}^{b,f}(\rho_1, \rho_2) + \left(\int_0^T \|u\|_{\mathbb{S}_{t,a}^1(\rho_3)}^q dt \right)^{1/q}.$$

Note that by (2.19) and Definition 3.5,

$$\|u\|_{\mathbb{S}_{T,a}^1(\rho_3)} \lesssim \|u\|_{\mathbb{S}_{T,a}^{2-\alpha}(\rho_3)}^{1/(2-\alpha)} \|u\|_{\mathbb{L}_T^\infty(\rho_3)}^{(1-\alpha)/(2-\alpha)},$$

which by Young's inequality implies that for any $\varepsilon > 0$, there is a constant $C_\varepsilon > 0$ such that

$$\|u\|_{\mathbb{S}_{T,a}^1(\rho_3)} \leq \varepsilon \|u\|_{\mathbb{S}_{T,a}^{2-\alpha}(\rho_3)} + C_\varepsilon \|u\|_{\mathbb{L}_T^\infty(\rho_3)}.$$

Substituting this into (4.38) and choosing ε small enough, we get

$$\|u\|_{\mathbb{S}_{T,a}^{2-\alpha}(\rho_3)} \lesssim \mathbb{A}_{T,\infty}^{b,f}(\rho_1, \rho_2) + \|u\|_{\mathbb{L}_T^\infty(\rho_3)},$$

which together with (4.39) and by Gronwall's inequality, we obtain (4.31).

(Uniqueness) Let u_1, u_2 be two paracontrolled solutions of PDE (4.1). By definition, it is easy to see that $u = u_1 - u_2$ is still a paracontrolled solution of (4.1) with $\varphi = f \equiv 0$. Thus by (4.31), we immediately have $u = 0$.

(Existence) Let $(b_n, f_n) \in L_T^\infty C_b^\infty \times L_T^\infty C_b^\infty$ be the approximation in Definition 3.15, and u_n be the corresponding solution of PDE (4.1). By the priori estimates (4.31), (4.7) and (3.40), we have the following uniform estimate:

$$(4.40) \quad \sup_n \left(\|u_n\|_{\mathbb{S}_{T,a}^{2-\alpha}(\rho_3)} + \|u_n^\sharp\|_{\mathbf{C}_{T,a}^{3-2\alpha}(\rho_4)} \right) < \infty.$$

By Lemma A.3, for any $\beta > \alpha$ and $\rho_5 \in \mathcal{P}_w$ with $\lim_{z \rightarrow \infty} (\rho_5/\rho_3)(z) = 0$, there are $u \in \mathbb{S}_{T,a}^{2-\alpha}(\rho_3)$ and a subsequence n_k such that

$$\lim_{k \rightarrow \infty} \|u_{n_k} - u\|_{\mathbb{S}_{T,a}^{2-\beta}(\rho_5)} = 0.$$

Moreover, let $u^\sharp := u - \nabla_v u \prec \mathcal{I}_\lambda b - \mathcal{I}_\lambda f$. By the above limit, (2.27) and (3.35), it is easy to see that for some $\rho_6 \in \mathcal{P}_w$,

$$\lim_{k \rightarrow \infty} \|u_{n_k}^\sharp - u^\sharp\|_{\mathbb{L}_T^\infty(\rho_6)} = 0,$$

which, together with (4.40), and by Fatou's lemma and the interpolation inequality (2.19), implies that $u^\sharp \in \mathbb{C}_{T,a}^{3-2\alpha}(\rho_4)$ and for any $\beta > \alpha$,

$$\lim_{k \rightarrow \infty} \|u_{n_k}^\sharp - u^\sharp\|_{\mathbb{C}_{T,a}^{3-2\beta}(\rho_6 \rho_4)} = 0.$$

By a standard limit procedure, one finds that u is a paracontrolled solution in the sense of Definition 4.1 (see [GIP15]). Finally, by the uniqueness of paracontrolled solutions, the full limit (4.32) holds. \square

5. Well-posedness of singular mean field equations. In this section we study the non-linear singular kinetic equations. Throughout this section we fix $T > 0$, $\alpha \in (\frac{1}{2}, \frac{2}{3})$, $\vartheta := \frac{9}{2-3\alpha}$ and

$$(5.1) \quad \kappa_0 < 0, \quad 0 \leq \kappa_1 \leq 1/(2\vartheta + 2),$$

and let

$$\kappa_2 := \kappa_1, \quad \kappa_3 := (2\vartheta + 2)\kappa_1, \quad \rho_i := \varrho^{\kappa_i}, \quad i = 0, 1, 2, 3,$$

where ϱ is given in (3.5). Consider the following nonlinear kinetic equation with distributional drift

$$(5.2) \quad \partial_t u = \Delta_v u - v \cdot \nabla_x u - W \cdot \nabla_v u - K * \langle u \rangle \cdot \nabla_v u, \quad u(0) = \varphi,$$

where $u : \mathbb{R}_+ \times \mathbb{R}^{2d} \rightarrow \mathbb{R}$ is a function of time variable t , position x and velocity v , $\langle u \rangle(t, x) := \int_{\mathbb{R}^d} u(t, x, v) dv$ stands for the mass, $K : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is a kernel function, and

$$K * \langle u \rangle(t, x) := \int_{\mathbb{R}^d} K(x - y) \langle u \rangle(t, y) dy.$$

Here we assume that $W(t, x, v)$ satisfies that

$$(5.3) \quad W \in \mathbb{B}_T^\alpha(\rho_1) \text{ has the approximation sequence } W_n \text{ with } \operatorname{div}_v W_n \equiv 0,$$

and

$$(5.4) \quad K \in \cup_{\beta > \alpha-1} \mathbf{C}^{\beta/3}.$$

REMARK 5.1. (i) For $\tilde{K}(x, v) = K(x)$, it is easy to see that

$$\tilde{K} \in \mathbf{C}_a^\beta \iff K \in \mathbf{C}^{\beta/3}, \quad \forall \beta \in \mathbb{R}.$$

Moreover, for $K(x) = |x|^{-r}$ with $r < (1 - \alpha)/3$, (5.4) holds.

(ii) Since $\operatorname{div}_v W \equiv 0$ and K does not depend on v , one can write (5.2) as the following divergence form:

$$(5.5) \quad \partial_t u = \Delta_v u - v \cdot \nabla_x u - \operatorname{div}_v((W + K * \langle u \rangle)u), \quad u(0) = \varphi.$$

In particular, when W and K are smooth, if φ is a probability density function, then so is the solution u .

To use the framework of the above sections we define the solution to (5.2) by the following transform: for $f \in \mathcal{S}'(\mathbb{R}^{2d})$, $\varphi \in \mathcal{S}(\mathbb{R}^{2d})$

$$\tau f(\varphi) := f(\tau\varphi) \quad \tau\varphi(x, v) := \varphi(x, -v).$$

It is easy to see this transform does not change Besov norm.

DEFINITION 5.2. We call $u \in \mathbb{S}_{T,a}^{2-\alpha}(\rho_3)$ a probability density paracontrolled solution to PDE (5.2) if τu is a paracontrolled solution to PDE (4.1) with $\lambda = 0$ and $b = \tau W + K * \langle u \rangle$ and initial value $\tau\varphi$

$$u \geq 0, \quad \int_{\mathbb{R}^{2d}} u(t, z) dz = 1, \quad t \in [0, T].$$

REMARK 5.3. (i) This definition should be equivalent to the definition using the semi-group associated with $\Delta_v - v \cdot \nabla_x$.

(ii) Let u be a probability density paracontrolled solution to PDE (5.2). Under (5.3) and (5.4), by Lemma 3.16, it is easy to see that $b = \tau W + K * \langle u \rangle \in \mathbb{B}_T^\alpha(\rho_1)$, whose approximation sequence can be taken as

$$b_n = \tau W_n + K_n * \langle u \rangle \quad \text{with } \operatorname{div}_v b_n \equiv 0,$$

where $K_n = K * \phi_n$ with ϕ_n being the usual mollifier.

For a density solution the nonlinear term can be bounded easily. To prove the existence of solutions we use smooth approximation and need to prove the convergence not only in the kinetic Hölder space but also in L^1 space since the nonlinear term contains a nonlocal interaction. The proof of the uniqueness part is more involved. To deal with the nonlinear term, we have to bound the difference of solutions in L^1 space which requires an uniform $L_t^2 L^1$ bound of the gradient of the solutions. To this end we use an entropy method and introduce the following entropy. For a probability density function f , one says that f has a finite entropy if

$$H(f) := \int_{\mathbb{R}^{2d}} f(z) \ln f(z) dz \in (-\infty, \infty).$$

The main result of this section is the following theorem.

THEOREM 5.4. Suppose that (5.1), (5.3) and (5.4) hold. Let $\gamma > 1 + \alpha$.

(Existence) For any probability density function $\varphi \in L^1(\rho_0) \cap \mathbf{C}_a^\gamma$, there exists at least a probability density paracontrolled solution $u \in \mathbb{S}_{T,a}^{2-\alpha}(\rho_3)$ to PDE (5.2). Moreover, there is a constant $C > 0$ such that for all $t \in [0, T]$,

$$(5.6) \quad \|u(t)\|_{L^1(\rho_0)} \leq C \|\varphi\|_{L^1(\rho_0)}$$

and if $|H(\varphi)| < \infty$, then it holds that

$$(5.7) \quad H(u(t)) + \|\nabla_v u\|_{L_t^2 L_z^1}^2 \leq H(\varphi),$$

and

$$(5.8) \quad |H(u(t))| + \|\nabla_v u\|_{L_t^2 L_z^1}^2 \leq H(\varphi) + C(\|\varphi\|_{L^1(\rho_0)} + 1).$$

(Stability) If in addition K is bounded, then for any $\varphi_1, \varphi_2 \in L^1(\rho_0) \cap \mathbf{C}_a^\gamma$ with $H(\varphi_1) < \infty$, and any probability density paracontrolled solutions u_1 and u_2 with initial values φ_1 and φ_2 , respectively, there is a constant $C > 0$ only depending on $\|K\|_{L^\infty}$, $\|\varphi_1\|_{L^1(\rho_0)}$, $H(\varphi_1)$ and $\|e^{-\rho_0}\|_{L^1}$ such that for all $t \in [0, T]$,

$$(5.9) \quad \|u_1(t) - u_2(t)\|_{L^1} \leq e^{Ct} \|\varphi_1 - \varphi_2\|_{L^1}.$$

REMARK 5.5. $\varphi \in L^1(\rho_0)$ is a moment requirement, i.e.,

$$\int_{\mathbb{R}^{2d}} |z|^{|\kappa_0|} \varphi(z) dz < \infty.$$

This is a common assumption in the entropy method (see [JW16]), which can be seen from the following Lemma 5.6.

We need the following elementary lemma.

LEMMA 5.6. *It holds that for any measurable $\phi, f \geq 0$, $\delta \in [0, 1)$ and $\rho \in \mathcal{P}_w$,*

$$\int \phi |f \ln(f + \delta)| \leq \int \phi f \ln(f + \delta) + 2 \left(\int \phi \rho f + \int \phi e^{-\rho} \right).$$

PROOF. By Young's inequality, we have

$$-r \ln(r + \delta) \leq -r \ln r \leq ar + e^{-a}, \quad \forall r \in [0, 1], \quad a \geq 0.$$

Hence,

$$|r \ln(r + \delta)| = r \ln(r + \delta) - 2r \ln(r + \delta) \mathbf{1}_{\{0 < r \leq 1 - \delta\}} \leq r \ln(r + \delta) + 2(ar + e^{-a}).$$

The desired estimate follows by taking $a = \rho$. \square

We recall the following result (cf. [RXZ21]).

LEMMA 5.7. *Let $b \in L_T^\infty C_b^\infty(\mathbb{R}^{2d})$ and let $Z_t^{z_0} = (X_t, V_t)$ be the unique solution of the following SDE:*

$$(5.10) \quad dX_t = V_t dt, \quad dV_t = \sqrt{2} dB_t + b(t, X_t, V_t) dt, \quad (X_0, V_0) = z_0 \in \mathbb{R}^{2d}.$$

Then for any initial probability measure μ_0 ,

$$\mu(t, dz) = \int_{\mathbb{R}^{2d}} \mathbf{P}(Z_t^{z_0} \in dz) \mu_0(dz_0)$$

is the unique solution to the following Fokker-Planck equation in the distributional sense:

$$\partial_t \mu = \Delta_v \mu - v \cdot \nabla_x \mu - \operatorname{div}_v(b\mu), \quad \mu(0) = \mu_0.$$

Now we derive the following a priori moment and entropy estimates. The proof is divided into three steps. First for given solution u we can find a linear approximation such that Theorem 4.7 can be applied. Next we prove (5.6) by a probabilistic method. Finally we use entropy method to prove (5.7) and (5.8).

LEMMA 5.8. *Under (5.3), let $u \in \mathbb{S}_{T,a}^{2-\alpha}(\rho_3)$ be a probability density paracontrolled solution of (5.2) with initial value $\varphi \in L^1(\rho_0) \cap \mathbf{C}_a^\gamma$. Then (5.6) holds. Moreover, if $H(\varphi) < \infty$ then (5.7) and (5.8) hold.*

PROOF. (STEP 1) Let $b_n \in L_T^\infty C_b^\infty(\mathbb{R}^{2d})$ be the approximation sequence as in Remark 5.3 and $\varphi_n = \varphi * \phi_n$ with ϕ_n being the usual mollifier. Since $b_n \in L_T^\infty C_b^\infty(\mathbb{R}^{2d})$, it is well known that there is a unique probability density solution $u_n \in L_T^\infty C_b^\infty(\mathbb{R}^{2d})$ to the following approximation Fokker-Planck equation:

$$(5.11) \quad \partial_t u_n = \Delta_v u_n - v \cdot \nabla_x u - \tau b_n \cdot \nabla_v u_n = \Delta_v u_n - v \cdot \nabla_x u_n - \operatorname{div}_v(\tau b_n u_n),$$

with $u_n(0) = \varphi_n$. It is easy to see that τu_n satisfies the following equation:

$$\partial_t \tau u_n = \Delta_v \tau u_n + v \cdot \nabla_x \tau u_n + b_n \cdot \nabla_v \tau u_n.$$

By (4.32) and definition of solutions, we have for some $\rho \in \mathcal{P}_w$ and $\beta \in (\alpha, 2)$,

$$\lim_{n \rightarrow \infty} \|\tau u_n - \tau u\|_{\mathbb{S}_{T,a}^{2-\beta}(\rho)} = 0,$$

which implies that

$$(5.12) \quad \lim_{n \rightarrow \infty} \|u_n - u\|_{\mathbb{S}_{T,a}^{2-\beta}(\rho)} = 0.$$

To show (5.6), (5.7) and (5.8), it suffices to show that for some $C > 0$ independent of n ,

$$(5.13) \quad \|u_n(t)\|_{L^1(\rho_0)} \lesssim_C \|\varphi_n\|_{L^1(\rho_0)} \lesssim_C \|\varphi\|_{L^1(\rho_0)},$$

and if $H(\varphi) < \infty$, then

$$(5.14) \quad H(u_n(t)) + \|\nabla_v u_n\|_{L_t^2 L^1}^2 \leq H(\varphi_n).$$

Indeed, it is easy to see that (5.13) implies (5.6) by Fatou's lemma. Now we show how to derive (5.7) and (5.8) from (5.14) and (5.13). First, since $r \mapsto r \log r$ is convex on $[0, \infty)$ and $H(\varphi) < \infty$, by Jensen's inequality, we have

$$(5.15) \quad H(\varphi_n) = H(\varphi * \phi_n) \leq H(\varphi),$$

and by the lower semi-continuity of $u \mapsto \|\nabla_v u\|_{L_t^2 L^1}$,

$$(5.16) \quad \|\nabla_v u\|_{L_t^2 L^1} \leq \liminf_{n \rightarrow \infty} \|\nabla_v u_n\|_{L_t^2 L^1}.$$

On the other hand, let $\kappa_0 < \kappa < 0$ and $\rho := \varrho^\kappa$. Recalling (3.5) and $\rho_0 = \varrho^{\kappa_0}$, for any $R > 0$, we have

$$\begin{aligned} \|u_n(t) - u(t)\|_{L^1(\rho)} &\leq \int |u_n(t, z) - u(t, z)| \cdot \mathbf{1}_{\{|z|_a \leq R\}} \cdot \rho(z) dz \\ &\quad + \int |u_n(t, z) - u(t, z)| \cdot \mathbf{1}_{\{|z|_a > R\}} \cdot \rho(z) dz \\ &\leq \int |u_n(t, z) - u(t, z)| \cdot \mathbf{1}_{\{|z|_a \leq R\}} \cdot \rho(z) dz \\ &\quad + C \sup_n \|u_n(t)\|_{L^1(\rho_0)} / R^{\kappa - \kappa_0}, \end{aligned}$$

which implies by first letting $n \rightarrow \infty$ and then $R \rightarrow \infty$,

$$(5.17) \quad \lim_{n \rightarrow \infty} \|u_n - u\|_{L_T^\infty L^1(\rho)} = 0.$$

Now we define the relative entropy for nonnegative measurable function f ,

$$(5.18) \quad H_\rho(f) := \int f \ln(fe^\rho) = H(f) + \|f\|_{L^1(\rho)}.$$

Since $r(\ln r - 1) \geq -1$ for $r \geq 0$, we have

$$\inf_n u_n(t) (\ln(u_n(t)e^\rho) - 1) \geq -e^{-\rho} \in L^1,$$

which by Fatou's lemma implies that

$$H_\rho(u(t)) - 1 \leq \liminf_{n \rightarrow \infty} \int u_n(t) e^\rho (\ln(u_n(t)e^\rho) - 1) e^{-\rho} = \liminf_{n \rightarrow \infty} H_\rho(u_n(t)) - 1.$$

This together with (5.17) and (5.18) yields

$$H(u(t)) \leq \lim_{n \rightarrow \infty} H(u_n(t)).$$

Combining this with (5.14)-(5.16), we obtain (5.7). Moreover, by (5.6) and (5.7) and Lemma 5.6, (5.8) follows.

(STEP 2) In this step we show (5.13) by showing a moment estimate of solution to (5.10) which is achieved by establishing a Krylov's type of estimate for the singular drift term. For simplicity, we drop the subscripts n below. By Lemma 5.7 one has

$$\|u(t)\|_{L^1(\rho_0)} = \int_{\mathbb{R}^{2d}} \mathbf{E} \rho_0(Z_t^{z_0}) \varphi(z_0) dz_0,$$

where $Z_t^{z_0} = (X_t, V_t)$ is the unique solution to SDE (5.10) with $b = \tau b_n$. Hence, to show (5.13), it suffices to prove that for some $C > 0$ independent of n ,

$$(5.19) \quad \mathbf{E} \rho_0(Z_t^{z_0}) \lesssim_C \rho_0(z_0), \quad \forall z_0 \in \mathbb{R}^{2d}.$$

By Itô's formula, we have

$$\mathbf{E} \rho_0(Z_t^{z_0}) = \rho_0(z_0) + \mathbf{E} \int_0^t (\Delta_v \rho_0 + v \cdot \nabla_x \rho_0)(Z_s^{z_0}) ds + \mathbf{E} \int_0^t (b \cdot \nabla_v \rho_0)(s, Z_s^{z_0}) ds.$$

Noting that by (3.7), for some $C_0 > 0$,

$$|\Delta_v \rho_0 + v \cdot \nabla_x \rho_0|(z) \leq C_0 \rho_0(z),$$

we obtain

$$\mathbf{E} \rho_0(Z_t^{z_0}) \leq \rho_0(z_0) + C_0 \mathbf{E} \int_0^t \rho_0(Z_s^{z_0}) ds + \mathbf{E} \int_0^t (b \cdot \nabla_v \rho_0)(s, Z_s^{z_0}) ds.$$

To estimate the last term, we use Theorem 4.7 to deduce a Krylov's type of estimate. More precisely, for fixed $t \in [0, T]$, let w^t be the unique smooth solution of the following backward PDE:

$$\partial_s w^t + (\Delta_v + v \cdot \nabla_x + b \cdot \nabla_v) w^t = b \cdot \nabla_v \rho_0, \quad w^t(t) = 0.$$

By Itô's formula again, we have

$$0 = \mathbf{E} w^t(t, Z_t^{z_0}) = w^t(0, z_0) + \mathbf{E} \int_0^t (b \cdot \nabla_v \rho_0)(s, Z_s^{z_0}) ds.$$

Hence,

$$(5.20) \quad \mathbf{E} \rho_0(Z_t^{z_0}) \leq \rho_0(z_0) + C_0 \mathbf{E} \int_0^t \rho_0(Z_s^{z_0}) ds - w^t(0, z_0).$$

Let $\beta \in (\alpha, 1)$ and $\rho_4 := (\rho_0 \varrho)^{-1}$. By (3.7) and (2.21), we have

$$\|\nabla_v \rho_0\|_{\mathbf{C}_a^\beta(\rho_4)} < \infty,$$

which by (2.29) yields that

$$\|b \cdot \nabla_v \rho_0\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_1 \rho_4)} \lesssim \|b\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_1)} \|\nabla_v \rho_0\|_{\mathbf{C}_a^\beta(\rho_4)} \lesssim \|b\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_1)}.$$

Moreover, by Lemma 3.18 we obtain

$$\begin{aligned} & \|b \circ \nabla_v \mathcal{J}_\lambda(b \cdot \nabla_v \rho_0)\|_{\mathbb{C}_{T,a}^{1-2\alpha}(\rho_1^2 \rho_4)} \leq \|(b \circ \nabla_v \mathcal{J}_\lambda b) \nabla_v \rho_0\|_{\mathbb{C}_{T,a}^{1-2\alpha}(\rho_1^2 \rho_4)} \\ & \quad + \|b \circ \nabla_v \mathcal{J}_\lambda(b \cdot \nabla_v \rho_0) - (b \circ \nabla_v \mathcal{J}_\lambda b) \nabla_v \rho_0\|_{\mathbb{C}_{T,a}^{1-2\alpha}(\rho_1^2 \rho_4)} \\ & \lesssim \|b \circ \nabla_v \mathcal{J}_\lambda b\|_{\mathbb{C}_{T,a}^{1-2\alpha}(\rho_1^2)} \|\nabla_v \rho_0\|_{\mathbf{C}_a^\beta(\rho_4)} + \|b\|_{\mathbb{C}_{T,a}^{-\alpha}(\rho_1)}^2 \|\nabla_v \rho_0\|_{\mathbf{C}_a^\beta(\rho_4)}. \end{aligned}$$

Since $(2\vartheta + 2)\kappa_1 \leq 1$ and $\rho_1 = \varrho^{\kappa_1}$, $\rho_4 = \varrho^{-\kappa_0-1}$, by Theorem 4.7 we have

$$\|w^t\|_{\mathbb{L}_T^\infty(\rho_0^{-1})} \lesssim \mathbb{A}_{T,\infty}^{b,b \cdot \nabla_v \rho_0}(\rho_1, \rho_1 \rho_4) < \infty,$$

which implies that for some $C_1 > 0$ independent of n and z_0 ,

$$|w^t(0, z_0)| \leq C_1 \rho_0(z_0).$$

Substituting this into (5.20) and by Gronwall's inequality we obtain (5.19).

(STEP 3) In this step we show (5.14) by entropy method. Recall χ in (3.20). For $\delta \in (0, 1)$ and $R \geq 1$, let

$$\beta_\delta(r) := r \ln(r + \delta), \quad \chi_R(x, v) := \chi\left(\frac{x}{R^3}, \frac{v}{R}\right).$$

Since u is a smooth solution of PDE (5.11), by the chain rule, it is easy to see that

$$\partial_t \beta_\delta(u) = \Delta_v \beta_\delta(u) - v \cdot \nabla_x \beta_\delta(u) - b \cdot \nabla_v \beta_\delta(u) - \beta_\delta''(u) |\nabla_v u|^2.$$

Multiplying both sides by χ_R , then integrating over $[0, t] \times \mathbb{R}^{2d}$ and by integration by parts and $\operatorname{div}_v b = 0$, we obtain

$$\begin{aligned} & \int \chi_R \beta_\delta(u(t)) - \int \chi_R \beta_\delta(\varphi) + \int_0^t \int \chi_R \beta_\delta''(u) |\nabla_v u|^2 \\ &= \int_0^t \int \left(\Delta_v \chi_R + v \cdot \nabla_x \chi_R + b \cdot \nabla_v \chi_R \right) \beta_\delta(u) \\ &\leq \|\Delta_v \chi_R + v \cdot \nabla_x \chi_R + b \cdot \nabla_v \chi_R\|_{\mathbb{L}_T^\infty} \int_0^t \int \chi_{2R} |\beta_\delta(u)| \\ (5.21) \quad &\leq C_\chi (1 + \|b\|_{L^\infty}) R^{-1} \int_0^t \int \chi_{2R} |\beta_\delta(u)|, \end{aligned}$$

where C_χ only depends on χ . For $m \in \mathbb{N}$, define

$$G_R^m(t) := \int \chi_{2^m R} |\beta_\delta(u(t))|.$$

Noting that $\beta_\delta'' \geq 0$, by Lemma 5.6, (5.21) and (5.13), we obtain

$$\begin{aligned} G_R^m(t) &\leq \int \chi_{2^m R} \beta_\delta(u(t)) + 2 \left(\int u(t) \rho_0 + \int e^{-\rho_0} \right) \\ &\leq \frac{C_b}{2^m R} \int_0^t G_R^{m+1}(s) ds + \int |\beta_\delta(\varphi)| + C(\|\varphi\|_{L^1(\rho_0)} + 1) \\ &\leq \frac{C_b}{R} \int_0^t G_R^{m+1}(s) ds + A_0, \end{aligned}$$

where $C_b := C_\chi(1 + \|b\|_{L^\infty})$ and

$$A_0 := \int |\beta_\delta(\varphi)| + C(\|\varphi\|_{L^1(\rho_0)} + 1) \leq \int |\varphi \ln \varphi| + 1 + C(\|\varphi\|_{L^1(\rho_0)} + 1) < \infty.$$

Here the first inequality is due to

$$(5.22) \quad |\beta_\delta(r)| \leq |r \log r| + r, \quad \delta \in (0, 1), \quad r \geq 0,$$

and the last inequality we used Lemma 5.6. By iteration, we obtain that for any $m \in \mathbb{N}$,

$$G_R^0(t) \leq A_0 \sum_{k=0}^{m-1} \frac{C_b^k t^k}{R^k k!} + \frac{C_b^m}{R^m} \int_0^t \int_0^{t_1} \cdots \int_0^{t_{m-1}} G_R^m(t_m) dt_m \cdots dt_1.$$

Since $u \in L_T^\infty C_b^\infty(\mathbb{R}^{2d})$, there is a constant $C_\delta > 0$ such that for any $R \geq 1$,

$$G_R^m(t) \leq C_\delta \int \chi_{2^m R} \leq C_\delta (2^m R)^{2d}.$$

Therefore,

$$G_R^0(t) \leq A_0 e^{C_b t/R} + \frac{C_b^m}{R^m} C_\delta (2^m R)^{2d} \frac{t^m}{m!},$$

which in turn implies that by first letting $m \rightarrow \infty$ and then $R \rightarrow \infty$,

$$(5.23) \quad \int |\beta_\delta(u(t))| = \lim_{R \rightarrow \infty} G_R^0(t) \leq A_0 = \int |\beta_\delta(\varphi)| + C(\|\varphi\|_{L^1(\rho_0)} + 1) < \infty.$$

Thus, by taking limits $R \rightarrow \infty$ on the both sides of (5.21), we obtain

$$\int \beta_\delta(u(t)) + \int_0^t \int \beta_\delta''(u(s)) |\nabla_v u(s)|^2 ds \leq \int \beta_\delta(\varphi).$$

By (5.23) (5.22) and Fatou's Lemma, we further have

$$(5.24) \quad \int |u(t) \ln(u(t))| \leq \int |\varphi \ln \varphi| + C(\|\varphi\|_{L^1(\rho_0)} + 1) < \infty,$$

Letting $\delta \downarrow 0$, by $\beta_\delta''(r) = \frac{1}{r+\delta} + \frac{\delta}{(r+\delta)^2}$ and Fatou's lemma,

$$(5.25) \quad \int u(t) \ln(u(t)) + \int_0^t \int \frac{|\nabla_v u(s)|^2}{u(s)} ds \leq \int \varphi \ln \varphi.$$

Here for the first and last term we used (5.24) (5.22) and dominated convergence theorem.

On the other hand, by Hölder's inequality, we have

$$\int_0^t \|\nabla_v u(s)\|_{L^1}^2 dt \leq \int_0^t \left(\|u(s)\|_{L^1} \int \frac{|\nabla_v u(s)|^2}{u(s)} ds \right) ds = \int_0^t \left(\int \frac{|\nabla_v u(s)|^2}{u(s)} ds \right) ds.$$

Substituting this into (5.25), we obtain (5.14). The proof is complete. \square

Now, we can give the proof of Theorem 5.4.

PROOF OF THEOREM 5.4. (Existence) By our definition of solutions it suffices to prove there exists a solution u to the following equation:

$$(5.26) \quad \partial_t u = \Delta_v u + v \cdot \nabla_x u + \tau W \cdot \nabla_v u + K * \langle u \rangle \cdot \nabla_v u, \quad u(0) = \tau \varphi.$$

Let $W_n \in L_T^\infty C_b^\infty(\mathbb{R}^{2d})$ be as in (5.3). Let $\phi_n(x) = n^d \phi_1(nx)$ be the usual modifier and $K_n := K * \phi_n \in C_b^\infty(\mathbb{R}^d)$. Since the coefficients are bounded and Lipschitz and $\operatorname{div}_v W_n = \operatorname{div}_v K_n = 0$, by standard fixed point argument, one can show that there is a unique smooth probability density solution u_n to the following PDE

$$(5.27) \quad \partial_t u_n = \Delta_v u_n + v \cdot \nabla_x u_n + (\tau W_n + K_n * \langle u_n \rangle) \cdot \nabla_v u_n, \quad u_n(0) = \tau \varphi_n.$$

Define

$$b_n(t, x, v) := \tau W_n(t, x, v) + K_n * \langle u_n \rangle(t, x).$$

Since for $\beta > (\alpha - 1)/3$

$$\|K_n * \langle u_n \rangle\|_{C^\beta} \leq \|K_n\|_{C^\beta} \|\langle u_n \rangle\|_{L^1} \leq \|K\|_{C^\beta} \|u_n\|_{L^1} \lesssim 1,$$

by (2.28), (3.30) and (3.34) and Remark 5.1 it is easy to see that

$$\|b_n \circ \nabla_{v \cdot \mathcal{J}}(K_n * \langle u_n \rangle)\|_{\mathbf{C}_a^{3\beta+1-\alpha}(\rho_1)} \lesssim \|b_n\|_{\mathbf{C}_a^{-\alpha}(\rho_1)} \|K_n * \langle u_n \rangle\|_{\mathbf{C}_a^{3\beta}} \lesssim 1,$$

where the implicit constant is independent of n . Thus, by definition we have

$$\sup_n \ell_T^{b_n}(\rho_1) \lesssim \sup_n \left(\ell_T^{W_n}(\rho_1) + \ell_T^{K_n * \langle u_n \rangle}(1) + \sum_{i=1}^d \mathbb{A}_{T,\infty}^{W_n, K_n^i * \langle u_n \rangle}(\rho_1, 1) \right) < \infty,$$

and by Theorem 4.7 and (4.7),

$$\sup_n \left(\|u_n\|_{\mathbb{S}_{T,a}^{2-\alpha}(\rho_3)} + \|u_n^\sharp\|_{\mathbb{C}_{T,a}^{3-2\alpha}(\rho_4)} \right) < \infty.$$

Thus, by Lemma A.3, there are $u \in \mathbb{S}_{T,a}^{2-\alpha}(\rho_3)$ and subsequence n_k such that for any $\beta > \alpha$ and $\rho_5 \in \mathcal{P}_w$ with $\lim_{z \rightarrow \infty} (\rho_5/\rho_3)(z) = 0$,

$$\lim_{k \rightarrow \infty} \|u_{n_k} - u\|_{\mathbb{S}_{T,a}^{2-\beta}(\rho_5)} = 0.$$

As in the proof of Theorem 4.7, one sees that $u^\sharp := u - P_t \varphi - \nabla_v u \prec \mathcal{J}b \in \mathbb{C}_{T,a}^{3-2\alpha}(\rho_4)$ and for some $\rho_6 \in \mathcal{P}_w$ and any $\beta > \alpha$,

$$\lim_{k \rightarrow \infty} \|u_{n_k}^\sharp - u^\sharp\|_{\mathbb{C}_{T,a}^{3-2\beta}(\rho_6)} = 0.$$

It is the same reason as in (5.17), we have

$$\lim_{k \rightarrow \infty} \|u_{n_k} - u\|_{L_T^\infty L^1} = 0.$$

In particular,

$$(5.28) \quad u \geq 0, \quad \int u(t, z) \equiv 1.$$

Since $K \in \mathbf{C}_a^\beta$ for some $\beta > \alpha - 1$,

$$\|K * \langle u_{n_k} \rangle - K * \langle u \rangle\|_{\mathbb{C}_{T,a}^\beta} \rightarrow 0 \quad \text{as } k \rightarrow \infty.$$

Let $\varepsilon = (\beta - \alpha + 1)/2 > 0$. By (2.24), we have

$$\|K_n * \langle u_n \rangle - K * \langle u_n \rangle\|_{\mathbb{C}_a^{\alpha-1+\varepsilon}} \lesssim n^{-\varepsilon} \|K\|_{\mathbf{C}^\beta} \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

which implies that

$$\|b_n \circ \nabla_{\mathcal{J}_\lambda} b_n - b \circ \nabla_{\mathcal{J}_\lambda} b\|_{\mathbb{C}_{T,a}^{1-2\alpha}(\rho_1^2)} \rightarrow 0.$$

Taking limits on both sides of (5.27), one sees that u is a probability density paracontrolled solution of PDE (5.26).

(Stability) By our definition it only suffices to prove the result of solution to (5.26). Let u_1, u_2 be two paracontrolled solutions of PDE (5.2) with the initial values φ_1 and φ_2 , respectively. For $i = 1, 2$, let u_i^n be the smooth approximation solution of the following linearized Fokker-Planck equation

$$\partial_t u_i^n = \Delta_v u_i^n + v \cdot \nabla_x u_i^n + (\tau W_n + K_n * \langle u_i \rangle) \cdot \nabla_v u_i^n, \quad u_i^n(0) = \varphi_i^n,$$

where $\varphi_i^n = \varphi_i * \phi_n$ and W_n is the approximation sequence in (5.3), $K_n = K * \phi_n$. By (4.32), we have for some $\rho \in \mathcal{P}_w$,

$$(5.29) \quad \lim_{n \rightarrow \infty} \|u_i^n - u_i\|_{\mathbb{L}_T^\infty(\rho)} = 0, \quad i = 1, 2.$$

Let

$$w_n := u_1^n - u_2^n, \quad w := u_1 - u_2,$$

and

$$b_n := \tau W_n + K_n * \langle u_2 \rangle, \quad f_n := K_n * \langle w \rangle \cdot \nabla_v u_1^n,$$

and for any $\delta > 0$,

$$\tilde{\beta}_\delta(r) := \sqrt{r^2 + \delta} - \sqrt{\delta}, \quad \chi_R(x, v) := \chi\left(\frac{x}{R^3}, \frac{v}{R}\right)$$

It is easy to see that

$$\partial_t w_n = \Delta_v w_n + v \cdot \nabla_x w_n + b_n \cdot \nabla_v w_n + f_n,$$

and similar as (5.21) by the chain rule and the integration by parts,

$$\begin{aligned} \partial_t \int \chi_R \tilde{\beta}_\delta(w_n) &= \int (\Delta_v \chi_R - v \cdot \nabla_x \chi_R) \tilde{\beta}_\delta(w_n) - \int \chi_R \tilde{\beta}_\delta''(w_n) |\nabla_v w_n|^2 \\ &\quad - \int (b_n \cdot \nabla_v \chi_R) \tilde{\beta}_\delta(w_n) + \int f_n \chi_R \tilde{\beta}_\delta'(w_n). \end{aligned}$$

Since $|\tilde{\beta}_\delta(r)| \leq |r|$, $|\tilde{\beta}_\delta'(r)| \leq 1$ and $\int |w_n| \leq 2$, there is a constant $C > 0$ independent of R such that

$$\partial_t \int \chi_R \tilde{\beta}_\delta(w_n) \leq C(R^{-2} + \|b_n\|_{L^\infty} R^{-1}) + \int |f_n| \chi_R.$$

Integrating both sides from 0 to t and letting $R \rightarrow \infty$ and $\delta \rightarrow 0$, we obtain

$$\|w_n(t)\|_{L^1} \leq \|w_n(0)\|_{L^1} + \int_0^t \|f_n\|_{L^1} ds.$$

Note that by Hölder's inequality,

$$\begin{aligned} \int_0^t \|f_n\|_{L^1} ds &\leq \int_0^t \|K_n * \langle w \rangle\|_{L^\infty} \|\nabla_v u_1^n\|_{L^1} ds \\ &\leq \|K\|_{L^\infty} \left(\int_0^t \|w\|_{L^1}^2 ds \right)^{\frac{1}{2}} \|\nabla_v u_1^n\|_{L_t^2 L^1}. \end{aligned}$$

Since τ does not change entropy and (5.14), (5.15) and (5.24) also hold for u_n which implies that

$$\|\nabla_v u_1^n\|_{L_t^2 L^1} \leq H(\varphi_1^n) - H(u_1^n(t)) \lesssim_C \int |\varphi_1 \ln \varphi_1| + (\|\varphi_1\|_{L^1(\rho_0)} + 1),$$

where C only depends on ρ_0 . Thus,

$$\|w_n(t)\|_{L^1} \leq \|w_n(0)\|_{L^1} + C_{K, \varphi_1, \rho_0} \left(\int_0^t \|w(s)\|_{L^1}^2 ds \right)^{\frac{1}{2}}.$$

Letting $n \rightarrow \infty$ and by (5.29) and Fatou's lemma,

$$\|w(t)\|_{L^1} \leq \|w(0)\|_{L^1} + C_{K, \varphi_1, \rho_0} \left(\int_0^t \|w(s)\|_{L^1}^2 ds \right)^{\frac{1}{2}},$$

which implies (5.9) by Gronwall's inequality. \square

6. Nonlinear martingale problem with singular drifts. In this section we consider the following nonlinear kinetic DDSDE with distributional drift:

$$(6.1) \quad dX_t = V_t dt, \quad dV_t = W(t, X_t, V_t) dt + (K * \mu_{X_t})(X_t) dt + \sqrt{2} dB_t,$$

where $W \in \mathbb{B}_T^\alpha(\varrho^\kappa)$ for some $\kappa > 0$ and $K(x) : \mathbb{R}^d \rightarrow \mathbb{R}^d$ satisfies that

$$K \in \cup_{\beta > \alpha-1} \mathbf{C}^{\beta/3}.$$

Here μ_{X_t} stands for the law of X_t in \mathbb{R}^d , and for a probability measure μ in \mathbb{R}^d ,

$$K * \mu(x) := \int_{\mathbb{R}^d} K(x-y) \mu(dy).$$

We first give the definition of solutions. Fix $T > 0$. Let \mathcal{C}_T be the space of all continuous functions from $[0, T]$ to \mathbb{R}^{2d} , and $\mathcal{P}(\mathcal{C}_T)$ the set of all probability measures over \mathcal{C}_T . Let \mathcal{B}_t be the natural σ -filtration, and z be the canonical process over \mathcal{C}_T , i.e., for $\omega \in \mathcal{C}_T$,

$$z_t(\omega) = (x_t(\omega), v_t(\omega)) = \omega_t.$$

As mentioned in the introduction, we define the martingale problem by using the linear version of the Kolomogorov backward equation. More precisely, for a continuous curve $\mu : [0, T] \rightarrow \mathcal{P}(\mathbb{R}^d)$ with respect to the weak convergence, define

$$\mathcal{L}_t^\mu := \Delta_v + v \cdot \nabla_x + (W(t) + K * \mu_t) \cdot \nabla_v.$$

As in Remark 5.3, it is easy to see that $W + K * \mu_t \in \mathbb{B}_T^\alpha(\varrho^\kappa)$. Let $f \in L_T^\infty C_b(\mathbb{R}^{2d})$ and $\varphi \in \mathbf{C}_a^\gamma(\mathbb{R}^{2d})$ for some $\gamma > 1 + \alpha$ and $\vartheta := \frac{9}{2-3\alpha}$. By Theorem 4.7, there exists a unique paracontrolled solution $u_f^\mu \in \mathbb{S}_{T;a}^{2-\alpha}(\varrho^{2(\vartheta+1)\kappa})$ to the following equation:

$$(6.2) \quad \partial_t u_f^\mu + \mathcal{L}_t^\mu u_f^\mu = f, \quad u_f^\mu(T) = \varphi.$$

For any $\delta > 0$, let $\mathcal{P}_\delta(\mathbb{R}^{2d})$ be the space of all probability measures ν on \mathbb{R}^{2d} with

$$\int_{\mathbb{R}^{2d}} \varrho(z)^{-\delta} \nu(dz) \asymp \int_{\mathbb{R}^{2d}} (1 + |z|_a^\delta) \nu(dz) < \infty.$$

We introduce the following notion about the martingale problem.

DEFINITION 6.1. (Martingale problem) Let $\delta > 0$. A probability measure $\mathbb{P} \in \mathcal{P}(\mathcal{C}_T)$ is called a martingale solution to SDE (6.1) starting from $\nu \in \mathcal{P}_\delta(\mathbb{R}^{2d})$, if $\mathbb{P} \circ Z_0^{-1} = \nu$ and for all $f \in L_T^\infty C_b(\mathbb{R}^{2d})$ and $\varphi \in \mathbf{C}_a^\gamma(\mathbb{R}^{2d})$ with some $\gamma > 1 + \alpha$,

$$M_t := u_f^\mu(t, z_t) - u_f^\mu(0, z_0) - \int_0^t f(s, z_s) ds$$

is a martingale under \mathbb{P} with respect to (\mathcal{B}_t) , where $\mu_t := \mathbb{P} \circ x_t^{-1}$ and u_f^μ is the paracontrolled solution to (6.2). The set of all martingale solutions \mathbb{P} associated with W, K and starting from ν is denoted by $\mathcal{M}_\nu(W, K)$.

REMARK 6.2. The moment assumption for ν is necessary for making sense of $\mathbb{E} u_f^\mu(0, z_0)$ since the solution u_f^μ lives in weighted spaces.

Our main result of this section is the following:

THEOREM 6.3. *Let $\alpha \in (\frac{1}{2}, \frac{2}{3})$ and $\vartheta := \frac{9}{2-3\alpha}$. Suppose that for some $\kappa \in (0, \frac{1}{2\vartheta+2}]$ and $\beta > \alpha - 1$,*

$$W \in \mathbb{B}_T^\alpha(\varrho^\kappa), \quad K \in \mathbf{C}^{\beta/3}.$$

Then for any $\nu \in \mathcal{P}_\delta(\mathbb{R}^{2d})$ with $\delta > (4\vartheta + 4)\kappa$, there exists at least one martingale solution $\mathbb{P} \in \mathcal{M}_\nu(W, K)$ to SDE (6.1). Moreover, if K is bounded measurable, then there exists at most one $\mathbb{P} \in \mathcal{M}_\nu(W, K)$.

The existence part of the proof follows by the convolution approximation. We prove the uniqueness first for the case $K = 0$ and then use Girsanov's transform. Let W_n be the approximation sequence of W , and $K_n = K * \phi_n$ with $\phi_n(x) = n^d \phi_1(nx)$ being the usual mollifier. We consider the following approximation SDE:

$$(6.3) \quad dX_t^n = V_t^n dt, \quad dV_t^n = W_n(t, Z_t^n) dt + (K_n * \mu_{X_t^n})(X_t^n) dt + \sqrt{2} dB_t,$$

where $Z^n = (X^n, V^n)$ and $\mathbf{P}^{-1} \circ Z_0^n = \nu$. Since W_n and K_n are globally Lipschitz, it is well-known that there is a unique strong solution Z^n to (6.3) (see [Wa18, Theorem 2.1]). We first establish the following uniform moment estimates for V_t^n by a PDE's method.

LEMMA 6.4. *Suppose $\delta > (4\vartheta + 4)\kappa$. For any $p \in (2, \frac{\delta}{(2\vartheta+2)\kappa}]$, there is a constant $C > 0$ such that for all $0 \leq s < t \leq T$,*

$$\sup_n \mathbf{E} |V_t^n - V_s^n|^p \lesssim_C (t - s)^{p/2}.$$

PROOF. By SDE (6.3), it suffices to prove that

$$(6.4) \quad \sup_n \mathbf{E} \left| \int_s^t b_n(r, Z_r^n) dr \right|^p \lesssim_C |t - s|^{(2-\alpha)p/2},$$

where

$$b_n(t, x, v) := W_n(t, x, v) + (K_n * \mu_{X_t^n})(x) \in L_T^\infty C_b^\infty(\mathbb{R}^{2d}).$$

Fix $t \in [0, T]$. Let u_n be the smooth solution to the following kinetic equation

$$\partial_t u_n = \Delta_v u_n + v \cdot \nabla_x u_n + b_n^t \cdot \nabla_v u_n - b_n^t, \quad u_n(0) = 0,$$

where $b_n^t(s, z) = b_n(t - s, z)$. By Theorem 4.7, for $\sigma := (2\vartheta + 2)\kappa$, we have

$$(6.5) \quad \sup_n \|u_n\|_{\mathbb{S}_{t,a}^{2-\alpha}(\varrho^\sigma)} < \infty.$$

Let

$$u_n^t(s) := u_n(t - s).$$

Then u_n^t satisfies the following equation

$$\partial_r u_n^t + \Delta_v u_n^t + v \cdot \nabla_x u_n^t + b_n \cdot \nabla_v u_n^t = b_n, \quad u_n^t(t) = 0.$$

By (6.3) and Itô's formula, one sees that

$$\begin{aligned} \int_s^t b_n(r, Z_r^n) dr &= u_n^t(t, Z_t^n) - u_n^t(s, Z_s^n) - \sqrt{2} \int_s^t \nabla_v u_n^t(r, Z_r^n) dB_r \\ &= u_n^t(t, \Gamma_{t-s} Z_s^n) - u_n^t(s, Z_s^n) - \sqrt{2} \int_s^t \nabla_v u_n^t(r, Z_r^n) dB_r, \end{aligned}$$

where the second step is due to

$$u_n^t(t, Z_t^n) = 0 = u_n^t(t, \Gamma_{t-s} Z_s^n).$$

By BDG's inequality, (3.19) and (3.30), we have for any $p \in (2, \frac{\delta}{(2\vartheta+2)\kappa}]$,

$$\begin{aligned} \mathbf{E} \left| \int_s^t b_n(r, Z_r^n) dr \right|^p &\lesssim (t-s)^{(2-\alpha)p/2} \|u_n\|_{\mathbb{S}_{t,a}^{2-\alpha}(\varrho^\sigma)}^p \mathbf{E} \varrho^{-p\sigma}(Z_s^n) \\ &\quad + \|\nabla_v u_n\|_{\mathbb{L}_t^\infty(\varrho^\sigma)}^p \mathbf{E} \left(\int_s^t \varrho^{-2\sigma}(Z_r^n) dr \right)^{p/2} \\ (6.6) \quad &\lesssim ((t-s)^{(2-\alpha)p/2} + (t-s)^{p/2}) \|u_n\|_{\mathbb{S}_{t,a}^{2-\alpha}(\varrho^\sigma)}^p \sup_{r \in [0,t]} \mathbf{E} \varrho^{-p\sigma}(Z_r^n). \end{aligned}$$

Finally, since $p\sigma \leq \delta$, as in showing (5.19), we have

$$(6.7) \quad \sup_n \sup_{s \in [0,t]} \mathbf{E} \varrho^{-\delta}(Z_s^n) \lesssim \int_{\mathbb{R}^{2d}} \varrho^{-\delta}(z) \nu(dz).$$

By (6.6), (6.7) and (6.5), we obtain (6.4). The proof is complete. \square

Now we give the following convergence result.

LEMMA 6.5. *Let $(\mu_n)_{n \in \mathbb{N}}$ be a family of probability measures on $C([0, T]; \mathbb{R}^d)$. Suppose that μ_n weakly converges to μ and $K \in \mathbf{C}^\beta$. Then for any $\beta_0 < \beta$, we have*

$$\lim_{n \rightarrow \infty} \|K_n * \mu_n - K * \mu\|_{L_T^\infty \mathbf{C}^{\beta_0}} = 0.$$

PROOF. It suffices to prove the result for β_0 satisfying $\beta - \beta_0 \in (0, 1)$. By Skorohod's representation theorem, there are a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ and random variables X_n, X with values in $C([0, T]; \mathbb{R}^d)$ so that

$$\lim_{n \rightarrow \infty} \sup_{s \in [0, T]} |X_n(s) - X(s)| = 0 \text{ a.s.,}$$

and

$$\tilde{\mathbb{P}} \circ (X_n)^{-1} = \mu_n, \quad \tilde{\mathbb{P}} \circ X^{-1} = \mu.$$

Let \mathcal{R}_j be the usual block operator with $a = (1, \dots, 1)$ in (2.4). By similar arguments as (2.25) we have for any $j \geq -1$ and $h \in \mathbb{R}^d$

$$(6.8) \quad \|\mathcal{R}_j K(\cdot + h) - \mathcal{R}_j K\|_{L^\infty} \lesssim |h|^{\beta-\beta_0} \|\mathcal{R}_j K\|_{\mathbf{C}^{\beta-\beta_0}} \lesssim 2^{-\beta_0 j} |h|^{\beta-\beta_0} \|K\|_{\mathbf{C}^\beta}.$$

From this, we derive that

$$\begin{aligned} \|\mathcal{R}_j K_m - \mathcal{R}_j K\|_{L^\infty} &\leq \sup_x \int_{\mathbb{R}^d} |\mathcal{R}_j K(x-y) - \mathcal{R}_j K(x)| \phi_m(y) dy \\ (6.9) \quad &\lesssim m^{-(\beta-\beta_0)} \|\mathcal{R}_j K\|_{\mathbf{C}^{\beta-\beta_0}} \lesssim 2^{-\beta_0 j} m^{-(\beta-\beta_0)} \|K\|_{\mathbf{C}^\beta}. \end{aligned}$$

Note that

$$\begin{aligned} |\mathcal{R}_j(K_n * \mu_n(s) - K * \mu(s))(x)| &= |\tilde{\mathbb{E}} \mathcal{R}_j K_n(x - X_n(s)) - \tilde{\mathbb{E}} \mathcal{R}_j K(x - X(s))| \\ &\leq \tilde{\mathbb{E}} |\mathcal{R}_j K_n(x - X_n(s)) - \mathcal{R}_j K(x - X_n(s))| \\ &\quad + \tilde{\mathbb{E}} |\mathcal{R}_j K(x - X_n(s)) - \mathcal{R}_j K(x - X(s))| \\ &=: \mathcal{J}_{n,j}^{(1)}(s, x) + \mathcal{J}_{n,j}^{(2)}(s, x). \end{aligned}$$

For $\mathcal{J}_{n,j}^{(1)}(s, x)$, by (6.9) we have

$$\|\mathcal{J}_{n,j}^{(1)}\|_{L_T^\infty L^\infty} \leq \|\mathcal{R}_j K_n - \mathcal{R}_j K\|_{L^\infty} \lesssim 2^{-\beta_0 j} n^{-(\beta-\beta_0)} \|K\|_{\mathbf{C}^\beta}.$$

For $\mathcal{J}_{n,j}^{(2)}(s, x)$, by the dominated convergence theorem and (6.8), we have

$$\begin{aligned} & \lim_{n \rightarrow \infty} \sup_j 2^{\beta_0 j} \|\mathcal{J}_{n,j}^{(2)}\|_{L_T^\infty L^\infty} \\ & \lesssim \tilde{\mathbb{E}} \left(\lim_{n \rightarrow \infty} \sup_{s \in [0, T]} \sup_j 2^{\beta_0 j} |\mathcal{R}_j K(x - X_n(s)) - \mathcal{R}_j K(x - X(s))| \right) \\ & \lesssim \tilde{\mathbb{E}} \left(\lim_{n \rightarrow \infty} \sup_{s \in [0, T]} |X_n(s) - X(s)|^{\beta-\beta_0} \right) \|K\|_{\mathbf{C}^\beta} = 0. \end{aligned}$$

From these two estimates, we derive the desired limit. \square

Now we can give the proof of Theorem 6.3.

PROOF OF THEOREM 6.3. (Existence) Let $\mathbb{P}_n = \mathbf{P} \circ Z^n$ be the law of Z^n in $(\mathcal{C}_T, \mathcal{B}_T)$. By Lemma 6.4 and Kolmogorov's criterion, we have for each $\varepsilon > 0$,

$$\lim_{\delta \rightarrow 0} \sup_n \mathbf{P} \left(\sup_{s \neq t \in [0, T], |t-s| \leq \delta} |V_t^n - V_s^n| > \varepsilon \right) = 0.$$

Since $X_t^n = \int_0^t V_s^n ds + X_0$ and

$$\lim_{R \rightarrow \infty} \sup_n \mathbf{P}(|Z_0^n| > R) = \lim_{R \rightarrow \infty} \nu\{z : |z| > R\} = 0,$$

it is easy to see that for each $\varepsilon > 0$,

$$\lim_{\delta \rightarrow 0} \sup_n \mathbf{P} \left(\sup_{s \neq t \in [0, T], |t-s| \leq \delta} |Z_t^n - Z_s^n| > \varepsilon \right) = 0.$$

Thus $(\mathbb{P}_n)_{n \in \mathbb{N}}$ is tight in \mathcal{C}_T .

Let \mathbb{P} be any accumulation point of $(\mathbb{P}_n)_{n \in \mathbb{N}}$. Without loss of generality, we assume \mathbb{P}_n weakly converges to \mathbb{P} . Let

$$\mu_n := \mathbb{P}_n \circ X^{-1}, \quad \mu := \mathbb{P} \circ X^{-1}.$$

Let ϕ_n be the usual mollifier in \mathbb{R}^{2d} and define

$$f_n(t, z) := f(t, \cdot) * \phi_n(z), \quad \varphi_n(z) := \varphi * \phi_n(z)$$

and

$$b_n := W_n + K_n * \mu_n, \quad b := W + K * \mu.$$

Since $b_n, f_n \in L_T^\infty C_b^\infty(\mathbb{R}^{2d})$, it is well known that there is a smooth solution $u_n \in L_T^\infty C_b^\infty(\mathbb{R}^{2d})$ to the following PDE:

$$(6.10) \quad \partial_t u_n + (\Delta_v + v \cdot \nabla_x + b_n \cdot \nabla_v) u_n = f_n, \quad u_n(T) = \varphi_n.$$

Now we define two functionals on \mathcal{C}_T :

$$M_t^n := M_t^n(z) := u_n(t, z_t) - u_n(0, z_0) - \int_0^t f_n(s, z_s) ds$$

and

$$M_t := M_t(z) := u_f^\mu(t, z_t) - u_f^\mu(0, z_0) - \int_0^t f(s, z_s) ds.$$

We want to show that for any $0 \leq s < t \leq T$ and \mathcal{B}_s -measurable bounded continuous functional G_s on \mathcal{C}_T ,

$$(6.11) \quad \mathbb{E}^\mathbb{P}(M_t G_s) = \mathbb{E}^\mathbb{P}(M_s G_s).$$

For each $n \in \mathbb{N}$, by (6.3), (6.10) and Itô's formula,

$$M_t^n(Z^n) = \int_0^t \nabla_v u_n(s, Z_s^n) dB_s$$

is a \mathbf{P} -martingale. Hence,

$$\mathbb{E}^{\mathbb{P}_n}(M_t^n G_s) = \mathbf{E}(M_t^n(Z^n) G_s(Z^n)) = \mathbf{E}(M_s^n(Z^n) G_s(Z^n)) = \mathbb{E}^{\mathbb{P}_n}(M_s^n G_s).$$

Thus, to show (6.11), it suffices to show that

$$(6.12) \quad \lim_{n \rightarrow \infty} \mathbb{E}^{\mathbb{P}_n}(M_t^n G_s) = \mathbb{E}^\mathbb{P}(M_t G_s).$$

Note that by Lemma 6.5, for $\gamma \in (\alpha - 1, \beta)$,

$$\lim_{n \rightarrow \infty} \|K_n * \mu_n - K * \mu\|_{\mathcal{C}_{T,a}^\gamma} = 0,$$

which by Lemma 3.16 implies that

$$(b, f) \in \mathbb{B}_T^\alpha(\varrho^\kappa, 1) \text{ with approximation sequence } (b_n, f_n).$$

Thus by Theorem 4.7, for any $\sigma > (2 + 2\vartheta)\kappa$,

$$(6.13) \quad \sup_n \|u_n\|_{\mathbb{L}_T^\infty(\varrho^\sigma)} < \infty, \quad \lim_{n \rightarrow \infty} \|u_n - u_f^\mu\|_{\mathbb{L}_T^\infty(\varrho^\sigma)} = 0.$$

Moreover, by (6.7) we have for any $\delta > (4 + 4\vartheta)\kappa$,

$$\sup_n \mathbb{E}^{\mathbb{P}_n}(\varrho^{-\delta}(z_t) + \varrho^{-\delta}(z_0)) < \infty.$$

Note that

$$|M_t^n - M_t| \leq \|u_n - u_f^\mu\|_{\mathbb{L}_T^\infty(\varrho^\sigma)} \left(\varrho^{-\sigma}(z_t) + \varrho^{-\sigma}(z_0) \right) + \int_0^t |f_n - f|(s, z_s) ds.$$

Since for each $s \in [0, t]$, $(\mathbb{P}_n \circ z_s^{-1})_{n \in \mathbb{N}}$ is tight, and for any $R > 0$,

$$\lim_{n \rightarrow \infty} \sup_{|z| \leq R} |f_n(s, z) - f(s, z)| = 0,$$

it is easy to see that

$$\begin{aligned} & \lim_{n \rightarrow \infty} \mathbb{E}^{\mathbb{P}_n} \int_0^t |f_n - f|(s, z_s) ds \\ & \leq \int_0^t \lim_{n \rightarrow \infty} \sup_{|z| \leq R} |f_n(s, z) - f(s, z)| ds + \frac{C}{R^\delta} \int_0^t \sup_n \mathbb{E}^{\mathbb{P}_n} \varrho^{-\delta}(z_s) ds. \end{aligned}$$

Thus, by (6.13),

$$(6.14) \quad \lim_{n \rightarrow \infty} |\mathbb{E}^{\mathbb{P}_n}(M_t^n G_s) - \mathbb{E}^\mathbb{P}(M_t G_s)| \leq \|G_s\|_\infty \lim_{n \rightarrow \infty} \mathbb{E}^{\mathbb{P}_n} |M_t^n - M_t| = 0.$$

Moreover, since M_t is a continuous functional on \mathcal{C}_T , and

$$\sup_n \mathbb{E}^{\mathbb{P}_n} |M_t G_s|^2 \stackrel{(6.13)}{\lesssim} \left[\sup_n \mathbb{E}^{\mathbb{P}_n} \left(\varrho^\delta(z_t) + \varrho^\delta(z_0) \right) + \|f\|_{L^\infty}^2 \right] \|G\|_{L^\infty} \stackrel{(6.7)}{<} \infty,$$

it is easy to see that

$$\lim_{n \rightarrow \infty} \mathbb{E}^{\mathbb{P}_n} (M_t G_s) = \mathbb{E}^{\mathbb{P}} (M_t G_s).$$

Combining the above calculations, we get (6.12). Thus, we complete the proof of the existence of a martingale solution.

(Uniqueness) First of all, we show the uniqueness for linear SDE, i.e., $K = 0$. Let $\mathbb{P}_1, \mathbb{P}_2 \in \mathcal{M}_\nu(W, 0)$ be two solutions of the martingale problem. Let $f \in L_T^\infty C_b(\mathbb{R}^{2d})$ and let u be the unique paracontrolled solution to (6.2) with $u(t) = 0$. By Definition 6.1, we have

$$\int_{\mathbb{R}^{2d}} u(0, z) \nu(dz) = -\mathbb{E}^{\mathbb{P}_i} \int_0^T f(s, Z_s) ds, \quad i = 1, 2,$$

which means that

$$\int_0^T \mathbb{E}^{\mathbb{P}_1} f(s, Z_s) ds = \int_0^T \mathbb{E}^{\mathbb{P}_2} f(s, Z_s) ds.$$

Hence, for any $f \in C_b(\mathbb{R}^{2d})$ and $t \in [0, T]$,

$$\mathbb{E}^{\mathbb{P}_1} f(Z_t) = \mathbb{E}^{\mathbb{P}_2} f(Z_t).$$

From this, by a standard way (see Theorem 4.4.3 in [EK86]), we derive that

$$\mathbb{P}_1 = \mathbb{P}_2.$$

For general nonlinear SDE, we use Girsanov's transformation method. Let $\mathbb{P}_1, \mathbb{P}_2 \in \mathcal{M}_\nu(W, K)$ be two solutions of the martingale problem. Let W_n, K_n be the approximations of W and K as above. We consider the following approximation of linearized SDEs: for $i = 1, 2$,

$$(6.15) \quad dX_t^{i,n} = V_t^{i,n} dt, \quad dV_t^{i,n} = W_n(Z_t^{i,n}) dt + (K_n * \mu_t^i)(X_t^{i,n}) + \sqrt{2} dB_t,$$

where $\mu_t^i := \mathbb{P}_i \circ x_t^{-1}$. As in the proof of the existence part, and due to the uniqueness of linear SDEs, for $i = 1, 2$, the law of $Z^{i,n} = (X^{i,n}, V^{i,n})$ weakly converges to \mathbb{P}_i as $n \rightarrow \infty$. In particular, for any $\varphi \in C_b(\mathbb{R}^d)$,

$$\mathbb{E} \varphi(X_t^{i,n}) \rightarrow \mathbb{E}^{\mathbb{P}_i} \varphi(x_t), \quad i = 1, 2.$$

On the other hand, we define

$$A_t^{i,n} := \exp \left\{ -\frac{1}{\sqrt{2}} \int_0^t (K_n * \mu_s^i)(X_s^{i,n}) dB_s - \frac{1}{4} \int_0^t |(K_n * \mu_s^i)(X_s^{i,n})|^2 ds \right\}.$$

Since

$$(6.16) \quad \|K_n * \mu_s^i\|_{L^\infty} \leq \|K\|_{L^\infty},$$

by Girsanov's theorem, under the new probability measure $Q^{i,n} := A_T^{i,n} \mathbf{P}$, for $t \in [0, T]$

$$\tilde{B}_t^{i,n} := \frac{1}{\sqrt{2}} \int_0^t (K_n * \mu_s^i)(X_s^{i,n}) ds + B_t$$

is still a Brownian motion, and

$$dX_t^{i,n} = V_t^{i,n} dt, \quad dV_t^{i,n} = W_n(Z_t^{i,n}) dt + \sqrt{2} d\tilde{B}_t^{i,n}.$$

Since the above SDE admits a unique weak solution, we have

$$Q^{1,n} \circ (Z^{1,n})^{-1} = Q^{2,n} \circ (Z^{2,n})^{-1}.$$

Thus, for any $\varphi \in C_b(\mathbb{R}^d)$,

$$\mathbf{E}\varphi(X_t^{1,n}) = \mathbf{E}(A_T^{1,n}\varphi(X_t^{1,n})/A_T^{1,n}) = \mathbf{E}(A_T^{2,n}\varphi(X_t^{2,n})Y_T^n)$$

and

$$|\mathbf{E}\varphi(X_t^{1,n}) - \mathbf{E}\varphi(X_t^{2,n})| \leq \|\varphi\|_{L^\infty} |\mathbf{E}A_T^{2,n}Y_T^n - 1|,$$

where

$$Y_T^n := \exp \left\{ \frac{1}{\sqrt{2}} \int_0^T (K_n * \mu_s^1)(X_s^{2,n}) dB_s + \frac{1}{4} \int_0^T |(K_n * \mu_s^1)(X_s^{2,n})|^2 ds \right\}.$$

On the other hand, since $t \mapsto A_t^{2,n}, Y_t^n$ are two semimartingales, by the product Itô formula, we obtain

$$\begin{aligned} A_T^{2,n}Y_T^n - 1 &= \int_0^T A_s^{2,n}dY_s^n + \int_0^T Y_s^n dA_s^{2,n} + [A^{2,n}, Y^n]_T \\ &= \frac{1}{\sqrt{2}} \int_0^T A_s^{2,n}Y_s^n (K_n * \mu_s^1 - K_n * \mu_s^2)(X_s^{2,n}) dB_s \\ &\quad + \frac{1}{4} \int_0^T A_s^{2,n}Y_s^n (|(K_n * \mu_s^1)(X_s^{2,n})|^2 - |(K_n * \mu_s^2)(X_s^{2,n})|^2) ds \\ &\quad + \frac{1}{4} \int_0^T A_s^{2,n}Y_s^n |(K_n * \mu_s^1)(X_s^{2,n}) - (K_n * \mu_s^2)(X_s^{2,n})|^2 ds. \end{aligned} \tag{6.17}$$

By (6.16), it follows from BDG's inequality and Hölder's inequality that for any $p \geq 2$,

$$\begin{aligned} \mathbb{E}|A_t^{2,n}Y_t^n|^p &\lesssim 1 + \|K\|_{L^\infty} \mathbb{E} \left(\int_0^t |A_s^{2,n}Y_s^n|^2 ds \right)^{p/2} + \|K\|_{L^\infty} \int_0^t \mathbb{E}|A_s^{2,n}Y_s^n|^p ds \\ &\lesssim 1 + \|K\|_{L^\infty} \int_0^t \mathbb{E}|A_s^{2,n}Y_s^n|^p ds, \end{aligned}$$

which by Gronwall's inequality implies that

$$\sup_n \sup_{s \in [0, T]} \mathbb{E}|A_s^{2,n}Y_s^n|^p < \infty.$$

Hence, by (6.17), BDG's inequality and Hölder's inequality, we arrive at

$$\mathbf{E}|A_T^{2,n}Y_T^n - 1| \lesssim \left(\int_0^T \|\mu_s^2 - \mu_s^1\|_{TV}^2 ds \right)^{1/2} + \int_0^T \|\mu_s^2 - \mu_s^1\|_{TV} ds,$$

where $\|\cdot\|_{TV}$ stands for the total variation norm of a signed measure. Combining the above calculations, we obtain that for all $\varphi \in C_b(\mathbb{R}^d)$,

$$\begin{aligned} |\mathbb{E}^{\mathbb{P}^1}\varphi(x_T) - \mathbb{E}^{\mathbb{P}^2}\varphi(x_T)| &= \lim_{n \rightarrow \infty} |\mathbf{E}\varphi(X_T^{1,n}) - \mathbf{E}\varphi(X_T^{2,n})| \\ &\lesssim \|\varphi\|_\infty \left(\int_0^T \|\mu_s^2 - \mu_s^1\|_{TV}^2 ds \right)^{1/2}, \end{aligned}$$

which in turn implies that

$$\|\mu_T^2 - \mu_T^1\|_{TV}^2 \lesssim \int_0^T \|\mu_s^2 - \mu_s^1\|_{TV}^2 ds.$$

By Gronwall's inequality, $\mu_t^1 = \mu_t^2$. Finally, we use the same argument as the uniqueness for linear equations to derive the uniqueness for nonlinear SDEs. \square

7. Existence of renormalized pairs in probabilistic sense. In this section we perform the construction of stochastic objects, i.e. renormalized pairs of the stochastic kinetic equations by probabilistic calculations. We state the main result in Subsection 7.1. In Subsection 7.2 we give examples of Gaussian noise satisfying the general assumptions. In the last subsection we give the proof of the main result.

7.1. Statement of main result. Let μ be a symmetric tempered measure on \mathbb{R}^{2d} , that is, for some $l \in \mathbb{N}$,

$$(7.1) \quad \int_{\mathbb{R}^{2d}} (1 + |\xi|)^{-l} \mu(d\xi) < \infty.$$

Let $L_{\mathbb{C}}^2(\mu)$ be the complex-valued Hilbert space with inner product

$$\langle f, g \rangle_{L_{\mathbb{C}}^2(\mu)} := \int_{\mathbb{R}^{2d}} f(\zeta) \overline{g(\zeta)} \mu(d\zeta) < \infty.$$

Let \mathbb{H} be the completion of $\mathcal{S}(\mathbb{R}^{2d})$ with respect to the inner product

$$\langle f, g \rangle_{\mathbb{H}} := \langle \hat{f}, \hat{g} \rangle_{L_{\mathbb{C}}^2(\mu)}.$$

DEFINITION 7.1. Let X be a Gaussian field on \mathbb{H} , i.e., X is a continuous linear operator from \mathbb{H} to $L^2(\Omega, \mathbf{P})$, and for each $f \in \mathbb{H}$, $X(f)$ is a real-valued Gaussian random variable with mean zero and variance $\|f\|_{\mathbb{H}}^2$. In particular,

$$(7.2) \quad \mathbf{E}(X(f)X(g)) = \int_{\mathbb{R}^{2d}} \hat{f}(\zeta) \hat{g}(-\zeta) \mu(d\zeta).$$

We call X the Gaussian noise with spectral measure μ (see [SV97]).

The following result is the main result of this section.

THEOREM 7.2. Suppose that μ is a Radon measure and satisfies

$$(S) \quad \mu(d\xi, d\eta) = \mu(d\xi, -d\eta) = \mu(-d\xi, d\eta),$$

and for some $\beta \in (\frac{1}{2}, \frac{2}{3})$,

$$(A^\beta) \quad \sup_{\zeta' \in \mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} \frac{\mu(d\zeta)}{(1 + |\zeta' + \zeta|_a)^{2\beta}} < \infty.$$

Let $W = (X_1, \dots, X_d)$ be d -independent Gaussian noise with common spectral measure μ . Then for any $\kappa > 0$ and $\alpha > \beta$, it holds that

$$\mathbf{P}\{\omega : W(\cdot, \omega) \in \mathbb{B}_T^\alpha(\varrho^\kappa)\} = 1.$$

REMARK 7.3. (i) Condition (A^β) implies that for any $\sigma, \gamma \geq 0$ with $\sigma + \gamma = 2\beta$,

$$(7.3) \quad \sup_{\zeta' \in \mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} \frac{\mu(d\zeta)}{(1 + |\zeta' + \zeta|_a)^\sigma (1 + |\zeta|_a)^\gamma} < \infty.$$

Indeed, it follows by the simple observation:

$$\begin{aligned} & \left(\int_{|\zeta|_a > |\zeta' + \zeta|_a} + \int_{|\zeta|_a \leq |\zeta' + \zeta|_a} \right) \frac{\mu(d\zeta)}{(1 + |\zeta' + \zeta|_a)^\sigma (1 + |\zeta|_a)^\gamma} \\ & \leq \int_{\mathbb{R}^{2d}} \frac{\mu(d\zeta)}{(1 + |\zeta' + \zeta|_a)^{2\beta}} + \int_{\mathbb{R}^{2d}} \frac{\mu(d\zeta)}{(1 + |\zeta|_a)^{2\beta}}. \end{aligned}$$

(ii) The symmetric assumption of μ in the second variable η allows us to use some cancellation to show the convergence in (7.23) below (see (7.22) below). In the classical case by symmetry the terms in the 0th Wiener chaos are zero. However, in the kinetic setting these terms are not zero but converge in the required weighted Besov space after subtracting suitable renormalization terms, which by symmetry assumption of μ are zero.

Let $\varphi \in \mathcal{S}(\mathbb{R}^{2d})$ be a symmetric function and define

$$X_\varphi(z) := X(\varphi(z - \cdot)).$$

Then by Lemma B.1 in appendix, $z \mapsto X_\varphi(z)$ has a smooth version.

Let $\varphi, \varphi' \in \mathcal{S}(\mathbb{R}^{2d})$ be two symmetric functions. For $H \in \mathcal{S}(\mathbb{R}^{4d})$, define

$$(7.4) \quad (X_\varphi \otimes X_{\varphi'})(H) := \int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} H(z, z') X_\varphi(z) X_{\varphi'}(z') dz dz'.$$

The following result is standard by the properties of Gaussian fields (see [SV97]). For the reader's convenience, we put the proof in Appendix B.

LEMMA 7.4. For any $H \in \mathcal{S}(\mathbb{R}^{4d})$, it holds that

$$(7.5) \quad \mathbf{E}\left((X_\varphi \otimes X_{\varphi'})(H)\right) = \int_{\mathbb{R}^{2d}} \hat{H}(\zeta, -\zeta) \hat{\varphi}(\zeta) \hat{\varphi}'(\zeta) \mu(d\zeta),$$

and

$$(7.6) \quad \text{Var}\left((X_\varphi \otimes X_{\varphi'})(H)\right) = 2 \int_{\mathbb{R}^{4d}} |(\text{Sym } \hat{H}_{\varphi, \varphi'})(\zeta, \zeta')|^2 \mu(d\zeta) \mu(d\zeta'),$$

where $\hat{H}_{\varphi, \varphi'}(\zeta, \zeta') := \hat{H}(\zeta, \zeta') \hat{\varphi}(\zeta) \hat{\varphi}'(\zeta')$ and

$$(7.7) \quad (\text{Sym } \hat{H}_{\varphi, \varphi'})(\zeta, \zeta') := (\hat{H}_{\varphi, \varphi'}(\zeta, \zeta') + \hat{H}_{\varphi, \varphi'}(\zeta', \zeta))/2.$$

If we do Wiener chaos decomposition for $(X_\varphi \otimes X_{\varphi'})(H)$ (see [Nua06, Ch.1]), $I_0 := \mathbf{E}\left((X_\varphi \otimes X_{\varphi'})(H)\right)$ corresponds to the term in the 0th Wiener chaos and $I_2 := (X_\varphi \otimes X_{\varphi'})(H) - \mathbf{E}\left((X_\varphi \otimes X_{\varphi'})(H)\right)$ gives the term in the second Wiener chaos.

REMARK 7.5. If X, Y are two independent Gaussian fields with the same spectral measure, then $\mathbf{E}((X_\varphi \otimes Y_{\varphi'})(H)) = 0$ and

$$(7.8) \quad \mathbf{E}\left((X_\varphi \otimes Y_{\varphi'})(H)\right)^2 = \int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} |\hat{H}(\zeta, \zeta')|^2 |\hat{\varphi}(\zeta)|^2 |\hat{\varphi}'(\zeta')|^2 \mu(d\zeta) \mu(d\zeta').$$

7.2. *Examples for (\mathbf{A}^β) .* In this subsection we provide three examples for condition (\mathbf{A}^β) to illustrate our result. We need the following simple lemma.

LEMMA 7.6. *For $\beta_1, \beta_2 \in [0, d)$ and $\gamma_1, \gamma_2 \geq 0$ with*

$$\gamma_1 + \beta_1 > d, \quad 3\beta_1 + \beta_2 + \gamma_2 > 4d,$$

it holds that

$$(7.9) \quad \sup_{\xi' \in \mathbb{R}^d} \int_{\mathbb{R}^d} \frac{d\xi}{|\xi|^{\beta_1} (1 + |\xi + \xi'|)^{\gamma_1}} < \infty,$$

and for $\zeta = (\xi, \eta) \in \mathbb{R}^{2d}$,

$$(7.10) \quad \sup_{\zeta' \in \mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} \frac{d\zeta}{|\xi|^{\beta_1} |\eta|^{\beta_2} (1 + |\zeta + \zeta'|_a)^{\gamma_2}} < \infty.$$

PROOF. For (7.9), we have

$$\begin{aligned} \int_{\mathbb{R}^d} \frac{d\xi}{|\xi|^{\beta_1} (1 + |\xi + \xi'|)^{\gamma_1}} &= \left(\int_{|\xi + \xi'| \leq |\xi|} + \int_{|\xi + \xi'| > |\xi|} \right) \frac{d\xi}{|\xi|^{\beta_1} (1 + |\xi + \xi'|)^{\gamma_1}} \\ &\leq \int_{|\xi + \xi'| \leq |\xi|} \frac{d\xi}{|\xi + \xi'|^{\beta_1} (1 + |\xi + \xi'|)^{\gamma_1}} + \int_{|\xi + \xi'| > |\xi|} \frac{d\xi}{|\xi|^{\beta_1} (1 + |\xi|)^{\gamma_1}} \\ &\leq \int_{\mathbb{R}^d} \frac{d\xi}{|\xi + \xi'|^{\beta_1} (1 + |\xi + \xi'|)^{\gamma_1}} + \int_{\mathbb{R}^d} \frac{d\xi}{|\xi|^{\beta_1} (1 + |\xi|)^{\gamma_1}} = 2 \int_{\mathbb{R}^d} \frac{d\xi}{|\xi|^{\beta_1} (1 + |\xi|)^{\gamma_1}}, \end{aligned}$$

which is finite by $\gamma_1 + \beta_1 > d$ and $\beta_1 < d$.

Next we show (7.10) by (7.9). Let

$$\theta := \frac{3(d - \beta_1)}{3(d - \beta_1) + (d - \beta_2)} = \frac{3(d - \beta_1)}{4d - 3\beta_1 - \beta_2} \in (0, 1).$$

Since $\gamma_2 > 4d - 3\beta_1 - \beta_2$, we have

$$(7.11) \quad \beta_1 + \theta\gamma_2/3 > d, \quad \beta_2 + (1 - \theta)\gamma_2 > d.$$

By $|\xi|^{1/3} \vee |\eta| \leq |\zeta|_a$, we have

$$\begin{aligned} \int_{\mathbb{R}^{2d}} \frac{d\zeta}{|\xi|^{\beta_1} |\eta|^{\beta_2} (1 + |\zeta + \zeta'|_a)^{\gamma_2}} &\leq \int_{\mathbb{R}^d} \frac{d\xi}{|\xi|^{\beta_1} (1 + |\xi + \xi'|^{1/3})^{\theta\gamma_2}} \\ &\quad \times \int_{\mathbb{R}^d} \frac{d\eta}{|\eta|^{\beta_2} (1 + |\eta + \eta'|)^{(1-\theta)\gamma_2}}, \end{aligned}$$

which in turn gives (7.10) by (7.11) and (7.9). \square

EXAMPLE 7.7. Fix $\beta \in (\frac{1}{2}, \frac{2}{3})$ and $\gamma \in (d - \frac{2}{3}\beta, d)$. Let

$$\mu(d\xi, d\eta) = |\xi|^{-\gamma} d\xi d\eta,$$

where $d\xi$ is the Lebesgue measure on \mathbb{R}^d and $\delta_0(d\eta)$ is the Dirac measure on \mathbb{R}^d concentrated at 0. By (7.9), one sees that (\mathbf{A}^β) holds. In this case, it is well known that for some $c_{d,\gamma} > 0$ (see [St70, p117, Lemma 2]),

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

In particular, for any $f, g \in \mathcal{S}(\mathbb{R}^{2d})$,

$$\begin{aligned} \mathbf{E}\left(X(f)X(g)\right) &= \int_{\mathbb{R}^{2d}} \hat{f}(\zeta)\hat{g}(-\zeta)\mu(d\zeta) = \int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} f(z)g(z')\hat{\mu}(z-z')dz'dz \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(x,v)dv\right) \left(\int_{\mathbb{R}^d} g(x',v)dv\right) \frac{c_{d,\gamma}dx dx'}{|x-x'|^{d-\gamma}}. \end{aligned}$$

Fix $\varphi \in \mathcal{S}(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} \varphi = 1$. For any $f \in \mathcal{S}(\mathbb{R}^d)$, if we define

$$X_1(f) := X(\tilde{f}), \quad \tilde{f}(x,v) := f(x)\varphi(v),$$

then for any $f, g \in \mathcal{S}(\mathbb{R}^d)$,

$$\mathbf{E}\left(X_1(f)X_1(g)\right) = c_{d,\gamma} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x)g(x') \frac{dx dx'}{|x-x'|^{d-\gamma}} = \int_{\mathbb{R}^d} \hat{f}(\xi)\hat{g}(-\xi) \frac{d\xi}{|\xi|^\gamma},$$

where the right hand side is just the inner product of homogenous Bessel potential space $\dot{\mathbb{H}}^{-\gamma}$ in \mathbb{R}^d (see [BCD11]). In particular, $X_1(f)$ can be extended to all $f \in \dot{\mathbb{H}}^{-\gamma}$. This corresponds to the noise independent of v variable. Let $d = 1$ and define

$$B_\gamma(y) := \left(X_1(\mathbf{1}_{[0,y]})\mathbf{1}_{y \geq 0} - X_1(\mathbf{1}_{[y,0]})\mathbf{1}_{y < 0}\right) \gamma^{1/2}(1+\gamma)^{1/2}(2c_{d,\gamma})^{-1/2}.$$

By the elementary calculation, we have

$$\mathbf{E}\left(B_\gamma(y)B_\gamma(y')\right) = \frac{1}{2}(|y|^{1+\gamma} + |y'|^{1+\gamma} - |y-y'|^{1+\gamma}).$$

Hence, $B_\gamma(y)$ is a fractional Brownian motion with Hurst parameter $H = \frac{1+\gamma}{2} \in (1 - \frac{\beta}{3}, 1)$, and for any $g \in \mathcal{S}(\mathbb{R})$,

$$X_1(g) = -\bar{c}_{d,\gamma} \int_{\mathbb{R}} g'(y)B_\gamma(y)dy.$$

In other words, $X_1 = \bar{c}_{d,\gamma}B'_\gamma$ in the distributional sense.

EXAMPLE 7.8. For $\beta \in (\frac{1}{2}, \frac{2}{3})$ and $0 \leq \gamma \in (d - 2\beta, d)$, let

$$\mu(d\xi, d\eta) = |\eta|^{-\gamma} \delta_0(d\xi) d\eta.$$

By (7.9), one sees that (\mathbf{A}^β) holds. When $d = 1$ and $\gamma = 0$, we have

$$\hat{\mu}(x,v) = \delta_0(dv)$$

and

$$\mathbf{E}\left(X(f)X(g)\right) = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(x,v)dx\right) \left(\int_{\mathbb{R}^d} g(y,v)dy\right) dv.$$

In particular, for $\varphi \in \mathcal{S}(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} \varphi = 1$, if we define

$$X_2(f) := X(\tilde{f}), \quad \tilde{f}(x,v) := \varphi(x)f(v)$$

then X_2 is independent of x and is a space white noise on \mathbb{R} . As Example 7.7, for general $\gamma \in [0, 1)$, X_2 corresponds to the derivative of a fractional Brownian motion with Hurst parameter $H = \frac{1+\gamma}{2} \in [\frac{1}{2}, 1)$.

EXAMPLE 7.9. For $\beta \in (\frac{1}{2}, \frac{2}{3})$ and $\gamma_1, \gamma_2 \in [0, d]$ with $3\gamma_1 + \gamma_2 > 4d - 2\beta$, let

$$\mu(d\xi, d\eta) = |\xi|^{-\gamma_1} |\eta|^{-\gamma_2} d\xi d\eta,$$

By (7.10), one sees that (\mathbf{A}^β) holds. When $\gamma_1 \gamma_2 \neq 0$, we have

$$\hat{\mu}(x, v) = c_{d,\gamma} |x|^{\gamma_1-d} |v|^{\gamma_2-d}.$$

When $\gamma_2 = 0$, since $\beta \in (\frac{1}{2}, \frac{2}{3})$, we have

$$(4d - 2\beta)/3 < \gamma_1 < d \Rightarrow d < 2\beta \Rightarrow d = 1,$$

and

$$\mathbf{E}\left(X(f)X(g)\right) = c_{1,\gamma} \int_{\mathbb{R}^3} f(x, v) g(y, v) \frac{dv dx dy}{|x - y|^{\gamma_1-1}}.$$

In particular, one can regard W being white in v -direction and colored in x -direction. In general W is the generalized derivative of a fractional Brownian sheet with $H_i = \frac{\gamma_i+1}{2}$ satisfying $3H_1 + H_2 > 4 - \beta$.

7.3. *Proof of Theorem 7.2.* Recall the definition of $(\phi_j^a)_{j \geq -1}$ in (2.3). We could assume that ϕ_{-1}^a has the following symmetry

$$\phi_{-1}^a(\xi, -\eta) = \phi_{-1}^a(\xi, \eta), \quad (\xi, \eta) \in \mathbb{R}^{2d}.$$

We introduce the following function for later use:

$$(7.12) \quad \psi(\zeta, \zeta') := \sum_{|i-j| \leq 1} \phi_i^a(\zeta) \phi_j^a(\zeta'), \quad \zeta, \zeta' \in \mathbb{R}^{2d}.$$

In particular, by the symmetry of ϕ_{-1}^a in η , for any $\zeta = (\xi, \eta), \zeta' = (\xi', \eta') \in \mathbb{R}^{2d}$,

$$(7.13) \quad \psi((\xi, \eta), (\xi', \eta')) = \psi((\xi, -\eta), (\xi', \eta')) = \psi((\xi, \eta), (\xi', -\eta')).$$

Now we recall some notations used before. Let $z = (x, v) \in \mathbb{R}^{2d}$ and $\zeta = (\xi, \eta)$. For $t \in \mathbb{R}$, we define

$$\Gamma_t z := (x + tv, v), \quad \hat{\Gamma}_t \zeta := (\xi, \eta + t\xi),$$

and for a function f on \mathbb{R}^{2d} and $y, z \in \mathbb{R}^{2d}$,

$$(\Gamma_t f)(z) := f(\Gamma_t z), \quad (\tau_y f)(z) := f(z - y).$$

Clearly,

$$\Gamma_t \Gamma_{-t} z = z, \quad \langle \Gamma_t z, \zeta \rangle = \langle z, \hat{\Gamma}_t \zeta \rangle,$$

and

$$(7.14) \quad (f * g)(z) = \int_{\mathbb{R}^{2d}} \tau_y f(z) g(y) dy$$

and

$$\widehat{\Gamma_t f}(\zeta) = \hat{\Gamma}_{-t} \hat{f}(\zeta), \quad \Gamma_t(f * g) = (\Gamma_t f) * (\Gamma_t g).$$

Recalling (3.2), we have for some $c_0 > 0$,

$$(7.15) \quad \hat{p}_s(\xi, \eta) = e^{-s|\eta|^2 - s^3|\xi|^2/3 - s^2\langle \xi, \eta \rangle} \leq e^{-c_0(s^3|\xi|^2 + s|\eta|^2)}.$$

Now let φ be a smooth probability density function with compact support and symmetric in the variable v . For $\varepsilon \in (0, 1)$, let

$$\varphi_\varepsilon(z) := \varepsilon^{-2d} \varphi(z/\varepsilon), \quad X_\varepsilon(z) := X_{\varphi_\varepsilon}(z) = X(\varphi_\varepsilon(z - \cdot)).$$

To prove Theorem 7.2 it suffices to prove

$$X \in \mathbf{C}_a^{-\alpha}(\rho^\kappa), \quad \mathbf{P} - a.s.$$

and

$$X \circ \nabla_v \mathcal{J} X \in C([0, T], \mathbf{C}_a^{1-2\alpha}(\rho^\kappa)), \quad \mathbf{P} - a.s.$$

Now we consider them separately.

(i) **Regularity of X .** As in (B.1), by the hypercontractivity of Gaussian random variables, for any $\bar{\alpha} \in (\beta, \beta + 1)$, we have

$$\begin{aligned} \mathbf{E} |\mathcal{R}_j^a X_\varepsilon(z) - \mathcal{R}_j^a X(z)|^p &\lesssim \left(\mathbf{E} |\mathcal{R}_j^a X_\varepsilon(z) - \mathcal{R}_j^a X(z)|^2 \right)^{p/2} \\ &= \left(\int_{\mathbb{R}^{2d}} |\phi_j^a(\zeta)|^2 |\hat{\varphi}_\varepsilon(\zeta) - 1|^2 \mu(d\zeta) \right)^{p/2} \\ &\stackrel{(B.4)}{\lesssim} 2^{\bar{\alpha} p j} \left(\int_{\mathbb{R}^{2d}} \frac{|\hat{\varphi}_\varepsilon(\zeta) - 1|^2}{(1 + |\zeta|_a)^{2\bar{\alpha}}} \mu(d\zeta) \right)^{p/2}, \end{aligned}$$

where the implicit constant does not depend on z . Noting that

$$|\hat{\varphi}_\varepsilon(\zeta) - 1| = |\hat{\varphi}(\varepsilon\zeta) - 1| \lesssim \varepsilon^{(\bar{\alpha}-\beta)/3} |\zeta|_a^{\bar{\alpha}-\beta},$$

by definition, we have for any $\alpha' > \bar{\alpha}$ and $p > 4d/\kappa$,

$$\begin{aligned} \mathbf{E} \|X_\varepsilon - X\|_{\mathbf{B}_{p,p}^{-\alpha',a}(\varrho^\kappa)}^p &= \sum_j 2^{-\alpha' p j} \int_{\mathbb{R}^{2d}} \mathbf{E} |\mathcal{R}_j^a X_\varepsilon(z) - \mathcal{R}_j^a X(z)|^p |\varrho^\kappa(z)|^p dz \\ (7.16) \quad &\lesssim \left(\int_{\mathbb{R}^{2d}} \frac{\varepsilon^{2(\bar{\alpha}-\beta)/3}}{(1 + |\zeta|_a)^{2\bar{\alpha}}} \mu(d\zeta) \right)^{p/2} \int_{\mathbb{R}^{2d}} |\varrho^\kappa(z)|^p dz, \end{aligned}$$

which, by (A ^{β}), converges to zero as $\varepsilon \rightarrow 0$. Furthermore, for $\alpha > \bar{\alpha}$, by Besov's embedding Theorem 2.6, for p large enough, we have

$$\lim_{\varepsilon \rightarrow 0} \mathbf{E} \|X_\varepsilon - X\|_{\mathbf{C}_a^{-\alpha}(\varrho^\kappa)}^p = 0.$$

(ii) **Regularity of $X \circ \nabla_v \mathcal{J} X$.** Since X is independent of t , by Lemma 3.17 we only need to show

$$\mathbf{E} \sup_{0 \leq s < t \leq T} \|X \circ \nabla_v \mathcal{J} X(t) - X \circ \nabla_v \mathcal{J} X(s)\|_{\mathbf{C}_a^{1-2\alpha}(\varrho^\kappa)} < \infty.$$

We represent $\mathcal{R}_\ell^a(X_\varepsilon \circ \nabla_v \mathcal{J} X_\varepsilon(t))$ in terms of $(X_\varepsilon \otimes X_\varepsilon)(H_t^\ell)$ as given in the following lemma.

LEMMA 7.10. *For any $t \geq 0$ and $\ell \geq -1$, we have*

$$(7.17) \quad \mathcal{R}_\ell^a(X_\varepsilon \circ \nabla_v \mathcal{J} X_\varepsilon(t)) = (X_\varepsilon \otimes X_\varepsilon)(H_t^\ell),$$

where

$$H_t^\ell(y, y') := \sum_{|i-j| \leq 1} \int_0^t \mathcal{R}_\ell^a(\tau_{y'} \check{\phi}_i^a \cdot \pi_{\Gamma-s} y (\mathcal{R}_j^a \nabla_v \Gamma_s p_s)) ds.$$

Moreover, for $\zeta = (\xi, \eta) \in \mathbb{R}^{2d}$, we have

$$(7.18) \quad \widehat{H}_t^\ell(\zeta, \zeta') = -i \int_0^t e^{-i\langle \cdot, \hat{\Gamma}_s \zeta + \zeta' \rangle} \phi_\ell^a(\hat{\Gamma}_s \zeta + \zeta') \psi(\hat{\Gamma}_s \zeta, \zeta') (\eta + s\xi) \hat{p}_s(\zeta) ds,$$

where ψ is defined by (7.12).

PROOF. By definition, we have

$$\begin{aligned} \mathcal{R}_i^a X_\varepsilon \cdot (\mathcal{R}_j^a \nabla_v \mathcal{J} X_\varepsilon)(t) &= \int_0^t (\check{\phi}_i^a * X_\varepsilon) \cdot (\mathcal{R}_j^a \nabla_v \Gamma_s p_s) * (\Gamma_s X_\varepsilon) ds \\ &= \int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} \left(\int_0^t \tau_{y'} \check{\phi}_i^a \cdot \tau_{\Gamma_{-s} y} (\mathcal{R}_j^a \nabla_v \Gamma_s p_s) ds \right) X_\varepsilon(y) X_\varepsilon(y') dy dy'. \end{aligned}$$

which implies (7.17). For (7.18), letting $h := \mathcal{R}_j^a (\nabla_v \Gamma_s p_s)$, by easy calculations, we have

$$\begin{aligned} &\int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} e^{-i\langle \zeta, y + \zeta', y' \rangle} \check{\phi}_\ell^a * (\tau_{y'} \check{\phi}_i^a \cdot \tau_{\Gamma_{-s} y} h) dy dy' \\ &= e^{-i\langle \cdot, \hat{\Gamma}_s \zeta + \zeta' \rangle} \phi_\ell^a(\hat{\Gamma}_s \zeta + \zeta') \phi_i^a(\zeta') \hat{h}(-\hat{\Gamma}_s \zeta), \end{aligned}$$

where for $\zeta = (\xi, \eta) \in \mathbb{R}^{2d}$,

$$\hat{h}(\zeta) = \phi_j^a(\zeta)(i\eta)(\hat{\Gamma}_{-s} \hat{p}_s)(\zeta).$$

Thus we obtain (7.18) by $\hat{\Gamma}_{-s} \hat{\Gamma}_s \zeta = \zeta$. \square

REMARK 7.11. Notice that for each $y, y' \in \mathbb{R}^{2d}$, $H_t^\ell(y, y')$ is a \mathbb{R}^d -valued function of z . In expressions (7.17) and (7.18), we have suppressed the variable z for simplicity. Without further declaration, we also use such a convention below.

For simplicity of notations, we write

$$M_t^\varepsilon(z) := (X_\varepsilon \circ \nabla_v \mathcal{J} X_\varepsilon(t))(z)$$

and

$$(7.19) \quad G_{t,s}^{\varepsilon,\varepsilon'}(z) := M_t^\varepsilon(z) - M_t^{\varepsilon'}(z) - M_s^\varepsilon(z) + M_s^{\varepsilon'}(z).$$

Below we drop the variable z . It is easy to see $M_t^\varepsilon = \mathbb{E}M_t^\varepsilon + M_t^\varepsilon - \mathbb{E}M_t^\varepsilon$ as the Wiener chaos decomposition for M_t^ε with $\mathbb{E}M_t^\varepsilon$ in the 0th Wiener chaos and $M_t^\varepsilon - \mathbb{E}M_t^\varepsilon$ in the second Wiener chaos. In the following we consider them separately.

Terms in the 0th Wiener chaos First we have the following estimates for the terms in the 0th Wiener chaos. This terms are not zero as the classical case. After subtracting formally divergent terms (see $\mathcal{J}_{22,\ell}^t$ below) which are zero by symmetry, the terms in the 0th Wiener chaos converge in the corresponding spaces. Note that by (7.17),

$$(7.20) \quad \mathcal{R}_\ell^a M_t^\varepsilon = (X_\varepsilon \otimes X_\varepsilon)(H_t^\ell).$$

and by (7.5),

$$\mathcal{R}_\ell^a \mathbb{E}M_t^\varepsilon = \mathbb{E}\mathcal{R}_\ell^a M_t^\varepsilon = \int_{\mathbb{R}^{2d}} \widehat{H}_t^\ell(\zeta, -\zeta) \hat{\varphi}_\varepsilon^2(\zeta) \mu(d\zeta) =: \Lambda_t^{\ell,\varepsilon}.$$

This corresponds to the zeroth Wiener chaos of random field $(X_\varepsilon \otimes X_\varepsilon)(H_t^\ell)$.

By (7.18) we make the following decomposition:

$$\begin{aligned}
\widehat{H}_t^\ell(\zeta, -\zeta) &= -i \int_0^t e^{-i\langle \cdot, \hat{\Gamma}_s \zeta - \zeta \rangle} \phi_\ell^a(\hat{\Gamma}_s \zeta - \zeta) \left(\psi(\hat{\Gamma}_s \zeta, -\zeta) - \psi(\zeta, -\zeta) \right) \eta \hat{p}_s(\zeta) ds \\
&\quad - i \int_0^t e^{-i\langle \cdot, \hat{\Gamma}_s \zeta - \zeta \rangle} \phi_\ell^a(\hat{\Gamma}_s \zeta - \zeta) \psi(\zeta, -\zeta) \eta \hat{p}_s(\zeta) ds \\
&\quad - i \int_0^t e^{-i\langle \cdot, \hat{\Gamma}_s \zeta - \zeta \rangle} \phi_\ell^a(\hat{\Gamma}_s \zeta - \zeta) \psi(\hat{\Gamma}_s \zeta, -\zeta) s \xi \hat{p}_s(\zeta) ds \\
&=: \mathcal{J}_{1,\ell}^t(\zeta) + \mathcal{J}_{2,\ell}^t(\zeta) + \mathcal{J}_{3,\ell}^t(\zeta).
\end{aligned}$$

For $\mathcal{J}_{1,\ell}^t(\zeta)$, noting that for $\zeta = (\xi, \eta)$,

$$(7.21) \quad \hat{\Gamma}_s \zeta - \zeta = (0, s\xi),$$

by (B.4) and (B.6) with $\gamma = 2\alpha - 1$, we have

$$\begin{aligned}
\|\mathcal{J}_{1,\ell}^t(\zeta)\|_{L^\infty} &\leq \int_0^t |\phi_\ell^a(\hat{\Gamma}_s \zeta - \zeta)| |\psi(\hat{\Gamma}_s \zeta, -\zeta) - \psi(\zeta, -\zeta)| |\eta| \hat{p}_s(\zeta) ds \\
&\lesssim 2^{(2\alpha-1)\ell} \int_0^t (1 + |s\xi|)^{1-2\alpha} \frac{|s\xi|^{2\alpha-1}}{(1 + |\zeta|_a)^{2\alpha-1}} |\eta| \hat{p}_s(\zeta) ds \\
&\lesssim 2^{(2\alpha-1)\ell} (1 + |\zeta|_a)^{2-2\alpha} \int_0^t \hat{p}_s(\zeta) ds \\
&\stackrel{(7.15), (B.7)}{\lesssim} 2^{(2\alpha-1)\ell} (1 + |\zeta|_a)^{-2\alpha}.
\end{aligned}$$

For $\mathcal{J}_{3,\ell}^t(\zeta)$, since $|\psi| \lesssim 1$, by (7.21), we have

$$\begin{aligned}
\|\mathcal{J}_{3,\ell}^t(\zeta)\|_{L^\infty} &\lesssim 2^{(2\alpha-1)\ell} \int_0^t (1 + |s\xi|)^{1-2\alpha} |s\xi| \hat{p}_s(\zeta) ds \\
&\lesssim 2^{(2\alpha-1)\ell} \int_0^t |s\xi|^{2-2\alpha} \hat{p}_s(\zeta) ds \\
&\stackrel{(7.15), (B.7)}{\lesssim} 2^{(2\alpha-1)\ell} (1 + |\zeta|_a)^{-2\alpha}.
\end{aligned}$$

For $\mathcal{J}_{2,\ell}^t(\zeta)$, by we can write

$$\begin{aligned}
\mathcal{J}_{2,\ell}^t(\zeta) &= -i \int_0^t e^{-i\langle \cdot, \hat{\Gamma}_s \zeta - \zeta \rangle} \phi_\ell^a(\hat{\Gamma}_s \zeta - \zeta) \psi(\zeta, -\zeta) \eta e^{-s|\eta|^2 - s^3|\xi|^2/3} (e^{-s^2\langle \xi, \eta \rangle} - 1) ds \\
&\quad - i \int_0^t e^{-i\langle \cdot, \hat{\Gamma}_s \zeta - \zeta \rangle} \phi_\ell^a(\hat{\Gamma}_s \zeta - \zeta) \psi(\zeta, -\zeta) \eta e^{-s|\eta|^2 - s^3|\xi|^2/3} ds \\
&=: \mathcal{J}_{21,\ell}^t(\zeta) + \mathcal{J}_{22,\ell}^t(\zeta).
\end{aligned}$$

For $\mathcal{J}_{21,\ell}^t(\zeta)$, noting that

$$|e^{-s^2\langle \xi, \eta \rangle} - 1| \leq |s\xi| |s\eta| e^{s^2|\xi||\eta|},$$

by (7.15) and (B.7), we have

$$\|\mathcal{J}_{21,\ell}^t(\zeta)\|_{L^\infty} \lesssim 2^{(2\alpha-1)\ell} \int_0^t (1 + |s\xi|)^{1-2\alpha} |s\xi| |s\eta| |\eta| e^{-c_0(s^3|\xi|^2 + s|\eta|^2)} ds$$

$$\begin{aligned}
&\lesssim 2^{(2\alpha-1)\ell} \int_0^t |s\xi|^{2-2\alpha} |s\eta| |\eta| e^{-c_0(s^3|\xi|^2+s|\eta|^2)} ds \\
&\lesssim 2^{(2\alpha-1)\ell} |\xi|^{2-2\alpha} |\eta|^2 \int_0^t s^{3-2\alpha} e^{-c_0(s^3|\xi|^2+s|\eta|^2)} ds \\
&\stackrel{(B.7)}{\lesssim} 2^{(2\alpha-1)\ell} (1 + |\zeta|_a)^{-2\alpha}.
\end{aligned}$$

On the other hand, by (7.21) and (7.13), one sees that

$$\mathcal{J}_{22,\ell}^t(\xi, -\eta) = -\mathcal{J}_{22,\ell}^t(\xi, \eta).$$

Since $\mu(d\xi, -d\eta) = \mu(d\xi, d\eta)$ and φ_ε is symmetric w.r.t. v variable, we have

$$(7.22) \quad \int_{\mathbb{R}^{2d}} \mathcal{J}_{22,\ell}^t(\zeta) \hat{\varphi}_\varepsilon^2(\zeta) \mu(d\zeta) \equiv 0.$$

Thus, we get

$$\Lambda_t^{\ell,\varepsilon} = \int_{\mathbb{R}^{2d}} \left(\mathcal{J}_{1,\ell}^t(\zeta) + \mathcal{J}_{21,\ell}^t(\zeta) + \mathcal{J}_{3,\ell}^t(\zeta) \right) \hat{\varphi}_\varepsilon^2(\zeta) \mu(d\zeta),$$

and

$$\begin{aligned}
\|\Lambda_t^{\ell,\varepsilon} - \Lambda_t^{\ell,\varepsilon'}\|_{L^\infty} &\lesssim \int_{\mathbb{R}^{2d}} \left(\|\mathcal{J}_{1,\ell}^t(\zeta)\|_{L^\infty} + \|\mathcal{J}_{21,\ell}^t(\zeta)\|_{L^\infty} + \|\mathcal{J}_{3,\ell}^t(\zeta)\|_{L^\infty} \right) \\
&\quad \times |\hat{\varphi}_\varepsilon^2(\zeta) - \hat{\varphi}_{\varepsilon'}^2(\zeta)| \mu(d\zeta) \\
&\lesssim 2^{(2\alpha-1)\ell} \int_{\mathbb{R}^{2d}} (1 + |\zeta|_a)^{-2\alpha} |\hat{\varphi}_\varepsilon^2(\zeta) - \hat{\varphi}_{\varepsilon'}^2(\zeta)| \mu(d\zeta).
\end{aligned}$$

By the dominated convergence theorem, we obtain

$$(7.23) \quad \lim_{\varepsilon, \varepsilon' \rightarrow 0} \sup_{\ell \geq -1} \sup_{t \in [0, T]} 2^{(1-2\alpha)\ell} \|\Lambda_t^{\ell,\varepsilon} - \Lambda_t^{\ell,\varepsilon'}\|_{L^\infty} = 0,$$

where the norm $\|\cdot\|_{L^\infty}$ is with respect to variable z . Thus, we have

$$\lim_{\varepsilon, \varepsilon' \rightarrow 0} \sup_{t \in [0, T]} \|\mathbf{E}M_t^\varepsilon - \mathbf{E}M_t^{\varepsilon'}\|_{\mathbf{C}_a^{1-2\alpha}} = 0.$$

Terms in the second Wiener chaos By Kolmogorov's continuity criterion and Besov's embedding Theorem 2.6, it suffices to show that for some $\delta > 0$, and any $\alpha > \beta$ and $p \geq 2$,

$$\lim_{\varepsilon, \varepsilon' \rightarrow 0} \sup_{0 \leq s < t \leq T} (t-s)^{-\delta p} \mathbf{E} \left(\|G_{t,s}^{\varepsilon, \varepsilon'} - \mathbf{E}G_{t,s}^{\varepsilon, \varepsilon'}\|_{\mathbf{B}_{p,p}^{1-2\alpha,a}(\varrho^\kappa)}^p \right) = 0.$$

Since $G_{t,s}^{\varepsilon, \varepsilon'} - \mathbf{E}G_{t,s}^{\varepsilon, \varepsilon'}$ belongs to the second Wiener chaos space, as in (7.16), we only need to show that

$$(7.24) \quad \lim_{\varepsilon, \varepsilon' \rightarrow 0} \sup_{0 \leq s < t \leq T} \sup_{\ell \geq -1} (t-s)^{-\delta} 2^{(2-4\alpha)\ell} \|\text{Var}(\mathcal{R}_\ell^a G_{t,s}^{\varepsilon, \varepsilon'})\|_{L^\infty} = 0.$$

Noting that by (7.19) and (7.20),

$$\begin{aligned}
\mathcal{R}_\ell^a G_{t,s}^{\varepsilon, \varepsilon'} &= (X_\varepsilon \otimes X_\varepsilon)(H_t^\ell - H_s^\ell) - (X_{\varepsilon'} \otimes X_{\varepsilon'})(H_t^\ell - H_s^\ell) \\
&= (X_{\varphi_\varepsilon - \varphi_{\varepsilon'}} \otimes X_{\varphi_\varepsilon})(H_t^\ell - H_s^\ell) + (X_{\varphi_{\varepsilon'}} \otimes X_{\varphi_\varepsilon - \varphi_{\varepsilon'}})(H_t^\ell - H_s^\ell),
\end{aligned}$$

by (7.6), we have

$$\begin{aligned}
\text{Var}(\mathcal{R}_\ell^a G_{t,s}^{\varepsilon,\varepsilon'}) &\leq 2\text{Var}((X_{\varphi_\varepsilon - \varphi_{\varepsilon'}} \otimes X_{\varphi_\varepsilon})(H_t^\ell - H_s^\ell)) \\
&\quad + 2\text{Var}((X_{\varphi_{\varepsilon'}} \otimes X_{\varphi_\varepsilon - \varphi_{\varepsilon'}})(H_t^\ell - H_s^\ell)) \\
&= 4 \int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} |\text{Sym}((\widehat{H}_t^\ell - \widehat{H}_s^\ell)K_{\varepsilon,\varepsilon'}^{(1)})(\zeta, \zeta')|^2 \mu(d\zeta) \mu(d\zeta') \\
&\quad + 4 \int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} |\text{Sym}((\widehat{H}_t^\ell - \widehat{H}_s^\ell)K_{\varepsilon,\varepsilon'}^{(2)})(\zeta, \zeta')|^2 \\
&\leq 4 \int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} |((\widehat{H}_t^\ell - \widehat{H}_s^\ell)K_{\varepsilon,\varepsilon'}^{(1)})(\zeta, \zeta')|^2 \mu(d\zeta) \mu(d\zeta') \\
&\quad + 4 \int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} |((\widehat{H}_t^\ell - \widehat{H}_s^\ell)K_{\varepsilon,\varepsilon'}^{(2)})(\zeta, \zeta')|^2 \mu(d\zeta) \mu(d\zeta'),
\end{aligned} \tag{7.25}$$

where

$$K_{\varepsilon,\varepsilon'}^{(1)}(\zeta, \zeta') := (\hat{\varphi}_\varepsilon(\zeta) - \hat{\varphi}_{\varepsilon'}(\zeta))\hat{\varphi}_\varepsilon(\zeta')$$

and

$$K_{\varepsilon,\varepsilon'}^{(2)}(\zeta, \zeta') := \hat{\varphi}_{\varepsilon'}(\zeta)(\hat{\varphi}_\varepsilon(\zeta') - \hat{\varphi}_{\varepsilon'}(\zeta')).$$

For any $\theta \in (0, 1)$, we have

$$|K_{\varepsilon,\varepsilon'}^{(1)}(\zeta, \zeta')| \leq |(\hat{\varphi}_\varepsilon - \hat{\varphi}_{\varepsilon'})(\zeta)| \lesssim |\varepsilon - \varepsilon'|^{\theta/3} |\zeta|_a^\theta$$

and

$$|K_{\varepsilon,\varepsilon'}^{(2)}(\zeta, \zeta')| \leq |(\hat{\varphi}_\varepsilon - \hat{\varphi}_{\varepsilon'})(\zeta')| \lesssim |\varepsilon - \varepsilon'|^{\theta/3} |\zeta'|_a^\theta.$$

Moreover, by (7.18) we clearly have

$$\|(\widehat{H}_t^\ell - \widehat{H}_s^\ell)K_{\varepsilon,\varepsilon'}^{(2)}(\zeta, \zeta')\|_{L^\infty} \lesssim |\varepsilon - \varepsilon'|^{\theta/3} \int_s^t \Phi_r^\ell(\zeta, \zeta') |\zeta'|_a^\theta |\eta + r\xi| \hat{p}_r(\zeta) dr,$$

and

$$\Phi_r^\ell(\zeta, \zeta') := |\phi_\ell^a(\hat{\Gamma}_r \zeta + \zeta')| |\psi(\hat{\Gamma}_r \zeta, \zeta')|.$$

Let $\sigma, \gamma \geq 0$ with $\sigma + \gamma = 2\beta$. Noting that by (B.4), (B.5) and (7.3),

$$\|\Phi_r^\ell(\zeta, \cdot)\| \cdot |a|_{L^2(\mu)}^\theta \lesssim \int_{\mathbb{R}^{2d}} \frac{2^{\sigma\ell} (1 + |\hat{\Gamma}_s \zeta|_a)^{\gamma+2\theta}}{(1 + |\hat{\Gamma}_s \zeta + \zeta'|_a)^\sigma (1 + |\zeta'|_a)^\gamma} \mu(d\zeta') \lesssim 2^{\sigma\ell} (1 + |\hat{\Gamma}_s \zeta|_a)^{\gamma+2\theta},$$

we have by Minkowski's inequality,

$$\begin{aligned}
\left\| \|(\widehat{H}_t^\ell - \widehat{H}_s^\ell)K_{\varepsilon,\varepsilon'}^{(2)}(\zeta, \cdot)\|_{L^\infty} \right\|_{L^2(\mu)} &\lesssim \int_s^t |\varepsilon - \varepsilon'|^{\theta/3} \|\Phi_r^\ell(\zeta, \cdot)\|_{L^2(\mu)} |\eta + r\xi| \hat{p}_r(\zeta) dr \\
&\lesssim |\varepsilon - \varepsilon'|^{\theta/3} 2^{\frac{\sigma\ell}{2}} \int_s^t (1 + |\hat{\Gamma}_r \zeta|_a)^{\frac{\gamma+2\theta}{2}} |\eta + r\xi| \hat{p}_r(\zeta) dr.
\end{aligned} \tag{7.26}$$

Since $|\eta| \vee |\xi|^{1/3} \leq |\zeta|_a$, by $\hat{\Gamma}_r \zeta = (\xi, \eta + r\xi)$, we have

$$\begin{aligned}
&(1 + |\hat{\Gamma}_r \zeta|_a)^{\frac{\gamma+2\theta}{2}} |\eta + r\xi| \\
&\lesssim (1 + |\zeta|_a)^{\frac{\gamma+2\theta}{2}} |\eta| + (r|\xi|)^{\frac{\gamma+2\theta}{2}} |\eta| + (1 + |\zeta|_a)^{\frac{\gamma+2\theta}{2}} |r\xi| + (r|\xi|)^{\frac{\gamma+2\theta}{2}+1} \\
&\lesssim (1 + |\zeta|_a)^{\frac{\gamma+2\theta}{2}+1} + r^{\frac{\gamma+2\theta}{2}} |\zeta|_a^{\frac{3(\gamma+2\theta)}{2}+1} + r(1 + |\zeta|_a)^{\frac{\gamma+2\theta}{2}+3} + r^{\frac{\gamma+2\theta}{2}+1} |\zeta|_a^{\frac{3(\gamma+2\theta)}{2}+3}.
\end{aligned}$$

If we choose $\sigma = 4\alpha - 2$ for some $\alpha > \beta$, then

$$\frac{\gamma}{2} - 1 = \beta - 2\alpha < -\beta.$$

Thus, by (7.15) and (B.7), for θ small enough there is a $\delta > 0$ such that for all $0 \leq s < t \leq T$ and $\zeta = (\xi, \eta) \in \mathbb{R}^{2d}$,

$$\int_s^t (1 + |\hat{\Gamma}_r \zeta|_a)^{\frac{\gamma+\theta}{2}} |\eta + r\xi| \hat{p}_r(\zeta) dr \lesssim (t-s)^{\delta/2} (1 + |\zeta|_a)^{-\beta}.$$

Substituting this into (7.26), we get

$$\left\| \left(\widehat{H}_t^\ell - \widehat{H}_s^\ell \right) K_{\varepsilon, \varepsilon'}^{(2)}(\zeta, \cdot) \right\|_{L^\infty}^2 \lesssim |\varepsilon - \varepsilon'|^{2\theta/3} 2^{(4\alpha-2)\ell} (t-s)^\delta (1 + |\zeta|_a)^{-2\beta}.$$

For the term containing $K^{(1)}$ we have the similar bounds. Substituting these into the left hand side of (7.25), we obtain the regularity of the term in the second Wiener chaos. Thus we complete the proof.

APPENDIX A: CHARACTERIZATIONS FOR $\mathbf{B}_{p,q}^{s,a}(\rho)$

In this appendix, we provide a detailed proof for Theorem 2.7. First of all, we prepare two useful lemmas for later use.

LEMMA A.1. *For any $\alpha > 0$, there is a constant $C = C(d, a, m, \alpha) > 0$ such that for all $\lambda > 0$,*

$$(A.1) \quad \int_{|h|_a \leq \lambda} |h|_a^{\alpha-a \cdot m} dh \lesssim_C \lambda^\alpha, \quad \int_{|h|_a > \lambda} |h|_a^{-\alpha-a \cdot m} dh \lesssim_C \lambda^{-\alpha}.$$

PROOF. Let $h = (h_1, \dots, h_n) \in \mathbb{R}^N$ with $h_i \in \mathbb{R}^{m_i}$. Define a transform $h \rightarrow \tilde{h}$ by

$$\tilde{h} := (\tilde{h}_1, \dots, \tilde{h}_n), \quad \tilde{h}_i := |h_i|^{\frac{1}{a_i}-1} h_i.$$

Clearly, for each $i = 1, \dots, n$,

$$|h_i| = |\tilde{h}_i|^{a_i}, \quad h_i = |\tilde{h}_i|^{a_i-1} \tilde{h}_i$$

and

$$|\det(\partial h_i / \partial \tilde{h}_i)| \leq a_i^{m_i} |\tilde{h}_i|^{a_i m_i - m_i} \leq a_i^{m_i} |\tilde{h}_1|^{a_i m_i - m_i},$$

where $\partial h_i / \partial \tilde{h}_i$ stands for the Jacobian matrix of the inverse transform $\tilde{h}_i \rightarrow h_i$, and $|\tilde{h}|_1 := \sum_{i=1}^n |\tilde{h}_i|$. Thus by the change of variable,

$$\begin{aligned} \int_{|h|_a \leq \lambda} |h|_a^{\alpha-a \cdot m} dh &= \int_{|\tilde{h}|_1 \leq \lambda} |\tilde{h}_1|^{\alpha-a \cdot m} \prod_{i=1}^n |\det(\partial h_i / \partial \tilde{h}_i)| d\tilde{h} \\ &\leq \prod_{i=1}^n a_i^{m_i} \int_{|\tilde{h}|_1 \leq \lambda} |\tilde{h}_1|^{\alpha-N} d\tilde{h} \lesssim \lambda^\alpha, \end{aligned}$$

where $N = m_1 + \dots + m_n$, and

$$\int_{|h|_a > \lambda} |h|_a^{-\alpha-a \cdot m} dh \leq \prod_{i=1}^n a_i^{m_i} \int_{|\tilde{h}|_1 > \lambda} |\tilde{h}_1|^{-\alpha-N} d\tilde{h} \lesssim \lambda^{-\alpha}.$$

The proof is complete. \square

By the following lemma we can estimate the norm in $\mathbf{B}_{p,q}^{s,a}(\rho)$ by duality.

LEMMA A.2. *Let $\rho \in \mathcal{W}$ with $\rho^{-1} \in \mathcal{W}$, $s \in \mathbb{R}$ and $p, q, p', q' \in [1, \infty]$ with $1/p + 1/p' = 1/q + 1/q' = 1$.*

(i) *For any $\varphi \in \mathcal{S}$ and $f \in \mathbf{B}_{p,q}^{s,a}(\rho)$, it holds that*

$$|\langle f, \varphi \rangle| \leq \|f\|_{\mathbf{B}_{p,q}^{s,a}(\rho)} \|\varphi\|_{\mathbf{B}_{p',q'}^{-s,a}(\rho^{-1})}.$$

(ii) *There is a constant $C = C(\rho, d, s, p, q) > 0$ such that for any $f \in \mathbf{B}_{p,q}^{s,a}(\rho)$,*

$$\|f\|_{\mathbf{B}_{p,q}^{s,a}(\rho)} \leq C \sup_{\varphi \in \mathcal{S}} \langle f, \varphi \rangle / \|\varphi\|_{\mathbf{B}_{p',q'}^{-s,a}(\rho^{-1})}.$$

PROOF. (i) By (2.5) and Hölder's inequality, we have

$$\begin{aligned} \langle f, \varphi \rangle &= \sum_{j \geq -1} \langle \mathcal{R}_j^a f, \tilde{\mathcal{R}}_j^a \varphi \rangle \leq \sum_{j \geq -1} \|\mathcal{R}_j^a f\|_{L^p(\rho)} \|\tilde{\mathcal{R}}_j^a \varphi\|_{L^{p'}(\rho^{-1})} \\ &\leq \|f\|_{\mathbf{B}_{p,q}^{s,a}(\rho)} \|\varphi\|_{\mathbf{B}_{p',q'}^{-s,a}(\rho^{-1})}. \end{aligned}$$

(ii) We follow the proof in [BCD11]. For $M \in \mathbb{N}$, let

$$U_M^{q'} := \left\{ (c_j)_{j \in \mathbb{N}} : \sum_{j \leq M} |c_j|^{q'} \leq 1, c_j = 0, j > M \right\}.$$

By the definition of $\mathbf{B}_{p,q}^{s,a}(\rho)$, we have

$$\begin{aligned} \|f\|_{\mathbf{B}_{p,q}^{s,a}(\rho)} &= \lim_{M \rightarrow \infty} \left(\sum_{j \leq M} 2^{jsq} \|\mathcal{R}_j^a f\|_{L^p(\rho)}^q \right)^{1/q} \\ &= \lim_{M \rightarrow \infty} \sup_{(c_j) \in U_M^{q'}} \sum_{j \leq M} c_j 2^{js} \|\mathcal{R}_j^a f\|_{L^p(\rho)}. \end{aligned}$$

Fix $\varepsilon > 0$ and $(c_j) \in U_M^{q'}$. Since

$$\|g\|_{L^p} = \sup_{h \in \mathcal{S}} \langle g, h \rangle / \|h\|_{L^{p'}},$$

for any $j \leq M$, there is a $\psi_j \in \mathcal{S}$ with $\|\psi_j\|_{L^{p'}} = 1$ such that

$$\begin{aligned} \|\mathcal{R}_j^a f\|_{L^p(\rho)} &\leq \int_{\mathbb{R}^N} \rho(x) \mathcal{R}_j^a f(x) \psi_j(x) dx + \frac{\varepsilon 2^{-js}}{(|c_j| + 1)(j^2 + 1)} \\ &= \int_{\mathbb{R}^N} f(x) \mathcal{R}_j^a(\rho \psi_j)(x) dx + \frac{\varepsilon 2^{-js}}{(|c_j| + 1)(j^2 + 1)}. \end{aligned}$$

Now, if we define $\varphi_M^{(c_j)} \in \mathcal{S}$ by

$$\varphi_M^{(c_j)}(x) := \sum_{j \leq M} c_j 2^{js} \mathcal{R}_j^a(\rho \psi_j)(x),$$

then

$$(A.2) \quad \|f\|_{\mathbf{B}_{p,q}^{s,a}(\rho)} \leq \lim_{M \rightarrow \infty} \sup_{(c_j) \in U_M^{q'}} \langle f, \varphi_M^{(c_j)} \rangle + \sum_{j \geq -1} \frac{\varepsilon}{j^2 + 1}.$$

Note that for by (2.14),

$$\begin{aligned}
 \|\varphi_M^{(c_j)}\|_{\mathbf{B}_{p',q'}^{-s,a}(\rho^{-1})}^{q'} &= \sum_{k \geq -1} 2^{-kq's} \left\| \sum_{j \leq M, |j-k| \leq 3} c_j 2^{js} \mathcal{R}_k^a \mathcal{R}_j^a(\rho \psi_j) \right\|_{L^{p'}(\rho^{-1})}^{q'} \\
 (A.3) \quad &\lesssim \sum_{j \leq M} c_j^{q'} \|\rho \psi_j\|_{L^{p'}(\rho^{-1})}^{q'} = \sum_{j \leq M} c_j^{q'} \|\psi_j\|_{L^{p'}}^{q'} = \sum_{j \leq M} c_j^{q'} \leq 1.
 \end{aligned}$$

Hence, by (A.2) and (A.3),

$$\|f\|_{\mathbf{B}_{p,q}^{s,a}(\rho)} \leq C \sup_{\varphi \in \mathcal{S}} \langle f, \varphi \rangle / \|\varphi\|_{\mathbf{B}_{p',q'}^{-s,a}(\rho^{-1})} + \sum_{j \geq -1} \frac{\varepsilon}{j^2 + 1}.$$

The proof is complete by letting $\varepsilon \rightarrow 0$. □

Now we can give the proof of Theorem 2.7.

PROOF OF THEOREM 2.7. (i) In this step we prove

$$(A.4) \quad \|f\|_{\tilde{\mathbf{B}}_{p,q}^{s,a}(\rho)} \lesssim \|f\|_{\mathbf{B}_{p,q}^{s,a}(\rho)}.$$

For simplicity, we set $M := [s] + 1$. Note that by (2.11),

$$\|\delta_h \mathcal{R}_j^a f\|_{L^p(\rho)} \lesssim (1 + |h|_a^\kappa) \sum_{i=1}^n \|\delta_{h_i} \mathcal{R}_j^a f\|_{L^p(\rho)},$$

where for $h = (h_1, \dots, h_n)$ and $x = (x_1, \dots, x_n)$,

$$\delta_{h_i} f(x) := f(\dots, x_{i-1}, x_i + h_i, x_{i+1}, \dots) - f(\dots, x_{i-1}, x_i, x_{i+1}, \dots).$$

By induction, one sees that

$$\|\delta_h^{(M)} \mathcal{R}_j^a f\|_{L^p(\rho)} \lesssim (1 + |h|_a^{M\kappa}) \sum_{i_1=1}^n \dots \sum_{i_M=1}^n \|\delta_{h_{i_1}} \dots \delta_{h_{i_M}} \mathcal{R}_j^a f\|_{L^p(\rho)}.$$

Let $|h|_a \leq 1$. By (2.11) and Bernstein's inequality (2.13), we have

$$\begin{aligned}
 \|\delta_{h_{i_1}} \dots \delta_{h_{i_M}} \mathcal{R}_j^a f\|_{L^p(\rho)} &\lesssim |h_{i_1}| \|\nabla_{x_{i_1}} \delta_{h_{i_2}} \dots \delta_{h_{i_M}} \mathcal{R}_j^a f\|_{L^p(\rho)} \\
 &\lesssim \dots \dots \dots \\
 &\lesssim |h_{i_1}| \dots |h_{i_M}| \|\nabla_{x_{i_1}} \dots \nabla_{x_{i_M}} \mathcal{R}_j^a f\|_{L^p(\rho)} \\
 &\lesssim |h_{i_1}| 2^{a_{i_1}j} \dots |h_{i_M}| 2^{a_{i_M}j} \|\mathcal{R}_j^a f\|_{L^p(\rho)} \\
 &\lesssim (2^j |h|_a)^{a_{i_1} + \dots + a_{i_M}} \|\mathcal{R}_j^a f\|_{L^p(\rho)}.
 \end{aligned}$$

Moreover, by (2.11), we clearly have

$$\|\delta_{h_{i_1}} \dots \delta_{h_{i_M}} \mathcal{R}_j^a f\|_{L^p(\rho)} \lesssim \|\mathcal{R}_j^a f\|_{L^p(\rho)}.$$

Hence,

$$\begin{aligned}
 \|\delta_{h_{i_1}} \dots \delta_{h_{i_M}} \mathcal{R}_j^a f\|_{L^p(\rho)} &\lesssim ((2^j |h|_a)^{a_{i_1} + \dots + a_{i_M}} \wedge 1) \|\mathcal{R}_j^a f\|_{L^p(\rho)} \\
 &\lesssim ((2^j |h|_a)^M \wedge 1) \|\mathcal{R}_j^a f\|_{L^p(\rho)},
 \end{aligned}$$

where the second inequality is due to $a_{i_1} + \dots + a_{i_M} \geq M$. Thus we obtain

$$(A.5) \quad \|\delta_h^{(M)} \mathcal{R}_j^a f\|_{L^p(\rho)} \lesssim ((2^j |h|_a)^M \wedge 1) \|\mathcal{R}_j^a f\|_{L^p(\rho)} = ((2^j |h|_a)^M \wedge 1) 2^{-sj} c_j,$$

where

$$c_j := 2^{sj} \|\mathcal{R}_j^a f\|_{L^p(\rho)}.$$

For $q = \infty$, we have

$$\|\delta_h^{(M)} f\|_{L^p(\rho)} \leq \sum_j \|\delta_h^{(M)} \mathcal{R}_j^a f\|_{L^p(\rho)} \lesssim \sum_j ((2^j |h|_a)^M \wedge 1) 2^{-sj} c_j \lesssim |h|_a^s \|f\|_{\mathbf{B}_{p,\infty}^{s,a}(\rho)}.$$

Next we assume $q \in [1, \infty)$. For $h \in \mathbb{R}^N$ with $|h|_a \leq 1$, we choose $j_h \in \mathbb{N}$ such that

$$(A.6) \quad |h|_a^{-1} \leq 2^{j_h} \leq 2|h|_a^{-1}.$$

Then by (A.5),

$$\begin{aligned} \|\delta_h^{(M)} f\|_{L^p(\rho)} &\leq \sum_{j \geq -1} \|\delta_h^{(M)} \mathcal{R}_j^a f\|_{L^p(\rho)} \lesssim \sum_{j \geq -1} ((2^j |h|_a)^M \wedge 1) 2^{-sj} c_j \\ &\leq |h|_a^M \sum_{j < j_h} 2^{(M-s)j} c_j + \sum_{j \geq j_h} 2^{-sj} c_j =: I_1(h) + I_2(h). \end{aligned}$$

For $I_1(h)$, by Hölder's inequality, we have

$$\begin{aligned} I_1^q(h) &\leq |h|_a^{qM} \left(\sum_{j < j_h} 2^{(M-s)j} \right)^{q-1} \sum_{j < j_h} 2^{(M-s)j} c_j^q \\ &\leq |h|_a^{M-s(1-q)} \sum_{j < j_h} 2^{(M-s)j} c_j^q. \end{aligned}$$

Thus by (A.6), Fubini's theorem and (A.1),

$$\begin{aligned} \int_{|h|_a \leq 1} |h|_a^{-sq} I_1^q(h) \frac{dh}{|h|_a^{a \cdot m}} &\leq \int_{|h|_a \leq 1} |h|_a^{M-s} \sum_{j < j_h} 2^{(M-s)j} c_j^q \frac{dh}{|h|_a^{a \cdot m}} \\ &\leq \sum_{j \geq -1} 2^{(M-s)j} c_j^q \int_{|h|_a \leq 2^{-j}} |h|_a^{M-s-a \cdot m} dh \\ &\lesssim \sum_{j \geq -1} 2^{(M-s)j} c_j^q 2^{-(M-s)j} = \|f\|_{\mathbf{B}_{p,q}^{s,a}(\rho)}^q. \end{aligned}$$

Similarly, one can show

$$\int_{|h|_a \leq 1} |h|_a^{-sq} I_2^q(h) \frac{dh}{|h|_a^{a \cdot m}} \lesssim \|f\|_{\mathbf{B}_{p,q}^{s,a}(\rho)}^q.$$

Moreover, for $s > 0$, we clearly have

$$\|f\|_{L^p(\rho)} \lesssim \|f\|_{\mathbf{B}_{p,q}^{s,a}(\rho)}.$$

Thus we obtain (A.4).

(ii) In this step we prove the converse part of (A.4). For $j \geq 0$, since $\int_{\mathbb{R}^N} \check{\phi}_j^a(h) dh = (2\pi)^{d/2} \phi_j^a(0) = 0$, by (2.20) and the change of variable, we have

$$\begin{aligned} \int_{\mathbb{R}^N} \check{\phi}_j^a(h) \delta_h^{(M)} f(x) dh &= \sum_{k=0}^M (-1)^{M-k} \binom{M}{k} \int_{\mathbb{R}^N} \check{\phi}_j^a(h) f(x + kh) dh \\ &= \sum_{k=1}^M (-1)^{M-k} \binom{M}{k} \int_{\mathbb{R}^N} \check{\phi}_j^a(h) f(x + kh) dh \\ &= \sum_{k=1}^M (-1)^{M-k} \binom{M}{k} \int_{\mathbb{R}^N} \phi_j^a(k \cdot) \check{\phi}_j^a(h) f(x + h) dh. \end{aligned}$$

In particular, if we define for $j \geq -1$,

$$\phi_j^{a,M}(\xi) := (-1)^{M+1} \sum_{k=1}^M (-1)^{M-k} \binom{M}{k} \phi_j^a(k\xi),$$

then

$$(-1)^{M+1} \int_{\mathbb{R}^N} \check{\phi}_j^a(h) \delta_h^{(M)} f(x) dh = [\phi_j^{a,M}]^* f(x) =: \mathcal{R}_j^{a,M} f(x),$$

and for $j \geq 0$,

$$\|\mathcal{R}_j^{a,M} f\|_{L^p(\rho)} \leq \int_{\mathbb{R}^N} |\check{\phi}_j^a(h)| \|\delta_h^{(M)} f\|_{L^p(\rho)} dh = I_j^0 + I_j^1 + I_j^2,$$

where

$$\begin{aligned} I_j^0 &:= \int_{|h|_a > 1} |\check{\phi}_j^a(h)| \|\delta_h^{(M)} f\|_{L^p(\rho)} dh, \\ I_j^1 &:= \int_{|h|_a \leq 2^{-j}} |\check{\phi}_j^a(h)| \|\delta_h^{(M)} f\|_{L^p(\rho)} dh, \\ I_j^2 &:= \int_{2^{-j} < |h|_a \leq 1} |\check{\phi}_j^a(h)| \|\delta_h^{(M)} f\|_{L^p(\rho)} dh. \end{aligned}$$

For I_j^0 , by (2.20) and (2.11), there is a $\kappa > 0$ such that

$$\|\delta_h^{(M)} f\|_{L^p(\rho)} \lesssim (1 + |h|_a^\kappa) \|f\|_{L^p(\rho)},$$

which implies that

$$\begin{aligned} I_j^0 &\lesssim \|f\|_{L^p(\rho)} \int_{|h|_a > 1} |\check{\phi}_j^a(h)| (1 + |h|_a^\kappa) dh \\ &= \|f\|_{L^p(\rho)} \int_{|h|_a > 2^j} |\check{\phi}_0^a(h)| (1 + 2^{-j\kappa} |h|_a^\kappa) dh \\ &\leq \|f\|_{L^p(\rho)} 2^{-j(s+1)} \int_{|h|_a > 2^j} |\check{\phi}_0^a(h)| (1 + |h|_a^\kappa) |h|_a^{s+1} dh \\ &\lesssim 2^{-j(s+1)} \|f\|_{L^p(\rho)}, \end{aligned}$$

and

$$\sum_{j \geq 0} 2^{sqj} (I_j^0)^q \lesssim \|f\|_{L^p(\rho)}^q.$$

For I_j^1 , by Hölder's inequality and change of variable, we have

$$\begin{aligned} (I_j^1)^q &\leq \left(\int_{|h|_a \leq 2^{-j}} |\check{\phi}_j^a(h)|^{q'} dh \right)^{q-1} \int_{|h|_a \leq 2^{-j}} \|\delta_h^{(M)} f\|_{L^p(\rho)}^q dh \\ &= 2^{a \cdot mj} \left(\int_{|h|_a \leq 1} |\check{\phi}_0^a(h)|^{q'} dh \right)^{q-1} \int_{|h|_a \leq 2^{-j}} \|\delta_h^{(M)} f\|_{L^p(\rho)}^q dh. \end{aligned}$$

Thus, by Fubini's theorem,

$$\begin{aligned} \sum_{j \geq 0} 2^{sqj} (I_j^1)^q &\lesssim \sum_{j \geq 0} 2^{sqj+a \cdot mj} \int_{|h|_a \leq 2^{-j}} \|\delta_h^{(M)} f\|_{L^p(\rho)}^q dh \\ &= \int_{|h|_a \leq 1} \sum_{j \geq 0, |h|_a \leq 2^{-j}} 2^{sqj+a \cdot mj} \|\delta_h^{(M)} f\|_{L^p(\rho)}^q dh \\ &\lesssim \int_{|h|_a \leq 1} |h|_a^{-sq-a \cdot m} \|\delta_h^{(M)} f\|_{L^p(\rho)}^q dh. \end{aligned}$$

For I_j^2 , by Hölder's inequality with respect to measure $\frac{dh}{|h|_a^{a \cdot m}}$, we also have

$$\begin{aligned} (I_j^2)^q &= 2^{-Mqj} \left(\int_{2^{-j} < |h|_a \leq 1} |2^{aj} h|_a^{a \cdot m + M} |\check{\phi}_0^a(2^{aj} h)| \frac{\|\delta_h^{(M)} f\|_{L^p(\rho)}}{|h|_a^M} \frac{dh}{|h|_a^{a \cdot m}} \right)^q \\ &\leq 2^{-Mqj} \left(\int_{|h|_a \geq 1} (|h|_a^{a \cdot m + M} |\check{\phi}_0^a(h)|)^{q'} \frac{dh}{|h|_a^{a \cdot m}} \right)^{q-1} \int_{2^{-j} \leq |h|_a \leq 1} \frac{\|\delta_h^{(M)} f\|_{L^p(\rho)}^q}{|h|_a^{Mq}} \frac{dh}{|h|_a^{a \cdot m}}. \end{aligned}$$

As above, by Fubini's theorem,

$$\begin{aligned} \sum_{j \geq 0} 2^{sqj} (I_j^2)^q &\lesssim \sum_{j \geq 0} 2^{(s-M)qj} \int_{2^{-j} \leq |h|_a \leq 1} \frac{\|\delta_h^{(M)} f\|_{L^p(\rho)}^q}{|h|_a^{Mq}} \frac{dh}{|h|_a^{a \cdot m}} \\ &\lesssim \int_{|h|_a \leq 1} \frac{\|\delta_h^{(M)} f\|_{L^p(\rho)}^q}{|h|_a^{sq}} \frac{dh}{|h|_a^{a \cdot m}}. \end{aligned}$$

For $j = -1$, estimate is easy. Hence,

$$\sum_{j \geq -1} 2^{sqj} \|\mathcal{R}_j^{a,M} f\|_{L^p(\rho)}^q \lesssim \int_{|h|_a \leq 1} \frac{\|\delta_h^{(M)} f\|_{L^p(\rho)}^q}{|h|_a^{sq}} \frac{dh}{|h|_a^{a \cdot m}} + \|f\|_{L^p(\rho)}^q.$$

On the other hand, noting that

$$\phi_j^{a,M}(\xi) = \phi_{-1}^{a,M}(2^{-a(j+1)}\xi) - \phi_{-1}^{a,M}(2^{-aj}\xi)$$

and

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

we have

$$\text{supp } \phi_j^{a,M} \subset B_{2^{j+1}}^a \setminus B_{(2^{j-1})/M}^a.$$

Thus, for any $i, j \geq -1$ with $|j - i| > \log_2 M + 2 =: \gamma$,

$$\mathcal{R}_j^{a,M} \mathcal{R}_i^a f(x) = 0.$$

Moreover, noting that for any $\xi \in \mathbb{R}^N$,

$$\begin{aligned} \sum_{j \geq -1} \phi_j^{a,M}(\xi) &= (-1)^{M+1} \sum_{k=1}^M \sum_{j \geq -1} (-1)^{M-k} \binom{M}{k} \phi_j^a(k\xi) \\ &= (-1)^{M+1} \sum_{k=1}^M (-1)^{M-k} \binom{M}{k} = 1, \end{aligned}$$

we have

$$\mathcal{R}_i^a f = \sum_{j \geq -1} \mathcal{R}_i^a \mathcal{R}_j^{a,M} f = \sum_{|j-i| \leq \gamma} \mathcal{R}_i^a \mathcal{R}_j^{a,M} f.$$

Therefore, by (2.13),

$$\begin{aligned} \sum_{i \geq -1} 2^{sqi} \|\mathcal{R}_i^a f\|_{L^p(\rho)}^q &\leq \sum_{i \geq -1} 2^{sqi} \sum_{|i-j| \leq \gamma} \|\mathcal{R}_i^a \mathcal{R}_j^{a,M} f\|_{L^p(\rho)}^q \\ &\lesssim \sum_{i \geq -1} 2^{sqi} \sum_{|i-j| \leq \gamma} \|\mathcal{R}_j^{a,M} f\|_{L^p(\rho)}^q \\ &\lesssim \sum_{j \geq -1} 2^{sqj} \|\mathcal{R}_j^{a,M} f\|_{L^p(\rho)}^q \lesssim \|f\|_{\tilde{\mathbf{B}}_{p,q}^{s,a}(\rho)}^q. \end{aligned}$$

(iii) In this step we prove the second equivalence in (2.21) for $s > 0$. For $s \in (0, 1)$ and $|h|_a \leq 1$, by (2.10) and (2.11), we have

$$\|[\delta_h, \rho] f\|_{L^p} \lesssim \|f \delta_h \rho\|_{L^p} \lesssim |h| \|f\|_{L^p(\rho)} \leq |h|_a \|f\|_{L^p(\rho)},$$

which implies that for $s \in (0, 1)$,

$$(A.7) \quad \|f\|_{\mathbf{B}_{p,q}^{s,a}(\rho)} \asymp \|f\|_{\tilde{\mathbf{B}}_{p,q}^{s,a}(\rho)} \asymp \|\rho f\|_{\tilde{\mathbf{B}}_{p,q}^{s,a}} \asymp \|\rho f\|_{\mathbf{B}_{p,q}^{s,a}}.$$

For $s \in [1, 2)$, we have

$$\|[\delta_h^{(2)}, \rho] f\|_{L^p} \lesssim \|f \delta_h^{(2)} \rho\|_{L^p} + \|\delta_h \rho \delta_h f\|_{L^p} \lesssim |h|_a^2 \|f \rho\|_{L^p} + |h|_a \|\rho \delta_h f\|_{L^p},$$

which in turn implies (A.7) for $s \in [1, 2)$ by definition and the equivalence for $s \in (0, 1)$. By induction one can show (A.7) for general $s \geq 2$.

(iv) In this step we show (2.22) for $s \leq 0$. For $s < 0$, by Lemma A.2 and the equivalence for $s > 0$ proven in step (iii), we have

$$\|f\|_{\mathbf{B}_{p,q}^{s,a}(\rho)} \lesssim \sup_{\varphi \in \mathcal{S}} \frac{|\langle f, \varphi \rangle|}{\|\varphi\|_{\mathbf{B}_{p',q'}^{-s,a}(\rho^{-1})}} \lesssim \sup_{\varphi \in \mathcal{S}} \frac{|\langle f, \varphi \rangle|}{\|\rho^{-1} \varphi\|_{\mathbf{B}_{p',q'}^{-s,a}}} \lesssim \|\rho f\|_{\mathbf{B}_{p,q}^{s,a}}.$$

For $s = 0$ and $q \in [1, \infty)$, we have

$$\|f\|_{\mathbf{B}_{p,q}^{0,a}(\rho)}^q = \sum_{j \geq -1} \|\mathcal{R}_j^a f\|_{L^p(\rho)}^q \leq \sum_{j \geq -1} \left(\sum_{k \geq -1} \|\mathcal{R}_j^a(\rho^{-1} \mathcal{R}_k^a(\rho f))\|_{L^p(\rho)} \right)^q \leq I_1 + I_2,$$

where

$$\begin{aligned} I_1 &:= \sum_{j \geq -1} \left(\sum_{k \leq j} \|\mathcal{R}_j^a(\rho^{-1} \mathcal{R}_k^a(\rho f))\|_{L^p(\rho)} \right)^q, \\ I_2 &:= \sum_{j \geq -1} \left(\sum_{k > j} \|\mathcal{R}_j^a(\rho^{-1} \mathcal{R}_k^a(\rho f))\|_{L^p(\rho)} \right)^q. \end{aligned}$$

Fix $\alpha \in (0, 1)$. By Hölder's inequality, we have

$$\begin{aligned}
I_1 &= \sum_{j \geq -1} \left(\sum_{k \leq j} 2^{\alpha k} 2^{-\alpha k} \|\mathcal{R}_j^a(\rho^{-1} \mathcal{R}_k^a(\rho f))\|_{L^p(\rho)} \right)^q \\
&\leq \sum_{j \geq -1} \left(\sum_{k \leq j} 2^{\alpha k q / (q-1)} \right)^{q-1} \sum_{k \leq j} 2^{-\alpha k q} \|\mathcal{R}_j^a(\rho^{-1} \mathcal{R}_k^a(\rho f))\|_{L^p(\rho)}^q \\
&\lesssim \sum_{j \geq -1} 2^{\alpha q j} \sum_{k \geq -1} 2^{-\alpha k q} \|\mathcal{R}_j^a(\rho^{-1} \mathcal{R}_k^a(\rho f))\|_{L^p(\rho)}^q \\
&= \sum_{k \geq -1} 2^{-\alpha k q} \sum_{j \geq -1} 2^{\alpha q j} \|\mathcal{R}_j^a(\rho^{-1} \mathcal{R}_k^a(\rho f))\|_{L^p(\rho)}^q.
\end{aligned}$$

Noting that

$$\sum_{j \geq -1} 2^{\alpha q j} \|\mathcal{R}_j^a(\rho^{-1} g)\|_{L^p(\rho)}^q = \|\rho^{-1} g\|_{\mathbf{B}_{p,q}^{\alpha,a}(\rho)}^q \lesssim \|g\|_{\mathbf{B}_{p,q}^{\alpha,a}},$$

we further have

$$I_1 \lesssim \sum_{k \geq -1} 2^{-\alpha k q} \|\mathcal{R}_k^a(\rho f)\|_{\mathbf{B}_{p,q}^{\alpha,a}}^q \lesssim \sum_{k \geq -1} \|\mathcal{R}_k^a(\rho f)\|_{L^p}^q = \|\rho f\|_{\mathbf{B}_{p,q}^{0,a}}^q.$$

Similarly, one can show

$$I_2 \lesssim \sum_{k \geq -1} 2^{\alpha k q} \|\mathcal{R}_k^a(\rho f)\|_{\mathbf{B}_{p,q}^{-\alpha,a}}^q \lesssim \|\rho f\|_{\mathbf{B}_{p,q}^{0,a}}^q.$$

Thus we get for $q \in [1, \infty)$,

$$\|f\|_{\mathbf{B}_{p,q}^{0,a}(\rho)}^q \lesssim \|\rho f\|_{\mathbf{B}_{p,q}^{0,a}}^q.$$

For $q = \infty$, it is similar. Moreover, for $s \leq 0$, by duality, we also have

$$\|\rho f\|_{\mathbf{B}_{p,q}^{s,a}} \lesssim \sup_{\varphi \in \mathcal{S}} \frac{|\langle \rho f, \varphi \rangle|}{\|\varphi\|_{\mathbf{B}_{p',q'}^{-s,a}}} \lesssim \sup_{\varphi \in \mathcal{S}} \frac{|\langle f, \rho \varphi \rangle|}{\|\rho \varphi\|_{\mathbf{B}_{p',q'}^{-s,a}(\rho^{-1})}} \lesssim \|f\|_{\mathbf{B}_{p,q}^{s,a}(\rho)}.$$

The proof is complete. \square

By (2.16) and characterization (2.21), the following compact embedding lemma is standard by Ascoli-Arzelà's lemma.

LEMMA A.3. *Let $T > 0$, $\rho_1, \rho_2 \in \mathcal{P}_w$ and $0 < \alpha_1 < \alpha_2 < 2$. If $\rho_1(z) = \varrho(z)^{-\kappa}$ for some $\kappa > 0$, then the following embedding is compact*

$$\mathbb{S}_{T,a}^{\alpha_2}(\rho_1 \rho_2) \hookrightarrow \mathbb{S}_{T,a}^{\alpha_1}(\rho_2).$$

PROOF. Let f_n be a bounded sequence of $\mathbb{S}_{T,a}^{\alpha_2}(\rho_1 \rho_2)$. For any $R \geq 1$, by (2.25), there is a constant $C = C(R, T) > 0$ such that for any $z_1, z_2 \in B_R^a$,

$$|f_n(t, z_1) - f_n(t, z_2)| \leq C |z_1 - z_2|^{\alpha_2/2}$$

and for any $z = (x, v) \in B_R^a$

$$\begin{aligned}
|f_n(t, z) - f_n(s, z)| &\leq |f_n(t, z) - f_n(s, \Gamma_{t-s} z)| + |f_n(s, \Gamma_{t-s} z) - f_n(s, z)| \\
&\lesssim_C |t - s|^{\alpha_2/2} + |(t - s)v|^{\alpha_2/6} \lesssim_C |t - s|^{\alpha_2/6}.
\end{aligned}$$

Hence, by Ascoli-Arzelà's theorem and a diagonalization method, there are a subsequence n_k and a continuous f such that for any $R \geq 1$,

$$(A.8) \quad \lim_{k \rightarrow \infty} \sup_{t \in [0, T]} \sup_{z \in B_R^a} |f_{n_k}(t, z) - f(t, z)| = 0.$$

In particular, $f \in \mathbb{S}_{T,a}^{\alpha_2}(\rho_1 \rho_2)$. It remains to show

$$(A.9) \quad \lim_{k \rightarrow \infty} \|f_{n_k} - f\|_{\mathbb{S}_{T,a}^{\alpha_1}(\rho_2)} = 0.$$

Note that by definition, for any $R \geq 1$,

$$\|\mathbf{1}_{\{|z|_a > R\}}(f_{n_k} - f)\|_{\mathbb{L}_T^\infty(\rho_2)} \leq \|f_{n_k} - f\|_{\mathbb{L}_T^\infty(\rho_1 \rho_2)} / (1 + R)^\kappa,$$

which together with (A.8) implies that

$$(A.10) \quad \lim_{k \rightarrow \infty} \|f_{n_k} - f\|_{\mathbb{L}_T^\infty(\rho_2)} = 0.$$

Since $(f_{n_k})_{k \in \mathbb{N}}$ is bounded in $\mathbb{S}_{T,a}^{\alpha_2}(\rho_2)$ and by the interpolation inequality (2.19),

$$\|f\|_{\mathbb{S}_{T,a}^{\alpha_1}(\rho_2)} \lesssim \|f\|_{\mathbb{S}_{T,a}^{\alpha_2}(\rho_2)}^{\alpha_1/\alpha_2} \|f\|_{\mathbb{L}_T^\infty(\rho_2)}^{1-\alpha_1/\alpha_2},$$

we get (A.9) by (A.10). The proof is complete. \square

APPENDIX B: PROOF OF LEMMA 7.4

In this section we collect some useful lemmas used in Section 7 and give the proof of Lemma 7.4.

LEMMA B.1. *For any $p \geq 2$ and $k \in \mathbb{N}$, we have*

$$\sup_{z \in \mathbb{R}^{2d}} \mathbf{E} |\nabla^k X_\varphi(z)|^p < \infty.$$

In particular, $z \mapsto X_\varphi(z)$ has a smooth version.

PROOF. Since W is a bounded linear operator from \mathbb{H} to $L^2(\Omega)$, we have

$$\nabla^k X_\varphi(z) = X(\nabla^k \varphi(z - \cdot)), \text{ a.s.}$$

By the hypercontractivity of Gaussian random variables and (7.2), we have

$$(B.1) \quad \mathbf{E} |\nabla^k X_\varphi(z)|^p \lesssim (\mathbf{E} |\nabla^k X_\varphi(z)|^2)^{p/2} = \left(\int_{\mathbb{R}^{2d}} |\widehat{\nabla^k \varphi}(\zeta)|^2 \mu(d\zeta) \right)^{p/2},$$

which is finite by $\varphi \in \mathcal{S}(\mathbb{R}^{2d})$ and (7.1). The proof is complete. \square

PROOF OF LEMMA 7.4. Note that by (7.2),

$$(B.2) \quad \mathbf{E} \left(X_\varphi(z) X_{\varphi'}(z') \right) = \int_{\mathbb{R}^{2d}} e^{i\zeta(z'-z)} \hat{\varphi}(\zeta) \hat{\varphi}'(\zeta) \mu(d\zeta) =: I_{\varphi, \varphi'}(z, z').$$

By (7.4) and Fubini's theorem, we have

$$(B.3) \quad \begin{aligned} \mathbf{E} \left((X_\varphi \otimes X_{\varphi'})(H) \right) &= \int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} H(z, z') I_{\varphi, \varphi'}(z, z') dz dz' \\ &= \int_{\mathbb{R}^{2d}} \hat{H}(\zeta, -\zeta) \hat{\varphi}(\zeta) \hat{\varphi}'(\zeta) \mu(d\zeta). \end{aligned}$$

Next we look at (7.6). Noting that for Gaussian random variables $(\xi_1, \xi_2, \xi_3, \xi_4)$,

$$\mathbf{E}(\xi_1 \xi_2 \xi_3 \xi_4) = \mathbf{E}(\xi_1 \xi_2) \mathbf{E}(\xi_3 \xi_4) + \mathbf{E}(\xi_1 \xi_3) \mathbf{E}(\xi_2 \xi_4) + \mathbf{E}(\xi_1 \xi_4) \mathbf{E}(\xi_2 \xi_3),$$

by Fubini's theorem again and (B.2), we have

$$\begin{aligned} \mathbf{E} \left((X_\varphi \otimes X_{\varphi'})(H) \right)^2 &= \mathbf{E} \left(\int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} H(z, z') X_\varphi(z) X_{\varphi'}(z') dz dz' \right)^2 \\ &= \int_{\mathbb{R}^{2d}} \cdots \int_{\mathbb{R}^{2d}} H(z, z') H(\bar{z}, \bar{z}') \mathbf{E} (X_\varphi(z) X_{\varphi'}(z') X_\varphi(\bar{z}) X_{\varphi'}(\bar{z}')) dz dz' d\bar{z} d\bar{z}' \\ &= \int_{\mathbb{R}^{2d}} \cdots \int_{\mathbb{R}^{2d}} H(z, z') H(\bar{z}, \bar{z}') \left(I_{\varphi, \varphi'}(z, z') I_{\varphi, \varphi'}(\bar{z}, \bar{z}') \right. \\ &\quad \left. + I_{\varphi, \varphi'}(z, \bar{z}) I_{\varphi', \varphi'}(z', \bar{z}') + I_{\varphi, \varphi'}(z, \bar{z}') I_{\varphi, \varphi'}(\bar{z}, z') \right) dz dz' d\bar{z} d\bar{z}'. \end{aligned}$$

Hence, by (B.3),

$$\begin{aligned} \text{Var} \left((X_\varphi \otimes X_{\varphi'})(H) \right) &= \mathbf{E} \left((X_\varphi \otimes X_{\varphi'})(H) \right)^2 - \left(\mathbf{E} \left((X_\varphi \otimes X_{\varphi'})(H) \right) \right)^2 \\ &= \int_{\mathbb{R}^{2d}} \cdots \int_{\mathbb{R}^{2d}} H(z, z') H(\bar{z}, \bar{z}') I_{\varphi, \varphi}(z, \bar{z}) I_{\varphi', \varphi'}(z', \bar{z}') dz dz' d\bar{z} d\bar{z}' \\ &\quad + \int_{\mathbb{R}^{2d}} \cdots \int_{\mathbb{R}^{2d}} H(z, z') H(\bar{z}, \bar{z}') I_{\varphi, \varphi'}(z, \bar{z}') I_{\varphi, \varphi'}(\bar{z}, z') dz dz' d\bar{z} d\bar{z}' \\ &= \int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} \hat{H}(\zeta, \zeta') \hat{H}(-\zeta, -\zeta') |\hat{\varphi}(\zeta)|^2 |\hat{\varphi}'(\zeta')|^2 \mu(d\zeta) \mu(d\zeta') \\ &\quad + \int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} \hat{H}(\zeta, -\zeta') \hat{H}(\zeta', -\zeta) \hat{\varphi}(\zeta) \hat{\varphi}(\zeta') \hat{\varphi}'(\zeta) \hat{\varphi}'(\zeta') \mu(d\zeta) \mu(d\zeta') \\ &= \int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} \hat{H}(\zeta, \zeta') \overline{\hat{H}(\zeta, \zeta')} |\hat{\varphi}(\zeta)|^2 |\hat{\varphi}'(\zeta')|^2 \mu(d\zeta) \mu(d\zeta') \\ &\quad + \int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} \hat{H}(\zeta, \zeta') \overline{\hat{H}(\zeta', \zeta)} \hat{\varphi}(\zeta) \hat{\varphi}(\zeta') \hat{\varphi}'(\zeta) \hat{\varphi}'(\zeta') \mu(d\zeta) \mu(d\zeta'), \end{aligned}$$

where the last step is due to the symmetry of $\hat{\varphi}$, $\hat{\varphi}'$ and μ . From this we get the desired equality (7.6). \square

Recall (7.12) and we have the following elementary lemmas.

LEMMA B.2. (i) For any $\gamma \in \mathbb{R}$, there is a constant $C > 0$ such that

$$(B.4) \quad |\phi_j^a(\zeta)| \lesssim_C 1 \wedge (2^{\gamma j} (1 + |\zeta|_a)^{-\gamma}), \quad j \geq -1, \quad \zeta \in \mathbb{R}^{2d},$$

and

$$(B.5) \quad |\psi(\zeta, \zeta')| \lesssim_C 1 \wedge ((1 + |\zeta|_a)^{-\gamma} (1 + |\zeta'|_a)^\gamma), \quad \zeta, \zeta' \in \mathbb{R}^{2d}.$$

(ii) For any $\gamma \in [0, 1]$, there is a constant $C > 0$ such that

$$(B.6) \quad |\psi(\zeta, \zeta') - \psi(\zeta, \zeta)| \lesssim_C |\zeta - \zeta'|_a^\gamma (1 + |\zeta|_a)^{-\gamma}, \quad \zeta, \zeta' \in \mathbb{R}^{2d}.$$

PROOF. (i) Note that

$$K_j := \text{supp} \phi_j^a \subset \{\zeta : 2^{j-1} \leq |\zeta|_a \leq 2^{j+1}\}, \quad j \geq 0.$$

For any $\gamma \in \mathbb{R}$, since for $j \geq 0$,

$$(1 + |\zeta|_a)^\gamma \mathbf{1}_{K_j} \lesssim 2^{j\gamma}, \quad (1 + |\zeta|_a)^\gamma \mathbf{1}_{\{|\zeta|_a \leq 1\}} \lesssim 1,$$

we have

$$|\phi_j^a(\zeta)| \leq \frac{(1 + |\zeta|_a)^\gamma}{(1 + |\zeta|_a)^\gamma} \mathbf{1}_{K_j}(\zeta) \lesssim \frac{2^{j\gamma}}{(1 + |\zeta|_a)^\gamma},$$

and

$$\begin{aligned} |\psi(\zeta, \zeta')| &\leq \sum_{|i-j| \leq 1} |\phi_i^a(\zeta)| |\phi_j^a(\zeta')| \lesssim \sum_{|i-j| \leq 1} \frac{2^{\gamma i} \mathbf{1}_{K_i}(\zeta)}{(1 + |\zeta|_a)^\gamma} \frac{(1 + |\zeta'|_a)^\gamma}{2^{\gamma j}} \\ &\lesssim \sum_{i \geq -1} \mathbf{1}_{K_i}(\zeta) \frac{(1 + |\zeta'|_a)^\gamma}{(1 + |\zeta|_a)^\gamma} \lesssim \frac{(1 + |\zeta'|_a)^\gamma}{(1 + |\zeta|_a)^\gamma}. \end{aligned}$$

(ii) Let $\gamma \in [0, 1]$. For $j \geq 0$, we have

$$|\phi_j^a(\zeta) - \phi_j^a(\zeta')| = |\phi_0^a(2^{-aj}\zeta) - \phi_0^a(2^{-aj}\zeta')| \lesssim |\zeta - \zeta'|_a^\gamma 2^{-j\gamma} \|\phi_0^a\|_{\mathbf{C}_a^\gamma}$$

and

$$|\phi_{-1}^a(\zeta) - \phi_{-1}^a(\zeta')| \lesssim |\zeta - \zeta'|_a^\gamma \|\phi_{-1}^a\|_{\mathbf{C}_a^\gamma}.$$

Thus, by (B.4),

$$|\psi(\zeta, \zeta') - \psi(\zeta, \zeta)| \lesssim |\zeta - \zeta'|_a^\gamma \sum_{j \geq -1} 2^{-j\gamma} \phi_j^a(\zeta) \lesssim \frac{|\zeta - \zeta'|_a^\gamma}{(1 + |\zeta|_a)^\gamma} \sum_{j \geq -1} \mathbf{1}_{K_j}(\zeta).$$

The proof is complete. \square

We also need the following simple lemma.

LEMMA B.3. *For any $T, \lambda > 0$, $\theta \in [0, 1]$ and $\gamma > 0$, there is a constant $C = C(T, \gamma, \theta, \lambda)$ such that for any $0 \leq s < t \leq T$ and $\zeta = (\xi, \eta) \in \mathbb{R}^{2d}$,*

$$(B.7) \quad \int_s^t r^{\gamma-1} e^{-\lambda(r^3|\xi|^2 + r|\eta|^2)} dr \lesssim_C |t - s|^{(\gamma \wedge 1)(1-\theta)} (1 + |\zeta|_a)^{-2\theta\gamma}.$$

PROOF. Note that

$$\int_0^t s^{\gamma-1} e^{-\lambda s^3|\xi|^2} ds \lesssim |\xi|^{-\frac{2\gamma}{3}}, \quad \int_0^t s^{\gamma-1} e^{-\lambda s|\eta|^2} ds \lesssim |\eta|^{-2\gamma},$$

and

$$\int_s^t r^{\gamma-1} dr = (t^\gamma - s^\gamma)/\gamma \lesssim (t - s)^{\gamma \wedge 1}.$$

Let $g(r, \zeta) := e^{-\lambda(r^3|\xi|^2 + r|\eta|^2)}$. For any $\theta \in [0, 1]$, we have

$$\begin{aligned} \int_s^t r^{\gamma-1} g(r, \zeta) dr &= \left(\int_s^t r^{\gamma-1} g(r, \zeta) dr \right)^{1-\theta} \left(\int_s^t r^{\gamma-1} g(r, \zeta) dr \right)^\theta \\ &\leq \left(\int_s^t r^{\gamma-1} dr \right)^{1-\theta} \left(\int_0^t s^{\gamma-1} g(s, \zeta) ds \right)^\theta \\ &\lesssim (t - s)^{(\gamma \wedge 1)(1-\theta)} \left(1 \wedge |\xi|^{-\frac{2\gamma}{3}} \wedge |\eta|^{-2\gamma} \right)^\theta, \end{aligned}$$

which in turn gives the result by $1 \vee |\xi|^{1/3} \vee |\eta| \asymp 1 + |\zeta|_a$. \square

Funding. X. Zhang is partially supported by NSFC (No. 12131019, 11731009). R.Z. and X.Z. are grateful to the financial supports by National Key R&D Program of China (No. 2022YFA1006300). R.Z. gratefully acknowledges financial support from the NSFC (No. 12271030) and BIT Science and Technology Innovation Program Project 2022CX01001. X.Z. is partially supported by National Key R&D Program of China (No. 2020YFA0712700), the NSFC (No. 12090014, 12288201) and the support by key Lab of Random Complex Structures and Data Science, Youth Innovation Promotion Association (2020003), Chinese Academy of Science. This work is also funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) 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”.

REFERENCES

- [AV04] Alexandre R. and Villani C. (2004). On the Landau approximation in plasma physics. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **21** 61–95.
- [Bo02] Bouchut F. (2002). Hypocoelliptic regularity in kinetic equations. *J. Math. Pures Appl.* (9). **81** 1135–1159.
- [BCD11] Bahouri H., Chemin J.-Y. and Danchin R. (2011). *Fourier analysis and nonlinear partial differential equations. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **343**. Springer, Heidelberg.
- [BG97] Bertini L. and Giacomin G. (1997). Stochastic Burgers and KPZ equations from particle systems. *Comm. Math. Phys.* **183** 571–607.
- [BL76] Bergh J. and Löfström J. (1976). *Interpolation spaces: an introduction* Springer-Verlag, Berlin Heidelberg New York.
- [BRTV98] Benachour S., Roynette B., Talay D. and Vallois P. (1998). Nonlinear self-stabilizing processes. I: Existence, invariant probability, propagation of chaos. *Stoch. Proc. Appl.* **75** 173–201.
- [BT97] Bossy M. and Talay D. (1997). A stochastic particle method for the McKean-Vlasov and the Burgers equation. *Math. Comp.* **66** 157–192.
- [Bon81] Bony J.-M. (1981). Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non linéaires. *Ann. Sci. École Norm. Sup. (4)* **14** 209–246.
- [CC18] Cannizzaro R. and Chouk K. (2018). Multidimensional SDEs with singular drift and universal construction of the polymer measure with white noise potential. *Ann. Probab.* **46** 1710–1763.
- [Ch17] Chaudru de Raynal P.E. (2017). Strong existence and uniqueness for degenerate SDE with Hölder drift. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 259–286.
- [CF16] Coghi M. and Flandoli F. (2016). Propagation of chaos for interacting particles subject to environmental noise. *Ann. Appl. Probab.* **26** 1407–1442.
- [CX10] Crisan D. and Xiong J. (2010). Approximate McKean-Vlasov representations for a class of SPDEs. *Stochastics* **82** 53–68.
- [CZ18] Chen Z.-Q. and Zhang X. (2018). L^p -maximal hypoelliptic regularity of nonlocal kinetic Fokker-Planck operators. *J. Math. Pures Appl.* (9). **116** 52–87.
- [C12] Corwin I. (2012). The Kardar-Parisi-Zhang equation and universality class. *Random Matrices Theory Appl.* **1** 1130001, 76.
- [Dac03] Dachkovski S. (2003). Anisotropic function spaces and related semi-linear hypoelliptic equations. *Math. Nachr.* **248/249** 40–61.
- [DD16] Delarue F. and Diel R. (2016). Rough paths and 1d SDE with a time dependent distributional drift: application to polymers. *Probab. Theory Related Fields* **165** 1–63.
- [Di96] Dintelman P. (1996). Fourier Multipliers between Weighted Anisotropic Function Spaces. Part I: Besov spaces. *Journal for Analysis and its Applications* **15** 579–601.
- [DM10] Delarue F. and Menozzi S. (2010). Density estimates for a random noise propagating through a chain of differential equations. *J. Funct. Anal.* **259** 1577–1630.
- [DW18] A. Debussche and H. Weber, The Schrödinger equation with spatial white noise potential, *Electron. J. Probab.* **23**, No. 28, 16, 2018.
- [EK86] Ethier N. and Kurtz G. (1986). *Markov Processes: Characterization and Convergence. Wiley series in probability and mathematical statistics.* Wiley.
- [Fu84] Funaki T. (1984). A certain class of diffusion processes associated with nonlinear parabolic equations. *Z. Wahrsch. Verw. Gebiete* **67** 331–348.

- [FHM14] Fournier N., Hauray M. and Mischler S. (2014). Propagation of chaos for the 2d viscous vortex model. *J. Eur. Math. Soc.* **16** 1423–1466.
- [GHa19] Gerencsér M. and Hairer M. (2019). A Solution Theory for Quasilinear Singular SPDEs. *Comm. Pure Appl. Math.* **72** 1983–2005.
- [GH19] Gubinelli M. and Hofmanová M. (2019). Global solutions to elliptic and parabolic Φ^4 models in Euclidean space. *Comm. Math. Phys.* **368** 1201–1266.
- [GH21] Gubinelli M. and Hofmanová M. (2021). A PDE construction of the Euclidean Φ^4 quantum field theory. *Comm. Math. Phys.* **384** 1–75.
- [GIP15] Gubinelli M., Imkeller P. and Perkowski N. (2015). Paracontrolled distributions and singular PDEs. *Forum Math. Pi* **3** e6–75.
- [GKO18] Gubinelli M., Koch H. and Oh T. (2018). Renormalization of the two-dimensional stochastic nonlinear wave equations. *Trans. Amer. Math. Soc.* **370** 7335–7359.
- [GKO18a] Gubinelli M., Koch H. and Oh T. (2018). Paracontrolled approach to the three-dimensional stochastic nonlinear wave equation with quadratic nonlinearity. *J. Eur. Math. Soc.*
- [GP17] Gubinelli M. and Perkowski N. (2017). KPZ reloaded. *Comm. Math. Phys.* **349** 165–269.
- [GP18] Gubinelli M. and Perkowski N. (2018). Energy solutions of KPZ are unique. *J. Amer. Math. Soc.* **31** 427–471.
- [Hai13] Hairer M. (2013). Solving the KPZ equation. *Ann. of Math. (2)* **178** 559–664.
- [Hai14] Hairer M. (2014). A theory of regularity structures. *Invent. Math.* **198** 269–504.
- [HL18] Hairer M. and Labbé C. (2018). Multiplicative stochastic heat equations on the whole space. *J. Eur. Math. Soc.* **20** 1005–1054.
- [HSS21] Hammersley W. R.P., Šiška D. and Szpruch Ł. (2021). Weak existence and uniqueness for McKean-Vlasov SDEs with common noise. *Ann. Probab.* **49** 527–555.
- [HWZ20] Hao Z., Wu M. and Zhang X. (2020). Schauder’s estimate for nonlocal kinetic equations and applications. *J. Math. Pures Appl. (9)* **140** 139–184.
- [HMP19] Huang L., Menozzi S. and Priola E. (2019). L^p -Estimates for degenerate non-local Kolmogorov operators. *J. Math. Pures Appl. (9)* **121** 162–215.
- [IS21] Imbert C. and Silvestre L. (2021). The Schauder estimate for kinetic integral equations. *Analysis and PDE* **14** 717–204.
- [JW16] Jabin P.-E. and Wang Z. (2016). Mean field limit and propagation of chaos for Vlasov systems with bounded forces. *J. Funct. Anal.* **271** 3588–3627.
- [JW18] Jabin P.-E. and Wang Z. (2018). Quantitative estimates of propagation of chaos for stochastic systems with $w^{-1,\infty}$ kernels. *Invent. Math.* **214** 523–591.
- [Ka56] Kac M. (1956). *Foundations of kinetic theory*. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, Berkeley and Los Angeles.
- [KPZ86] Kardar M., Parisi G. and Zhang Y. (1986). Dynamic scaling of growing interfaces. *Phys. Rev. Lett.* **56** 889–892.
- [Kol34] Kolmogorov A.N. (1934). Zufällige Bewegungen. *Ann. Math.* **35** 116–117.
- [KP22] Kremp H. and Perkowski N. (2022). Multidimensional SDE with distributional drift and Lévy noise. *Bernoulli* **28** 1757–1783.
- [La23] Lacker D. (2023). Hierarchies, entropy, and quantitative propagation of chaos for mean field diffusions. *Probability and Mathematical Physics* **4** 377–432.
- [Lan36] Landau L.D. (1937). The kinetic equation in the case of Coulomb interaction. *Zh. Eksper. i Teoret. Fiz* **7:2** 203–209 (Translated from German).
- [Lo05] Lorenzi L. (2005). Schauder estimates for degenerate elliptic and parabolic problems with unbounded coefficients in RN. *Differential Integral Equations* **18** 531–566.
- [McK67] McKean H. P. (1967). *Propagation of chaos for a class of non-linear parabolic equations*. In *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)*, pages 41–57. Air Force Office Sci. Res., Arlington, Va.
- [MV20] Mishura Y. S. and Veretennikov A. Y. (2020). Existence and uniqueness theorems for solutions of McKean-Vlasov stochastic equations. *Theor. Probability and Math. Statist.* **103** 59–101.
- [MW17] Mourrat J.-C. and Weber H. (2017). Global well-posedness of the dynamic Φ^4 model in the plane. *Ann. Probab.* **45** 2398–2476.
- [MW17a] Mourrat J.-C. and Weber H. (2017). The dynamic Φ^4_3 model comes down from infinity. *Comm. Math. Phys.* **356** 673–753.
- [Nua06] Nualart D. (2006). *The Malliavin Calculus and Related Topics, second ed., Probability and Its Applications* (New York), Springer-Verlag, Berlin.

- [Osa86] Osada H. (1986). Propagation of chaos for the two dimensional navier-stokes equation. *Proceedings of the Japan Academy, Series A, Mathematical Sciences* **62** 8–11.
- [OSSW18] Otto F., Sauer J., Smith S. and Weber H. (2018). Parabolic Equations with Rough Coefficients and Singular Forcing. Available in arXiv:1803.07884.
- [OSSW21] Otto F., Sauer J., Smith S. and Weber H. (2021). A priori bounds for quasi-linear SPDEs in the full sub-critical regime. Available in arXiv:2103.11039.
- [OW19] Otto F. and Weber H. (2019). Quasilinear SPDEs via Rough Paths. *Arch. Ration. Mech. Anal.* **232** 873–950.
- [Po04] Polidoro S. (2004). *Recent results on Kolmogorov equations and applications*, pp. 129-143 in *Proceedings of the Workshop on Second Order Subelliptic Equations and Applications* (Cortona, Italy, 2003), edited by I. Birindelli et al., Università degli Studi della Basilicata, Potenza.
- [PR19] Perkowski N. and Rosati T. C. (2019). The KPZ equation on the real line. *Electron. J. Probab.* **24** 1–56.
- [PR21] Perkowski N. and Rosati T. C. (2021). A rough super-Brownian motion. *Ann. Probab.* **49** 908–943.
- [Pr09] Priola E. (2009). Global Schauder estimates for a class of degenerate Kolmogorov equations. *Studia Math.* **194** 117–153.
- [R20] Rosenzweig M. (2020). The Mean-Field Limit of Stochastic Point Vortex Systems with Multiplicative Noise. Available in arXiv:2011.12180.
- [RXZ21] Röckner M., Xie L. and Zhang X. (2020). Superposition principle for non-local Fokker-Planck-Kolmogorov operators. *Probab. Theory and Related Fields* **178** 699–733.
- [RZ21] Röckner M. and Zhang X. (2021). Well-posedness of distribution dependent SDEs with singular drifts. *Bernoulli* **27** 1131–1158.
- [Szn91] Sznitman A.-S. (1991). *Topics in propagation of chaos*. In *École d’Été de Probabilités de Saint-Flour XIX/1989, volume 1464 of Lecture Notes in Math.*, pages 165-251. Springer, Berlin.
- [Ser20] Serfaty S. (2020). Mean field limit for coulomb-type flows. *Duke Math. J.* **169** 2887–2935.
- [St70] Stein E.M. (1970). *Singular Integrals and Differentiability Properties of Functions*. Princeton University Press.
- [SV97] Svante J. (1997). *Gaussian Hilbert Spaces*. Cambridge University Press.
- [Tri06] Triebel H. (2006). *Theory of function spaces III*. Basel, Birkhäuser.
- [Tri92] Triebel H. (1992). *Theory of function spaces II, volume 84 of Monographs in Mathematics*. Birkhäuser Verlag, Basel.
- [Vl68] Vlasov A. (1968). The vibrational properties of an electron gas. *Physics-Uspeski* **10** 721–733.
- [Wa18] Wang F.-Y. (2018). Distribution dependent SDEs for Landau type equations. *Stoch. Proc. Appl.* **128** 595–621.
- [WZ16] Wang F.-Y. and Zhang X. (2016). Degenerate SDE with Hölder-Dini drift and Non-Lipschitz noise coefficient. *SIAM J. Math. Anal.* **48** 2189–2222.
- [Zh18] Zhang X. (2018). Stochastic Hamiltonian flows with singular coefficients. *Sci. China Math.* **61** 1353–1384.
- [ZZ18] Zhang X. and Zhao G. (2018). Singular Brownian Diffusion Processes. *Commun. Math. Stat.* **6** 533–581.
- [ZZ21] Zhang X. and Zhang X. (2021). Cauchy problem of stochastic kinetic equations. Available in arXiv:2103.02267, to appear in *Ann. Appl. Probab.*
- [ZZZ22] Zhang X., Zhu R. and Zhu X. (2022). Singular HJB equations with applications to KPZ on the real line. *Probab. Theory Relat. Fields* **183** 789–869.