SECOND ORDER FRACTIONAL MEAN-FIELD SDES WITH SINGULAR KERNELS AND MEASURE INITIAL DATA

ZIMO HAO, MICHAEL RÖCKNER AND XICHENG ZHANG

ABSTRACT. In this paper we establish the local and global well-posedness of weak and strong solutions to second order fractional mean-field SDEs with singular/distribution interaction kernels and measure initial value, where the kernel can be Newton or Coulomb potential, Riesz potential, Biot-Savart law, etc. Moreover, we also show the stability, smoothness and the short time singularity and large time decay estimates of the distribution density. Our results reveal a phenomenon that for *nonlinear* mean-field equations, the regularity of the initial distribution could balance the singularity of the kernel. The precise relationship between the singularity of kernels and the regularity of initial values are calculated, which belongs to the subcritical regime in the scaling sense. In particular, our results provide a microscopic probabilistic explanation and establish a unified treatment for many physical models such as the fractional Vlasov-Poisson-Fokker-Planck system, the vorticity formulation of 2D-fractal Navier-Stokes equations, surface quasi-geostrophic models, fractional porous medium equation with viscosity, etc.

Keywords: Mean-field SDEs, Anisotropic Besov spaces, Coulomb potential, Riesz potential.

AMS 2010 Mathematics Subject Classification: 60H10, 35H10

CONTENTS

1. Introduction

In this paper we are concerned with the following second order mean-field stochastic differential equation in $\mathbb{R}^d$ (also called McKean-Vlasov or distribution-dependent SDE) driven by standard Brownian motion W_t,

$$\ddot{X}_t = (b_t * \mu_t)(X_t, \dot{X}_t) + \sqrt{2}\dot{W}_t, \quad (X_0, \dot{X}_0) \sim \mu_0, \tag{1.1}$$

or by an α-stable process $L_t^{(\alpha)}$ with $\alpha \in (1,2)$, whose infinitesimal generator is given by the usual fractional Laplacian $\Delta^{\alpha/2}$,

$$\ddot{X}_t = (b_t * \mu_t)(X_t, \dot{X}_t) + \dot{L}_t^{(\alpha)}, \quad (X_0, \dot{X}_0) \sim \mu_0, \tag{1.2}$$

where $\mu_0 \in \mathscr{P}(\mathbb{R}^{2d})$ is the initial probability distribution on $\mathbb{R}^{2d}$, μ_t stands for the joint law of $(X_t, \dot{X}_t)$, and b is a time-dependent $\mathbb{R}^d$-valued (Schwartz) distribution over the phase space $\mathbb{R}^{2d}$, $b_t * \mu_t$ is the usual convolution of two distributions. Here we require that $b * \mu \in L^1_{loc}(\mathbb{R}_+ \times \mathbb{R}^{2d})$. Below for the unification of notations, we shall write

$$L_t^{(2)} = \sqrt{2}W_t.$$

Note that if we introduce the velocity variable $V_t := \dot{X}_t$, then SDEs (1.1) and (1.2) can be written as the following first-order degenerate system:

$$\begin{cases} X_t = X_0 + \displaystyle\int_0^t V_s \mathrm{d}s, \\ V_t = V_0 + \displaystyle\int_0^t (b_s * \mu_s)(X_s, V_s)\mathrm{d}s + L_t^{(\alpha)}. \end{cases} \tag{1.3}$$

It should be noted that if $b(t, x, v) = b(t, v)$ does not depend on the position variable x, then SDE (1.3) reduces to the following first-order nondegenerate SDE

$$\mathrm{d}V_t = (b_t * \mu_t)(V_t) + \mathrm{d}L_t^{(\alpha)}. \tag{1.4}$$

Suppose that $\mu_t(\mathrm{d}x, \mathrm{d}v) = f(t, x, v)\mathrm{d}x\mathrm{d}v$. Then by Itô's formula, f solves the following nonlinear kinetic Fokker-Planck equation in the distributional sense:

$$\partial_t f + v \cdot \nabla_x f = \Delta_v^{\frac{\alpha}{2}} f - \mathrm{div}_v((b * f)f). \tag{1.5}$$

If $b(t, x, v) = b(t, v)$ does not depend on x, then $\rho(t, v) := \int_{\mathbb{R}^d} f(t, x, v)\mathrm{d}x$ solves the following nonlinear and nondegenerate Fokker-Planck equation:

$$\partial_t \rho = \Delta_v^{\frac{\alpha}{2}} \rho - \mathrm{div}_v((b * \rho)\rho). \tag{1.6}$$

The aim of this work is to show the weak and strong well-posedness of SDEs (1.3) and (1.4) with singular b for lowly regular initial values via studying PDEs (1.5) and (1.6). Especially, we find a precise relationship between the singularity of the kernel b and the regularity of μ_0 in terms of their Besov norms, which is almost sharp and belongs to the subcritical regime in the scaling sense. In particular, it allows kernels in (1.6) like $b(x) = \nabla|x|^{s-d}$ for some $s \in (2 - \alpha, d)$ and $\rho_0 \in L^1 \cap L^{p_0}$ for some $p_0 > \frac{d}{s-2+\alpha}$, see Remark 1.10 below.

1.1. **Background and motivation.** Consider the interacting many-body system of N-particles in $\mathbb{R}^d$ governed by Newtonian mechanics and stochastic noises,

$$\begin{cases} \dot{X}_t^{N,i} = V_t^{N,i}, & i = 1, 2, \cdots, N, \\ \dot{V}_t^{N,i} = \dfrac{1}{N} \displaystyle\sum_{j \neq i} K(X_t^{N,i} - X_t^{N,j}) + \sqrt{2\varepsilon}\dot{W}^{N,i}, \end{cases} \tag{1.7}$$

and subjected to i.i.d. initial value $(X_0^{N,i}, V_0^{N,i}) = (\xi_0^i, \eta_0^i)$, whose law is absolutely continuous or singular with respect to the Lebesgue measure, where $(X_t^{N,i}, V_t^{N,i})$ stands for the position and velocity of the ith-particle at time t, $K : \mathbb{R}^d \to \mathbb{R}^d$ models the pairwise interaction between different particles, $(\dot{W}^{N,i})_{i=1,\cdots,N}$ are the generalized derivatives of N independent Brownian motions, which models the collision between particles and a background medium, $\varepsilon > 0$ stands for the intensity of the noise, see [77] for the physical background. In this work we are mainly interested in the interaction kernel $K = A \cdot \nabla U$ (see [40]), where A is some $d \times d$-matrix and

$$U(x) = |x|^{-s}, \quad s \in (0, d) \quad \text{or} \quad U(x) = \ln|x|, \quad d = 1, 2. \tag{1.8}$$

Note that $U(x) = |x|^{2-d}$ for $d \geqslant 3$ or $U(x) = \ln|x|$ for $d = 2$ is the classical Newton or Coulomb kernel, $U(x) = |x|^{1-d}$ for $d \geqslant 2$ is called the Riesz kernel. Since it is hard to directly study the particle system (1.7) by observing the motion of each particle in physical models, one usually considers the empirical measure

$$\mu_t^N(\mathrm{d}x, \mathrm{d}v) := \frac{1}{N} \sum_{i=1}^{N} \delta_{(X_t^{N,i}, V_t^{N,i})}(\mathrm{d}x, \mathrm{d}v).$$

It is assumed that μ_t^N weakly converges to some measure of type $f(t, x, v)\mathrm{d}x\mathrm{d}v$, which is also called mean-field limit or propagation of chaos in Kac's sense (see [50], [40] and [18]), where the density f satisfies the following Vlasov-Poisson-Fokker-Planck equation (abbreviated as VPFP)

$$\partial_t f + v \cdot \nabla_x f = \varepsilon \Delta_v f - \mathrm{div}_v(K * \langle f \rangle f) = \varepsilon \Delta_v f - (K * \langle f \rangle) \cdot \nabla_v f. \tag{1.9}$$

Here $\langle f \rangle(t, x) := \int_{\mathbb{R}^d} f(t, x, v)\mathrm{d}v$ is the mass or charge density. PDE (1.9) can be regarded as a macroscopic description of the microscopic particle system (1.7).

In the three dimensional case, when $\varepsilon = 0$, that is, there is no viscosity in (1.9), and U is the Newton/Coulomb potential, which corresponds to gravitational/electrostatic interaction, PDE (1.9) is also called Vlasov-Poisson equation and describes the evolution of the distribution density of particles in the phase space $\mathbb{R}^6$ and appears in galactic dynamics and plasma physics (see [63] and references therein). No viscosity implies that PDE (1.9) is a time-reversible system in the sense that if f is a solution of (1.9), then $\tilde{f}(t, x, v) := f(-t, x, -v)$ is also a solution of (1.9). The existence of global smooth solutions to the Vlasov-Poisson equation was shown in [59], [66], [70] under some moment assumptions on the initial value. The existence of classical solutions to the VPFP equation was obtained in [10], [14], [64] by the same method as in [59]. All the results require that the initial value $f_0 \in L^1 \cap L^\infty$ has a finite second order moment in v, that is,

$$\int_{\mathbb{R}^{2d}} |v|^2 f_0(x, v)\mathrm{d}x\mathrm{d}v < \infty. \tag{1.10}$$

When the initial distribution f_0 is a measure, say in some Morrey space, and satisfies some smallness assumptions, the well-posedness of VPFP (1.9) was established in [14] (see the earlier work [86] for one dimensional case). On the other hand, from the physical viewpoint, the mean-field limit provides an explanation for Vlasov-Poisson equations from Newton's second law for the motion of N particles. Such a problem has been studied extensively in the context of PDEs (see [50] and [40]). When K is Lipschitz continuous, the propagation of chaos for particle system (1.7) was firstly shown by McKean [61]. When the interaction kernel K is bounded measurable, Jabin and Wang [51] showed the quantitative rate of the propagation of chaos via the entropy method. Based on the BBGKY method, Lacker [57] showed the optimal rate of convergence by the Girsanov

transform and deep analysis. For singular potentials U, there is also some cutoff version about the mean-field limits, see [47] and references therein. A complete review of mean field limits for Vlasov-Poisson equations is contained in [50] and [40].

From the probabilistic view point, it is natural to expect that the distributional density solution f of VPFP (1.9) is the law of the solution $Z_t = (X_t, V_t)$ to the following mean-field SDE, if a solution exists in Itô's sense:

$$\dot{X}_t = V_t, \quad \dot{V}_t = \int_{\mathbb{R}^d} K(X_t - y)\mu_{X_t}(\mathrm{d}y) + \sqrt{2\varepsilon}\dot{W}_t, \tag{1.11}$$

where μ_{X_t} stands for the probability distribution measure of X_t. As discussed above, although there are a lot of references about the solvability of PDE (1.9), there are few results about SDE (1.11), e.g., about its weak/strong well-posedness in the probabilistic sense. The difficulty of solving (1.11) lies in that DDSDE (1.11) is degenerate and the drift is too singular for $K = A \cdot \nabla U$ with U given in (1.8).

Besides the second order particle system (1.7), we are also interested in the following first order system that is related to the point vortex system and other physical models:

$$\dot{X}_t^{N,i} = \frac{1}{N} \sum_{j \neq i} K(X_t^{N,i} - X_t^{N,j}) + \sqrt{2\varepsilon}\dot{W}_t^{N,i}. \tag{1.12}$$

As above, it is usually assumed that the limit of the empirical measure $\mu_t^N(\mathrm{d}x) := \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{N,i}}(\mathrm{d}x)$ solves the following Fokker-Planck equation:

$$\partial_t \rho = \varepsilon\Delta\rho - \mathrm{div}((K * \rho)\rho), \tag{1.13}$$

where ρ is the distribution density of the solution to the following McKean-Vlasov SDE:

$$X_t = X_0 + \int_0^t (K * \mu_{X_s})(X_s)\mathrm{d}s + \sqrt{2\varepsilon}W_t. \tag{1.14}$$

For the above first order particle system with $K = A \cdot \nabla U$ as described in (1.8), Serfaty [72] established the conditional mean field convergence by modulated free energy. On the two dimensional torus, when K is the Biot-Savart law, Jabin and Wang [52] showed some quantitative rate of convergence of propagation of chaos for the point vortex system by the entropy method. When the interaction kernel is in L^p-space with $p > d$, the strong convergence of propagation of chaos was obtained recently in [43]. The purpose of the current work is not to prove the propagation of chaos, but concentrated on the solvability of SDEs (1.11) (and hence the corresponding nonlinear Fokker-Planck equations (1.9) and (1.13) respectively) and (1.14) with singular interaction kernels.

Before going further, we introduce more examples from physics for equation (1.13).

Example 1. Consider the following vorticity form of the 2D Navier-Stokes equation:

$$\partial_t w = \Delta w + (K_2 * w) \cdot \nabla w = \Delta w + \mathrm{div}((K_2 * w)w), \tag{1.15}$$

where K_2 is the Biot-Savart law given by

$$K_2(x_1, x_2) = (-x_2, x_1)/|x|^2.$$

If the initial vorticity w_0 is a finite signed measure, the existence of weak solutions to equation (1.15) was proved by Cottet [31] and Giga, Miyakawa and Osada [38]. The uniqueness is a more difficult problem and was partially proved in [38], [37]and then completely solved by Gallagher and Gallay [36]. The proofs in [38] and [36] strongly depend on the heat kernel estimates of $\partial_t - \Delta - U \cdot \nabla$ established by Osada in [65], where $U : \mathbb{R}_+ \times \mathbb{R}^2 \to \mathbb{R}^2$ is a divergence free vector field which satisfies

$$\|U(t)\|_{L^\infty(\mathbb{R}^2)} \leqslant C/\sqrt{t}, \quad \|\partial_2 U_1(t) - \partial_1 U_2(t)\|_{L^1(\mathbb{R}^2)} \leqslant C.$$

For the classical Navier-Stokes equation of velocity fields, there are also numerous papers about the initial velocity being in some homogeneous spaces (see [54], [39], [55], [13], [56]). It should be noted that there is an essential discrepancy between the vorticity formulation and the original

two-dimensional Navier-Stokes equations because if $u \in L^2(\mathbb{R}^2)$ and $w = \partial_2 u_1 - \partial_1 u_2 \in L^1(\mathbb{R}^2)$, then $\int_{\mathbb{R}^2} w dx = 0$. One of our main results shows that when the initial vorticity is in $\cup_{\varepsilon > 0} \mathbf{B}_\infty^{-3/2+\varepsilon}$, then there is a unique smooth solution to equation (1.15). In particular, it includes the white noise in $\mathbb{R}^2$ as the initial vorticity, which seems to be new.

Example 2. Consider the following surface quasi-geostrophic equation in $\mathbb{R}^2$ (abbreviated by SQG): for $\alpha \in (0,2)$,

$$\partial_t \theta = \Delta^{\frac{\alpha}{2}} \theta + (\mathscr{R} * \theta) \cdot \nabla \theta, \tag{1.16}$$

where $\mathscr{R} = (-\mathscr{R}_2, \mathscr{R}_1)$ and $\mathscr{R}_i$ is the Riesz transform

$$\mathscr{R}_i(x) = x_i/|x|^3, \ x \in \mathbb{R}^2, \ i = 1, 2.$$

The SQG equation appeared in the study of the mathematical theory of meteorology and oceanography [29], [46]. It is classified in subcritical/critical/supercritical cases according to the fractional power $\alpha \in (1,2)/\alpha = 1/\alpha \in (0,1)$. In [30], Constantin and Wu showed that in the subcritical case any solution with smooth initial value is smooth for all time. In both the critical case and supercritical cases, Chae and Lee [17] showed the well-posedness of solutions with initial conditions small in Besov spaces (see also [80]). In the critical case, Cafferalli and Vaseur [12] showed the global well-posedness for SQG (1.16) for large initial values. All these results require that the initial value is in L^2. We are interested in the well-posedness of (1.16) when the initial data are measures or certain Schwartz distributions.

Example 3. Let $d \geqslant 3$, $s \in [0,1]$ and $\varepsilon \geqslant 0$. Consider the following fractional porous medium equation in $\mathbb{R}^d$ with viscosity:

$$\partial_t u = \varepsilon \Delta u + \mathrm{div}((\nabla \Delta^{-s} u) u).$$

For $s \in (0,1]$, it is well known that the nonlocal operator $\Delta^{-s} u$ can be written as (see [73])

$$\Delta^{-s} u = K * u, \quad K(x) := c_{d,s} |x|^{2s-d},$$

where $c_{d,s} > 0$ is a normalizing constant. Thus, if we take

$$U(x) = |x|^{2s-d} \mathbb{1}_{s \in (0,1]} + \delta_0(\mathrm{d}x) \mathbb{1}_{s=0},$$

then for $s \in [0,1]$,

$$\nabla \Delta^{-s} u = \nabla U * u.$$

For $s = 0$ and $\varepsilon = 0$, it is the classical porous medium equation. For $s = 1$ and $\varepsilon = 0$, it was investigated in [19] and [34] for the evolution of the vortex density in a superconductor. For $s \in (0,1)$ and $\varepsilon = 0$, it is called fractional porous medium equation and was studied in [9] and [76]. For $\varepsilon > 0$ and $s \in [0,1]$, it was considered as the viscosity approximation of fractional porous medium equation in [71] and [25]. In particular, our results can be applied to the above equation for $\varepsilon > 0$. In fact, we shall consider a kinetic version, which seems not physically relevant, but appears interesting for mathematics.

1.2. **Main results.** In this section we provide explicit conditions on b and initial conditions so that the mean-field SDE (1.3) admits a unique weak/strong solution in the probabilistic sense.

To state our main results, we first introduce some notations used in this paper. For a muti-index $\boldsymbol{p} = (p_x, p_v) \in [1, \infty]^2$, we define

$$\|f\|_{\mathbb{L}^{\boldsymbol{p}}} := \|f\|_{\boldsymbol{p}} := \left(\int_{\mathbb{R}^d} \|f(\cdot, v)\|_{p_x}^{p_v} \mathrm{d}v \right)^{1/p_v}, \tag{1.17}$$

and introduce the scaling parameter $\boldsymbol{a}$ by

$$\boldsymbol{a} := (1 + \alpha, 1), \quad \alpha \in (0, 2).$$

For notational convenience, we shall write below

$$\frac{1}{\boldsymbol{p}} := \left(\frac{1}{p_x}, \frac{1}{p_v}\right), \quad \boldsymbol{a} \cdot \frac{d}{\boldsymbol{p}} := \frac{(1+\alpha)d}{p_x} + \frac{d}{p_v},$$

and for any $\boldsymbol{p}, \boldsymbol{q} \in [1, \infty]^2$,

$$\boldsymbol{p} \geqslant \boldsymbol{q} \Leftrightarrow p_x \geqslant q_x, \ p_v \geqslant q_v.$$

Moreover, we use bold letters to denote constant vectors in $\mathbb{R}^2$, for example,

$$\mathbf{1} := (1, 1), \quad \boldsymbol{d} := (d, d).$$

We remark that the above mixed $\mathbb{L}^{\boldsymbol{p}}$-norm was first introduced in [7]. Recently, it was used in an essential way in [43] to study the strong convergence of propagation of chaos for particle systems with singular interactions (see also [58], [84]). We emphasize that the order of (p_x, p_v) in definition (1.17) is important because the above $\mathbb{L}^{\boldsymbol{p}}$-norm is invariant under the group action $\Gamma_t f(x, v) = f(x - tv, v)$ associated with $\partial_t + v \cdot \nabla_v$. More precisely, for any $\boldsymbol{p} \in [1, \infty]^2$ and $f \in \mathbb{L}^{\boldsymbol{p}}$,

$$\|\Gamma_t f\|_{\boldsymbol{p}} = \left(\int_{\mathbb{R}^d} \|f(\cdot - tv, v)\|_{p_x}^{p_v} \mathrm{d}v\right)^{1/p_v} = \left(\int_{\mathbb{R}^d} \|f(\cdot, v)\|_{p_x}^{p_v} \mathrm{d}v\right)^{1/p_v} = \|f\|_{\boldsymbol{p}}. \tag{1.18}$$

Now if we let $Z_t := (X_t, V_t)$, then SDE (1.3) can be written in the following compact form:

$$Z_t = Z_0 + \int_0^t \mathbf{H}_s(Z_s)\mathrm{d}s + \sigma L_t^{(\alpha)}, \tag{1.19}$$

where $\sigma := (0, \mathbb{I})^* \in \mathbb{R}^{2d} \otimes \mathbb{R}^d$ and for $z = (x, v)$,

$$\mathbf{H}_t(z) := (v, H_t(z))^* \in \mathbb{R}^{2d}, \quad H_t(z) := (b_t * \mu_t)(z). \tag{1.20}$$

We first give the definitions of weak/strong solutions to mean-field SDEs (1.19).

Definition 1.1. *Let $\alpha \in (1, 2]$ and $\mathscr{G} := (\Omega, \mathscr{F}, \mathbf{P}; (\mathscr{F}_t)_{t \geqslant 0})$ be a stochastic basis. Let $Z := (Z_t)_{t \geqslant 0}$ and $L^{(\alpha)} := (L_t^{(\alpha)})_{t \geqslant 0}$ be two $\mathbb{R}^{2d}$ and $\mathbb{R}^d$-valued cádlág $\mathscr{F}_t$-adapted processes.*

(A) (Weak solutions) *For given distribution $\mu_0 \in \mathscr{P}(\mathbb{R}^{2d})$, we call $(Z, L^{(\alpha)}, \mathscr{G})$ a weak solution of SDE (1.19) with initial distribution μ_0 if*
  (i) $L^{(\alpha)}$ is an α-stable process with generator $\Delta^{\alpha/2}$ and $\mathbf{P} \circ Z_0^{-1} = \mu_0$.
  (ii) For each $t > 0$, $\mathbf{P} \circ Z_t^{-1}$ admits a C^∞ density $f(t, z)$ so that

$$\int_0^T \sup_{z \in \mathbb{R}^{2d}} |H_s(z)| \mathrm{d}s < \infty, \quad H_t(z) := (b_t * f(t))(z). \tag{1.21}$$

  (iii) For all $t > 0$

$$Z_t = Z_0 + \int_0^t \mathbf{H}_s(Z_s)\mathrm{d}s + \sigma L_t^{(\alpha)}, \quad \mathbf{P} - a.s.$$

(B) (Strong solutions) *Let $L^{(\alpha)}$ be an α-stable process with generator $\Delta^{\alpha/2}$ and*

$$\mathscr{F}_t^0 := \mathscr{F}_0 \vee \sigma\{L_s^{(\alpha)}, s \leqslant t\}.$$

  For given $Z_0 \in \mathscr{F}_0^0$, we call a $\mathbb{R}^{2d}$-value cádlág $\mathscr{F}_t^0$-adapted process $Z := (Z_t)_{t \geqslant 0}$ a strong solution of SDE (1.19) starting from Z_0 if (ii) and (iii) above hold.

Remark 1.2. The above definitions are standard (for example, see [53] and [35]), where we used the convention that the 2-stable process with generator Δ is just a Brownian motion with variance $2t$. We would like to mention that condition (1.21) guarantees the well-definedness of $\int_0^t \mathbf{H}_s(Z_s)\mathrm{d}s$.

Our first result is the following weak/strong well-posedness about the second order mean-field SDE (1.3) or (1.19).

Theorem 1.3. *Let* $\alpha \in (1, 2]$, $q_b \in (\frac{\alpha}{\alpha-1}, \infty]$ *and* $\boldsymbol{p}_0, \boldsymbol{\varrho}_0, \boldsymbol{\varrho}_1 \in [1, \infty]^2$ *with* $\mathbf{1} \leqslant \frac{1}{\boldsymbol{p}_0} + \frac{1}{\boldsymbol{\varrho}_0}$ *and* $\boldsymbol{\varrho}_0 \leqslant \boldsymbol{\varrho}_1$. *Let* $\bar{\boldsymbol{\varrho}}_1 \in [1, \infty]^2$ *with* $\frac{1}{\boldsymbol{\varrho}_1} + \frac{1}{\bar{\boldsymbol{\varrho}}_1} = \mathbf{1}$ *and* $\boldsymbol{p}_1 := \boldsymbol{p}_0 \wedge \bar{\boldsymbol{\varrho}}_1$. *Define*

$$\mathscr{A}_i := \boldsymbol{a} \cdot (\tfrac{d}{\boldsymbol{p}_i} + \tfrac{d}{\boldsymbol{\varrho}_i} - \boldsymbol{d}), \ i = 0, 1, \quad \mathbb{Z}_\beta := \mathbf{B}_{\boldsymbol{p}_0; \boldsymbol{a}}^\beta \cap \mathbf{B}_{\boldsymbol{p}_1; \boldsymbol{a}}^\beta, \ \beta \in \mathbb{R},$$

where the anisotropic Besov space $\mathbf{B}_{\boldsymbol{p}; \boldsymbol{a}}^\beta := \mathbf{B}_{\boldsymbol{p}; \boldsymbol{a}}^{\beta; \infty}$ *is defined in Definition 2.2 below.*

(A) (Weak well-posedness) *Suppose that for some* $\beta_0 \in (\frac{\alpha}{q_b} - \alpha, 0)$ *and* $\mathbb{R} \ni \beta_b \neq \mathscr{A}_i - \beta_0$,

$$0 < \mathscr{A}_0 - \beta_0 - \beta_b < \alpha - \tfrac{\alpha}{q_b} + \beta_0 \wedge (-1), \tag{1.22}$$

and $b = b_0 + b_1$, *where for some* $\kappa_b > 0$,

$$\kappa_b := \|b_0\|_{L^{q_b}(\mathbb{R}_+; \mathbf{B}_{\boldsymbol{\varrho}_0; \boldsymbol{a}}^{\beta_b})} + \|b_1\|_{L^{q_b}(\mathbb{R}_+; \mathbf{B}_{\boldsymbol{\varrho}_1; \boldsymbol{a}}^{\beta_b})} < \infty. \tag{1.23}$$

For any $\gamma_b > (\alpha - \alpha/q_b - 1 - \mathscr{A}_1) \vee 0$, *there is a constant* $C_0 = C_0(\alpha, d, \beta_0, \boldsymbol{p}_0, \beta_b, q_b, \gamma_b, \boldsymbol{\varrho}_0, \boldsymbol{\varrho}_1) > 0$ *such that for any* $T \geqslant 1$ *and initial distribution* $\mu_0 \in \mathbb{Z}_{\beta_0}$ *satisfying*

$$\|\mu_0\|_{\mathbb{Z}_{\beta_0}} \leqslant C_0 T^{-\frac{\gamma_b}{\alpha}} / \kappa_b, \tag{1.24}$$

there exists a unique weak solution Z *to mean field SDE* (1.19) *on* $[0, T]$ *in the sense of Definition 1.1 so that* $H \in \mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)$ *for some* $q \in (\frac{\alpha}{\alpha-1}, \infty)$ *and* $\beta > 0$, *where* H *is defined by* (1.20) *and* $\mathbf{C}_{\boldsymbol{a}}^\beta := \mathbf{B}_{\infty; \boldsymbol{a}}^\beta$ *is the anisotropic Hölder space.*

(B) (Strong well-posedness) *Suppose that for some* $\beta_0 \in (\frac{\alpha}{q_b} - \alpha, 0)$ *and* $\mathbb{R} \ni \beta_b \neq \mathscr{A}_i - \beta_0$,

$$0 < \mathscr{A}_0 - \beta_0 - \beta_b < \tfrac{3}{2}\alpha - \tfrac{\alpha}{q_b} - 2, \tag{1.25}$$

and $b = b_0 + b_1$, *where*

$$b_i \in L^{q_b}(\mathbb{R}_+; \mathbf{B}_{\boldsymbol{\varrho}_i; x, \boldsymbol{a}}^{s_0, s_1} \cap \mathbf{B}_{\boldsymbol{\varrho}_i; \boldsymbol{a}}^{\beta_b}), \quad i = 0, 1, \tag{1.26}$$

and $s_0 = 1 + \frac{\alpha}{2}$, $s_1 = \frac{\alpha}{2} - 1 + \beta_b$ *and the mixed Besov space* $\mathbf{B}_{\boldsymbol{\varrho}_i; x, \boldsymbol{a}}^{s_0, s_1}$ *is defined in Definition 2.2 below. Let* μ_0 *be the law of the initial random variable* Z_0. *Then, under* (1.24), *for any* $\gamma_b > (\alpha - \alpha/q_b - 1 - \mathscr{A}_1) \vee 0$ *and* $T > 0$, *there is a unique strong solution* Z *to mean field SDE* (1.19) *on* $[0, T]$ *in the sense of Definition 1.1 so that* $H \in \mathbb{L}_T^q(\mathbf{C}_x^\gamma \cap \mathbf{C}_{\boldsymbol{a}}^\beta)$ *for some* $q \in (\frac{\alpha}{\alpha-1}, \infty)$ *and* $\gamma > \frac{1+\alpha/2}{1+\alpha}$, $\beta > 1 - \frac{\alpha}{2}$, *where* $\mathbf{C}_x^\gamma$ *is the Hölder space along the x-direction.*

Moreover, we have:

(i) **(Smoothness of density and short time singularity estimates)** *For any* $t \in (0, T]$, Z_t *has a smooth density* f *with short time regularity*

$$\sup_{t \in (0, T]} \left(t^{\frac{\beta}{\alpha}} \|f(t)\|_{\mathbb{Z}_{\beta+\beta_0}} \right) < \infty, \quad \text{for any } \beta \geqslant 0.$$

(ii) **(Large time decay estimate)** *Suppose* $\boldsymbol{a} \cdot \frac{d}{\boldsymbol{\varrho}_1} > \alpha - \frac{\alpha}{q_b} - 1$ *and* $\boldsymbol{p}_0 \neq \mathbf{1}$. *Then we can take* $\gamma_b = 0$ *in* (1.24), *and there is a global weak solution to mean-field SDE* (1.19) *so that for any* $\beta \geqslant 0$ *and* $i = 0, 1$,

$$\sup_{t > 1} \left(t^{\frac{\boldsymbol{a} \cdot (\boldsymbol{d} - d/\boldsymbol{p}_i)}{\alpha}} \|f(t)\|_{\mathbf{B}_{\boldsymbol{p}_i; \boldsymbol{a}}^{\beta, 1}} \right) < \infty.$$

(iii) **(Weak stability)** *Suppose that the assumptions in* **(A)** *hold. Let* $\widetilde{b}$ *be another kernel satisfying* (1.23) *and* $\widetilde{\mu}_0 \in \mathbb{Z}_{\beta_0}$ *another initial distribution that satisfies* (1.24). *Let* $\widetilde{f}$ *be the distribution density of the solution* $\widetilde{Z}$ *corresponding to* $\widetilde{b}$ *and* $\widetilde{\mu}_0$. *Then for any* $\beta \geqslant 0$,

$$\sup_{t \in (0, T]} \left(t^{\frac{\beta}{\alpha}} \|f(t) - \widetilde{f}(t)\|_{\mathbb{Z}_{\beta+\beta_0}} \right) \lesssim_{C_\beta} \sum_{i=0, 1} \|b_i - \widetilde{b}_i\|_{\mathbb{L}_T^{q_b}(\mathbf{B}_{\boldsymbol{\varrho}_i; \boldsymbol{a}}^{\beta_b})} + \|\mu_0 - \widetilde{\mu}_0\|_{\mathbb{Z}_{\beta_0}}, \tag{1.27}$$

where $C_\beta = C_\beta(\alpha, d, \beta_0, \boldsymbol{p}_0, \beta_b, q_b, \boldsymbol{\varrho}_0, \boldsymbol{\varrho}_1, T, \kappa_b) > 0$.

(iv) **(Strong stability)** *Suppose that the assumptions in* **(B)** *hold. Let* $\widetilde{b}$ *be another kernel satisfying* (1.26) *and* $\widetilde{\mu}_0 \in \mathbb{Z}_{\beta_0}$ *another initial distribution that satisfies* (1.24). *Let* $\widetilde{Z}$ *be the strong solution of SDE* (1.19) *corresponding to* $\widetilde{b}$ *and* $\widetilde{Z}_0 \sim \widetilde{\mu}_0$. *Then*

$$\mathbb{E}\left(\sup_{t\in(0,T]}|Z_t - \widetilde{Z}_t|^2\right) \lesssim_C \mathbb{E}|Z_0 - \widetilde{Z}_0|^2 + \sum_{i=0,1}\|b_i - \widetilde{b}_i\|^2_{\mathbb{L}_T^{q_b}(\mathbf{B}^{\beta_b}_{\boldsymbol{\varrho}_i;\boldsymbol{a}})} + \|\mu_0 - \widetilde{\mu}_0\|^2_{\mathbb{Z}_{\beta_0}}, \qquad (1.28)$$

where $C = C(\alpha, d, \beta_0, \boldsymbol{p}_0, \beta_b, q_b, \boldsymbol{\varrho}_0, \boldsymbol{\varrho}_1, T, \kappa_b) > 0$.

(v) **(Global existence)** *If* $\boldsymbol{p}_0 = \mathbf{1}$ *or* $(\mathrm{div}_v b)^- = 0$ *and* $\boldsymbol{p}_0 = (p_0, p_0)$, $\boldsymbol{\varrho}_1 = (\varrho_1, \varrho_1)$, *then the smallness condition* (1.24) *can be dropped.*

Remark 1.4. If $\boldsymbol{\varrho}_0 = \boldsymbol{\varrho}_1$, then $\boldsymbol{p}_0 = \boldsymbol{p}_1$ since $\boldsymbol{p}_0 \leqslant \bar{\boldsymbol{\varrho}}_1$ by $\mathbf{1} \leqslant \frac{1}{\boldsymbol{p}_0} + \frac{1}{\boldsymbol{\varrho}_0}$. Moreover, $(\mathrm{div}_v b)^- = 0$ means that for any nonnegative $g \in C_c^\infty(\mathbb{R}^{2d})$,

$$\int_{\mathbb{R}^{2d}} \langle b(t, z), \nabla g(z)\rangle \mathrm{d}z \leqslant 0, \quad \text{for a.e. } t \in \mathbb{R}_+.$$

Remark 1.5. Let $\mu_0 \in \mathscr{P}(\mathbb{R}^{2d}) \cap \mathbb{Z}_{\beta_0}$. Define for $\varepsilon \in (0, 1)$,

$$\mu_0^\varepsilon(\mathrm{d}x, \mathrm{d}v) := \mu_0(\varepsilon^{1+\alpha}\mathrm{d}x, \varepsilon\mathrm{d}v).$$

Then $\mu_0^\varepsilon \in \mathscr{P}(\mathbb{R}^{2d}) \cap \mathbb{Z}_{\beta_0}$ and by (2.10) below, we have

$$\|\mu_0^\varepsilon\|_{\mathbf{B}^{\beta_0}_{\boldsymbol{p}_i;\boldsymbol{a}}} \leqslant 2\varepsilon^{\beta_0 + \boldsymbol{a}\cdot(\boldsymbol{d} - \frac{d}{\boldsymbol{p}_i})}\|\mu_0\|_{\mathbf{B}^{\beta_0}_{\boldsymbol{p}_i;\boldsymbol{a}}}, \quad i = 0, 1.$$

From this, one sees that (1.24) holds for μ_0^ε as long as $\beta_0 + \boldsymbol{a}\cdot(\boldsymbol{d} - \frac{d}{\boldsymbol{p}_i}) > 0$ and ε is small enough.

Remark 1.6. The condition $\beta_0 \in \{0, \mathscr{A}_i - \beta_b, i = 0, 1\}$ are purely technical and can be dropped since we can choose different $\beta_0' < \beta_0$ so that (1.22) still holds. This condition comes from the following Besov estimate: for $\boldsymbol{p} \leqslant \boldsymbol{p}'$ and $\beta' \neq \beta - \mathscr{A}$ with $\mathscr{A} := \boldsymbol{a}\cdot(\frac{\boldsymbol{d}}{\boldsymbol{p}} - \frac{\boldsymbol{d}}{\boldsymbol{p}'})$,

$$\|P_t f\|_{\mathbf{B}^{\beta',1}_{\boldsymbol{p}';\boldsymbol{a}}} \lesssim_C (1 \wedge t)^{-\frac{(\beta' - \beta + \mathscr{A})\vee 0}{\alpha}}\|f\|_{\mathbf{B}^{\beta,\infty}_{\boldsymbol{p};\boldsymbol{a}}}, \quad t > 0,$$

where P_t is the kinetic semigroup (see Lemma 2.16 below). Note that this estimate is stronger than the usual estimate $\|\nabla_v P_t f\|_{\boldsymbol{p}} \leqslant C(1 \wedge t)^{-\frac{1}{\alpha}}\|f\|_{\boldsymbol{p}}$. Indeed,

$$\|\nabla_v P_t f\|_{\boldsymbol{p}} \leqslant C\|P_t f\|_{\mathbf{B}^{1,1}_{\boldsymbol{p};\boldsymbol{a}}} \leqslant C(1 \wedge t)^{-\frac{1}{\alpha}}\|f\|_{\mathbf{B}^{0,\infty}_{\boldsymbol{p};\boldsymbol{a}}} \leqslant C(1 \wedge t)^{-\frac{1}{\alpha}}\|f\|_{\boldsymbol{p}}.$$

The above estimate plays a crucial role in our proofs.

Remark 1.7. Let $b : \mathbb{R}_+ \times \mathbb{R}^{2d} \times \mathbb{R}^{2d}$ be a measurable function with $|b(t, z, z')| \leqslant K_t(z - z')$, where K satisfies (1.23) with $-1 < \beta_b < 0$. If we replace $b * \mu$ by

$$(b \circledast \mu)(z) := \int_{\mathbb{R}^{2d}} b(z, z')\mu(\mathrm{d}z'),$$

then the weak well-posedness still holds (see Remark 4.9 below). In this case, it is of course not expected that (1.19) admits a smooth density except for the case where that $z \mapsto b(t, z, z')$ is smooth.

Remark 1.8. The stability estimates (1.27) and (1.28) provide us the possibility of showing the propagation of chaos for the cutoff kernel b. For instance, consider the kernel $K(x) = \nabla|x|^{s-d}$ and the cutoff kernel $K_N(x) := \nabla(|x| \vee N^{-\delta})^{s-d}$, where $s \in (0, d)$ and $\delta > 0$. For any $\varrho \in (\frac{d}{d-s}, \infty]$ and $\beta < \frac{d}{p} - d + s$, by (A.6) in Appendix A, one sees that

$$\|K - K_N\|_{\mathbf{B}^\beta_\varrho} \leqslant N^{-\delta(\frac{d}{p} - d - \beta + s)}.$$

Thus one can consider the regularized particle system (1.7) with $K = K_N$ (see [47]).

Now we sketch the main ingredient of the proof of Theorem 1.3. Let P_t be the kinetic semigroup associated with $\Delta_v^{\alpha/2} - v \cdot \nabla_x$ (see (2.22) for a definition). First of all we need to solve FPE (1.5) and obtain some precise asymptotic estimates in anisotropic Besov spaces. More precisely, by Duhamel's principle,

$$f(t) = P_t f_0 + \mathcal{H}_t^b(f, f),$$

where

$$\mathcal{H}_t^b(f, g) = -\int_0^t P_{t-s} \mathrm{div}_v((b_s * f_s)g_s)\mathrm{d}s.$$

The key point is to obtain a priori estimates for the bilinear form $(f, g) \mapsto \mathcal{H}^b(f, g)$ in suitable spaces, and then use the contraction mapping theorem. To achieve this goal, we first establish Schauder estimates for the non-homogenous kinetic equation $\partial_t u = (\Delta_v^{\alpha/2} - v \cdot \nabla_x)u + f$ in time-weighted spaces (see Theorem 2.18 below). To show the smoothness and large time decay estimate for the solution, by the semigroup property of P_t and shifting the time, the following simple observation is important (see Lemma 3.9 below):

$$\mathcal{H}_{t+s}^b(f, g) = P_s \mathcal{H}_t^b(f, g) + \mathcal{H}_s^{b^t}(f^t, g^t), \quad s, t > 0,$$

where $f^t(\cdot, x, v) = f(t + \cdot, x, v)$. After establishing the solvability of (1.5), to prove the weak well-posedness of McKean-Vlasov SDE (1.19), we shall freeze the distribution-dependent drift along the smooth solution of the nonlinear Fokker-Planck equation, and then utilize the uniqueness of linear Fokker-Planck equations. For the strong well-posedness, we use Zvonkin's transformation to kill the singular drift, where the freezed drift is Hölder continuous in the spatial variable, but has a singularity around the starting point with respect to the time variable.

As a corollary, we also consider the following first order non-degenerate mean-field SDE

$$\mathrm{d}X_t = (b_t * \mu_{X_t})(X_t) + \mathrm{d}L_t^{(\alpha)}, \ \alpha \in (1, 2]. \tag{1.29}$$

We have the following weak/strong well-posedness result.

Theorem 1.9. *Let* $\alpha \in (1, 2]$, $q_b \in (\frac{\alpha}{\alpha-1}, \infty]$ *and* $p_0, \varrho_0, \varrho_1 \in [1, \infty]$ *with* $1 \leqslant \frac{1}{p_0} + \frac{1}{\varrho_0}$ *and* $\varrho_0 \leqslant \varrho_1$. *Let* $\bar{\varrho}_1 \in [1, \infty]$ *with* $\frac{1}{\varrho_1} + \frac{1}{\bar{\varrho}_1} = 1$ *and* $p_1 := p_0 \wedge \bar{\varrho}_1$. *Define*

$$\mathscr{A}_i := \frac{d}{p_i} + \frac{d}{\varrho_i} - d, \ i = 0, 1, \quad \mathbb{Z}_\beta := \mathbf{B}_{p_0}^\beta \cap \mathbf{B}_{p_1}^\beta, \ \beta \in \mathbb{R},$$

where $\mathbf{B}_p^\beta = \mathbf{B}_p^{\beta,\infty}$ *is the isotropic Besov space defined in Definition 2.2 below.*
(A) (Weak well-posedness) *Suppose that for some* $\beta_0 \in (\frac{\alpha}{q_b} - \alpha, 0)$ *and* $\mathbb{R} \ni \beta_b \neq \mathscr{A}_i - \beta_0$,

$$0 < \mathscr{A}_0 - \beta_0 - \beta_b < \alpha - \frac{\alpha}{q_b} + \beta_0 \wedge (-1), \tag{1.30}$$

and $b = b_0 + b_1$, *where*

$$\kappa_b := \|b_0\|_{L^{q_b}(\mathbb{R}_+; \mathbf{B}_{\varrho_0}^{\beta_b})} + \|b_1\|_{L^{q_b}(\mathbb{R}_+; \mathbf{B}_{\varrho_1}^{\beta_b})} < \infty. \tag{1.31}$$

For any $\gamma_b > (\alpha - \alpha/q_b - 1 - \mathscr{A}_1) \vee 0$, *there is a constant* $C_0 = (\alpha, d, \beta_0, p_0, \beta_b, q_b, \varrho_0, \varrho_1) > 0$ *such that for any* $T \geqslant 1$ *and initial distribution* $\mu_0 \in \mathbb{Z}_{\beta_0}$ *satisfying*

$$\|\mu_0\|_{\mathbb{Z}_{\beta_0}} \leqslant C_0 T^{-\frac{\gamma_b}{\alpha}}/\kappa_b, \tag{1.32}$$

there exists a unique weak solution X *to mean field SDE* (1.29) *on* $[0, T]$ *in the sense of Definition 1.1 so that* $H \in \mathbb{L}_T^q(\mathbf{C}^\beta)$ *for some* $q \in (\frac{\alpha}{\alpha-1}, \infty)$ *and* $\beta > 0$, *where* $H_t(x) := (b_t * \mu_{X_t})(x)$ *and* $\mathbf{C}^\beta := \mathbf{B}_\infty^\beta$ *is the isotropic Hölder space.*
(B) (Strong well-posedness) *Suppose that for some* $\beta_0 \in (\frac{\alpha}{q_b} - \alpha, 0)$ *and* $\mathbb{R} \ni \beta_b \neq \mathscr{A}_i - \beta_0$,

$$0 < \mathscr{A}_0 - \beta_0 - \beta_b < \frac{3}{2}\alpha - \frac{\alpha}{q_b} - 2. \tag{1.33}$$

For any $\gamma_b > (\alpha - \alpha/q_b - 1 - \mathscr{A}_1) \vee 0$ *and* $T \geqslant 1$, *under* (1.31) *and* (1.32), *there is a unique strong solution* X *to mean field SDE* (1.29) *on* $[0, T]$ *in the sense of Definition 1.1 so that* $H \in \mathbb{L}_T^q(\mathbf{C}^\beta)$ *for some* $q \in (\frac{\alpha}{\alpha-1}, \infty)$ *and* $\beta > 1 - \frac{\alpha}{2}$, *where* $H_t(x) := (b_t * \mu_{X_t})(x)$

Moreover, we have the following:

(i) **(Smoothness of density and short time singularity estimates)** *For any $t \in (0, T]$, Z_t has a smooth density $\rho(t)$ with regularity that for any $\beta \geqslant 0$,*

$$\sup_{t \in (0,T]} \left(t^{\frac{\beta}{\alpha}} \|\rho(t)\|_{\mathbb{Z}_{\beta+\beta_0}} \right) < \infty.$$

(ii) **(Large time decay estimate)** *Suppose $\frac{d}{\varrho_1} > \alpha - \frac{\alpha}{q_b} - 1$ and $p_0 \neq 1$. Then we can take $\gamma_b = 0$ in (1.32), and there is a global weak solution to mean-field SDE (1.29) so that for any $\beta \geqslant 0$ and $i = 0, 1$,*

$$\sup_{t > 1} \left(t^{\frac{d-d/p_i}{\alpha}} \|\rho(t)\|_{\mathbf{B}_{p_i}^{\beta,1}} \right) < \infty.$$

(iii) **(Two-sided estimates)** *For any $T > 0$, there are constants $C_1, C_2, C_3 > 0$ such that for all $t \in (0, T) \times \mathbb{R}^d$,*

$$\frac{C_1^{-1}}{t^{d/2}} \int_{\mathbb{R}^d} e^{-\frac{|x-y|^2}{C_2 t}} \mu_0(\mathrm{d}x) \leqslant \rho(t, y) \leqslant \frac{C_1}{t^{d/2}} \int_{\mathbb{R}^d} e^{-\frac{C_2|x-y|^2}{t}} \mu_0(\mathrm{d}x), \quad \alpha = 2,$$

and

$$\int_{\mathbb{R}^d} \frac{C_3^{-1} t \mu_0(\mathrm{d}x)}{(t^{1/\alpha} + |x-y|)^{d+\alpha}} \leqslant \rho(t, y) \leqslant \int_{\mathbb{R}^d} \frac{C_3 t \mu_0(\mathrm{d}x)}{(t^{1/\alpha} + |x-y|)^{d+\alpha}}, \quad \alpha \in (1, 2).$$

(iv) **(Global existence)** *If $p_0 = 1$ or $(\mathrm{div}b)^- = 0$, then we can drop the smallness condition (1.32).*

Remark 1.10. When $p_0 = 1$, conditions (1.30) and (1.33) reduce to

$$\frac{d}{\varrho_0} + \frac{\alpha}{q_b} - \beta_b < \alpha - 1, \quad \frac{d}{\varrho_0} + \frac{\alpha}{q_b} - \beta_b < \frac{3\alpha}{2} - 2, \tag{1.34}$$

which coincide with **(C0)** and **(C0$_S$)** in [22]. In this case, under the same Besov regularity on b, **(A)** and **(B)** were established by Chaudru de Raynal, Jabir and Menozzi in [22]. We assume $p_0 \geqslant 1$, since we were interested to cover more examples from physics for which typical p_0 has to be taken in $(1, \infty]$. For example, if $b(x) = \nabla |x|^{s-d}$ for some $s \in (2 - \alpha, d)$ and $\mu_0 \in \cap_{\varepsilon > 0} \mathbf{B}_{p_0}^{-\varepsilon}$, where $p_0 > \frac{d}{s+\alpha-2}$, then the assumptions in Theorem 1.9 hold by (A.5) in Appendix A below. Moreover, it is well known that for each $x \in \mathbb{R}^d$, the Dirac measure $\delta_x \in \mathbf{B}_1^0$. Thus one sees that for $b = (\delta_{x_0}, \cdots, \delta_{x_d})$ and $\mu_0 \in \cap_{\varepsilon > 0} \mathbf{B}_{p_0}^{-\varepsilon}$, where $p_0 > \frac{d}{\alpha-1}$, the assumptions in Theorem 1.9 hold. In this case, SDE (1.29) is a density-dependent or Nemyskii's type SDEs.

Remark 1.11. In this remark we make simple regularity analysis to illustrate the sharpness of condition (1.30) for $\beta_0 \in [-1, 0)$, which belongs to the subcritical regime. If the strict inequality in (1.30) is an equality, then it is called critical. First of all, we consider the following SDE with linear drift

$$\mathrm{d}X_t = H_t(X_t)\mathrm{d}t + \mathrm{d}L_t^{(\alpha)}, \quad X_0 \sim \mu_0, \tag{1.35}$$

where $H \in L_t^q L_x^p$ for some $q, p \in [1, \infty]$. By simple scaling analysis, it is well known that for SDE (1.35) or the associated Kolmogorov equations to be uniquely solvable, even though μ_0 is smooth, we have to require (see [82])

$$\frac{d}{p} + \frac{\alpha}{q} < \alpha - 1. \tag{1.36}$$

In other words, for linear SDE (1.35), the regularity of initial values does not affect the singularity of H. Now we consider the following first-order DDSDE with convolution drift:

$$\mathrm{d}X_t = b_t * \mu_t(X_t)\mathrm{d}t + \mathrm{d}L_t^{(\alpha)}, \quad X_0 \sim \mu_0, \tag{1.37}$$

where μ_t is the law of X_t. Suppose that for some $\beta_0, \beta_b \in \mathbb{R}$ and $q_b, p_b, p_0 \in [1, \infty]$,

$$b \in L^{q_b}(\mathbb{R}_+; \mathbf{B}_{p_b}^{\beta_b}), \quad \mu_0 \in \mathbf{B}_{p_0}^{\beta_0}. \tag{1.38}$$

We want to seek reasonable conditions on q_b, β_b, p_b and β_0, p_0 so that SDE (1.37) is uniquely solvable. By Itô's formula, μ_t solves the Fokker-Planck equation in the distributional sense:

$$\partial_t \mu = \Delta^{\frac{\alpha}{2}} \mu - \operatorname{div}((b * \mu)\mu).$$

Let P_t be the semigroup associated with $\Delta^{\frac{\alpha}{2}}$. By Duhamel's formula,

$$\mu_t = P_t \mu_0 - \int_0^t P_{t-s} \operatorname{div}((b_s * \mu_s)\mu_s) \mathrm{d}s. \tag{1.39}$$

For any $p_1 \geqslant p_0$ and $\gamma \geqslant \frac{d}{p_1} - \frac{d}{p_0}$, by the regularization effect of P_t (see (2.33) below), we have

$$\|P_t \mu_0\|_{\mathbf{B}_{p_1}^{\gamma+\beta_0}} \leqslant C(1 \wedge t)^{-(\gamma + \frac{d}{p_0} - \frac{d}{p_1})/\alpha} \|\mu_0\|_{\mathbf{B}_{p_0}^{\beta_0}}.$$

In particular, for any $T > 0$, without considering the second term in (1.39),

$$\mu \text{ at most belongs to } L_T^{q_1}(\mathbf{B}_{p_1}^{\gamma+\beta_0}) \text{ for any } (q_1, p_1, \gamma) \in \mathscr{I}, \tag{1.40}$$

where

$$\mathscr{I} := \left\{ (q_1, p_1, \gamma) : p_1 \geqslant p_0, \ 0 \leqslant \gamma + \frac{d}{p_0} - \frac{d}{p_1} < \frac{\alpha}{q_1} \right\}. \tag{1.41}$$

Moreover, under (1.38), for any $(q_1, p_1, \gamma) \in \mathscr{I}$, by Young's convolution inequality,

$$H_t(x) := b_t * \mu_t(x) \in L_T^q(\mathbf{B}_p^\beta),$$

where (β, q, p) is defined by

$$\beta = \beta_b + \gamma + \beta_0, \quad \tfrac{1}{q} = \tfrac{1}{q_b} + \tfrac{1}{q_1}, \quad 1 + \tfrac{1}{p} = \tfrac{1}{p_b} + \tfrac{1}{p_1}. \tag{1.42}$$

After this analysis, in order to make sense for $\int_0^T H_s(X_s) \mathrm{d}s$ and to solve SDE (1.35), it is reasonable to assume $H_t(x) \in L_T^q(L^p)$ for some q, p satisfying (1.36). Thus by (1.42) with $\beta > 0$, (1.41) and (1.36), we have

$$-\beta_0 - \beta_b < \gamma < \tfrac{\alpha}{q_1} + \tfrac{d}{p_1} - \tfrac{d}{p_0} = \tfrac{\alpha}{q} - \tfrac{\alpha}{q_b} + d + \tfrac{d}{p} - \tfrac{d}{p_b} - \tfrac{d}{p_0} < \alpha - 1 - \tfrac{\alpha}{q_b} + d - \tfrac{d}{p_b} - \tfrac{d}{p_0}.$$

This is the case of (1.30) for $\beta_0 \in [-1, 0)$. However, for $\beta_0 \in (\frac{\alpha}{q_b} - \alpha, -1)$, at this moment, we do not know whether the following condition that comes from (1.30)

$$2\beta_0 - \beta_b < \alpha - \tfrac{\alpha}{q_b} + d - \tfrac{d}{p_b} - \tfrac{d}{p_0}$$

is sharp. Of course, for a smaller differentiability index β_0, the initial value could be more singular. We will study this in a forthcoming work.

1.3. Related works.
The study of kinetic equations has a long history. These are used to describe the evolution of the distribution density of a single particle that is characterized by Newton's second law in gas theory, aerodynamics, plasma physics and non-equilibrium statistical physics (see [15], [50],[40]). In 1872, Boltzmann put forward the fundamental kinetic equation of gas theory, which is a non-linear integro-differential equation, now named Boltzmann equation, that is,

$$\partial_t f + v \cdot \nabla_x f = Q(f, f),$$

where $Q(f, f)$ is the collision operator (see [16]). It describes the motion of the molecules as a certain random process determined by collisions between pairs of molecules (see [16]).

The two-sided estimates of the heat kernel of kinetic operator $\mathcal{K} := \Delta_v - v \cdot \nabla_x + b \cdot \nabla_x$ was studied in [32] for Lipschitz drift b and in [24] for Hölder drift b. The L^p-regularity estimate of kinetic operators $\mathcal{K}$ was shown in Bouchut [11]. For the fractional kinetic operator $\mathcal{K}^{(\alpha)} := \Delta_v^{\alpha/2} - v \cdot \nabla_x + b \cdot \nabla_x$, the L^p-maximal regularity estimates were established in [27] and [48] (see also [33]), the Schauder estimates were obtained in [44] and [49]. In fact, our proofs about the solvability of FPE (1.5) strongly depend on the method developed in [44].

Consider now the following McKean-Vlasov or distribution-dependent SDE

$$\mathrm{d}X_t = b(t, X_t, \mu_{X_t})\mathrm{d}t + \sigma(t, X_t, \mu_{X_t})\mathrm{d}W_t. \tag{1.43}$$

When b and σ are Lipschitz continuous in x and in the unknown distribution μ with respect to the Euclidean distance and Wasserstein metric, it is by now standard to show the existence and uniqueness of strong solutions to DDSDE (1.43) (cf. [74] and [18]). Moreover, under some one-sided Lipschitz assumptions, Wang [78] showed the strong well-posedness and some functional inequalities for DDSDE (1.43). For coefficients b, σ only measurable in the measure variable, hence also containing the case of distribution density dependent (also called Nemytskii-type) coefficients, existence and uniqueness of weak solutions was proved in a series of papers (see [2, 3, 4, 5, 6]) and in the case of multiplicative Lévy noise [67]. In [41], Hammersley, Siska and Szpruch proved the existence of weak solutions to SDE (1.43) on a domain with continuous and unbounded coefficients under Lyapunov-type conditions. Moreover, uniqueness is also obtained under some functional Lyapunov conditions. When the diffusion matrix is uniformly non-degenerate and are only measurable and of at most linear growth, by using the classical Krylov estimates, Mishura and Veretennikov [62] showed the existence of weak solutions. The uniqueness is also proved when the diffusion coefficient does not depend on the unknown distribution μ and is Lipschitz continuous in x and b is Lipschitz continuous with respect to μ. When $b(x, \mu) = \int_{\mathbb{R}^d} \bar{b}(x, y) \mu(\mathrm{d}x)$ with $|\bar{b}(x, y)| \leqslant h(x - y)$, where $h \in L^p(\mathbb{R}^d)$ for some $p > d$, and $\sigma(x, \mu) = \sigma(x)$ does not depend on the distribution and is uniformly elliptic and Hölder continuous, the well-posedness of DDSDE (1.43) was shown in [68]. More recently, Chaudru de Raynal, Jabir and Menozzi [22] studied the well-posedness to the following DDSDE with convolution drift

$$\mathrm{d}X_t = (b_t * \mu_{X_t})(X_t)\mathrm{d}t + \mathrm{d}L_t^{(\alpha)}.$$

Under (1.34), for $b \in L_T^{q_b}(\mathbf{B}_{p_b}^{\beta_b})$ and any initial distribution μ_0, by smoothing b and a duality method, they showed the weak/strong well-posedness to the above SDE. Therein they also obtained some Hölder regularity of the solution. Later, Han in [42] improved it to the smooth solution under some conditions of b.

Next consider the following linear second order kinetic SDE:

$$\ddot{X}_t = H(X_t, \dot{X}_t) + \dot{L}_t^{(\alpha)}, \quad (X_0, \dot{X}_0) \sim \mu_0, \tag{1.44}$$

where H is only a function of (t, x, v). When $\alpha = 2$ and H is $2/3 + \varepsilon$-Hölder continuous in x and ε-Hölder continuous in v, the strong well-posedness was established in [20] and [79]. The L^p-type drift was considered in [83]. In [21], Chaudru de Raynal and Menozzi showed the weak well-posedness when $H \in L_t^q L_{x,v}^p$, where $\frac{4d}{p} + \frac{2}{q} < 1$. For $\alpha \in (1, 2)$, the strong well-posedness for kinetic SDE (1.44) was established in [44] when $H \in L^\infty(C_{x,v}^{\gamma, \beta})$, where $\gamma \in (\frac{2+\alpha}{2(1+\alpha)}, 1)$ and $\beta > 1 - \frac{\alpha}{2}$. It seems that there are few results about the well-posedness of kinetic DDSDEs with singular interaction kernels. The aim of this paper is to fill this gap.

After the contents of this paper was presented in a talk of "The Non-local operators, probability and singularities" online seminar, Stephane Menozzi informed us in a very kind e-mail that he with P.E. Chaudru de Raynal and J.F. Jabir have a paper in preparation with related results and sent us a current (draft) version of it ([23]). We would like to thank Stephane very much for this. After a detailed discussion with Stephane, we emphasize the following differences between our and their work.

- The first and foremost difference concerns the motivation and the models in applications. In their paper, they continue the research from their very interesting previous work [22] and consider non-degenerate McKean-Vlasov equations where the initial data is a real function, i.e. $\beta_0 \geqslant 0$, while in our present paper, we consider degenerate kinetic equations with singular kernels and measure initial data, which can be deduced naturally from the second order N-particle systems (1.7). As a consequence, our results can be applied to non-degenerate equations (see Theorem 1.9).
- The key condition (1.30) is not the same as $(\mathbf{C}_1)$ in [23]. More precisely, there is only one β_b in our condition (1.30). But the condition $(\mathbf{C}_1)$ in [23] includes $2\beta_b$. Moreover, in [23] $\beta_0 \geqslant 0$ is required, whereas $\beta_0 < 0$ is allowed in our paper. It should be noted that our

condition is sharp in the scaling sense (see Remark 1.11). Furthermore, in the PDE part, we cover the critical case in Theorem 3.11, where the key point is that in our Schauder estimate Theorem 2.18 the norms in the left and right hand sides are $\mathbf{B}_{p'}^{\beta',1}$ and $\mathbf{B}_p^{\beta,\infty}$, respectively.

- Another important difference is in the methods of showing the well-posedness of the mean-field SDEs are quite different. In [23], the authors mollify the singular interaction kernel and show some uniform estimate for the solution to the nonlinear Fokker-Planck equation. Then, by the weak convergence method they obtain the existence of a solution in the sense of a martingale problem. In our work, we use the contraction mapping theorem to show the well-posedness of the nonlinear Fokker-Planck equation and then obtain the well-posedness of the linearized SDE. Then, based on the uniqueness of the linear Fokker-Planck equation, the solution to the linearized SDE is just the solution to the mean-field SDE (see Section 4.2 for more details).
- For concrete physical models like the Navier-Stokes equations, in [23], in order to avoid the non-integrability at infinity the authors cut off the kernel by an indicator function of a bounded domain, whereas, in our work, we consider the kernel being the sum of two singular kernels so that we can cover the physical interaction kernel in the whole space.
- In the current work, we also pay much attention to the smoothness, stability, and large-time behavior of solutions to nonlinear Fokker-Planck equations, which are not considered in [23]. In fact, there are many techniques developed to study these aspects. For instance, we use a time shift trick to obtain the smoothness and large time decay estimates of solutions to nonlinear Fokker-Planck equations. Moreover, by choosing suitable parameters, based on an interpolation technique, we can drop the smallness assumption of $\|u_0\|_m$ in the PDE's literature, which is important to study the related mean-field SDEs, see Remark 5.3 below.
- Finally, there are also some differences in the techniques. For instance, for the treatment of the convolution $b * u$, it is regarded as an inner product $\langle b(x - \cdot), u(\cdot) \rangle$ in [23]. In our paper, we use Young's inequality to estimate it. Moreover, we work with the kinetic operator and in anisotropic Besov spaces, which are more difficult to handle, since for instance $\mathscr{R}_j^a P_t \neq P_t \mathscr{R}_j^a$. To obtain the well-posedness of the equations with large initial data locally in time in the critical case, we show an asymptotic estimate for small time for the semigroup P_t, see Lemma 3.15.

Summarizing the above, in our opinion [23] is a very valuable contribution to the subject containing various interesting and different focal points in comparison to our paper.

1.4. **Structure of the paper.** This paper is organized as follows. In Section 2 we introduce the anisotropic Besov spaces as well as some of their important properties. Most of the results are well known and can be found in [75]. Moreover, we also prepare the basic Schauder estimates for the model kinetic equation in time-weighted spaces (see Theorem 2.18).

In Section 3, we study the global and local well-posedness for nonlinear kinetic Fokker-Planck equation (1.5) with small or large initial data. Moreover, we also investigate the smoothness, stability and large time decay estimates of the solution. More precisely, in Subsection 3.1, we prove our main result Theorem 3.11 about the global well-posedness for small initial values for the nonlinear kinetic Fokker-Planck equation (1.5). In Lemma 3.5, we establish a priori bilinear estimate for $\mathscr{H}_t^b(f,g)$, where Young's inequality for convolutions provides all the regularity induced by the kernel. In particular, the dependence on parameters is calculated carefully. Then in Lemma 3.9, we show the core estimate in the solution space. In Subsection 3.2, we consider large time decay estimates. In Subsection 3.3, we show the local well-posedness for large initial values in Theorem 3.17. In Subsection 3.4, we prove two results about the global well-posedness for (1.5) for arbitrary large initial value. The corresponding results for non-degenerate fractional Fokker-Planck equations are also presented.

Section 4 we devote to the study of mean-field kinetic SDEs and prove Theorems 1.3 and 1.9. In Subsection 4.1, we first show the weak and strong well-posedness for linear kinetic SDEs with Hölder drifts and driven by α-stable processes by Zvonkin's transformation. In particular, for strong well-posedness, we need to assume the drift to be in $L_T^q(\mathbf{C}_x^{\frac{\alpha+\beta}{1+\alpha}} \cap \mathbf{C}_v^\beta)$, where $\beta > 1 - \frac{\alpha}{2}$. Note that for $q = \infty$, this has already been obtained in [44]. In Subsection 4.2, we prove our main theorems by freezing the drifts.

In Section 5 we study the well-posedness of fractional Vlasov-Poisson-Fokker-Planck equation (1.9) with measures as initial data and the associated kinetic McKean-Vlasov SDEs. We also obtain an optimal large time decay estimate when the initial data are small. For Wiener noise in three dimensional case, when the initial data have finite second moment, using well-known results, we prove global existence of smooth solutions for large initial values. In the fractional case, the global well-posedness of (1.9) for large initial values seems to be open since there is no finite second order moment (see (1.10)), which is crucial to show global well-posedness in [10, 14, 64].

In Section 6, we apply our main results to the vorticity formulation of fractional Navier-Stokes and subcritical SQG equations. For two (and three) dimensional fractional Navier-Stokes equation, we show the global existence of smooth solutions when the initial vorticity is in $\mathbf{B}_\infty^{-(1+\alpha)/2+}$ (and small). Moreover, in the two dimensional case, when the initial vorticity is a probability measure, we show the weak and strong well-posedness for the associated mean-field SDEs. For the three dimensional case, we also obtain an optimal large time decay estimate. To the best of our knowledge, they are new, at least for fractional Navier-Stokes equations.

In Section 7 we present an application to fractional kinetic porous medium equations with viscosity and show the global existence of smooth solutions for large initial values.

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

$$a \wedge b := \max(a, b), \quad a \vee b := \min(a, b), \quad a^+ := 0 \vee a.$$

By $A \lesssim_C B$ and $A \asymp_C B$ or simply $A \lesssim B$ and $A \asymp B$, we respectively mean that for some constant $C \geqslant 1$,

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

Below we collect some frequently used notations for the readers' convenience.

- $\mathscr{P}(\mathbb{R}^N) :=$ set of all probability measures on $\mathbb{R}^N$.
- $\mathcal{M}(\mathbb{R}^N) :=$ set of all finite Radon measures on $\mathbb{R}^N$.
- $\|\cdot\|_m :=$ total variation norm.
- $\mathbf{B}_{p;\boldsymbol{a}}^{s;q} :=$ anisotropic Besov space.
- $\mathbf{B}_p^{s;q} :=$ isotropic Besov space.
- $\mathbf{C}_{\boldsymbol{a}}^s :=$ anisotropic Hölder-Zygmund space.
- $\mathscr{R}_j^{\boldsymbol{a}} :=$ anisotropic block operator.
- $P_t :=$ semigroup of $\Delta_v^{\alpha/2} - v \cdot \nabla_x$ or $\Delta^{\alpha/2}$ (see (2.22) below).
- $\mathbb{X}_\beta, \mathbb{Y}_\beta, \mathbb{Z}_\beta :=$ some subspaces of Besov spaces (see (3.8) below).
- $\mathbb{S}_\beta :=$ space of solutions (see (3.21) below).

2. ANISOTROPIC BESOV SPACE AND KINETIC SEMIGROUP

In this section we first recall the definition of anisotropic Besov space with mixed integrability indices as well as its basic properties (see [75] and [85]). Then we study the fractional kinetic semigroup and prove time-weighted Schauder's type estimates in anisotropic Besov spaces for fractional kinetic operators. In order to study the small time singularity and large time decay estimates, we introduce a time-weighted L^p-space by considering different weights for small time and large time.

2.1. **Anisotropic Besov spaces.** For an L^1-integrable function f in $\mathbb{R}^{2d}$, let $\hat{f}$ be the Fourier transform of f defined by

$$\hat{f}(\xi) := (2\pi)^{-d} \int_{\mathbb{R}^{2d}} e^{-i\xi \cdot z} f(z) dz, \quad \xi \in \mathbb{R}^{2d},$$

and $\check{f}$ the Fourier inverse transform of f defined by

$$\check{f}(z) := (2\pi)^{-d} \int_{\mathbb{R}^{2d}} e^{i\xi \cdot z} f(\xi) d\xi, \quad z \in \mathbb{R}^{2d}.$$

Recall $\boldsymbol{a} = (1 + \alpha, 1)$. For $z = (x, v)$ and $z' = (x', v')$ in $\mathbb{R}^{2d}$, we introduce the anisotropic distance

$$|z - z'|_{\boldsymbol{a}} := |x - x'|^{1/(1+\alpha)} + |v - v'|.$$

Note that $z \mapsto |z|_{\boldsymbol{a}}$ is not smooth. For $r > 0$ and $z \in \mathbb{R}^{2d}$, we also introduce the ball centered at z and with radius r with respect to the above distance as follows:

$$B_r^{\boldsymbol{a}}(z) := \{z' \in \mathbb{R}^{2d} : |z' - z|_{\boldsymbol{a}} \leqslant r\}, \quad B_r^{\boldsymbol{a}} := B_r^{\boldsymbol{a}}(0).$$

Let $\chi_0^{\boldsymbol{a}}$ be a symmetric C^∞-function on $\mathbb{R}^{2d}$ with

$$\chi_0^{\boldsymbol{a}}(\xi) = 1 \text{ for } \xi \in B_1^{\boldsymbol{a}} \text{ and } \chi_0^{\boldsymbol{a}}(\xi) = 0 \text{ for } \xi \notin B_2^{\boldsymbol{a}}.$$

We define

$$\phi_j^{\boldsymbol{a}}(\xi) := \begin{cases} \chi_0^{\boldsymbol{a}}(2^{-j\boldsymbol{a}}\xi) - \chi_0^{\boldsymbol{a}}(2^{-(j-1)\boldsymbol{a}}\xi), & j \geqslant 1, \\ \chi_0^{\boldsymbol{a}}(\xi), & j = 0, \end{cases}$$

where for $s \in \mathbb{R}$ and $\xi = (\xi_1, \xi_2)$,

$$2^{s\boldsymbol{a}}\xi = (2^{s(1+\alpha)}\xi_1, 2^s\xi_2).$$

Note that

$$\text{supp}(\phi_j^{\boldsymbol{a}}) \subset \{\xi : 2^{j-1} \leqslant |\xi|_{\boldsymbol{a}} \leqslant 2^{j+1}\}, \ j \geqslant 1, \ \text{supp}(\phi_0^{\boldsymbol{a}}) \subset B_2^{\boldsymbol{a}}, \tag{2.1}$$

and

$$\sum_{j \geqslant 0} \phi_j^{\boldsymbol{a}}(\xi) = 1, \quad \forall \xi \in \mathbb{R}^{2d}. \tag{2.2}$$

Let $\mathscr{S}$ be the space of all Schwartz functions on $\mathbb{R}^{2d}$ and $\mathscr{S}'$ the dual space of $\mathscr{S}$, called the tempered distribution space. For given $j \geqslant 0$, the dyadic anisotropic block operator $\mathscr{R}_j^{\boldsymbol{a}}$ is defined on $\mathscr{S}'$ by

$$\mathscr{R}_j^{\boldsymbol{a}} f(z) := (\phi_j^{\boldsymbol{a}} \hat{f})^{\check{}}(z) = \check{\phi}_j^{\boldsymbol{a}} * f(z), \tag{2.3}$$

where the convolution is understood in the distributional sense and by scaling,

$$\check{\phi}_j^{\boldsymbol{a}}(z) = 2^{(j-1)(2+\alpha)d} \check{\phi}_1^{\boldsymbol{a}}(2^{(j-1)\boldsymbol{a}}z), \quad j \geqslant 1. \tag{2.4}$$

For $j \in \mathbb{N}$, by definition it is easy to see that

$$\mathscr{R}_j^{\boldsymbol{a}} = \mathscr{R}_j^{\boldsymbol{a}} \widetilde{\mathscr{R}}_j^{\boldsymbol{a}}, \quad \text{where} \quad \widetilde{\mathscr{R}}_j^{\boldsymbol{a}} := \sum_{|i-j| \leqslant 2} \mathscr{R}_i^{\boldsymbol{a}}, \tag{2.5}$$

where we used the convention that $\mathscr{R}_i^{\boldsymbol{a}} := 0$ for $i < 0$. Moreover, by the symmetry of $\phi_j^{\boldsymbol{a}}$,

$$\langle \mathscr{R}_j^{\boldsymbol{a}} f, g \rangle = \langle f, \mathscr{R}_j^{\boldsymbol{a}} g \rangle, \quad f \in \mathscr{S}', \ g \in \mathscr{S}.$$

Similarly, we can define the isotropic block operator $\mathscr{R}_j f = \check{\phi}_j * f$ in $\mathbb{R}^d$, where

$$\text{supp}(\phi_j) \subset \{\xi : 2^{j-1} \leqslant |\xi| \leqslant 2^{j+1}\}, \ j \geqslant 1, \ \text{supp}(\phi_0) \subset B_2. \tag{2.6}$$

The following Bernstein inequality is standard and proven in [85]. We omit the details.

Lemma 2.1 (Bernstein's inequality). *For any $\boldsymbol{k} := (k_1, k_2) \in \mathbb{N}_0^2$, $\boldsymbol{p} \leqslant \boldsymbol{p}' \in [1, \infty]^2$, there is a constant $C = C(\boldsymbol{k}, \boldsymbol{p}, \boldsymbol{p}', \alpha, d) > 0$ such that for all $j \geqslant 0$,*

$$\|\nabla_x^{k_1} \nabla_v^{k_2} \mathscr{R}_j^{\boldsymbol{a}} f\|_{\boldsymbol{p}'} \lesssim_C 2^{j\boldsymbol{a} \cdot (\boldsymbol{k} + \frac{d}{\boldsymbol{p}} - \frac{d}{\boldsymbol{p}'})} \|\mathscr{R}_j^{\boldsymbol{a}} f\|_{\boldsymbol{p}}, \tag{2.7}$$

where $\| \cdot \|_{\boldsymbol{p}}$ is defined in (1.17).

Now we introduce the following anisotropic Besov spaces (cf. [75, Chapter 5]).

Definition 2.2. *Let $s \in \mathbb{R}$, $q \in [1, \infty]$ and $\boldsymbol{p} \in [1, \infty]^2$. The anisotropic Besov space is defined by*

$$\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{s,q} := \left\{ f \in \mathscr{S}' : \|f\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{s,q}} := \left(\sum_{j \geqslant 0} \left(2^{js} \|\mathscr{R}_j^{\boldsymbol{a}} f\|_{\boldsymbol{p}} \right)^q \right)^{1/q} < \infty \right\},$$

where $\| \cdot \|_{\boldsymbol{p}}$ is defined in (1.17). *Let $(s_0, s_1) \in \mathbb{R}^2$. The mixed Besov space is defined by*

$$\mathbf{B}_{\boldsymbol{p};x,\boldsymbol{a}}^{s_0,s_1} := \left\{ f \in \mathscr{S}' : \|f\|_{\mathbf{B}_{\boldsymbol{p};x,\boldsymbol{a}}^{s_0,s_1}} := \sup_{k,j \geqslant 0} 2^{\frac{ks_0}{1+\alpha}} 2^{js_1} \|\mathscr{R}_k^x \mathscr{R}_j^{\boldsymbol{a}} f\|_{\boldsymbol{p}} < \infty \right\},$$

where for a function $f : \mathbb{R}^{2d} \to \mathbb{R}$,

$$\mathscr{R}_k^x f(x, v) := \mathscr{R}_k f(\cdot, v)(x).$$

Moreover, we also define

$$\mathbf{B}_{\boldsymbol{p};x}^s := \mathbf{B}_{\boldsymbol{p};x}^{s,\infty} := \left\{ f \in \mathscr{S}' : \|f\|_{\mathbf{B}_{\boldsymbol{p};x}^s} := \sup_{k \geqslant 0} 2^{ks} \|\mathscr{R}_k^x f\|_{\boldsymbol{p}} < \infty \right\}.$$

Similarly, one defines the usual isotropic Besov spaces $\mathbf{B}_p^{s,q}$ in $\mathbb{R}^d$ in terms of isotropic block operators $\mathscr{R}_j$. If there is no confusion, we shall write

$$\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^s := \mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{s,\infty}, \quad \mathbf{B}_p^s := \mathbf{B}_p^{s,\infty}.$$

Remark 2.3. The mixed Besov space $\mathbf{B}_{\boldsymbol{p};x,\boldsymbol{a}}^{s_0,s_1}$ shall be used in the study of strong solutions of kinetic SDEs. Note that for any $f \in \mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{s,q}$, by (2.2) we have

$$f = \sum_{j \geqslant 0} \mathscr{R}_j^{\boldsymbol{a}} f \quad \text{in } \mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{s,q}. \tag{2.8}$$

Remark 2.4. Any finite Radon measure μ belongs to $\mathbf{B}_{1;\boldsymbol{a}}^{0,\infty}$. Indeed, by definition and (2.4),

$$\|\mu\|_{\mathbf{B}_{1;\boldsymbol{a}}^{0,\infty}} = \sup_{j \geqslant 0} \|\mathscr{R}_j^{\boldsymbol{a}} \mu\|_1 \leqslant \sup_{j \geqslant 0} \|\check{\phi}_j^{\boldsymbol{a}}\|_1 \|\mu\|_m = \left(\|\check{\phi}_0^{\boldsymbol{a}}\|_1 \vee \|\check{\phi}_1^{\boldsymbol{a}}\|_1 \right) \|\mu\|_m < \infty, \tag{2.9}$$

where $\|\mu\|_m$ stands for the norm of total variation. Moreover, for any $\varepsilon = 2^{-m} \in (0, 1)$ with $m \in \mathbb{N}$, if we define

$$\mu_\varepsilon(\mathrm{d}z) := \mu(\mathrm{d}(\varepsilon^{\boldsymbol{a}} z)),$$

then for $j \geqslant 1$, by scaling $\check{\phi}_j^{\boldsymbol{a}}(\varepsilon^{-\boldsymbol{a}} z) = \varepsilon^{(2+\alpha)d} \check{\phi}_{m+j}^{\boldsymbol{a}}(z)$, it is easy to see that

$$\mathscr{R}_j^{\boldsymbol{a}} \mu_\varepsilon(z) = \varepsilon^{(2+\alpha)d} (\mathscr{R}_{j+m}^{\boldsymbol{a}} \mu)(\varepsilon^{\boldsymbol{a}} z).$$

Hence, for any $\boldsymbol{p} \in [1, \infty]^2$,

$$\|\mathscr{R}_j^{\boldsymbol{a}} \mu_\varepsilon\|_{\boldsymbol{p}} = \varepsilon^{s + \boldsymbol{a} \cdot (\boldsymbol{d} - \frac{d}{\boldsymbol{p}})} 2^{ms} \|\mathscr{R}_{j+m}^{\boldsymbol{a}} \mu\|_{\boldsymbol{p}}, \ j \geqslant 1.$$

Moreover, for $j = 0$,

$$\mathscr{F}(\check{\phi}_0^{\boldsymbol{a}}(\varepsilon^{-\boldsymbol{a}} z))(\xi) = \varepsilon^{(2+\alpha)d} \phi_0^{\boldsymbol{a}}(\varepsilon^{\boldsymbol{a}} \xi) \sum_{j=0}^m \phi_j^{\boldsymbol{a}}(\xi),$$

which implies that

$$\|\mathscr{R}_0^{\boldsymbol{a}}\mu_\varepsilon\|_{\boldsymbol{p}} \leqslant \varepsilon^{\boldsymbol{a}\cdot(\boldsymbol{d}-\frac{d}{\boldsymbol{p}})} \sum_{j=0}^m \|\mathscr{R}_j^{\boldsymbol{a}}\mu\|_{\boldsymbol{p}} \leqslant 2\varepsilon^{\boldsymbol{a}\cdot(\boldsymbol{d}-\frac{d}{\boldsymbol{p}})+s\mathbb{1}_{s<0}} \|\mu\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{s,\infty}}.$$

Therefore, for any $s \in \mathbb{R}\backslash\{0\}$ and $q \in [1,\infty]$,

$$\|\mu_\varepsilon\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{s,q}} \leqslant 2 \left(\varepsilon^{s+\boldsymbol{a}\cdot(\boldsymbol{d}-\frac{d}{\boldsymbol{p}})} + \mathbb{1}_{s>0}\varepsilon^{\boldsymbol{a}\cdot(\boldsymbol{d}-\frac{d}{\boldsymbol{p}})} \right) \|\mu\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{s,q}}. \tag{2.10}$$

For a function $f : \mathbb{R}^{2d} \to \mathbb{R}$, the first-order difference operator is defined by

$$\delta_h^{(1)} f(z) := \delta_h f(z) := f(z+h) - f(z), \quad z, h \in \mathbb{R}^{2d}.$$

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

$$\delta_h^{(M+1)} f(z) = \delta_h \circ \delta_h^{(M)} f(z).$$

The following characterization is well-known (cf. [85] and [45, Theorem 2.7]).

Proposition 2.5. *For $s > 0$, $q \in [1,\infty]$ and $\boldsymbol{p} \in [1,\infty]^2$, an equivalent norm of $\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{s,q}$ is given by*

$$\|f\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{s,q}} \asymp \left(\int_{|h|_{\boldsymbol{a}}\leqslant 1} \left(\frac{\|\delta^{([s]+1)}f\|_{\boldsymbol{p}}}{|h|_{\boldsymbol{a}}^s} \right)^q \frac{\mathrm{d}h}{|h|_{\boldsymbol{a}}^{(2+\alpha)d}} \right)^{1/q} + \|f\|_{\boldsymbol{p}}, \tag{2.11}$$

where $[s]$ is the integer part of s. In particular, $\mathbf{C}_{\boldsymbol{a}}^s := \mathbf{B}_{\infty;\boldsymbol{a}}^{s,\infty}$ is the anisotropic Hölder-Zygmund space, and for $s \in (0,1)$, there is a constant $C = C(\alpha, d, s) > 0$ such that

$$\|f\|_{\mathbf{C}_{\boldsymbol{a}}^s} \asymp_C \|f\|_\infty + \sup_{z \neq z'} |f(z) - f(z')|/|z - z'|_{\boldsymbol{a}}^s.$$

Similarly, the mixed and x-direction Hölder-Zygmund spaces are defined by

$$\mathbf{C}_{x,\boldsymbol{a}}^{s_0,s_1} := \mathbf{B}_{\infty;x,\boldsymbol{a}}^{s_0,s_1}, \quad \mathbf{C}_x^s := \mathbf{B}_{\infty;x}^s.$$

The following lemma provides important properties about Besov spaces. Although the proofs are standard, for the readers' convenience, we provide detailed proofs in the appendix.

Lemma 2.6. *(i) For any $\boldsymbol{p} \in [1,\infty]^2$, $s' > s$ and $q \in [1,\infty]$, it holds that*

$$\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{0,1} \hookrightarrow \mathbb{L}^{\boldsymbol{p}} \hookrightarrow \mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{0,\infty}, \quad \mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{s',\infty} \hookrightarrow \mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{s,1} \hookrightarrow \mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{s,q}. \tag{2.12}$$

(ii) For any $\beta, \beta_1, \beta_2 \in \mathbb{R}$, $q, q_1, q_2 \in [1,\infty]$ and $\boldsymbol{p}, \boldsymbol{p}_1, \boldsymbol{p}_2 \in [1,\infty]^2$ with

$$\beta = \beta_1 + \beta_2, \quad 1 + \frac{1}{\boldsymbol{p}} = \frac{1}{\boldsymbol{p}_1} + \frac{1}{\boldsymbol{p}_2}, \quad \frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2},$$

it holds that

$$\|f * g\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta,q}} \leqslant 5\|f\|_{\mathbf{B}_{\boldsymbol{p}_1;\boldsymbol{a}}^{\beta_1,q_1}} \|g\|_{\mathbf{B}_{\boldsymbol{p}_2;\boldsymbol{a}}^{\beta_2,q_2}}. \tag{2.13}$$

(iii) For any $s > 0$, $q \in [1,\infty]$ and $\boldsymbol{p}, \boldsymbol{p}_1, \boldsymbol{p}_2 \in [1,\infty]^2$ with $\frac{1}{\boldsymbol{p}} = \frac{1}{\boldsymbol{p}_1} + \frac{1}{\boldsymbol{p}_2}$, there is a constant $C = C(\alpha, d, s, q, \boldsymbol{p}_1, \boldsymbol{p}_2) > 0$ such that for any $(f, g) \in \mathbf{B}_{\boldsymbol{p}_1;\boldsymbol{a}}^{s,q} \times \mathbf{B}_{\boldsymbol{p}_2;\boldsymbol{a}}^{s,q}$,

$$\|fg\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{s,q}} \lesssim_C \|f\|_{\mathbf{B}_{\boldsymbol{p}_1;\boldsymbol{a}}^{s,q}} \|g\|_{\boldsymbol{p}_2} + \|f\|_{\boldsymbol{p}_1} \|g\|_{\mathbf{B}_{\boldsymbol{p}_2;\boldsymbol{a}}^{s,q}}. \tag{2.14}$$

(iv) Let $\boldsymbol{p}, \boldsymbol{p}_1, \boldsymbol{p}_2 \in [1,\infty]^2$ and $s, s_0, s_1 \in \mathbb{R}$, $\theta \in [0,1]$, $q, q_1, q_2 \in [1,\infty]$. Suppose that

$$\frac{1}{q} = \frac{1-\theta}{q_1} + \frac{\theta}{q_2}, \quad \frac{1}{\boldsymbol{p}} \leqslant \frac{1-\theta}{\boldsymbol{p}_1} + \frac{\theta}{\boldsymbol{p}_2}, \quad s - \boldsymbol{a} \cdot \frac{d}{\boldsymbol{p}} = (1-\theta)\left(s_1 - \boldsymbol{a} \cdot \frac{d}{\boldsymbol{p}_1}\right) + \theta\left(s_2 - \boldsymbol{a} \cdot \frac{d}{\boldsymbol{p}_2}\right). \tag{2.15}$$

Then there is a constant $C = C(\boldsymbol{p}, \boldsymbol{p}_1, \boldsymbol{p}_2, s, s_1, s_2, \theta) > 0$ such that

$$\|f\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{s,q}} \lesssim_C \|f\|_{\mathbf{B}_{\boldsymbol{p}_1;\boldsymbol{a}}^{s_1,q_1}}^{1-\theta} \|f\|_{\mathbf{B}_{\boldsymbol{p}_2;\boldsymbol{a}}^{s_2,q_2}}^{\theta}. \tag{2.16}$$

In particular, for $1 \leqslant \boldsymbol{p}_1 \leqslant \boldsymbol{p} \leqslant \infty$, $q \in [1,\infty]$ and $s = s_1 + \boldsymbol{a} \cdot \left(\frac{d}{\boldsymbol{p}} - \frac{d}{\boldsymbol{p}_1}\right)$,

$$\|f\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{s,q}} \lesssim_C \|f\|_{\mathbf{B}_{\boldsymbol{p}_1;\boldsymbol{a}}^{s_1,q}}. \tag{2.17}$$

(v) *For any $\varepsilon > 0$ and $\boldsymbol{p} \in [1, \infty]^2$, $(s_0, s_1) \in \mathbb{R}^2$, we have the following embeddings:*

$$\mathbf{B}_{\boldsymbol{p};x,\boldsymbol{a}}^{s_0-\varepsilon,s_1+\varepsilon} \hookrightarrow \mathbf{B}_{\boldsymbol{p};x,\boldsymbol{a}}^{s_0,s_1} \hookrightarrow \mathbf{B}_{\boldsymbol{p};x,\boldsymbol{a}}^{s_0+\varepsilon,s_1-\varepsilon}, \tag{2.18}$$

and for $s_1 > 0$,

$$\mathbf{B}_{\boldsymbol{p};x,\boldsymbol{a}}^{s_0,s_1} \hookrightarrow \mathbf{B}_{\boldsymbol{p};x}^{s_0+s_1}. \tag{2.19}$$

(vi) *For $\theta \in [0,1]$, $\gamma, \beta \in \mathbb{R}$ and $\boldsymbol{p} \in [1, \infty]^2$, there is a constant $C = C(d, \theta, \boldsymbol{p}) > 0$ such that*

$$\|f\|_{\mathbf{B}_{\boldsymbol{p};x,\boldsymbol{a}}^{\theta\gamma,(1-\theta)\beta}} \lesssim \|f\|_{\mathbf{B}_{\boldsymbol{p};x}^{\gamma/(1+\alpha)}}^{\theta} \|f\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta}}^{1-\theta}. \tag{2.20}$$

Remark 2.7. Inequality (2.16) is a version of the classical Gagliardo-Nirenberg inequality in Besov spaces. Inequality (2.17) is the classical Sobolev inequality.

2.2. **Time-weighted Schauder estimates.** In this subsection we prove the basic time-weighted Schauder estimates for the model kinetic equation: $\partial_t u = (\Delta_v - v \cdot \nabla_x)u + f$. For this goal, we need to make a detailed study for the kinetic semigroup. Here much attention is paid on the large time estimate of kinetic semigroup in Besov spaces.

For $\alpha \in (0, 2]$, let $p_t(z) = p_t(x, v)$ be the distributional density of $Z_t := (\int_0^t L_s^{(\alpha)} \mathrm{d}s, L_t^{(\alpha)})$, where $L_t^{(\alpha)}$ is the standard α-stable process as in the introduction. Since $(L_{\lambda t}^{(\alpha)})_{t \geqslant 0} \overset{(d)}{=} \lambda^{\frac{1}{\alpha}} (L_t^{(\alpha)})_{t \geqslant 0}$ for $\lambda > 0$, it is easy to see that $p_t(x, v)$ enjoys the following scaling property

$$p_t(x, v) = t^{-\frac{(2+\alpha)d}{\alpha}} p_1(t^{-\frac{1+\alpha}{\alpha}} x, t^{-\frac{1}{\alpha}} v), \quad t > 0. \tag{2.21}$$

The kinetic semigroup of operator $\Delta_v^{\alpha/2} - v \cdot \nabla_x$ is given by

$$P_t f(z) := \mathbb{E} f(\Gamma_t z + Z_t) = \Gamma_t(p_t * f)(z) = (\Gamma_t p_t * \Gamma_t f)(z) = \int_{\mathbb{R}^d} \Gamma_t p_t(z - z') \Gamma_t f(z') \mathrm{d}z', \tag{2.22}$$

where

$$\Gamma_t f(z) := f(\Gamma_t z) := f(x - tv, v).$$

The following lemma is originally proven in [44] (see also [85]).

Lemma 2.8. *For any $l \geqslant 0$, there is a constant $C = C(l, \alpha, d) > 0$ such that*

$$\|\mathscr{R}_j^{\boldsymbol{a}} p_t\|_1 + \|\mathscr{R}_j^{\boldsymbol{a}} \Gamma_t p_t\|_1 \lesssim_C (t2^{j\alpha})^{-l} \wedge 1, \ \forall j \geqslant 1, t > 0.$$

Remark 2.9. The above estimate only holds for high frequency operator $\mathscr{R}_j^{\boldsymbol{a}}$, $j \geqslant 1$.

As an easy consequence we have

Lemma 2.10. *For any $\boldsymbol{p} \in [1, \infty]^2$, $k \in \mathbb{N}_0$ and $\beta \in \mathbb{R}$, there is a constant $C > 0$ such that*

$$\|\nabla_v^k p_t\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta}} \lesssim_C t^{\frac{\boldsymbol{a} \cdot (d/\boldsymbol{p} - \boldsymbol{d}) - k}{\alpha}} (1 + t^{-\frac{\beta \vee 0}{\alpha}}), \ t > 0.$$

Proof. By scaling (2.21) and the change of variables, it is easy to see that for any $m \in \mathbb{N}_0$,

$$\|\nabla_v^m p_t\|_{\boldsymbol{p}} = t^{\frac{\boldsymbol{a} \cdot (d/\boldsymbol{p} - \boldsymbol{d}) - m}{\alpha}} \|\nabla_v^m p_1\|_{\boldsymbol{p}} \lesssim_C t^{\frac{\boldsymbol{a} \cdot (d/\boldsymbol{p} - \boldsymbol{d}) - m}{\alpha}}. \tag{2.23}$$

Without loss of generality, we assume $\beta \geqslant 0$. By definition, Bernstein's inequality (2.7) and Lemma 2.8 with $l = \frac{\beta + k + \boldsymbol{a} \cdot (\boldsymbol{d} - d/\boldsymbol{p})}{\alpha}$, we have

$$\begin{aligned}
\|\nabla_v^k p_t\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta}} &\leqslant \|\mathscr{R}_0^{\boldsymbol{a}} \nabla_v^k p_t\|_{\boldsymbol{p}} + \sup_{j \geqslant 1} \left(2^{j\beta} \|\mathscr{R}_j^{\boldsymbol{a}} \nabla_v^k p_t\|_{\boldsymbol{p}}\right) \\
&\lesssim \|\nabla_v^k p_t\|_{\boldsymbol{p}} + \sup_{j \geqslant 1} 2^{j(\beta + k + \boldsymbol{a} \cdot (\boldsymbol{d} - d/\boldsymbol{p}))} \|\mathscr{R}_j^{\boldsymbol{a}} p_t\|_1 \\
&\lesssim t^{\frac{\boldsymbol{a} \cdot (d/\boldsymbol{p} - \boldsymbol{d}) - k}{\alpha}} + t^{\frac{\boldsymbol{a} \cdot (d/\boldsymbol{p} - \boldsymbol{d}) - k - \beta}{\alpha}}.
\end{aligned}$$

The proof is complete. $\qquad \square$

The following lemma is proved in [44, Lemma 6.7]. For the readers' convenience, we reproduce the proof here.

Lemma 2.11. *For $t > 0$ and $j \geqslant 1$, we define*

$$\Theta_j^t := \left\{ \ell \in \mathbb{N}_0 \,\big|\, 2^\ell \leqslant 2^{\alpha+2}(2^j + t2^{j(1+\alpha)}), \ 2^j \leqslant 2^{\alpha+2}(2^\ell + t2^{\ell(1+\alpha)}) \right\}.$$

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

$$\langle \mathcal{R}_j^{\boldsymbol{a}} f, \Gamma_t \mathcal{R}_\ell^{\boldsymbol{a}} g \rangle = 0.$$

(ii) For any $0 \neq \beta \in \mathbb{R}$, it holds that

$$\sum_{\ell \in \Theta_j^t} 2^{-\ell\beta} \lesssim 2^{-j\beta} (1 + (t2^{j\alpha}))^{|\beta|}. \tag{2.24}$$

Proof. (i) By Parsavel's identity of Fourier's transform, we have

$$\langle \mathcal{R}_j^{\boldsymbol{a}} f, \Gamma_t \mathcal{R}_\ell^{\boldsymbol{a}} g \rangle = \int_{\mathbb{R}^{2d}} \widehat{\mathcal{R}_j^{\boldsymbol{a}} f}(\xi_1, \xi_2) \overline{\widehat{\Gamma_t \mathcal{R}_\ell^{\boldsymbol{a}} g}}(\xi_1, \xi_2) \mathrm{d}\xi_1 \mathrm{d}\xi_2$$

$$= \int_{\mathbb{R}^{2d}} \phi_j^{\boldsymbol{a}}(\xi_1, \xi_2) f(\xi_1, \xi_2) \phi_\ell^{\boldsymbol{a}}(\xi_1, \xi_2 - t\xi_1) g(\xi_1, \xi_2 - t\xi_1) \mathrm{d}\xi_1 \mathrm{d}\xi_2.$$

Note that

$$\mathrm{supp} \phi_j^{\boldsymbol{a}} \subset \{\xi : 2^{j-1} \leqslant |\xi|_{\boldsymbol{a}} \leqslant 2^{j+1}\}.$$

For fixed $j \in \mathbb{N}$ and $t > 0$, it is easy to see that $\langle \mathcal{R}_j^{\boldsymbol{a}} f, \Gamma_t \mathcal{R}_\ell^{\boldsymbol{a}} g \rangle = 0$ if and only if

$$\{\xi : 2^{j-1} \leqslant |\xi|_{\boldsymbol{a}} \leqslant 2^{j+1}\} \cap \{\xi = (\xi_1, \xi_2) : 2^{\ell-1} \leqslant |(\xi_1, \xi_2 - t\xi_1)|_{\boldsymbol{a}} \leqslant 2^{\ell+1}\} = \emptyset,$$

which is equivalent to say $\ell \notin \Theta_j^t$.

(ii) Fix $\beta > 0$. Note that for $\ell \in \Theta_j^t$ with $\ell \leqslant j$,

$$2^{-\ell} \leqslant 2^{\alpha+2} 2^{-j} (1 + t2^{\ell\alpha}) \leqslant 2^{\alpha+2} 2^{-j} (1 + t2^{j\alpha}) =: D \Rightarrow \ell \geqslant -\ln D / \ln 2.$$

Hence,

$$\sum_{\ell \in \Theta_j^t} 2^{-\ell\beta} = \sum_{\ell \in \Theta_j^t; \ell \geqslant j} 2^{-\ell\beta} + \sum_{\ell \in \Theta_j^t; \ell < j} 2^{-\ell\beta} \leqslant \sum_{\ell \geqslant j} 2^{-\ell\beta} + \sum_{-\ln D/\ln 2 \leqslant \ell} 2^{-\ell\beta}$$

$$\leqslant \frac{2^{-j\beta}}{1 - 2^{-\beta}} + \frac{D^\beta}{1 - 2^{-\beta}} \lesssim 2^{-j\beta} (1 + t2^{j\alpha})^\beta.$$

Similarly, for $D := 2^{\alpha+2} 2^j (1 + t2^{j\alpha})$, we have

$$\sum_{\ell \in \Theta_j^t} 2^{\ell\beta} = \sum_{\ell \in \Theta_j^t; \ell \geqslant j} 2^{\ell\beta} + \sum_{\ell \in \Theta_j^t; \ell < j} 2^{\ell\beta} \leqslant \sum_{\ell \leqslant \ln D/\ln 2} 2^{\ell\beta} + \sum_{\ell < j} 2^{\ell\beta}$$

$$\lesssim D^\beta + 2^{j\beta} \lesssim 2^{j\beta} (1 + t2^{j\alpha})^\beta.$$

The proof is complete. □

Using the above two lemmas, we can show the following crucial lemma.

Lemma 2.12. *Let $\boldsymbol{p}, \boldsymbol{p}' \in [1, \infty]^2$ with $\boldsymbol{p} \leqslant \boldsymbol{p}'$. Write $\mathscr{A} := \boldsymbol{a} \cdot (\frac{d}{\boldsymbol{p}} - \frac{d}{\boldsymbol{p}'})$. Recall $\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^\beta := \mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta,\infty}$.*

*(i) (**High frequency part of** $P_t f$) For any $\beta \in \mathbb{R}$ and $\gamma \geqslant 0$, there is a constant $C = C(\alpha, d, \boldsymbol{p}, \boldsymbol{p}', \beta, \gamma) > 0$ such that for any $f \in \mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^\beta$ and all $j \geqslant 1$ and $t > 0$,*

$$\|\mathcal{R}_j^{\boldsymbol{a}} P_t f\|_{\boldsymbol{p}'} \lesssim_C 2^{j(\mathscr{A}-\beta)} ((2^{j\alpha}t)^{-\gamma} \wedge 1) \|f\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^\beta}. \tag{2.25}$$

*(ii) (**Low frequency part of** $P_t f$) For any $\beta \in \mathbb{R}$, there is a constant $C = C(\alpha, d, \boldsymbol{p}, \boldsymbol{p}', \beta) > 0$ such that for any $f \in \mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^\beta$ and all $t > 0$,*

$$\|\mathcal{R}_0^{\boldsymbol{a}} P_t f\|_{\boldsymbol{p}'} \lesssim_C (1 \vee t)^{-\frac{\mathscr{A}}{\alpha}} \|f\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^\beta}, \tag{2.26}$$

and for any $F \in (\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^\beta)^d$ and all $t > 0$,

$$\|\mathcal{R}_0^{\boldsymbol{a}} P_t \mathrm{div}_v F\|_{\boldsymbol{p}'} \lesssim_C (1 \vee t)^{-\frac{\mathscr{A}+1}{\alpha}} \|F\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^\beta}. \tag{2.27}$$

Proof. (i) By the interpolation theorem, we may assume $\beta \neq 0$. Note that by (i) of Lemma 2.11,

$$\mathscr{R}_j^{\boldsymbol{a}} P_t f = \mathscr{R}_j^{\boldsymbol{a}}\Big(\Gamma_t p_t * \Gamma_t f\Big) \overset{(2.8)}{=} \sum_{\ell \geqslant 0} \mathscr{R}_j^{\boldsymbol{a}}\Gamma_t p_t * \Gamma_t \mathscr{R}_\ell^{\boldsymbol{a}} f = \sum_{\ell \in \Theta_j^t} \mathscr{R}_j^{\boldsymbol{a}}\Gamma_t p_t * \Gamma_t \mathscr{R}_\ell^{\boldsymbol{a}} f. \tag{2.28}$$

By Bernstein's inequality (2.7) and Young's inequality, we have

$$\|\mathscr{R}_j^{\boldsymbol{a}} P_t f\|_{\boldsymbol{p}'} \lesssim 2^{j\mathscr{A}}\|\mathscr{R}_j^{\boldsymbol{a}} P_t f\|_{\boldsymbol{p}} \leqslant 2^{j\mathscr{A}}\|\mathscr{R}_j^{\boldsymbol{a}}\Gamma_t p_t\|_1 \sum_{\ell \in \Theta_j^t}\|\Gamma_t \mathscr{R}_\ell^{\boldsymbol{a}} f\|_{\boldsymbol{p}} \overset{(1.18)}{=} 2^{j\mathscr{A}}\|\mathscr{R}_j^{\boldsymbol{a}}\Gamma_t p_t\|_1 \sum_{\ell \in \Theta_j^t}\|\mathscr{R}_\ell^{\boldsymbol{a}} f\|_{\boldsymbol{p}}$$

$$\leqslant 2^{j\mathscr{A}}\|\mathscr{R}_j^{\boldsymbol{a}}\Gamma_t p_t\|_1 \sum_{\ell \in \Theta_j^t} 2^{-\ell\beta}\|f\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^\beta} \overset{(2.24)}{\lesssim} \|\mathscr{R}_j^{\boldsymbol{a}}\Gamma_t p_t\|_1 2^{j(\mathscr{A}-\beta)}(1+(t2^{j\alpha}))^{|\beta|}\|f\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^\beta},$$

which together with Lemma 2.8 with $l = \gamma$ and $l = \gamma + |\beta|$ yields (2.25).

(ii) We only prove (2.27). Noting that by definitions,

$$\mathscr{R}_0^{\boldsymbol{a}} P_t \mathrm{div}_v F = \check{\phi}_0^{\boldsymbol{a}} * \Gamma_t(p_t * \mathrm{div}_v F) = \Gamma_t(\Gamma_{-t}\check{\phi}_0^{\boldsymbol{a}} * (p_t * \mathrm{div}_v F)),$$

we have

$$\|\mathscr{R}_0^{\boldsymbol{a}} P_t \mathrm{div}_v F\|_{\boldsymbol{p}'} \overset{(1.18)}{=} \|\Gamma_{-t}\check{\phi}_0^{\boldsymbol{a}} * (p_t * \mathrm{div}_v F)\|_{\boldsymbol{p}'}. \tag{2.29}$$

By Young's inequality and integration by parts, we have for $\varepsilon > 0$,

$$\|\mathscr{R}_0^{\boldsymbol{a}} P_t \mathrm{div}_v F\|_{\boldsymbol{p}'} \leqslant \|\Gamma_{-t}\check{\phi}_0^{\boldsymbol{a}}\|_1 \|\nabla_v p_t * F\|_{\boldsymbol{p}'} \lesssim \|\check{\phi}_0^{\boldsymbol{a}}\|_1 \|\nabla_v p_t * F\|_{\mathbf{B}_{\boldsymbol{p}';\boldsymbol{a}}^\varepsilon},$$

which, by (2.13) and Lemma 2.10, implies that for $\boldsymbol{q} \in [1,\infty]$ being defined by $1 + \frac{1}{\boldsymbol{p}'} = \frac{1}{\boldsymbol{q}} + \frac{1}{\boldsymbol{p}}$ and $t \geqslant 1$,

$$\|\mathscr{R}_0^{\boldsymbol{a}} P_t \mathrm{div}_v F\|_{\boldsymbol{p}'} \lesssim \|\nabla_v p_t\|_{\mathbf{B}_{\boldsymbol{q};\boldsymbol{a}}^{\varepsilon-\beta}}\|F\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^\beta} \lesssim t^{\frac{\boldsymbol{a}\cdot(d/\boldsymbol{q}-\boldsymbol{d})-1}{\alpha}}\|F\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^\beta} = t^{-\frac{\mathscr{A}+1}{\alpha}}\|F\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^\beta}. \tag{2.30}$$

Moreover, for $t \in (0,1]$ and $\varepsilon > 0$, by (2.29) and (2.13),

$$\|\mathscr{R}_0^{\boldsymbol{a}} P_t \mathrm{div}_v F\|_{\boldsymbol{p}'} = \|p_t * (\Gamma_{-t}\check{\phi}_0^{\boldsymbol{a}} * \mathrm{div}_v F)\|_{\boldsymbol{p}'} \leqslant \|p_t\|_1 \|\Gamma_{-t}\check{\phi}_0^{\boldsymbol{a}} * \mathrm{div}_v F\|_{\boldsymbol{p}'}$$

$$= \|(\nabla_v \Gamma_{-t}\check{\phi}_0^{\boldsymbol{a}}) * F\|_{\boldsymbol{p}'} \lesssim \|(\nabla_v \Gamma_{-t}\check{\phi}_0^{\boldsymbol{a}}) * F\|_{\mathbf{B}_{\boldsymbol{p}';\boldsymbol{a}}^\varepsilon}$$

$$\lesssim \|\nabla_v \Gamma_{-t}\check{\phi}_0^{\boldsymbol{a}}\|_{\mathbf{B}_{\boldsymbol{q};\boldsymbol{a}}^{\varepsilon-\beta}}\|F\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^\beta}.$$

Since $\check{\phi}_0^{\boldsymbol{a}} \in \mathscr{S}$, by the chain rule, we have

$$\sup_{t \in [0,1]}\|\nabla_v \Gamma_{-t}\check{\phi}_0^{\boldsymbol{a}}\|_{\mathbf{B}_{\boldsymbol{q};\boldsymbol{a}}^{\varepsilon-\beta}} \lesssim \sup_{t \in [0,1]} \sum_{j \leqslant |\varepsilon-\beta|+1}\|\nabla_v^j \Gamma_{-t}\check{\phi}_0^{\boldsymbol{a}}\|_{\boldsymbol{q}} < \infty.$$

Hence,

$$\sup_{t \in [0,1]}\|\mathscr{R}_0^{\boldsymbol{a}} P_t \mathrm{div}_v F\|_{\boldsymbol{p}'} \lesssim \|F\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^\beta},$$

which together with (2.30) yields (2.27). $\qquad\square$

Remark 2.13. The differentiability index β in (2.26) and (2.27) does not make any contribution about the decay rate of large time of $\|\mathscr{R}_0^{\boldsymbol{a}} P_t \mathrm{div}_v F\|_{\boldsymbol{p}'}$. This is natural since we are using the non-homoegenous norm in the right hand side.

Fix $\alpha \in (0,2]$. For given $\boldsymbol{\gamma} = (\gamma_0, \gamma_1) \in \mathbb{R}^2$, $q \in [1,\infty]$ and a Banach space $\mathbb{B}$, let $\mathbb{L}_{\boldsymbol{\gamma}}^q(\mathbb{B})$ be the space of all functions $f : (0,\infty) \to \mathbb{B}$ with finite norm

$$\|f\|_{\mathbb{L}_{\boldsymbol{\gamma}}^q(\mathbb{B})} := \left(\int_0^\infty \big(\|f(t)\|_{\mathbb{B}}(1 \wedge t)^{\frac{\gamma_0}{\alpha}}(1 \vee t)^{\frac{\gamma_1}{\alpha}}\big)^q \mathrm{d}t\right)^{1/q} < \infty.$$

If γ_0, γ_1 are positive, then γ_0 and γ_1 stand for the small time singularity and large time decay weights, respectively.

We shall use the following relation for two weights $\boldsymbol{\gamma}, \boldsymbol{\gamma}' \in \mathbb{R}^2$.

Definition 2.14. *For $\boldsymbol{\gamma} = (\gamma_0, \gamma_1)$ and $\boldsymbol{\gamma}' = (\gamma_0', \gamma_1') \in \mathbb{R}^2$, one says $\boldsymbol{\gamma}' \gtreqqless \boldsymbol{\gamma}$ if $\gamma_0' \geqslant \gamma_0$ and $\gamma_1' \leqslant \gamma_1$, and $\boldsymbol{\gamma}' \gtreqless \boldsymbol{\gamma}$ if $\gamma_0' \geqslant \gamma_0$ and $\gamma_1' < \gamma_1$. In particular, if $\boldsymbol{\gamma}' \gtreqqless \boldsymbol{\gamma}$, then*

$$\|f\|_{\mathbb{L}^q_{\boldsymbol{\gamma}'}(\mathbb{B})} \leqslant \|f\|_{\mathbb{L}^q_{\boldsymbol{\gamma}}(\mathbb{B})}. \tag{2.31}$$

The following proposition is simple, but quite useful for proving the large time decay estimates and the smoothness of solutions by shifting the time variable (see Lemma 3.9 below).

Proposition 2.15. *For any $\boldsymbol{\gamma}, \boldsymbol{\gamma}' \in [0, \infty)^2$, it holds that*

$$\|f(t + \cdot)\|_{\mathbb{L}^q_{\boldsymbol{\gamma}'}(\mathbb{B})} \leqslant (1 \wedge t)^{-\frac{\gamma_0}{\alpha}} (1 \vee t)^{-\frac{\gamma_1}{\alpha}} \|f\|_{\mathbb{L}^q_{\boldsymbol{\gamma}+\boldsymbol{\gamma}'}(\mathbb{B})}, \quad \forall t > 0. \tag{2.32}$$

Proof. Noting that for $\boldsymbol{\gamma} = (\gamma_0, \gamma_1), \boldsymbol{\gamma}' = (\gamma_0', \gamma_1') \in [0, \infty)^2$ and $s, t > 0$,

$$(1 \wedge s)^{\frac{\gamma_0'}{\alpha}} \leqslant (1 \wedge (t + s))^{\frac{\gamma_0 + \gamma_0'}{\alpha}} (1 \wedge t)^{-\frac{\gamma_0}{\alpha}}$$

and

$$(1 \vee s)^{\frac{\gamma_0'}{\alpha}} \leqslant (1 \vee (t + s))^{\frac{\gamma_0 + \gamma_0'}{\alpha}} (1 \vee t)^{-\frac{\gamma_0}{\alpha}},$$

by definition, we have

$$\|f(t + \cdot)\|_{\mathbb{L}^q_{\boldsymbol{\gamma}'}(\mathbb{B})} = \left(\int_0^\infty \left(\|f(t + s)\|_{\mathbb{B}} (1 \wedge s)^{\frac{\gamma_0'}{\alpha}} (1 \vee s)^{\frac{\gamma_1'}{\alpha}} \right)^q \mathrm{d}s \right)^{1/q}$$

$$\leqslant \left(\int_t^\infty \left(\|f(s)\|_{\mathbb{B}} (1 \wedge s)^{\frac{\gamma_0 + \gamma_0'}{\alpha}} (1 \vee s)^{\frac{\gamma_1 + \gamma_1'}{\alpha}} \right)^q \mathrm{d}s \right)^{1/q} (1 \wedge t)^{-\frac{\gamma_0}{\alpha}} (1 \vee t)^{-\frac{\gamma_1}{\alpha}}$$

$$\leqslant (1 \wedge t)^{-\frac{\gamma_0}{\alpha}} (1 \vee t)^{-\frac{\gamma_1}{\alpha}} \|f\|_{\mathbb{L}^q_{\boldsymbol{\gamma}+\boldsymbol{\gamma}'}(\mathbb{B})}.$$

The proof is complete. $\qquad\square$

The following lemma provides useful estimates about the kinetic semigroup P_t.

Lemma 2.16. *Let $\boldsymbol{p}_0, \boldsymbol{p}_1, \boldsymbol{p}' \in [1, \infty]^2$ with $\boldsymbol{p}_0, \boldsymbol{p}_1 \leqslant \boldsymbol{p}'$. Write $\mathscr{A}_i := \boldsymbol{a} \cdot (\frac{d}{\boldsymbol{p}_i} - \frac{d}{\boldsymbol{p}'})$, $i = 0, 1$. For any $\beta, \beta' \in \mathbb{R}$, there are constants $C_i = C_i(\alpha, d, \beta, \beta', \boldsymbol{p}_i, \boldsymbol{p}') > 0$, $i = 0, 1$ such that for all $t > 0$,*

$$\mathbb{1}_{\beta' \neq \beta - \mathscr{A}_0} \|P_t f\|_{\mathbf{B}^{\beta',1}_{\boldsymbol{p}';\boldsymbol{a}}} + \|P_t f\|_{\mathbf{B}^{\beta'}_{\boldsymbol{p}';\boldsymbol{a}}} \lesssim_{C_0} (1 \wedge t)^{-\frac{(\beta'-\beta+\mathscr{A}_0)\vee 0}{\alpha}} \|f\|_{\mathbf{B}^{\beta}_{\boldsymbol{p}_0;\boldsymbol{a}}}, \tag{2.33}$$

where $\mathbf{B}^{\beta}_{\boldsymbol{p};\boldsymbol{a}} = \mathbf{B}^{\beta,\infty}_{\boldsymbol{p};\boldsymbol{a}}$, and for $t \geqslant 1$,

$$\|P_t f\|_{\mathbf{B}^{\beta',1}_{\boldsymbol{p}';\boldsymbol{a}}} \lesssim_{C_1} t^{-\frac{\mathscr{A}_1}{\alpha}} \|f\|_{\mathbf{B}^{\beta}_{\boldsymbol{p}_1;\boldsymbol{a}}}. \tag{2.34}$$

Proof. (i) Let $\beta_0 := \beta' - \beta + \mathscr{A}_0 \neq 0$. By (2.26) and (2.25) with $\gamma > \frac{\beta_0 \vee 0}{\alpha}$, we have for all $t > 0$,

$$\|P_t f\|_{\mathbf{B}^{\beta',1}_{\boldsymbol{p}';\boldsymbol{a}}} = \|\mathscr{R}^{\boldsymbol{a}}_0 P_t f\|_{\boldsymbol{p}'} + \sum_{j \geqslant 1} 2^{j\beta'} \|\mathscr{R}^{\boldsymbol{a}}_j P_t f\|_{\boldsymbol{p}'}$$

$$\lesssim \|f\|_{\mathbf{B}^{\beta}_{\boldsymbol{p}_0;\boldsymbol{a}}} + \sum_{j \geqslant 1} 2^{j\beta_0} ((2^{j\alpha} t)^{-\gamma} \wedge 1) \|f\|_{\mathbf{B}^{\beta}_{\boldsymbol{p}_0;\boldsymbol{a}}}.$$

Since $\beta_0 \neq 0$ and $\gamma\alpha > \beta_0 \vee 0$, by Lemma A.1 we have

$$\sum_{j \geqslant 1} 2^{j\beta_0} ((2^{j\alpha} t)^{-\gamma} \wedge 1) \lesssim t^{-\frac{\beta_0 \vee 0}{\alpha}}, \ t > 0.$$

Thus we get

$$\|P_t f\|_{\mathbf{B}^{\beta',1}_{\boldsymbol{p}';\boldsymbol{a}}} \lesssim \|f\|_{\mathbf{B}^{\beta}_{\boldsymbol{p}_0;\boldsymbol{a}}} + t^{-\frac{\beta_0 \vee 0}{\alpha}} \|f\|_{\mathbf{B}^{\beta}_{\boldsymbol{p}_0;\boldsymbol{a}}} \leqslant 2(1 \wedge t)^{-\frac{(\beta'-\beta+\mathscr{A}_0)\vee 0}{\alpha}} \|f\|_{\mathbf{B}^{\beta}_{\boldsymbol{p}_0;\boldsymbol{a}}}.$$

If $\beta_0 = 0$, we directly have

$$\|P_t f\|_{\mathbf{B}^{\beta'}_{\boldsymbol{p}';\boldsymbol{a}}} \leqslant \sup_{j \geqslant 0} 2^{j\beta'} \|\mathscr{R}^{\boldsymbol{a}}_j P_t f\|_{\boldsymbol{p}'} \lesssim \|f\|_{\mathbf{B}^{\beta}_{\boldsymbol{p}_0;\boldsymbol{a}}}.$$

Combining the above two estimates, we obtain (2.33).

(ii) Since $\mathscr{A}_1 = \boldsymbol{a} \cdot (\frac{\boldsymbol{d}}{\boldsymbol{p}_1} - \frac{d}{\boldsymbol{p}'})$, by (2.26) and (2.25) with $\gamma > \frac{\mathscr{A}_1}{\alpha}$ large enough, we have for $t \geqslant 1$,

$$\|P_t f\|_{\mathbf{B}^{\beta',1}_{\boldsymbol{p}';\boldsymbol{a}}} = \|\mathscr{R}^{\boldsymbol{a}}_0 P_t f\|_{\boldsymbol{p}'} + \sum_{j \geqslant 1} 2^{j\beta'} \|\mathscr{R}^{\boldsymbol{a}}_j P_t f\|_{\boldsymbol{p}'}$$
$$\lesssim t^{-\frac{\mathscr{A}_1}{\alpha}} \|f\|_{\mathbf{B}^{\beta}_{\boldsymbol{p}_1;\boldsymbol{a}}} + t^{-\gamma} \sum_{j \geqslant 1} 2^{j(\beta' + \mathscr{A}_1 - \alpha\gamma)} \|f\|_{\mathbf{B}^{\beta}_{\boldsymbol{p}_1;\boldsymbol{a}}} \lesssim t^{-\frac{\mathscr{A}_1}{\alpha}} \|f\|_{\mathbf{B}^{\beta}_{\boldsymbol{p}_1;\boldsymbol{a}}}.$$

The proof is complete. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Remark 2.17. Estimates (2.33) and (2.34) provide short time and large time estimates of the kinetic semigroup, respectively. In particular, the rate of large time decay only depends on the integrability indices. Note that

$$\|P_t f\|_{\boldsymbol{p}} \overset{(2.12)}{\lesssim} \|P_t f\|_{\mathbf{B}^{0,1}_{\boldsymbol{p};\boldsymbol{a}}} \overset{(2.34)}{\lesssim} \|f\|_{\mathbf{B}^{0,\infty}_{\boldsymbol{p};\boldsymbol{a}}} \overset{(2.12)}{\lesssim}_C \|f\|_{\boldsymbol{p}}, \quad t \geqslant 1.$$

It is the same reason that (2.33) is stronger than the classical estimate:

$$\|\nabla_v P_t f\|_{\boldsymbol{p}} \lesssim (1 \wedge t)^{-\frac{1}{\alpha}} \|f\|_{\boldsymbol{p}}.$$

Moreover, for any $\beta, \beta' \in \mathbb{R}$ and $\boldsymbol{p} \in [1, \infty]^2$, noting that by (2.33),

$$\|P_t f\|_{\mathbf{B}^{\beta+\beta'}_{\boldsymbol{p};\boldsymbol{a}}} \lesssim (1 \wedge t)^{-\frac{\beta' \vee 0}{\alpha}} \|f\|_{\mathbf{B}^{\beta}_{\boldsymbol{p};\boldsymbol{a}}}, \quad t > 0,$$

by the real interpolation theorem (see [8, Theorem 6.4.5, (1)]), we have for any $q \in [1, \infty]$,

$$\|P_t f\|_{\mathbf{B}^{\beta+\beta',q}_{\boldsymbol{p};\boldsymbol{a}}} \lesssim_C (1 \wedge t)^{-\frac{\beta' \vee 0}{\alpha}} \|f\|_{\mathbf{B}^{\beta,q}_{\boldsymbol{p};\boldsymbol{a}}}, \quad t > 0. \tag{2.35}$$

For $f \in \mathcal{S}'$ and $t > 0$, let

$$u(t, x, v) := \mathscr{I}_t f(x, v) := \int_0^t P_{t-s} f(s, x, v) \mathrm{d}s.$$

By the definition (2.22) of P_t, it is easy to see that in the distributional sense,

$$\partial_t u = \Delta_v^{\alpha/2} u - v \cdot \nabla_x u + f.$$

We now show the following basic Schauder estimate in time-weighted spaces.

Theorem 2.18. (Time-weighted Schauder's estimate) Let $\alpha \in (0, 2)$, $q \in (1, \infty]$ and $\boldsymbol{p}, \boldsymbol{p}' \in [1, \infty]^2$ with $\boldsymbol{p} \leqslant \boldsymbol{p}'$. Write

$$\mathscr{Q} := \alpha - \frac{\alpha}{q}, \quad \mathscr{A} := \boldsymbol{a} \cdot (\frac{d}{\boldsymbol{p}} - \frac{d}{\boldsymbol{p}'}).$$

Recall $\mathbf{B}^{\beta}_{\boldsymbol{p};\boldsymbol{a}} := \mathbf{B}^{\beta,\infty}_{\boldsymbol{p};\boldsymbol{a}}$. Suppose that $\theta \in (0, \mathscr{Q})$ and $\boldsymbol{\gamma} = (\gamma_0, \gamma_1)$, $\boldsymbol{\gamma}' = (\gamma'_0, \gamma'_1) \in \mathbb{R}^2$ satisfy

$$\boldsymbol{\gamma} \in [0, \mathscr{Q})^2, \quad \boldsymbol{\gamma}' \gtrless \boldsymbol{\gamma} - \big(\theta, (\mathscr{Q} - 1 - \mathscr{A}) \vee 0\big), \tag{2.36}$$

see Definition 2.14 for notation $\gtrless$. Let $\Theta := (\alpha, d, q, \boldsymbol{p}, \boldsymbol{p}', \boldsymbol{\gamma}, \boldsymbol{\gamma}', \theta)$ be the parameter set and

$$\beta' := \mathscr{Q} - 1 - \mathscr{A} - \theta.$$

For any $\beta \in \mathbb{R}$, there is a constant $C = C(\beta, \Theta) > 0$ such that for any $F \in (\mathbb{L}^q_{\boldsymbol{\gamma}}(\mathbf{B}^{\beta}_{\boldsymbol{p};\boldsymbol{a}}))^d$,

$$\|\mathscr{I} \mathrm{div}_v F\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}'}(\mathbf{B}^{\beta+\beta',1}_{\boldsymbol{p}';\boldsymbol{a}})} \lesssim_C \|F\|_{\mathbb{L}^q_{\boldsymbol{\gamma}}(\mathbf{B}^{\beta}_{\boldsymbol{p};\boldsymbol{a}})}. \tag{2.37}$$

Moreover, if $\theta = 0$ or $\mathscr{Q}$, then

$$\|\mathscr{I} \mathrm{div}_v F\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}'}(\mathbf{B}^{\beta+\beta'}_{\boldsymbol{p}';\boldsymbol{a}})} \lesssim_C \|F\|_{\mathbb{L}^q_{\boldsymbol{\gamma}}(\mathbf{B}^{\beta}_{\boldsymbol{p};\boldsymbol{a}})}. \tag{2.38}$$

Proof. First of all, by (2.27), Hölder's inequality and Lemma A.2, we have

$$\|\mathscr{R}_0^{\boldsymbol{a}}\mathscr{I}_t\mathrm{div}_vF\|_{\boldsymbol{p}'} \lesssim \int_0^t \|\mathscr{R}_0^{\boldsymbol{a}}P_{t-s}\mathrm{div}_vF_s\|_{\boldsymbol{p}'}\mathrm{d}s \lesssim \int_0^t (1\vee(t-s))^{-\frac{\mathscr{A}+1}{\alpha}}\|F_s\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta}}\mathrm{d}s$$

$$\lesssim \left(\int_0^t \left((1\wedge(t-s)^{-\frac{\mathscr{A}+1}{\alpha}})(1\wedge s)^{-\frac{\gamma_0}{\alpha}}(1\vee s)^{-\frac{\gamma_1}{\alpha}}\right)^r \mathrm{d}s\right)^{\frac{1}{r}}\|F\|_{\mathbb{L}_{\boldsymbol{\gamma}}^q(\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta})}$$

$$\lesssim (1\wedge t)^{\frac{1}{r}-\frac{\gamma_0}{\alpha}}(1\vee t)^{\frac{1}{r}-\frac{\gamma_1}{\alpha}}\ell_\vartheta^{1/r}(t)\|F\|_{\mathbb{L}_{\boldsymbol{\gamma}}^q(\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta})}, \tag{2.39}$$

where $\vartheta := r(\mathscr{A}+1)/\alpha$ and

$$\ell_\vartheta(t) := (1\vee t)^{-\vartheta\wedge 1}(1+\mathbb{1}_{\vartheta=1}\ln(1\vee t)).$$

Noting that by (2.36) and $\frac{\alpha}{r} = \alpha - \frac{\alpha}{q} = \mathscr{Q}$,

$$\alpha(\tfrac{1}{r}-\tfrac{\gamma_1}{\alpha}-\tfrac{\mathscr{A}+1}{\alpha}\wedge\tfrac{1}{r}) = \mathscr{Q}-\gamma_1-(\mathscr{A}+1)\wedge\mathscr{Q} = (\mathscr{Q}-1-\mathscr{A})\vee 0 - \gamma_1 < -\gamma_1',$$

we have

$$(1\vee t)^{\frac{1}{r}-\frac{\gamma_1}{\alpha}}\ell_\vartheta^{1/r}(t) = (1\vee t)^{\frac{(\mathscr{Q}-1-\mathscr{A})\vee 0-\gamma_1}{\alpha}}(1+\mathbb{1}_{\vartheta=1}\ln(1\vee t))^{\frac{1}{r}} \lesssim (1\vee t)^{-\frac{\gamma_1'}{\alpha}}.$$

Substituting this into (2.39), we obtain

$$\|\mathscr{R}_0^{\boldsymbol{a}}\mathscr{I}_t\mathrm{div}_vF\|_{\boldsymbol{p}'} \lesssim (1\wedge t)^{\frac{\mathscr{Q}-\gamma_0}{\alpha}}(1\vee t)^{-\frac{\gamma_1'}{\alpha}}\|F\|_{\mathbb{L}_{\boldsymbol{\gamma}}^q(\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta})}. \tag{2.40}$$

Next, let $\frac{1}{q}+\frac{1}{r}=1$. By (2.25), Hölder's inequality and Lemma A.2, we have for any $j\geqslant 1$,

$$\|\mathscr{R}_j^{\boldsymbol{a}}\mathscr{I}_t\mathrm{div}_vF\|_{\boldsymbol{p}'} \leqslant \int_0^t \|\mathscr{R}_j^{\boldsymbol{a}}P_{t-s}\mathrm{div}_vF_s\|_{\boldsymbol{p}'}\mathrm{d}s \lesssim 2^{j(\mathscr{A}+1-\beta)}\int_0^t \left[((t-s)2^{j\alpha})^{-2}\wedge 1\right]\|\mathrm{div}_vF_s\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta-1}}\mathrm{d}s$$

$$\leqslant 2^{j(\mathscr{A}+1-\beta)}\left[\int_0^t \left([((t-s)2^{j\alpha})^{-2}\wedge 1](1\wedge s)^{-\frac{\gamma_0}{\alpha}}(1\vee s)^{-\frac{\gamma_1}{\alpha}}\right)^r \mathrm{d}s\right]^{\frac{1}{r}}\|F\|_{\mathbb{L}_{\boldsymbol{\gamma}}^q(\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta})}$$

$$\lesssim 2^{j(\mathscr{A}+1-\beta)}((2^{j\alpha}t)^{-\frac{1}{r}}\wedge 1)(1\wedge t)^{\frac{1}{r}-\frac{\gamma_0}{\alpha}}(1\vee t)^{\frac{1}{r}-\frac{\gamma_1}{\alpha}}\|F\|_{\mathbb{L}_{\boldsymbol{\gamma}}^q(\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta})}. \tag{2.41}$$

Since $\frac{\alpha}{r}=\mathscr{Q}$, for any $\theta\in(0,\mathscr{Q})$, by Lemma A.1 we get

$$\sum_{j\geqslant 1} 2^{j(\beta+\mathscr{Q}-1-\mathscr{A}-\theta)}\|\mathscr{R}_j^{\boldsymbol{a}}\mathscr{I}_t\mathrm{div}_vF\|_{\boldsymbol{p}'}$$

$$\lesssim \sum_{j\geqslant 1} 2^{j(\mathscr{Q}-\theta)}((2^{j\alpha}t)^{-\frac{1}{r}}\wedge 1)(1\wedge t)^{\frac{1}{r}-\frac{\gamma_0}{\alpha}}(1\vee t)^{\frac{1}{r}-\frac{\gamma_1}{\alpha}}\|F\|_{\mathbb{L}_{\boldsymbol{\gamma}}^q(\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta})}$$

$$\lesssim t^{-\frac{\mathscr{Q}-\theta}{\alpha}}(1\wedge t)^{\frac{\mathscr{Q}-\gamma_0}{\alpha}}(1\vee t)^{\frac{\mathscr{Q}-\gamma_1}{\alpha}}\|F\|_{\mathbb{L}_{\boldsymbol{\gamma}}^q(\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta})}$$

$$= (1\wedge t)^{\frac{\theta-\gamma_0}{\alpha}}(1\vee t)^{\frac{\theta-\gamma_1}{\alpha}}\|F\|_{\mathbb{L}_{\boldsymbol{\gamma}}^q(\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta})}.$$

Since $\gamma_1'<\gamma_1$, one can choose $\theta'\in(0,\theta]$ so that $\theta'-\gamma_1\leqslant -\gamma_1'$ and for $t\geqslant 1$,

$$\sum_{j\geqslant 1} 2^{j(\beta+\mathscr{Q}-1-\mathscr{A}-\theta')}\|\mathscr{R}_j^{\boldsymbol{a}}\mathscr{I}_t\mathrm{div}_vF\|_{\boldsymbol{p}'} \lesssim t^{\frac{\theta'-\gamma_1}{\alpha}}\|F\|_{\mathbb{L}_{\boldsymbol{\gamma}}^q(\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta})} \leqslant t^{-\frac{\gamma_1'}{\alpha}}\|F\|_{\mathbb{L}_{\boldsymbol{\gamma}}^q(\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta})}.$$

Thus, we further have

$$\sum_{j\geqslant 1} 2^{j(\beta+\mathscr{Q}-1-\mathscr{A}-\theta)}\|\mathscr{R}_j^{\boldsymbol{a}}\mathscr{I}_t\mathrm{div}_vF\|_{\boldsymbol{p}'} \lesssim (1\wedge t)^{\frac{\theta-\gamma_0}{\alpha}}(1\vee t)^{-\frac{\gamma_1'}{\alpha}}\|F\|_{\mathbb{L}_{\boldsymbol{\gamma}}^q(\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta})},$$

which, together with (2.40) and $\theta<\mathscr{Q}$, yields

$$\|\mathscr{I}_t\mathrm{div}_vF\|_{\mathbf{B}_{\boldsymbol{p}';\boldsymbol{a}}^{\beta+\mathscr{Q}-1-\mathscr{A}-\theta,1}} = \|\mathscr{R}_0^{\boldsymbol{a}}\mathscr{I}_t\mathrm{div}_vF\|_{\boldsymbol{p}'} + \sum_{j\geqslant 1} 2^{j(\beta+\mathscr{Q}-1-\mathscr{A}-\theta)}\|\mathscr{R}_j^{\boldsymbol{a}}\mathscr{I}_t\mathrm{div}_vF\|_{\boldsymbol{p}'}$$

$$\lesssim (1\wedge t)^{\frac{\theta-\gamma_0}{\alpha}}(1\vee t)^{-\frac{\gamma_1'}{\alpha}}\|F\|_{\mathbb{L}_{\boldsymbol{\gamma}}^q(\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta})}.$$

Moreover, by (2.41) we directly have

$$\|\mathcal{R}_j^{\boldsymbol{a}} \mathscr{I}_t \mathrm{div}_v F\|_{\boldsymbol{p}'} \lesssim \begin{cases} 2^{j(\mathscr{A}+1-\beta-\mathscr{Q})}(1 \wedge t)^{-\frac{\gamma_0}{\alpha}}(1 \vee t)^{-\frac{\gamma_1}{\alpha}}\|F\|_{\mathbb{L}_{\boldsymbol{\gamma}}^q(\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta})}, \\ 2^{j(\mathscr{A}+1-\beta)}(1 \wedge t)^{\frac{\mathscr{Q}-\gamma_0}{\alpha}}(1 \vee t)^{\frac{\mathscr{Q}-\gamma_1}{\alpha}}\|F\|_{\mathbb{L}_{\boldsymbol{\gamma}}^q(\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta})}, \end{cases}$$

which together with (2.40) gives (2.38) for $\theta = 0$ and $\theta = \mathscr{Q}$ by definition. $\qquad\square$

Remark 2.19. From the proof, one sees that if $\mathscr{A} \neq \mathscr{Q} - 1$, then $\vartheta \neq 1$ and we can take

$$\boldsymbol{\gamma}' = \boldsymbol{\gamma} - \big(\theta, (\mathscr{Q} - 1 - \mathscr{A}) \vee 0\big).$$

Remark 2.20. About the regularity index $\beta' = \alpha - \frac{\alpha}{q} - 1 - \mathscr{A} - \theta$ in (2.37), α is the gain from the kinetic semigroup, and $(\frac{\alpha}{q}, 1, \mathscr{A}, \theta)$ are, respectively, the loss from the time integrablility of f, the divergence, Sobolev's embedding and the singularity of small time weight.

3. Well-posedness of nonlinear kinetic Fokker-Planck equation

In this section we aim to establish the global well-posedness to the following Fokker-Planck equation with small initial value:

$$\partial_t u = (\Delta_v^{\frac{\alpha}{2}} - v \cdot \nabla_x)u - \mathrm{div}_v((b * u)u), \ u(0) = u_0. \tag{3.1}$$

Firstly we introduce the definition of weak solutions to the above equation.

Definition 3.1. *Let $u_0 \in \mathcal{S}'$ and $u \in L_{loc}^1([0, \infty); L^p(\mathbb{R}^{2d}))$ for some $p \in [1, \infty]$. We call u a weak solution of* (3.1) *with initial value u_0 if*

$$(b * u)u \in L_{loc}^1([0, \infty); \mathcal{S}'),$$

and for any $\varphi \in C_c^\infty(\mathbb{R}^{2d})$ and $t > 0$,

$$\langle u_t, \varphi \rangle = \langle u_0, \varphi \rangle + \int_0^t \Big[\langle u_s, (\Delta_v^{\alpha/2} + v \cdot \nabla_x)\varphi \rangle + \langle (b * u_s)u_s, \nabla_v \varphi \rangle \Big] \mathrm{d}s, \tag{3.2}$$

where $\langle u_t, \varphi \rangle := \int_{\mathbb{R}^{2d}} u_t(z)\varphi(z)\mathrm{d}z = \int u_t \varphi$.

Remark 3.2. For $\alpha \in (0, 2)$, note that up to a constant $c_{d,\alpha} > 0$ (see [69]),

$$\Delta_v^{\alpha/2} u(v) = \lim_{\varepsilon \downarrow 0} \int_{|w| \geqslant \varepsilon} \frac{u(v+w) - u(v)}{|w|^{d+\alpha}} \mathrm{d}w = \frac{1}{2} \int_{\mathbb{R}^d} \frac{u(v+w) + u(v-w) - 2u(v)}{|w|^{d+\alpha}} \mathrm{d}w. \tag{3.3}$$

From this representation, it is easy to see that for $\varphi \in C_c^\infty(\mathbb{R}^{2d})$,

$$\Delta_v^{\alpha/2} \varphi \in (L^1 \cap L^\infty)(\mathbb{R}^{2d}).$$

In particular, $\int_0^t \langle u_s, \Delta_v^{\alpha/2} \varphi \rangle \mathrm{d}s$ is well-defined for $u \in L_{loc}^1([0, \infty); L^p(\mathbb{R}^{2d}))$. Moreover, by Duhamel's formula, (3.2) is equivalent to the following integral equation

$$u_t = P_t u_0 - \int_0^t P_{t-s} \mathrm{div}_v((b_s * u_s)u_s)\mathrm{d}s. \tag{3.4}$$

3.1. **Global well-posedness of small initial value.** In this subsection we prove a basic result about the global well-posedness for (3.1) with small initial data based on the contraction mapping theorem. First of all, we introduce the main indices and their assumptions.

(H$_0$) Let $\alpha \in (1, 2]$, $q_b \in (\frac{\alpha}{\alpha-1}, \infty]$ and $\boldsymbol{p}_0, \boldsymbol{\varrho}_0, \boldsymbol{\varrho}_1 \in [1, \infty]^2$ with $\boldsymbol{\varrho}_0 \leqslant \boldsymbol{\varrho}_1$ and $\mathbf{1} \leqslant \frac{1}{\boldsymbol{p}_0} + \frac{1}{\boldsymbol{\varrho}_0}$. Let $\bar{\boldsymbol{\varrho}}_0, \bar{\boldsymbol{\varrho}}_1 \in [1, \infty]^2$ be defined by $\frac{1}{\boldsymbol{\varrho}_0} + \frac{1}{\bar{\boldsymbol{\varrho}}_0} = \mathbf{1}$ and $\frac{1}{\boldsymbol{\varrho}_1} + \frac{1}{\bar{\boldsymbol{\varrho}}_1} = \mathbf{1}$. Let $\boldsymbol{p}_1 := \boldsymbol{p}_0 \wedge \bar{\boldsymbol{\varrho}}_1$ and

$$\mathscr{Q} := \alpha - \frac{\alpha}{q_b}, \quad \mathscr{A}_i := \boldsymbol{a} \cdot (\frac{d}{\boldsymbol{p}_i} + \frac{d}{\boldsymbol{\varrho}_i} - \boldsymbol{d}) = \boldsymbol{a} \cdot (\frac{d}{\boldsymbol{p}_i} - \frac{d}{\bar{\boldsymbol{\varrho}}_i}), \ i = 0, 1.$$

Let $\beta_0 \in (-\mathscr{Q}, 0)$ and $\beta_b \in \mathbb{R}$. Suppose that $\beta_0 \neq \mathscr{A}_i - \beta_b, i = 0, 1$ and

$$0 < \mathscr{A}_0 - \beta_0 - \beta_b \leqslant \mathscr{Q} - 1 \quad \text{and} \quad \mathscr{A}_0 - 2\beta_0 - \beta_b < \mathscr{Q}. \tag{3.5}$$

and $\boldsymbol{\gamma}_b := (\gamma_{b0}, \gamma_{b1})$,

$$\gamma_{b0} = 0, \quad \gamma_{b1} \in ((\mathcal{Q} - 1 - \mathscr{A}_1) \vee 0, (\mathcal{Q} - \mathscr{A}_1) \vee 1). \tag{3.6}$$

Remark 3.3. Note that $(\beta_0, \boldsymbol{p}_0, \boldsymbol{p}_1)$ is referred to the regularity of initial data u_0 and $(q_b, \gamma_b, \beta_b, \boldsymbol{\varrho}_0, \boldsymbol{\varrho}_1)$ is referred to the regularity of kernel function b. Since $\boldsymbol{\varrho}_0 \leqslant \boldsymbol{\varrho}_1$ and $\frac{1}{\boldsymbol{p}_0} \geqslant 1 - \frac{1}{\boldsymbol{\varrho}_0} = \frac{1}{\bar{\boldsymbol{\varrho}}_0}$, we have

$$\frac{1}{\boldsymbol{p}_1} - \frac{1}{\bar{\boldsymbol{\varrho}}_1} = \frac{1}{\boldsymbol{p}_0 \wedge \bar{\boldsymbol{\varrho}}_1} - \frac{1}{\bar{\boldsymbol{\varrho}}_1} \leqslant \frac{d}{\boldsymbol{p}_0} - \frac{d}{\bar{\boldsymbol{\varrho}}_0} \Rightarrow \mathscr{A}_1 \leqslant \mathscr{A}_0.$$

Moreover, condition (3.5) is equivalent to the following conditions:

$$\begin{cases} 0 < \mathscr{A}_0 - \beta_0 - \beta_b \leqslant \mathcal{Q} - 1, & \beta_0 \in (-1, 0), \\ 0 < \mathscr{A}_0 - \beta_0 - \beta_b < \mathcal{Q} + \beta_0, & \beta_0 \in (-\mathcal{Q}, -1]. \end{cases}$$

For simplicity of notations, in the following we shall use the following parameter set

$$\Theta := \left(\alpha, d, \beta_0, \boldsymbol{p}_0, q_b, \gamma_b, \beta_b, \boldsymbol{\varrho}_0, \boldsymbol{\varrho}_1\right). \tag{3.7}$$

Recalling $\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta} := \mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta,\infty}$, with the parameters in $(\mathbf{H}_0)$, we introduce three basic spaces for later use: for $\beta \geqslant 0$,

$$\mathbb{X}_\beta := \bigcap_{i=0,1} \mathbf{B}_{\bar{\boldsymbol{\varrho}}_i;\boldsymbol{a}}^{\beta-\beta_b,1}, \quad \mathbb{Y}_\beta := \bigcap_{i=0,1} \mathbf{B}_{\boldsymbol{p}_i;\boldsymbol{a}}^{\beta,1}, \quad \mathbb{Z}_\beta := \bigcap_{i=0,1} \mathbf{B}_{\boldsymbol{p}_i;\boldsymbol{a}}^{\beta}. \tag{3.8}$$

Note that the dual space of $\mathbb{X}_\beta$ is given by

$$\mathbb{X}_\beta' = \mathbf{B}_{\boldsymbol{\varrho}_0;\boldsymbol{a}}^{\beta_b-\beta} + \mathbf{B}_{\boldsymbol{\varrho}_1;\boldsymbol{a}}^{\beta_b-\beta}.$$

More precisely, for any $b \in \mathbb{X}_\beta'$, it can be written as

$$b = b_0 + b_1, \quad b_i \in \mathbf{B}_{\boldsymbol{\varrho}_i;\boldsymbol{a}}^{\beta_b-\beta}, i = 0, 1.$$

By Lemma 2.16, we have the following estimates about the kinetic semigroup P_t in $\mathbb{X}_\beta$, $\mathbb{Y}_\beta$ and $\mathbb{Z}_\beta$.

Lemma 3.4. *Under $(\mathbf{H}_0)$, for any $\beta \geqslant 0$, there is a constant $C = C(\beta, \Theta) > 0$ such that for all $t > 0$,*

$$\|P_t f\|_{\mathbb{X}_\beta} \lesssim_C (1 \wedge t)^{-\frac{\beta+\mathscr{A}_0-\beta_0-\beta_b}{\alpha}} \|f\|_{\mathbb{Z}_{\beta_0}}, \quad \|P_t f\|_{\mathbb{Y}_\beta} \lesssim_C (1 \wedge t)^{-\frac{\beta-\beta_0}{\alpha}} \|f\|_{\mathbb{Z}_{\beta_0}}, \tag{3.9}$$

and for all $t \geqslant 1$,

$$\|P_t f\|_{\mathbb{X}_\beta} \lesssim_C t^{-\frac{\mathscr{A}_1}{\alpha}} \|f\|_{\mathbb{Z}_{\beta_0}}, \quad \|P_t f\|_{\mathbb{Y}_\beta} \lesssim_C \|f\|_{\mathbb{Z}_{\beta_0}}. \tag{3.10}$$

Moreover, we also have

$$\|P_t f\|_{\mathbb{X}_\beta} \lesssim_C (1 \wedge t)^{-\frac{\beta}{\alpha}} \|f\|_{\mathbb{X}_0}, \quad \|P_t f\|_{\mathbb{Y}_\beta} \lesssim_C (1 \wedge t)^{-\frac{\beta}{\alpha}} \|f\|_{\mathbb{Y}_0}. \tag{3.11}$$

Proof. For $i = 0, 1$, since $\beta_0 \neq 0, \mathscr{A}_i - \beta_b$, by (2.33), it is easy to see that

$$\|P_t f\|_{\mathbf{B}_{\boldsymbol{p}_i;\boldsymbol{a}}^{\beta,1}} \lesssim (1 \wedge t)^{-\frac{\beta-\beta_0}{\alpha}} \|f\|_{\mathbf{B}_{\boldsymbol{p}_i;\boldsymbol{a}}^{\beta_0}},$$

and for any $0 \leqslant \beta \neq \beta_0 + \beta_b - \mathscr{A}_i$,

$$\|P_t f\|_{\mathbf{B}_{\bar{\boldsymbol{\varrho}}_i;\boldsymbol{a}}^{\beta-\beta_b,1}} \lesssim (1 \wedge t)^{-\frac{\mathscr{A}_i+\beta-\beta_b-\beta_0}{\alpha}} \|f\|_{\mathbf{B}_{\boldsymbol{p}_i;\boldsymbol{a}}^{\beta_0}} \leqslant (1 \wedge t)^{-\frac{\beta+\mathscr{A}_0-\beta_0-\beta_b}{\alpha}} \|f\|_{\mathbf{B}_{\boldsymbol{p}_i;\boldsymbol{a}}^{\beta_0}},$$

where the second inequality is due to $\mathscr{A}_1 \leqslant \mathscr{A}_0$. Moreover, by interpolation inequality (2.16), the above inequality holds for all $\beta \geqslant 0$. Thus we get (3.9). Similarly, estimate (3.10) follows by (2.34) and $\mathscr{A}_1 \leqslant \mathscr{A}_0$. Estimate (3.11) follows by (2.35). $\square$

Now we introduce the following bilinear form for later use:

$$\mathcal{H}_t^b(u,w)(x,v) := -\int_0^t P_{t-s}\mathrm{div}_v((b_s * u_s)w_s)(x,v)\mathrm{d}s. \tag{3.12}$$

Note that by the change of variables and the semigroup property of $(P_t)_{t\geqslant 0}$,

$$\mathcal{H}_{t+s}^b(u,w) = P_s\mathcal{H}_t^b(u,w) + \mathcal{H}_s^{b^t}(u^t,w^t), \quad s,t > 0, \tag{3.13}$$

where for a function f,

$$f^t(s) = f(t+s).$$

The following lemma provides important estimates about the above bilinear form, which are consequences of Theorem 2.18 and Lemma 2.6.

Lemma 3.5. *Suppose* $(\mathbf{H}_0)$. *Let* $\theta \in (0, \mathcal{Q})$, $\boldsymbol{p}, \boldsymbol{p}' \in [1,\infty]^2$ *with* $\boldsymbol{p} \leqslant \boldsymbol{p}'$ *and write* $\mathscr{A} := \boldsymbol{a}\cdot(\frac{d}{\boldsymbol{p}} - \frac{d}{\boldsymbol{p}'})$. *For any* $\beta \geqslant 0$ *and* $\boldsymbol{\gamma}, \bar{\boldsymbol{\gamma}}, \boldsymbol{\gamma}', \boldsymbol{\gamma}_b \in \mathbb{R}^2$ *satisfying* $\boldsymbol{\gamma} + \bar{\boldsymbol{\gamma}} + \boldsymbol{\gamma}_b \in [0, \mathcal{Q})^2$ *and*

$$\boldsymbol{\gamma}' \gtrless \boldsymbol{\gamma} + \bar{\boldsymbol{\gamma}} + \boldsymbol{\gamma}_b - (\theta, (\mathcal{Q} - 1 - \mathscr{A}) \vee 0),$$

there is a constant $C = C(\beta, \theta, \boldsymbol{\gamma}, \bar{\boldsymbol{\gamma}}, \boldsymbol{\gamma}_b, \boldsymbol{\gamma}', \Theta) > 0$ *such that for all* $b \in \mathbb{L}_{\boldsymbol{\gamma}_b}^{q_b}(\mathbb{X}_0')$, $u \in \mathbb{L}_{\bar{\boldsymbol{\gamma}}}^\infty(\mathbb{X}_\beta)$ *and* $w \in \mathbb{L}_{\boldsymbol{\gamma}}^\infty(\mathbb{L}^{\boldsymbol{p}} \cap \mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^\beta)$,

$$\|\mathcal{H}^b(u,w)\|_{\mathbb{L}_{\boldsymbol{\gamma}'}^\infty(\mathbf{B}_{\boldsymbol{p}';\boldsymbol{a}}^{\beta+\beta',1})} \lesssim_C \|b\|_{\mathbb{L}_{\boldsymbol{\gamma}_b}^{q_b}(\mathbb{X}_0')} \Big[\|u\|_{\mathbb{L}_{\bar{\boldsymbol{\gamma}}}^\infty(\mathbb{X}_0)}\|w\|_{\mathbb{L}_{\boldsymbol{\gamma}}^\infty(\mathbb{L}^{\boldsymbol{p}})}\mathbb{1}_{\beta=0}$$

$$+ \big(\|u\|_{\mathbb{L}_{\bar{\boldsymbol{\gamma}}}^\infty(\mathbb{X}_\beta)}\|w\|_{\mathbb{L}_{\boldsymbol{\gamma}}^\infty(\mathbb{L}^{\boldsymbol{p}})} + \|u\|_{\mathbb{L}_{\bar{\boldsymbol{\gamma}}}^\infty(\mathbb{X}_0)}\|w\|_{\mathbb{L}_{\boldsymbol{\gamma}}^\infty(\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^\beta)}\big)\mathbb{1}_{\beta>0}\Big], \tag{3.14}$$

where $\beta' := \mathcal{Q} - 1 - \mathscr{A} - \theta$.

Proof. For any $\beta \geqslant 0$, by Theorem 2.18, we have

$$\|\mathcal{H}^b(u,w)\|_{\mathbb{L}_{\boldsymbol{\gamma}'}^\infty(\mathbf{B}_{\boldsymbol{p}';\boldsymbol{a}}^{\beta+\beta',1})} \lesssim \|(b*u)w\|_{\mathbb{L}_{\boldsymbol{\gamma}+\bar{\boldsymbol{\gamma}}+\boldsymbol{\gamma}_b}^{q_b}(\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^\beta)}. \tag{3.15}$$

Since $b = b_0 + b_1 \in \mathbb{X}_0' = \mathbf{B}_{\boldsymbol{\varrho}_0;\boldsymbol{a}}^{\beta_b} + \mathbf{B}_{\boldsymbol{\varrho}_1;\boldsymbol{a}}^{\beta_b}$, by (2.12) and (2.13), we have for $i = 0, 1$,

$$\|(b_i*u)w\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^0} \lesssim \|(b_i*u)w\|_{\mathbb{L}^{\boldsymbol{p}}} \leqslant \|b_i*u\|_{\mathbb{L}^\infty}\|w\|_{\mathbb{L}^{\boldsymbol{p}}}$$

$$\lesssim \|b_i*u\|_{\mathbf{B}_{\infty;\boldsymbol{a}}^{0,1}}\|w\|_{\mathbb{L}^{\boldsymbol{p}}} \lesssim \|b_i\|_{\mathbf{B}_{\boldsymbol{\varrho}_i;\boldsymbol{a}}^{\beta_b}}\|u\|_{\mathbf{B}_{\bar{\boldsymbol{\varrho}}_i;\boldsymbol{a}}^{-\beta_b,1}}\|w\|_{\mathbb{L}^{\boldsymbol{p}}}. \tag{3.16}$$

Hence,

$$\|(b*u)w\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^0} \lesssim \|b\|_{\mathbb{X}_0'}\|u\|_{\mathbb{X}_0}\|w\|_{\mathbb{L}^{\boldsymbol{p}}}$$

and

$$\|(b*u)w\|_{\mathbb{L}_{\boldsymbol{\gamma}+\bar{\boldsymbol{\gamma}}+\boldsymbol{\gamma}_b}^{q_b}(\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^0)} \lesssim \|b\|_{\mathbb{L}_{\boldsymbol{\gamma}_b}^{q_b}(\mathbb{X}_0')}\|u\|_{\mathbb{L}_{\bar{\boldsymbol{\gamma}}}^\infty(\mathbb{X}_0)}\|w\|_{\mathbb{L}_{\boldsymbol{\gamma}}^\infty(\mathbb{L}^{\boldsymbol{p}})}. \tag{3.17}$$

For $\beta > 0$, by (iii) and (ii) of Lemma 2.6, we have for $i = 0, 1$,

$$\|(b_i*u)w\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^\beta} \lesssim \|b_i*u\|_{\mathbf{B}_{\infty;\boldsymbol{a}}^\beta}\|w\|_{\mathbb{L}^{\boldsymbol{p}}} + \|b_i*u\|_{\mathbb{L}^\infty}\|w\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^\beta}$$

$$\lesssim \|b_i\|_{\mathbf{B}_{\boldsymbol{\varrho}_i;\boldsymbol{a}}^{\beta_b}}\Big(\|u\|_{\mathbf{B}_{\bar{\boldsymbol{\varrho}}_i;\boldsymbol{a}}^{\beta-\beta_b,1}}\|w\|_{\mathbb{L}^{\boldsymbol{p}}} + \|u\|_{\mathbf{B}_{\bar{\boldsymbol{\varrho}}_i;\boldsymbol{a}}^{-\beta_b,1}}\|w\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^\beta}\Big).$$

Hence,

$$\|(b*u)w\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^\beta} \lesssim \|b\|_{\mathbb{X}_0'}\Big(\|u\|_{\mathbb{X}_\beta}\|w\|_{\mathbb{L}^{\boldsymbol{p}}} + \|u\|_{\mathbb{X}_0}\|w\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^\beta}\Big)$$

and

$$\|(b*u)w\|_{\mathbb{L}_{\boldsymbol{\gamma}+\bar{\boldsymbol{\gamma}}+\boldsymbol{\gamma}_b}^{q_b}(\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^\beta)} \lesssim \|b\|_{\mathbb{L}_{\boldsymbol{\gamma}_b}^{q_b}(\mathbb{X}_0')}\Big(\|u\|_{\mathbb{L}_{\bar{\boldsymbol{\gamma}}}^\infty(\mathbb{X}_\beta)}\|w\|_{\mathbb{L}_{\boldsymbol{\gamma}}^\infty(\mathbb{L}^{\boldsymbol{p}})} + \|u\|_{\mathbb{L}_{\bar{\boldsymbol{\gamma}}}^\infty(\mathbb{X}_0)}\|w\|_{\mathbb{L}_{\boldsymbol{\gamma}}^\infty(\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^\beta)}\Big). \tag{3.18}$$

Substituting (3.17) and (3.18) into (3.15), we obtain (3.14). $\qquad\square$

Remark 3.6. If we assume that $-1 < \beta_b < 0$ and $|b_t(x,y)| \leqslant K_t(x-y)$ for some $K \in \mathbb{L}^{q_b}_{\gamma_b}(\mathbb{X}'_0)$, and $b * u$ is replaced by $b \circledast u = \int_{\mathbb{R}^d} b(x,y)u(y)\mathrm{d}y$, then (3.14) still holds for $\beta = 0$. Indeed, for the corresponding term $b \circledast u$ in (3.16), we can estimate it as follows:

$$\|b \circledast u\|_{\mathbb{L}^\infty} \leqslant \|K * |u|\|_{\mathbb{L}^\infty} \lesssim \sum_{i=0,1} \|K_i * |u|\|_{\mathbf{B}^{0,1}_{\infty;\mathbf{a}}} \lesssim \sum_{i=0,1} \|K_i\|_{\mathbf{B}^{\beta_b}_{\varrho_i;\mathbf{a}}} \||u|\|_{\mathbf{B}^{-\beta_b,1}_{\bar{\varrho}_i;\mathbf{a}}}.$$

Since $-1 < \beta_b < 0$, by characterization (2.11) and $||a|-|b|| \leqslant |a-b|$, it is easy to see that

$$\||u|\|_{\mathbf{B}^{-\beta_b,1}_{\bar{\varrho}_i;\mathbf{a}}} \lesssim \|u\|_{\mathbf{B}^{-\beta_b,1}_{\varrho_i;\mathbf{a}}}.$$

Hence,

$$\|b \circledast u\|_{\mathbb{L}^\infty} \lesssim \|K\|_{\mathbb{X}'_0} \|u\|_{\mathbb{X}_0}.$$

Remark 3.7. It shall be seen below that estimate (3.14) for $\beta = 0$ is enough to show the existence and uniqueness of a weak solution for equation (3.4). While, estimate (3.14) for $\beta > 0$ shall be used to show the smoothness of solutions.

The following lemma is a direct application of Lemma 3.5.

Lemma 3.8. Suppose $(\mathbf{H}_0)$. Let $\beta \geqslant 0$ and $\theta \in (0, \mathscr{Q})$. Let $b \in \mathbb{L}^{q_b}_{\gamma_b}(\mathbb{X}'_0)$ and $\boldsymbol{\gamma}, \bar{\boldsymbol{\gamma}} \in \mathbb{R}^2$ satisfy

$$\boldsymbol{\gamma} + \bar{\boldsymbol{\gamma}} + \boldsymbol{\gamma}_b \in [0, \mathscr{Q})^2.$$

(i) Let $\beta' := \mathscr{Q} - 1 - \theta$. For any $\boldsymbol{\gamma}' \in \mathbb{R}^2$ with

$$\boldsymbol{\gamma}' \gtrless \boldsymbol{\gamma} + \bar{\boldsymbol{\gamma}} + \boldsymbol{\gamma}_b - (\theta, \mathscr{Q} - 1),$$

there is a $C = C(\Theta, \beta, \theta, \boldsymbol{\gamma}, \bar{\boldsymbol{\gamma}}, \boldsymbol{\gamma}') > 0$ such that for all $u \in \mathbb{L}^\infty_{\bar{\boldsymbol{\gamma}}}(\mathbb{X}_\beta)$ and $w \in \mathbb{L}^\infty_{\boldsymbol{\gamma}}(\mathbb{Y}_\beta)$,

$$\|\mathscr{H}^b(u,w)\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}'}(\mathbb{Y}_{\beta+\beta'})} \lesssim_C \|b\|_{\mathbb{L}^{q_b}_{\gamma_b}(\mathbb{X}'_0)} \big(\|u\|_{\mathbb{L}^\infty_{\bar{\boldsymbol{\gamma}}}(\mathbb{X}_\beta)} \|w\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}}(\mathbb{Y}_0)} + \|u\|_{\mathbb{L}^\infty_{\bar{\boldsymbol{\gamma}}}(\mathbb{X}_0)} \|w\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}}(\mathbb{Y}_\beta)}\big). \tag{3.19}$$

(ii) Let $\beta' := \mathscr{Q} - 1 - \mathscr{A}_0 + \beta_b - \theta$. For any $\boldsymbol{\gamma}' \in \mathbb{R}^2$ with

$$\boldsymbol{\gamma}' \gtrless \boldsymbol{\gamma} + \bar{\boldsymbol{\gamma}} + \boldsymbol{\gamma}_b - (\theta, (\mathscr{Q} - 1 - \mathscr{A}_1) \vee 0),$$

there is a $C = C(\Theta, \beta, \theta, \boldsymbol{\gamma}, \bar{\boldsymbol{\gamma}}, \boldsymbol{\gamma}') > 0$ such that for all $u \in \mathbb{L}^\infty_{\bar{\boldsymbol{\gamma}}}(\mathbb{X}_\beta)$ and $w \in \mathbb{L}^\infty_{\boldsymbol{\gamma}}(\mathbb{Y}_\beta)$,

$$\|\mathscr{H}^b(u,w)\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}'}(\mathbb{X}_{\beta+\beta'})} \lesssim_C \|b\|_{\mathbb{L}^{q_b}_{\gamma_b}(\mathbb{X}'_0)} \big(\|u\|_{\mathbb{L}^\infty_{\bar{\boldsymbol{\gamma}}}(\mathbb{X}_\beta)} \|w\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}}(\mathbb{Y}_0)} + \|u\|_{\mathbb{L}^\infty_{\bar{\boldsymbol{\gamma}}}(\mathbb{X}_0)} \|w\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}}(\mathbb{Y}_\beta)}\big). \tag{3.20}$$

Proof. (i) For $i = 0, 1$, by (3.14) with $(\boldsymbol{p}', \boldsymbol{p}) = (\boldsymbol{p}_i, \boldsymbol{p}_i)$, we have

$$\|\mathscr{H}^b(u,w)\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}'}(\mathbf{B}^{\beta+\beta',1}_{\boldsymbol{p}_i;\mathbf{a}})} \lesssim \|b\|_{\mathbb{L}^{q_b}_{\gamma_b}(\mathbb{X}'_0)} \Big(\|u\|_{\mathbb{L}^\infty_{\bar{\boldsymbol{\gamma}}}(\mathbb{X}_\beta)} \|w\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}}(\mathbf{B}^{0,1}_{\boldsymbol{p}_i})} + \|u\|_{\mathbb{L}^\infty_{\bar{\boldsymbol{\gamma}}}(\mathbb{X}_0)} \|w\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}}(\mathbf{B}^{\beta,1}_{\boldsymbol{p}_i})}\Big).$$

Thus we obtain (3.19) by the definition of $\mathbb{Y}_\beta$ in (3.8).

(ii) For $i = 0, 1$, noting that by $\mathscr{A}_1 \leqslant \mathscr{A}_0$,

$$\mathscr{Q} - 1 - \mathscr{A}_1 - \theta \geqslant \mathscr{Q} - 1 - \mathscr{A}_0 - \theta = \beta' - \beta_b,$$

and

$$\boldsymbol{\gamma}' \gtrless \boldsymbol{\gamma} + \bar{\boldsymbol{\gamma}} + \boldsymbol{\gamma}_b - (\theta, (\mathscr{Q} - 1 - \mathscr{A}_i) \vee 0),$$

by (3.14) with $(\boldsymbol{p}', \boldsymbol{p}) = (\bar{\boldsymbol{\varrho}}_i, \boldsymbol{p}_i)$, we have

$$\|\mathscr{H}^b(u,w)\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}'}(\mathbf{B}^{\beta+\beta'-\beta_b,1}_{\bar{\varrho}_i;\mathbf{a}})} \lesssim \|\mathscr{H}^b(u,w)\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}'}(\mathbf{B}^{\beta+\mathscr{Q}-1-\mathscr{A}_i-\theta,1}_{\bar{\varrho}_i;\mathbf{a}})}$$
$$\lesssim \|b\|_{\mathbb{L}^{q_b}_{\gamma_b}(\mathbb{X}'_0)} \Big(\|u\|_{\mathbb{L}^\infty_{\bar{\boldsymbol{\gamma}}}(\mathbb{X}_\beta)} \|w\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}}(\mathbf{B}^{0,1}_{\boldsymbol{p}_i})} + \|u\|_{\mathbb{L}^\infty_{\bar{\boldsymbol{\gamma}}}(\mathbb{X}_0)} \|w\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}}(\mathbf{B}^{\beta,1}_{\boldsymbol{p}_i})}\Big),$$

which implies (3.20) by the definition of $\mathbb{X}_\beta$ in (3.8). $\qquad\square$

Before stating and proving our main result, under $(\mathbf{H}_0)$, we introduce the following time-weighted spaces where the solutions of equation (3.1) shall stay: for $\beta \geqslant 0$,

$$\mathbb{S}_\beta := \mathbb{L}^\infty_{\boldsymbol{\ell}_\beta}(\mathbb{X}_\beta) \cap \mathbb{L}^\infty_{\boldsymbol{\gamma}_\beta}(\mathbb{Y}_\beta), \tag{3.21}$$

where

$$\boldsymbol{\ell}_\beta := \big(\beta + \mathscr{A}_0 - \beta_0 - \beta_b, (\mathscr{Q}-1) \wedge \mathscr{A}_1\big), \quad \boldsymbol{\gamma}_\beta := \big(\beta - \beta_0, 0\big). \tag{3.22}$$

It should be noticed that the norm $\|\cdot\|_{\mathbb{S}_\beta}$ does not have the monotonicity with respect to β, i.e.,

$$\|f\|_{\mathbb{S}_0} \not\lesssim C\|f\|_{\mathbb{S}_\beta}, \quad \beta > 0.$$

But the following interpolation inequality holds: for any $\theta \in [0,1]$ and $\beta > 0$,

$$\|f\|_{\mathbb{S}_{\theta\beta}} \lesssim \|f\|_{\mathbb{S}_0}^{1-\theta}\|f\|_{\mathbb{S}_\beta}^\theta. \tag{3.23}$$

Indeed, by the interpolation inequality (2.16), we have $\|f\|_{\mathbb{X}_{\theta\beta}} \lesssim \|f\|_{\mathbb{X}_\beta}^\theta \|f\|_{\mathbb{X}_0}^{1-\theta}$ and

$$\begin{aligned}
\|f\|_{\mathbb{L}^\infty_{\boldsymbol{\ell}_{\theta\beta}}(\mathbb{X}_{\theta\beta})} &= \sup_{t>0} \left(\|f\|_{\mathbb{X}_{\theta\beta}}(1 \wedge t)^{\theta\beta + \mathscr{A}_0 - \beta_0 - \beta_b}(1 \vee t)^{(\mathscr{Q}-1)\wedge\mathscr{A}_1} \right) \\
&\lesssim \sup_{t>0} \left(\|f\|_{\mathbb{X}_\beta}(1 \wedge t)^{\beta + \mathscr{A}_0 - \beta_0 - \beta_b}(1 \vee t)^{(\mathscr{Q}-1)\wedge\mathscr{A}_1} \right)^\theta \\
&\quad \times \sup_{t>0} \left(\|f\|_{\mathbb{X}_0}(1 \wedge t)^{\mathscr{A}_0 - \beta_0 - \beta_b}(1 \vee t)^{(\mathscr{Q}-1)\wedge\mathscr{A}_1} \right)^{1-\theta} \\
&= \|f\|_{\mathbb{L}^\infty_{\boldsymbol{\ell}_\beta}(\mathbb{X}_\beta)}^\theta \|f\|_{\mathbb{L}^\infty_{\boldsymbol{\ell}_0}(\mathbb{X}_0)}^{1-\theta}.
\end{aligned}$$

Similarly,

$$\|f\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}_{\theta\beta}}(\mathbb{Y}_{\theta\beta})} \lesssim \|f\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}_\beta}(\mathbb{Y}_\beta)}^\theta \|f\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}_0}(\mathbb{Y}_0)}^{1-\theta}.$$

Combining the above two estimates, we get (3.23).

Let

$$\Lambda := (\mathscr{Q}-1) \wedge (\mathscr{Q}-1-\mathscr{A}_0 + \beta_b) > 0.$$

Now we use Lemma 3.8 to show the following crucial lemma.

Lemma 3.9. *Under* $(\mathbf{H}_0)$*, for any* $\beta \geqslant 0$ *and* $\beta' \in [0, \Lambda)$*, there is a constant* $C = C(\beta', \beta, \Theta) > 0$ *such that for all* $b \in \mathbb{L}^{q_b}_{\boldsymbol{\gamma}_b}(\mathbb{X}'_0)$ *and* $u, w \in \mathbb{S}_0 \cap \mathbb{S}_\beta$,

$$\|\mathscr{H}^b(u,w)\|_{\mathbb{S}_{\beta+\beta'}} \lesssim_C \|b\|_{\mathbb{L}^{q_b}_{\boldsymbol{\gamma}_b}(\mathbb{X}'_0)}\big(\|u\|_{\mathbb{S}_\beta} + \|u\|_{\mathbb{S}_0}\big)\big(\|w\|_{\mathbb{S}_\beta} + \|w\|_{\mathbb{S}_0}\big) =: \mathscr{U}, \tag{3.24}$$

where $\mathbb{S}_\beta$ *is defined in* (3.21).

Proof. We divide the proof into three steps.

(*Step 1*) In this step we first show (3.24) for $\beta = \beta' = 0$. Observe that

$$\gamma_{b1} \in ((\mathscr{Q}-1-\mathscr{A}_1) \vee 0, (\mathscr{Q}-\mathscr{A}_1) \vee 1) = (\mathscr{Q}-1-(\mathscr{Q}-1)\wedge\mathscr{A}_1, \mathscr{Q}-(\mathscr{Q}-1)\wedge\mathscr{A}_1).$$

By the definition (3.22) and the assumption (3.5), it is easy to see that

$$\boldsymbol{\gamma}_0 + \boldsymbol{\ell}_0 + \boldsymbol{\gamma}_b = (\mathscr{A}_0 - 2\beta_0 - \beta_b, (\mathscr{Q}-1)\wedge\mathscr{A}_1 + \gamma_{b1}) \in [0, \mathscr{Q})^2.$$

Since $\boldsymbol{\gamma}_0 \gtrless \boldsymbol{\gamma}_0 + \boldsymbol{\ell}_0 + \boldsymbol{\gamma}_b - (\mathscr{Q}-1, \mathscr{Q}-1)$, by (i) of Lemma 3.8 with the choice of $\theta = \mathscr{Q}-1$, $\beta = 0$ and $\boldsymbol{\gamma} = \boldsymbol{\gamma}_0, \bar{\boldsymbol{\gamma}} = \boldsymbol{\ell}_0, \boldsymbol{\gamma}' = \boldsymbol{\gamma}_0$, we have

$$\|\mathscr{H}^b(u,w)\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}_0}(\mathbb{Y}_0)} \lesssim \|b\|_{\mathbb{L}^{q_b}_{\boldsymbol{\gamma}_b}(\mathbb{X}'_0)}\|u\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}_0}(\mathbb{X}_0)}\|w\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}_0}(\mathbb{Y}_0)},$$

Moreover, for $\theta := \mathscr{Q}-1-\mathscr{A}_0 + \beta_b$, since $\boldsymbol{\ell}_0 \gtrless \boldsymbol{\gamma}_0 + \boldsymbol{\ell}_0 + \boldsymbol{\gamma}_b - (\theta, (\mathscr{Q}-1-\mathscr{A}_1) \vee 0)$, by (ii) of Lemma 3.8 with the choice of the above θ, $\beta = 0$ and $\boldsymbol{\gamma} = \boldsymbol{\gamma}_0, \bar{\boldsymbol{\gamma}} = \boldsymbol{\gamma}' = \boldsymbol{\ell}_0$, we have

$$\|\mathscr{H}^b(u,w)\|_{\mathbb{L}^\infty_{\boldsymbol{\ell}_0}(\mathbb{X}_0)} \lesssim \|b\|_{\mathbb{L}^{q_b}_{\boldsymbol{\gamma}_b}(\mathbb{X}'_0)}\|u\|_{\mathbb{L}^\infty_{\boldsymbol{\ell}_0}(\mathbb{X}_0)}\|w\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}_0}(\mathbb{Y}_0)}.$$

Hence,

$$\|\mathscr{H}^b(u,w)\|_{\mathbb{S}_0} = \|\mathscr{H}^b(u,w)\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}_0}(\mathbb{Y}_0)} + \|\mathscr{H}^b(u,w)\|_{\mathbb{L}^\infty_{\boldsymbol{\ell}_0}(\mathbb{X}_0)} \lesssim \|b\|_{\mathbb{L}^{q_b}_{\boldsymbol{\gamma}_b}(\mathbb{X}'_0)}\|u\|_{\mathbb{S}_0}\|w\|_{\mathbb{S}_0}. \tag{3.25}$$

(*Step 2*) Let $\beta \geqslant 0$. For fixed $\beta' \in [0, \mathcal{Q} - 1)$, let
$$\theta = \mathcal{Q} - 1 - \beta' \in (0, \mathcal{Q}).$$
By (i) of Lemma 3.8 with the choice of the above θ, $\boldsymbol{\gamma} = \bar{\boldsymbol{\gamma}} = 0$ and $\boldsymbol{\gamma}' = (-\theta, -\ell_{01})$, we have
$$\|\mathcal{H}_{\cdot}^{b^t}(u^t, w^t)\|_{\mathbb{L}^{\infty}_{\boldsymbol{\gamma}'}(\mathbb{Y}_{\beta+\beta'})} \lesssim \|b\|_{\mathbb{L}^{q_b}_{\boldsymbol{\gamma}_b}(\mathbb{X}'_0)} \Big(\|u^t\|_{\mathbb{L}^{\infty}_0(\mathbb{X}_\beta)} \|w^t\|_{\mathbb{L}^{\infty}_0(\mathbb{Y}_0)} + \|u^t\|_{\mathbb{L}^{\infty}_0(\mathbb{X}_0)} \|w^t\|_{\mathbb{L}^{\infty}_0(\mathbb{Y}_\beta)} \Big).$$
Note that by (2.32) and (3.22),
$$\|u^t\|_{\mathbb{L}^{\infty}_0(\mathbb{X}_\beta)} \|w^t\|_{\mathbb{L}^{\infty}_0(\mathbb{Y}_0)} \lesssim (1 \wedge t)^{-\frac{\ell_{\beta 0} + \gamma_{00}}{\alpha}} (1 \vee t)^{-\frac{\ell_{01}}{\alpha}} \|u\|_{\mathbb{L}^{\infty}_{\boldsymbol{\ell}_\beta}(\mathbb{X}_\beta)} \|w\|_{\mathbb{L}^{\infty}_{\boldsymbol{\gamma}_0}(\mathbb{Y}_0)}, \tag{3.26}$$
and
$$\|u^t\|_{\mathbb{L}^{\infty}_0(\mathbb{X}_0)} \|w^t\|_{\mathbb{L}^{\infty}_0(\mathbb{Y}_\beta)} \lesssim (1 \wedge t)^{-\frac{\ell_{00} + \gamma_{\beta 0}}{\alpha}} (1 \vee t)^{-\frac{\ell_{01}}{\alpha}} \|u\|_{\mathbb{L}^{\infty}_{\boldsymbol{\ell}_0}(\mathbb{X}_0)} \|w\|_{\mathbb{L}^{\infty}_{\boldsymbol{\gamma}_\beta}(\mathbb{Y}_\beta)}. \tag{3.27}$$
Thanks to $\ell_{\beta 0} + \gamma_{00} = \ell_{00} + \gamma_{\beta 0} = \ell_{00} + \beta - \beta_0$, we further have
$$\|\mathcal{H}_t^{b^t}(u^t, w^t)\|_{\mathbb{Y}_{\beta+\beta'}} \lesssim (1 \wedge t)^{\frac{\mathcal{Q} - 1 - \beta'}{\alpha}} (1 \vee t)^{\frac{\ell_{01}}{\alpha}} \|\mathcal{H}_{\cdot}^{b^t}(u^t, w^t)\|_{\mathbb{L}^{\infty}_{\boldsymbol{\gamma}'}(\mathbb{Y}_{\beta+\beta'})}$$
$$\lesssim (1 \wedge t)^{\frac{\mathcal{Q} - 1 - \beta' - \ell_{00} - \beta + \beta_0}{\alpha}} \mathscr{U} \leqslant (1 \wedge t)^{-\frac{\beta' + \beta - \beta_0}{\alpha}} \mathscr{U},$$
where $\mathscr{U}$ is defined by (3.24). On the other hand, by (3.11) and (3.25), we have
$$\|P_t \mathcal{H}_t^b(u, w)\|_{\mathbb{Y}_{\beta+\beta'}} \lesssim (1 \wedge t)^{-\frac{\beta+\beta'}{\alpha}} \|\mathcal{H}_t^b(u, w)\|_{\mathbb{Y}_0} \lesssim (1 \wedge t)^{-\frac{\beta+\beta'+\gamma_{00}}{\alpha}} \|\mathcal{H}^b(u, w)\|_{\mathbb{L}^{\infty}_{\boldsymbol{\gamma}_0}(\mathbb{Y}_0)}$$
$$\lesssim (1 \wedge t)^{-\frac{\beta+\beta'+\gamma_{00}}{\alpha}} \|b\|_{\mathbb{L}^{q_b}_{\boldsymbol{\gamma}_b}(\mathbb{X}'_0)} \|u\|_{\mathbb{S}_0} \|w\|_{\mathbb{S}_0} \leqslant (1 \wedge t)^{-\frac{\beta+\beta'+\gamma_{00}}{\alpha}} \mathscr{U}.$$
Combining the above two estimates and by (3.13) with $s = t$, we obtain
$$\|\mathcal{H}_{2t}^b(u, w)\|_{\mathbb{Y}_{\beta+\beta'}} \lesssim (1 \wedge t)^{-\frac{\beta+\beta'+\gamma_{00}}{\alpha}} \mathscr{U}. \tag{3.28}$$
(*Step 3*) Let $\beta \geqslant 0$. For fixed $\beta' \in [0, \mathcal{Q} - 1 - \mathscr{A}_0 + \beta_b)$, let
$$\theta := \mathcal{Q} - 1 - \mathscr{A}_0 + \beta_b - \beta' \in (0, \mathcal{Q}).$$
By (ii) of Lemma 3.8 with the choice of the above θ, $\boldsymbol{\gamma} = \bar{\boldsymbol{\gamma}} = 0$ and $\boldsymbol{\gamma}' = (-\theta, 0)$, we have
$$\|\mathcal{H}_{\cdot}^{b^t}(u^t, w^t)\|_{\mathbb{L}^{\infty}_{\boldsymbol{\gamma}'}(\mathbb{X}_{\beta+\beta'})} \lesssim \|b\|_{\mathbb{L}^{q_b}_{\boldsymbol{\gamma}_b}(\mathbb{X}'_0)} \Big(\|u^t\|_{\mathbb{L}^{\infty}_0(\mathbb{X}_\beta)} \|w^t\|_{\mathbb{L}^{\infty}_0(\mathbb{Y}_0)} + \|u^t\|_{\mathbb{L}^{\infty}_0(\mathbb{X}_0)} \|w^t\|_{\mathbb{L}^{\infty}_0(\mathbb{Y}_\beta)} \Big).$$
By (3.26) and (3.27), we further have
$$\|\mathcal{H}_t^{b^t}(u^t, w^t)\|_{\mathbb{X}_{\beta+\beta'}} \lesssim (1 \wedge t)^{\frac{\mathcal{Q} - 1 - \mathscr{A}_0 + \beta_b - \beta'}{\alpha}} \|\mathcal{H}_{\cdot}^{b^t}(u^t, w^t)\|_{\mathbb{L}^{\infty}_{\boldsymbol{\gamma}'}(\mathbb{X}_{\beta+\beta'})}$$
$$\lesssim (1 \wedge t)^{\frac{\mathcal{Q} - 1 - \mathscr{A}_0 + \beta_b - \beta' - \ell_{00} - \beta + \beta_0}{\alpha}} (1 \vee t)^{-\frac{\ell_{01}}{\alpha}} \mathscr{U}$$
$$\lesssim (1 \wedge t)^{-\frac{\ell_{00} + \beta + \beta'}{\alpha}} (1 \vee t)^{-\frac{\ell_{01}}{\alpha}} \mathscr{U}.$$
Moreover, by (3.11) we have
$$\|P_t \mathcal{H}_t^b(u, w)\|_{\mathbb{X}_{\beta+\beta'}} \lesssim (1 \wedge t)^{-\frac{\beta+\beta'}{\alpha}} \|\mathcal{H}_t^b(u, w)\|_{\mathbb{X}_0}$$
$$\lesssim (1 \wedge t)^{-\frac{\beta+\beta'+\ell_{00}}{\alpha}} (1 \vee t)^{-\frac{\ell_{01}}{\alpha}} \|\mathcal{H}^b(u, w)\|_{\mathbb{L}^{\infty}_{\boldsymbol{\ell}_0}(\mathbb{X}_0)}$$
$$\overset{(3.25)}{\lesssim} (1 \wedge t)^{-\frac{\beta+\beta'+\ell_{00}}{\alpha}} (1 \vee t)^{-\frac{\ell_{01}}{\alpha}} \|b\|_{\mathbb{L}^{q_b}_{\boldsymbol{\gamma}_b}(\mathbb{X}'_0)} \|u\|_{\mathbb{S}_0} \|w\|_{\mathbb{S}_0}.$$
Hence, by (3.13) with $s = t$ again,
$$\|\mathcal{H}_{2t}^b(u, w)\|_{\mathbb{X}_{\beta+\beta'}} \lesssim (1 \wedge t)^{-\frac{\beta+\beta'+\ell_{00}}{\alpha}} (1 \vee t)^{-\frac{\ell_{01}}{\alpha}} \mathscr{U},$$
which together with (3.28) yields (3.24). $\qquad \square$

Remark 3.10. In the situation of Remark 3.6, (3.24) still holds for $\beta = 0$.

Now we can state and prove our main result of this section.

Theorem 3.11. *Under* $(\mathbf{H_0})$, *there are two constants* $C_0, C_1 > 0$ *only depending on* Θ *such that for any* $b \in \mathbb{L}^{q_b}_{\gamma_b}(\mathbb{X}'_0)$ *and* $u_0 \in \mathbb{Z}_{\beta_0}$ *satisfying*

$$\|u_0\|_{\mathbb{Z}_{\beta_0}} \|b\|_{\mathbb{L}^{q_b}_{\gamma_b}(\mathbb{X}'_0)} \leqslant C_0, \tag{3.29}$$

there exists a unique weak solution to Fokker-Planck equation (3.1) *such that*

$$\|u\|_{\mathbb{S}_0} \leqslant C_1 \|u_0\|_{\mathbb{Z}_{\beta_0}}. \tag{3.30}$$

Moreover, we have the following conclusions:

(i) **(Smoothness)** *For any* $\beta \geqslant 0$, *there is a constant* $C_\beta = C_\beta(\Theta, \|b\|_{\mathbb{L}^{q_b}_{\gamma_b}(\mathbb{X}'_0)}) > 0$ *such that*

$$\|u\|_{\mathbb{S}_\beta} \leqslant C_\beta. \tag{3.31}$$

(ii) **(Stability)** *For another* $\widetilde{b} \in \mathbb{L}^{q_b}_{\gamma_b}(\mathbb{X}'_0)$ *and* $\widetilde{u}_0 \in \mathbb{Z}_{\beta_0}$ *satisfying* (3.29), *let* $\widetilde{u}$ *be the unique solution corresponding to* $\widetilde{b}$ *and* $\widetilde{u}_0$. *Then for any* $\beta \geqslant 0$, *there is a constant* $C_\beta > 0$ *only depending on* β, Θ *and* $\|b\|_{\mathbb{L}^{q_b}_{\gamma_b}(\mathbb{X}'_0)}, \|\widetilde{b}\|_{\mathbb{L}^{q_b}_{\gamma_b}(\mathbb{X}'_0)}$ *such that*

$$\|u - \widetilde{u}\|_{\mathbb{S}_\beta} \leqslant C_\beta \big(\|u_0 - \widetilde{u}_0\|_{\mathbb{Z}_{\beta_0}} + \|b - \widetilde{b}\|_{\mathbb{L}^{q_b}_{\gamma_b}(\mathbb{X}'_0)} \big). \tag{3.32}$$

Proof. We use the contraction mapping theorem to prove the existence and uniqueness of a weak solution. First of all, by (3.9) and (3.10), we have for any $\beta \geqslant 0$,

$$\|P.u_0\|_{\mathbb{S}_\beta} \lesssim \|u_0\|_{\mathbb{Z}_{\beta_0}}. \tag{3.33}$$

Thus, by (3.33) and (3.24), we can define a map $\mathcal{U} : \mathbb{S}_0 \to \mathbb{S}_0$ by

$$\mathcal{U}(u)(t) := P_t u_0 + \mathcal{H}^b_t(u, u) \tag{3.34}$$

so that for some constants $C_2, C_3 > 0$ only depending on Θ and for all $u \in \mathbb{S}_0$,

$$\|\mathcal{U}(u)\|_{\mathbb{S}_0} \leqslant C_2 \|u_0\|_{\mathbb{Z}_{\beta_0}} + C_3 \|b\|_{\mathbb{L}^{q_b}_{\gamma_b}(\mathbb{X}'_0)} \|u\|^2_{\mathbb{S}_0}, \tag{3.35}$$

and for all $u, \widetilde{u} \in \mathbb{S}_0$,

$$\|\mathcal{U}(u) - \mathcal{U}(\widetilde{u})\|_{\mathbb{S}_0} = \|\mathcal{H}^b(u, u) - \mathcal{H}^b(\widetilde{u}, \widetilde{u})\|_{\mathbb{S}_0} \leqslant C_3 \|b\|_{\mathbb{L}^{q_b}_{\gamma_b}(\mathbb{X}'_0)} \Big(\|u\|_{\mathbb{S}_0} + \|\widetilde{u}\|_{\mathbb{S}_0} \Big) \|u - \widetilde{u}\|_{\mathbb{S}_0}. \tag{3.36}$$

Now let us choose

$$\lambda = 4C_2 \|u_0\|_{\mathbb{Z}_{\beta_0}} / 3 = C_1 \|u_0\|_{\mathbb{Z}_{\beta_0}},$$

and suppose

$$\|u_0\|_{\mathbb{Z}_{\beta_0}} \|b\|_{\mathbb{L}^{q_b}_{\gamma_b}(\mathbb{X}'_0)} \leqslant 3/(16C_3 C_2) =: C_0.$$

Then it is easy to see that

$$C_2 \|u_0\|_{\mathbb{Z}_{\beta_0}} + C_3 \|b\|_{\mathbb{L}^{q_b}_{\gamma_b}(\mathbb{X}'_0)} \lambda^2 \leqslant \lambda, \quad 2C_3 \|b\|_{\mathbb{L}^{q_b}_{\gamma_b}(\mathbb{X}'_0)} \lambda \leqslant 1/2, \tag{3.37}$$

and by (3.35) and (3.36), $\mathcal{U}$ is a contraction mapping on the closed ball

$$\mathscr{B}^\lambda := \Big\{ u \in \mathbb{S}_0 : \|u\|_{\mathbb{S}_0} \leqslant \lambda \Big\}.$$

Therefore, there is a unique fixed point $u \in \mathscr{B}^\lambda$ so that for all $t > 0$,

$$u(t) = \mathcal{U}(u)(t) = P_t u_0 + \mathcal{H}^b_t(u, u). \tag{3.38}$$

(**Smoothness**) In this step we show the smoothness of the solution by induction and (3.24). Suppose that we have shown (3.31) for some $\beta \geqslant 0$. Then by (3.33) and (3.24), we have for any $\beta' \in [0, \Lambda)$,

$$\|u\|_{\mathbb{S}_{\beta+\beta'}} \leqslant \|P.u_0\|_{\mathbb{S}_{\beta+\beta'}} + \|\mathcal{H}^b(u, u)\|_{\mathbb{S}_{\beta+\beta'}} \lesssim \|u_0\|_{\mathbb{Z}_{\beta_0}} + \|b\|_{\mathbb{L}^{q_b}_{\gamma_b}(\mathbb{X}'_0)} \big(\|u\|_{\mathbb{S}_\beta} + \|u\|_{\mathbb{S}_0} \big)^2 < \infty.$$

(**Stability**) Let $\widetilde{u}$ be the unique solution corresponding to $(\widetilde{b}, \widetilde{u}_0)$. By (3.38) we have

$$u(t) - \widetilde{u}(t) = P_t u_0 - P_t \widetilde{u}_0 + \mathcal{H}_t^b(u, u) - \mathcal{H}_t^{\widetilde{b}}(\widetilde{u}, \widetilde{u}).$$

By symmetry, we may assume $\|\widetilde{u}_0\|_{\mathbb{Z}_{\beta_0}} \leqslant \|u_0\|_{\mathbb{Z}_{\beta_0}}$. By (3.36) and (3.37), we have

$$\|\mathcal{H}^b(u, u) - \mathcal{H}^b(\widetilde{u}, \widetilde{u})\|_{\mathbb{S}_0} \leqslant C_3 \|b\|_{\mathbb{L}_{\gamma_b}^{q_b}(\mathbb{X}_0')} \Big(\|u\|_{\mathbb{S}_0} + \|\widetilde{u}\|_{\mathbb{S}_0} \Big) \|u - \widetilde{u}\|_{\mathbb{S}_0} \leqslant \tfrac{1}{2} \|u - \widetilde{u}\|_{\mathbb{S}_0},$$

and

$$\|\mathcal{H}^b(\widetilde{u}, \widetilde{u}) - \mathcal{H}^{\widetilde{b}}(\widetilde{u}, \widetilde{u})\|_{\mathbb{S}_0} = \|\mathcal{H}^{b-\widetilde{b}}(\widetilde{u}, \widetilde{u})\|_{\mathbb{S}_0} \leqslant C_3 \|b - \widetilde{b}\|_{\mathbb{L}_{\gamma_b}^{q_b}(\mathbb{X}_0')} \|\widetilde{u}\|_{\mathbb{S}_0}^2.$$

Thus by (3.33),

$$\|u - \widetilde{u}\|_{\mathbb{S}_0} \leqslant C_2 \|u_0 - \widetilde{u}_0\|_{\mathbb{Z}_{\beta_0}} + \tfrac{1}{2} \|u - \widetilde{u}\|_{\mathbb{S}_0} + C_3 \|b - \widetilde{b}\|_{\mathbb{L}_{\gamma_b}^{q_b}(\mathbb{X}_0')},$$

which implies that

$$\|u - \widetilde{u}\|_{\mathbb{S}_0} \leqslant 2C_2 \|u_0 - \widetilde{u}_0\|_{\mathbb{Z}_{\beta_0}} + 2C_3 \|b - \widetilde{b}\|_{\mathbb{L}_{\gamma_b}^{q_b}(\mathbb{X}_0')}.$$

Next we use induction and (3.24) to show (3.32) for any $\beta > 0$. Suppose that we have shown (3.32) for some $\beta \geqslant 0$. Then by (3.33) and (3.24), we have for any $\beta' \in [0, \Lambda)$,

$$\begin{aligned}
\|u - \widetilde{u}\|_{\mathbb{S}_{\beta+\beta'}} &\leqslant \|P_\cdot(u_0 - \widetilde{u}_0)\|_{\mathbb{S}_{\beta+\beta'}} + \|\mathcal{H}^b(u, u) - \mathcal{H}^{\widetilde{b}}(\widetilde{u}, \widetilde{u})\|_{\mathbb{S}_{\beta+\beta'}} \\
&\lesssim \|u_0 - \widetilde{u}_0\|_{\mathbb{Z}_{\beta_0}} + \|\mathcal{H}^b(u, u) - \mathcal{H}^b(\widetilde{u}, \widetilde{u})\|_{\mathbb{S}_{\beta+\beta'}} + \|\mathcal{H}^{b-\widetilde{b}}(\widetilde{u}, \widetilde{u})\|_{\mathbb{S}_{\beta+\beta'}} \\
&\lesssim \|u_0 - \widetilde{u}_0\|_{\mathbb{Z}_{\beta_0}} + \|b\|_{\mathbb{L}_{\gamma_b}^{q_b}(\mathbb{X}_0')} \Big(\|u - \widetilde{u}\|_{\mathbb{S}_\beta} \|u\|_{\mathbb{S}_0} + \|u\|_{\mathbb{S}_\beta} \|u - \widetilde{u}\|_{\mathbb{S}_0} \\
&\quad + \|\widetilde{u}\|_{\mathbb{S}_\beta} \|u - \widetilde{u}\|_{\mathbb{S}_0} + \|u - \widetilde{u}\|_{\mathbb{S}_\beta} \|\widetilde{u}\|_{\mathbb{S}_0} \Big) + \|b - \widetilde{b}\|_{\mathbb{L}_{\gamma_b}^{q_b}(\mathbb{X}_0')} \|\widetilde{u}\|_{\mathbb{S}_\beta} \|\widetilde{u}\|_{\mathbb{S}_0},
\end{aligned}$$

which, by the induction hypothesis and (3.31), in turn implies that

$$\|u - \widetilde{u}\|_{\mathbb{S}_{\beta+\beta'}} \lesssim \|u_0 - \widetilde{u}_0\|_{\mathbb{Z}_{\beta_0}} + \|b - \widetilde{b}\|_{\mathbb{L}_{\gamma_b}^{q_b}(\mathbb{X}_0')}.$$

Thus we complete the proof. $\qquad\qquad\square$

Remark 3.12. Estimate (3.32) provides the continuous dependence of the solutions with respect to the kernel function b and initial value u_0. We assume $b = b_0 + b_1 \in \mathbb{L}_0^{q_b}(\mathbb{X}_0')$, where b_0 has low integrability and b_1 has high integrability, which is motivated by example (A.5). Moreover, we require $\gamma_b > (\mathcal{Q} - 1 - \mathscr{A}_1) \vee 0$, which means that b has some time decay property as $t \to \infty$. This of course rules out the *time-independent* b. However, when considering the finite time interval, we can drop this assumption. Indeed, for fixed $T \geqslant 1$, by definition it is easy to see that

$$\|b \mathbb{1}_{[0,T]}\|_{\mathbb{L}_{\gamma_b}^{q_b}(\mathbb{X}_0')} \leqslant T^{\frac{\gamma_b}{\alpha}} \|b\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')}. \tag{3.39}$$

Thus, we can use Theorem 3.11 to $b \mathbb{1}_{[0,T]}$. In the next subsection, we shall consider the large time decay estimate under extra assumptions about the initial value u_0 and integrability index ϱ_1.

Remark 3.13. Remark 3.6 is applicable to Theorem 3.11 except that (3.31) and (3.32) only hold for $\beta \in [0, \Lambda)$ since (3.24) still holds for $\beta = 0$ under the assumptions of Remark 3.6.

3.2. **Large time decay estimates.** In this subsection we drop the condition $\gamma_b > (\mathcal{Q} - 1 - \mathscr{A}_1) \vee 0$ by assuming extra assumptions about the initial value and integrability index ϱ_1. More precisely, we have

Theorem 3.14. *In addition* $(\mathbf{H}_0)$, *we assume* $\boldsymbol{a} \cdot \frac{d}{\varrho_1} > \mathcal{Q} - 1$, $\boldsymbol{p}_0 \neq 1$ *and* $b \in \mathbb{L}_0^{q_b}(\mathbb{X}_0')$. *For any* $u_0 \in \mathbb{Z}_{\beta_0} \cap \mathbf{B}_{1;\boldsymbol{a}}^{\beta_0}$, *there are two constants* $\delta \geqslant 1$ *and* $C_0 > 0$ *only depending on* Θ *such that if*

$$\|u_0\|_{\mathbb{Z}_{\beta_0}} \|b\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')} \Big(\big(\|b\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')} \|u_0\|_{\mathbf{B}_{1;\boldsymbol{a}}^{\beta_0}} \big) \vee 1 \Big)^\delta \leqslant C_0, \tag{3.40}$$

then there is a unique global weak solution u *to FPE* (3.1) *such that*

$$\sup_{t>0} \Big((1 \wedge t)^{-\frac{\beta_0}{\alpha}} \big(\|u(t)\|_{\mathbb{Y}_0} + \|u(t)\|_{\mathbf{B}_{1;\boldsymbol{a}}^{0,1}} \big) + (1 \wedge t)^{\frac{\mathscr{A}_0 - \beta_0 - \beta_b}{\alpha}} \|u(t)\|_{\mathbb{X}_0} \Big) < \infty. \tag{3.41}$$

Moreover, we have the following conclusions:

(i) **(Smoothness)** *For any $\beta \geqslant 0$, it holds that*

$$\sup_{t>0} \left((1 \wedge t)^{\frac{\beta-\beta_0}{\alpha}} \big(\|u(t)\|_{\mathbb{Y}_\beta} + \|u(t)\|_{\mathbf{B}_{1;a}^{\beta,1}} \big) + (1 \wedge t)^{\frac{\beta+\mathscr{A}_0-\beta_0-\beta_b}{\alpha}} \|u(t)\|_{\mathbb{X}_\beta} \right) < \infty. \tag{3.42}$$

(ii) **(Stability)** *For another $\widetilde{b} \in \mathbb{L}_0^{q_b}(\mathbb{X}_0')$ and $\widetilde{u}_0 \in \mathbb{Z}_{\beta_0}$ satisfying the same condition* (3.40), *let $\widetilde{u}$ be the unique solution corresponding to $\widetilde{b}$ and $\widetilde{u}_0$. For any $\beta \geqslant 0$, there is a constant $C_\beta > 0$ only depending on $\Theta, \|b\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')}, \|\widetilde{b}\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')}$ such that*

$$\sup_{t>0} \left((1 \wedge t)^{\frac{\beta-\beta_0}{\alpha}} \|u(t) - \widetilde{u}(t)\|_{\mathbb{Y}_\beta} \right) \leqslant C_\beta \big(\|u_0 - \widetilde{u}_0\|_{\mathbb{Z}_{\beta_0}} + \|b - \widetilde{b}\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')} \big). \tag{3.43}$$

(iii) **(Large time decay estimate)** *For any $\beta \geqslant 0$ and $i = 0, 1$, it holds that*

$$\sup_{t \geqslant 1} \left(t^{\frac{a \cdot (d - d/p_i)}{\alpha}} \|u(t)\|_{\mathbf{B}_{p_i;a}^{\beta,1}} \right) < \infty. \tag{3.44}$$

Proof. The proof is similar to Theorem 3.11. But we need to carefully treat the large time decay estimates. Since $a \cdot \frac{d}{\varrho_1} > \mathscr{Q} - 1$ and $p_0 \neq 1$,

$$\mathscr{Q} - 1 - (\mathscr{Q} - 1) \wedge \mathscr{A}_1 = (\mathscr{Q} - 1 - \mathscr{A}_1) \vee 0 < a \cdot \frac{d}{\varrho_1} - \mathscr{A}_1 = a \cdot (d - \frac{d}{p_1}),$$

one can fix two large time decay weights $\gamma_1, \ell_1 \in [0, \infty)$ with $\gamma_1 + \ell_1 \in [0, \mathscr{Q})$ and

$$\gamma_1 \in ((\mathscr{Q} - 1 - \mathscr{A}_1) \vee 0, a \cdot (d - \tfrac{d}{p_1})), \quad \ell_1 \in (\mathscr{Q} - 1, a \cdot \tfrac{d}{\varrho_1}). \tag{3.45}$$

Let the weights $\boldsymbol{\gamma}_\beta$ and $\boldsymbol{\ell}_\beta$ in the definition (3.21) of solution space $\mathbb{S}_\beta$ be replaced by

$$\boldsymbol{\gamma}_\beta := (\beta - \beta_0, \gamma_1), \quad \boldsymbol{\ell}_\beta := (\beta + \mathscr{A}_0 - \beta_0 - \beta_b, \ell_1). \tag{3.46}$$

In the following we divide the proof into three steps.

(*Step 1*) In this step we show that for any $\beta \geqslant 0$, there are constants $C_1, C_2, \delta_0, \delta_1 > 0$ only depending on Θ such that for any $T \geqslant 1$,

$$\|P.u_0\|_{\mathbb{S}_\beta} \leqslant C_1 T^{\delta_0} \|u_0\|_{\mathbb{Z}_{\beta_0}} + C_2 T^{-\delta_1} \|u_0\|_{\mathbf{B}_{1;a}^{\beta_0}}. \tag{3.47}$$

As in proving (3.10), by (2.34) and $p_1 \leqslant p_0$, we have

$$\sup_{t \geqslant 1} \left(t^{\frac{\gamma_1}{\alpha}} \|P_t u_0\|_{\mathbb{Y}_\beta} \right) \leqslant \sup_{t \in [1,T]} \left(t^{\frac{\gamma_1}{\alpha}} \|P_t u_0\|_{\mathbb{Y}_\beta} \right) + \sup_{t \geqslant T} \left(t^{\frac{\gamma_1}{\alpha}} \|P_t u_0\|_{\mathbb{Y}_\beta} \right)$$

$$\lesssim T^{\frac{\gamma_1}{\alpha}} \|u_0\|_{\mathbb{Z}_{\beta_0}} + \sup_{t \geqslant T} \left(t^{\frac{\gamma_1 - a \cdot (d - d/p_1)}{\alpha}} \|u_0\|_{\mathbf{B}_{1;a}^{\beta_0}} \right)$$

$$\leqslant T^{\frac{\gamma_1}{\alpha}} \|u_0\|_{\mathbb{Z}_{\beta_0}} + T^{\frac{\gamma_1 - a \cdot (d - d/p_1)}{\alpha}} \|u_0\|_{\mathbf{B}_{1;a}^{\beta_0}},$$

and by $\bar{\varrho}_1 \leqslant \bar{\varrho}_0$,

$$\sup_{t \geqslant 1} \left(t^{\frac{\ell_1}{\alpha}} \|P_t u_0\|_{\mathbb{X}_\beta} \right) \leqslant \sup_{t \in [1,T]} \left(t^{\frac{\ell_1}{\alpha}} \|P_t u_0\|_{\mathbb{X}_\beta} \right) + \sup_{t \geqslant T} \left(t^{\frac{\ell_1}{\alpha}} \|P_t u_0\|_{\mathbb{X}_\beta} \right)$$

$$\lesssim T^{\frac{\ell_1}{\alpha}} \|u_0\|_{\mathbb{Z}_{\beta_0}} + \sup_{t \geqslant T} \left(t^{\frac{\ell_1 - a \cdot d/\varrho_1}{\alpha}} \|u_0\|_{\mathbf{B}_{1;a}^{\beta_0}} \right)$$

$$\leqslant T^{\frac{\ell_1}{\alpha}} \|u_0\|_{\mathbb{Z}_{\beta_0}} + T^{\frac{\ell_1 - a \cdot d/\varrho_1}{\alpha}} \|u_0\|_{\mathbf{B}_{1;a}^{\beta_0}}.$$

Moreover, by (3.9), we also have

$$\sup_{t \in (0,1]} \left(t^{\frac{\gamma_{\beta_0}}{\alpha}} \|P_t u_0\|_{\mathbb{Y}_\beta} + t^{\frac{\ell_{\beta_0}}{\alpha}} \|P_t u_0\|_{\mathbb{X}_\beta} \right) \lesssim \|u_0\|_{\mathbb{Z}_{\beta_0}}.$$

Combining the above estimates and by definition (3.21), we obtain (3.47).

(*Step 2*) In this step we show the existence and uniqueness of a global solution under some smallness condition (3.40). With the weights (3.46), by Lemma 3.8, it is completely the same as Lemma 3.9 to derive that for any $\beta \geqslant 0$ and $\beta' \in [0, \Lambda)$,

$$\|\mathcal{H}^b(u,w)\|_{\mathbb{S}_{\beta+\beta'}} \lesssim \|b\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')} \big(\|u\|_{\mathbb{S}_\beta} + \|u\|_{\mathbb{S}_0} \big) \big(\|w\|_{\mathbb{S}_\beta} + \|w\|_{\mathbb{S}_0} \big). \tag{3.48}$$

Thus by (3.47) and (3.48), we can define a map $\mathcal{U} : \mathbb{S}_0 \to \mathbb{S}_0$ by

$$\mathcal{U}(u)(t) := P_t u_0 + \mathcal{H}_t^b(u,u)$$

so that for some constants $C_1, C_2, C_3 > 0$ only depending on Θ and for all $u \in \mathbb{S}_0$,

$$\|\mathcal{U}(u)\|_{\mathbb{S}_0} \leqslant C_1 T^{\delta_0} \|u_0\|_{\mathbb{Z}_{\beta_0}} + C_2 T^{-\delta_1} \|u_0\|_{\mathbf{B}_{1;a}^{\beta_0}} + C_3 \|b\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')} \|u\|_{\mathbb{S}_0}^2,$$

and for all $u, \widetilde{u} \in \mathbb{S}_0$,

$$\|\mathcal{U}(u) - \mathcal{U}(\widetilde{u})\|_{\mathbb{S}_0} = \|\mathcal{H}^b(u,u) - \mathcal{H}^b(\widetilde{u},\widetilde{u})\|_{\mathbb{S}_0} \leqslant C_3 \|b\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')} \big(\|u\|_{\mathbb{S}_0} + \|\widetilde{u}\|_{\mathbb{S}_0} \big) \|u - \widetilde{u}\|_{\mathbb{S}_0}.$$

Now let us choose $T \geqslant 1$ large enough and be fixed so that

$$C_2 T^{-\delta_1} \|u_0\|_{\mathbf{B}_{1;a}^{\beta_0}} \leqslant 1/(12 C_3 \|b\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')}).$$

Suppose that for $\delta = \delta_0/\delta_1$ and some $C_0 = C_0(\Theta) > 0$,

$$\|u_0\|_{\mathbb{Z}_{\beta_0}} \leqslant \frac{T^{-\delta_0}}{12 C_1 C_3 \|b\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')}} \leqslant \frac{C_0}{\|b\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')} ((\|b\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')} \|u_0\|_{\mathbf{B}_{1;a}^{\beta_0}} \vee 1)^\delta},$$

where we have used $T \geqslant 1$ in the second inequality. It is easy to see that for $\lambda = \frac{1}{4 C_3 \|b\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')}}$,

$$C_1 T^{\delta_0} \|u_0\|_{\mathbb{Z}_{\beta_0}} + C_2 T^{-\delta_1} \|u_0\|_{\mathbf{B}_{1;a}^{\beta_0}} + C_3 \|b\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')} \lambda^2 \leqslant \lambda.$$

In particular,

$$\mathcal{U} \text{ is a contraction mapping on } \mathscr{B}^\lambda := \big\{ u \in \mathbb{S}_0 : \|u\|_{\mathbb{S}_0} \leqslant \lambda \big\}. \tag{3.49}$$

On the other hand, letting

$$\boldsymbol{p} = \boldsymbol{p}' = \mathbf{1}, \quad \boldsymbol{\gamma}' = \boldsymbol{\gamma} = (-\beta_0, 0), \quad \theta = \mathscr{Q} - 1, \quad \bar{\boldsymbol{\gamma}} = \boldsymbol{\ell}_0,$$

by (3.14) with the above choice of parameters, we also have

$$\|\mathcal{H}^b(u,w)\|_{\mathbb{L}_{\boldsymbol{\gamma}}^\infty(\mathbf{B}_{1;a}^{0,1})} \lesssim \|b\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')} \|u\|_{\mathbb{L}_{\boldsymbol{\ell}_0}^\infty(\mathbb{X}_0)} \|w\|_{\mathbb{L}_{\boldsymbol{\gamma}}^\infty(\mathbf{B}_{1;a}^{0,1})}. \tag{3.50}$$

Moreover, by (2.33) and (2.34),

$$\|P_\cdot u_0\|_{\mathbb{L}_{\boldsymbol{\gamma}}^\infty(\mathbf{B}_{1;a}^{0,1})} \lesssim \|u_0\|_{\mathbf{B}_{1;a}^{\beta_0}}.$$

Hence, there are two constants $C_4, C_5 > 0$ such that for all $u, \widetilde{u} \in \mathbb{S}_0 \cap \mathbb{L}_{\boldsymbol{\gamma}}^\infty(\mathbf{B}_{1;a}^{0,1})$,

$$\|\mathcal{U}(u)\|_{\mathbb{L}_{\boldsymbol{\gamma}}^\infty(\mathbf{B}_{1;a}^{0,1})} \leqslant C_4 \|u_0\|_{\mathbf{B}_{1;a}^{\beta_0}} + C_5 \|b\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')} \|u\|_{\mathbb{S}_0} \|u\|_{\mathbb{L}_{\boldsymbol{\gamma}}^\infty(\mathbf{B}_{1;a}^{0,1})}, \tag{3.51}$$

and

$$\|\mathcal{U}(u) - \mathcal{U}(\widetilde{u})\|_{\mathbb{L}_{\boldsymbol{\gamma}}^\infty(\mathbf{B}_{1;a}^{0,1})} \leqslant C_5 \|b\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')} \big(\|u - \widetilde{u}\|_{\mathbb{S}_0} \|u\|_{\mathbb{L}_{\boldsymbol{\gamma}}^\infty(\mathbf{B}_{1;a}^{0,1})} + \|\widetilde{u}\|_{\mathbb{S}_0} \|u - \widetilde{u}\|_{\mathbb{L}_{\boldsymbol{\gamma}}^\infty(\mathbf{B}_{1;a}^{0,1})} \big). \tag{3.52}$$

Without loss of generality, we may assume $C_5 \leqslant C_3$, where C_3 is the same as above. Otherwise, we may replace above C_3 by bigger C_5. Thus by (3.49), (3.51) and (3.52), it is standard to derive that there is a unique $u \in \mathbb{S}_0 \cap \mathbb{L}_{\boldsymbol{\gamma}}^\infty(\mathbf{B}_{1;a}^{0,1})$ such that

$$u(t) = \mathcal{U}(u)(t) = P_t u_0 + \mathcal{H}_t^b(u,u)$$

and

$$\|u\|_{\mathbb{L}_{\boldsymbol{\gamma}}^\infty(\mathbf{B}_{1;a}^{0,1})} \lesssim \|u_0\|_{\mathbf{B}_{1;a}^{\beta_0}}. \tag{3.53}$$

Moreover, the proofs of smoothness and stability are completely the same as Theorem 3.11.

(*Step 3*) In this step we prove the optimal large time decay estimate (3.44) by shifting time variable as used in Lemma 3.9. Let

$$\eta_i := \boldsymbol{a} \cdot \left(\boldsymbol{d} - \tfrac{d}{\boldsymbol{p}_i}\right), \ i = 0, 1.$$

Note that by (3.4) and (3.13),

$$u(t+s) = P_s u(t) + \mathscr{H}^{b^t}_s(u^t, u^t), \quad s, t > 0,$$

where $u^t(s) := u(t+s)$. Thus for fixed $\beta \geqslant 0$ and any $t \geqslant 1$ and $i = 0, 1$,

$$\|u(2t)\|_{\mathbf{B}^{\beta,1}_{\boldsymbol{p}_i;\boldsymbol{a}}} \leqslant \|P_t u(t)\|_{\mathbf{B}^{\beta,1}_{\boldsymbol{p}_i;\boldsymbol{a}}} + \|\mathscr{H}^{b^t}_t(u^t, u^t)\|_{\mathbf{B}^{\beta,1}_{\boldsymbol{p}_i;\boldsymbol{a}}}.$$

For the first term, by (2.34) and (3.53) we have

$$\|P_t u(t)\|_{\mathbf{B}^{\beta,1}_{\boldsymbol{p}_i;\boldsymbol{a}}} \lesssim t^{-\frac{\eta_i}{\alpha}} \|u(t)\|_{\mathbf{B}^0_{1;\boldsymbol{a}}} \lesssim t^{-\frac{\eta_i}{\alpha}} \|u_0\|_{\mathbf{B}^{\beta_0}_{1;\boldsymbol{a}}}, \quad t \geqslant 1.$$

For the second term, by (3.14) with $\theta = \mathscr{Q} - 1$ and

$$\bar{\boldsymbol{\gamma}} = \boldsymbol{\gamma} = \boldsymbol{\gamma}_b = 0, \quad \boldsymbol{\gamma}' = (-\theta, \gamma'_1), \ \gamma'_1 \in (-\ell_1, 1 - \mathscr{Q}),$$

we derive that for any $t \geqslant 1$ and $i = 0, 1$,

$$\|\mathscr{H}^{b^t}_\cdot(u^t, u^t)\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}'}(\mathbf{B}^{\beta,1}_{\boldsymbol{p}_i;\boldsymbol{a}})} \lesssim \|b^t\|_{\mathbb{L}^{q_b}_0(\mathbb{X}'_0)} \|u^t\|_{\mathbb{L}^\infty_0(\mathbb{X}_\beta)} \|u^t\|_{\mathbb{L}^\infty_0(\mathbf{B}^{\beta,1}_{\boldsymbol{p}_i;\boldsymbol{a}})}$$

$$\overset{(2.32)}{\lesssim} \|b\|_{\mathbb{L}^{q_b}_0(\mathbb{X}'_0)} t^{-\frac{\ell_1}{\alpha}} \|u\|_{\mathbb{L}^\infty_{\boldsymbol{\ell}_\beta}(\mathbb{X}_\beta)} \|u^t\|_{\mathbb{L}^\infty_0(\mathbf{B}^{\beta,1}_{\boldsymbol{p}_i;\boldsymbol{a}})}.$$

Combining the above two estimates, we obtain that for any $t \geqslant 1$,

$$\|u(2t)\|_{\mathbf{B}^{\beta,1}_{\boldsymbol{p}_i;\boldsymbol{a}}} \lesssim t^{-\frac{\eta_i}{\alpha}} \|u_0\|_{\mathbf{B}^{\beta_0}_{1;\boldsymbol{a}}} + t^{-\frac{\gamma'_1+\ell_1}{\alpha}} \|u^t\|_{\mathbb{L}^\infty_0(\mathbf{B}^{\beta,1}_{\boldsymbol{p}_i;\boldsymbol{a}})}, \ i = 0, 1.$$

In particular, there are two constants $C_6 = C_6(\beta, \Theta) > 0$ and $C_7 = C_7(\beta, \Theta, \|b\|_{\mathbb{L}^{q_b}_0(\mathbb{X}'_0)}) > 0$ such that for any $t \geqslant 2$,

$$\|u(t)\|_{\mathbf{B}^{\beta,1}_{\boldsymbol{p}_i;\boldsymbol{a}}} \leqslant C_6 t^{-\frac{\eta_i}{\alpha}} \|u_0\|_{\mathbf{B}^{\beta_0}_{1;\boldsymbol{a}}} + C_7 t^{-\frac{\gamma'_1+\ell_1}{\alpha}} \|u^{t/2}\|_{\mathbb{L}^\infty_0(\mathbf{B}^{\beta,1}_{\boldsymbol{p}_i;\boldsymbol{a}})}, \ i = 0, 1.$$

Since $\gamma'_1 + \ell_1 > 0$, one can choose $T_0 = T_0(\beta, \Theta, \|b\|_{\mathbb{L}^{q_b}_0(\mathbb{X}'_0)}, \gamma'_1, \ell_1) \geqslant 2$ large enough so that

$$C_7 T_0^{-\frac{\gamma'_1+\ell_1}{\alpha}} \leqslant 2^{-1-\frac{\eta_i}{\alpha}}, \quad i = 0, 1,$$

and so,

$$\sup_{t \geqslant T_0} \left(t^{\frac{\eta_i}{\alpha}} \|u(t)\|_{\mathbf{B}^{\beta,1}_{\boldsymbol{p}_i;\boldsymbol{a}}}\right) \leqslant C_6 \|u_0\|_{\mathbf{B}^{\beta_0}_{1;\boldsymbol{a}}} + 2^{-1-\frac{\eta_i}{\alpha}} \sup_{t \geqslant T_0} \left(t^{\frac{\eta_i}{\alpha}} \|u^{t/2}\|_{\mathbb{L}^\infty_0(\mathbf{B}^{\beta,1}_{\boldsymbol{p}_i;\boldsymbol{a}})}\right)$$

$$= C_6 \|u_0\|_{\mathbf{B}^{\beta_0}_{1;\boldsymbol{a}}} + 2^{-1-\frac{\eta_i}{\alpha}} \sup_{t \geqslant T_0} \left(t^{\frac{\eta_i}{\alpha}} \sup_{s \geqslant t/2} \|u(s)\|_{\mathbf{B}^{\beta,1}_{\boldsymbol{p}_i;\boldsymbol{a}}}\right)$$

$$\leqslant C_6 \|u_0\|_{\mathbf{B}^{\beta_0}_{1;\boldsymbol{a}}} + \frac{1}{2} \sup_{t \geqslant T_0} \sup_{s \geqslant t/2} \left(s^{\frac{\eta_i}{\alpha}} \|u(s)\|_{\mathbf{B}^{\beta,1}_{\boldsymbol{p}_i;\boldsymbol{a}}}\right)$$

$$= C_6 \|u_0\|_{\mathbf{B}^{\beta_0}_{1;\boldsymbol{a}}} + \frac{T_0^{\frac{\eta_i}{\alpha}}}{2} \sup_{s \in [\frac{T_0}{2}, T_0]} \|u(s)\|_{\mathbf{B}^{\beta,1}_{\boldsymbol{p}_i;\boldsymbol{a}}} + \frac{1}{2} \sup_{s \geqslant T_0} \left(s^{\frac{\eta_i}{\alpha}} \|u(s)\|_{\mathbf{B}^{\beta,1}_{\boldsymbol{p}_i;\boldsymbol{a}}}\right).$$

Hence,

$$\sup_{t \geqslant T_0} \left(t^{\frac{\eta_i}{\alpha}} \|u(t)\|_{\mathbf{B}^{\beta,1}_{\boldsymbol{p}_i;\boldsymbol{a}}}\right) \leqslant 2C_6 \|u_0\|_{\mathbf{B}^{\beta_0}_{1;\boldsymbol{a}}} + T_0^{\frac{\eta_i}{\alpha}} \sup_{s \in [\frac{T_0}{2}, T_0]} \|u(s)\|_{\mathbf{B}^{\beta,1}_{\boldsymbol{p}_i;\boldsymbol{a}}} \leqslant C_8.$$

The large time decay estimate now follows. $\qquad\qquad\square$

3.3. **Local well-posedness of large initial value.** In this section we study the local well-posedness for FPE (3.1) with large initial values. First of all we show the following asymptotic estimate of short time for $P_t f$ in Besov spaces.

Lemma 3.15. *Let* $\beta, \beta' \in \mathbb{R}$ *and* $\boldsymbol{p}, \boldsymbol{p}' \in [1, \infty]^2$ *with* $\boldsymbol{p} \leqslant \boldsymbol{p}'$. *Let* $\mathscr{A} := \boldsymbol{a} \cdot (\frac{d}{\boldsymbol{p}} - \frac{d}{\boldsymbol{p}'})$. *Suppose that*

$$\beta < 0, \quad \mathscr{A} + \beta' - \beta > 0.$$

For any $f \in \cup_{q \in [1,\infty)} \mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta;q}$, *it holds that*

$$\lim_{t \downarrow 0} \left(t^{\frac{\mathscr{A} + \beta' - \beta}{\alpha}} \|P_t f\|_{\mathbf{B}_{\boldsymbol{p}';\boldsymbol{a}}^{\beta',1}} \right) = 0$$

Proof. By the Bernstein inequality (2.7) and (2.28), we have

$$\|\mathscr{R}_j^{\boldsymbol{a}} P_t f\|_{\boldsymbol{p}'} \lesssim 2^{j\mathscr{A}} \|\mathscr{R}_j^{\boldsymbol{a}} P_t f\|_{\boldsymbol{p}} \leqslant 2^{j\mathscr{A}} \sum_{\ell \in \Theta_j^t} \|\mathscr{R}_j^{\boldsymbol{a}} \Gamma_t p_t * \Gamma_t \mathscr{R}_\ell^{\boldsymbol{a}} f\|_{\boldsymbol{p}}.$$

Here and below the implicit constant does not depend on $t \in (0, 1)$. Thus by Young's inequality,

$$t^{\frac{\mathscr{A}+\beta'-\beta}{\alpha}} \sum_{j \geqslant 1} 2^{j\beta'} \|\mathscr{R}_j^{\boldsymbol{a}} P_t f\|_{\boldsymbol{p}'} \lesssim t^{\frac{\mathscr{A}+\beta'-\beta}{\alpha}} \sum_{j \geqslant 1} \sum_{\ell \in \Theta_j^t} 2^{j(\mathscr{A}+\beta')} \|\mathscr{R}_j^{\boldsymbol{a}} \Gamma_t p_t\|_1 \|\mathscr{R}_\ell^{\boldsymbol{a}} f\|_{\boldsymbol{p}} = \sum_{\ell \geqslant 0} c_{\ell,t} 2^{\ell\beta} \|\mathscr{R}_\ell^{\boldsymbol{a}} f\|_{\boldsymbol{p}},$$

where

$$c_{\ell,t} := t^{\frac{\mathscr{A}+\beta'-\beta}{\alpha}} \sum_{j \geqslant 1} 2^{j(\mathscr{A}+\beta')} 2^{-\ell\beta} \mathbb{1}_{\ell \in \Theta_j^t} \|\mathscr{R}_j^{\boldsymbol{a}} \Gamma_t p_t\|_1.$$

By the definition of Θ_j^t, Lemmas 2.8 and A.1, for $l \geqslant \frac{2(\mathscr{A}+\beta'-\beta)}{\alpha}$, we have

$$c_{\ell,t} = \sum_{j \geqslant 1} (t 2^{j\alpha})^{\frac{\mathscr{A}+\beta'}{\alpha}} (t 2^{\ell\alpha})^{-\frac{\beta}{\alpha}} \mathbb{1}_{\ell \in \Theta_j^t} \|\mathscr{R}_j^{\boldsymbol{a}} \Gamma_t p_t\|_1$$

$$\lesssim \sum_{j \geqslant 1} (t 2^{j\alpha})^{\frac{\mathscr{A}+\beta'}{\alpha}} (t 2^{j\alpha} (1 + t 2^{j\alpha}))^{-\frac{\beta}{\alpha}} ((t 2^{j\alpha})^{-l} \wedge 1) \lesssim 1,$$

and

$$\sum_{\ell \geqslant 1} c_{\ell,t} = t^{\frac{\mathscr{A}+\beta'-\beta}{\alpha}} \sum_{j \geqslant 1} 2^{j(\mathscr{A}+\beta')} \sum_{\ell \in \Theta_j^t} 2^{-\ell\beta} \|\mathscr{R}_j^{\boldsymbol{a}} \Gamma_t p_t\|_1$$

$$\overset{(2.24)}{\lesssim} t^{\frac{\mathscr{A}+\beta'-\beta}{\alpha}} \sum_{j \geqslant 1} 2^{j(\mathscr{A}+\beta'-\beta)} (1 + (t 2^{j\alpha}))^{-\beta} ((t 2^{j\alpha})^{-l} \wedge 1) \lesssim 1.$$

Hence, for any $q \in [1, \infty)$, by Hölder's inequality,

$$t^{\frac{\mathscr{A}+\beta'-\beta}{\alpha}} \sum_{j \geqslant 1} 2^{j\beta'} \|\mathscr{R}_j^{\boldsymbol{a}} P_t f\|_{\boldsymbol{p}'} \lesssim \left[\sum_{\ell \geqslant 0} c_{\ell,t} (2^{\ell\beta} \|\mathscr{R}_\ell^{\boldsymbol{a}} f\|_{\boldsymbol{p}})^q \right]^{\frac{1}{q}} \left[\sum_{\ell \geqslant 1} c_{\ell,t} \right]^{1-\frac{1}{q}} \lesssim \left[\sum_{\ell \geqslant 0} c_{\ell,t} (2^{\ell\beta} \|\mathscr{R}_\ell^{\boldsymbol{a}} f\|_{\boldsymbol{p}})^q \right]^{\frac{1}{q}}.$$

In particular, if $f \in \mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta_0,q}$, then by the dominated convergence theorem,

$$\lim_{t \downarrow 0} \left(t^{\frac{\mathscr{A}+\beta'-\beta}{\alpha}} \sum_{j \geqslant 1} 2^{j\beta'} \|\mathscr{R}_j^{\boldsymbol{a}} P_t f\|_{\boldsymbol{p}'} \right) \lesssim \left[\sum_{\ell \geqslant 0} \lim_{t \downarrow 0} c_{\ell,t} (2^{\ell\beta} \|\mathscr{R}_\ell^{\boldsymbol{a}} f\|_{\boldsymbol{p}})^q \right]^{\frac{1}{q}} = 0. \qquad (3.54)$$

On the other hand, by (2.26), for $t \in (0, 1)$,

$$\|\mathscr{R}_0^{\boldsymbol{a}} P_t f\|_{\boldsymbol{p}'} \lesssim \|f\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta}},$$

which together with (3.54) yields the desired limit. $\qquad \square$

The following lemma is a direct application of the above lemma.

Lemma 3.16. *Under* $(\mathbf{H}_0)$, *for any* $f \in \cup_{q\in[1,\infty)}(\mathbf{B}_{\boldsymbol{p}_0;\boldsymbol{a}}^{\beta_0;q} \cap \mathbf{B}_{\boldsymbol{p}_1;\boldsymbol{a}}^{\beta_0;q})$, *it holds that*

$$\lim_{t\downarrow 0}\left(t^{\frac{\mathscr{A}_0-\beta_0-\beta_b}{\alpha}}\|P_tf\|_{\mathbb{X}_0} + t^{-\frac{\beta_0}{\alpha}}\|P_tf\|_{\mathbb{Y}_0}\right) = 0. \tag{3.55}$$

Now we can show the following local well-posedness for FPE (3.1) with large initial value.

Theorem 3.17. *Suppose that* $(\mathbf{H}_0)$ *and one of the following two conditions hold:*

(i) $b \in \mathbb{L}_0^{q_b}(\mathbb{X}_0')$ *for some* $q_b \in [1,\infty)$ *and* $u_0 \in \mathbb{Z}_{\beta_0} = \mathbf{B}_{\boldsymbol{p}_0;\boldsymbol{a}}^{\beta_0;\infty} \cap \mathbf{B}_{\boldsymbol{p}_1;\boldsymbol{a}}^{\beta_0;\infty}$;

(ii) $b \in \mathbb{L}_0^{\infty}(\mathbb{X}_0')$ *and* $u_0 \in \widetilde{\mathbb{Z}}_{\beta_0}$, *where* $\widetilde{\mathbb{Z}}_{\beta_0} := \cup_{q\in[1,\infty)}(\mathbf{B}_{\boldsymbol{p}_0;\boldsymbol{a}}^{\beta_0;q} \cap \mathbf{B}_{\boldsymbol{p}_1;\boldsymbol{a}}^{\beta_0;q}) \subset \mathbb{Z}_{\beta_0}$.

Then there is a small time $T_0 > 0$ *depending only on* Θ *and the corresponding norms of* b *and* u_0 *so that* (3.1) *admits a unique smooth solution* u *on* $(0,T_0]$ *and for any* $\beta \geqslant 0$,

$$\|u\mathbb{1}_{[0,T_0]}\|_{\mathbb{S}_\beta} < \infty,$$

where $\mathbb{S}_\beta$ *is defined in* (3.21). *Moreover, for another* $(\widetilde{b},\widetilde{u}_0)$ *satisfying the above (i) or (ii), let* $\widetilde{u}$ *be the unique local solution corresponding to* $(\widetilde{b},\widetilde{u}_0)$ *on the time interval* $[0,\widetilde{T}_0]$. *Then for any* $\beta \geqslant 0$,

$$\|(u-\widetilde{u})\mathbb{1}_{[0,T_0\wedge\widetilde{T}_0]}\|_{\mathbb{S}_\beta} \leqslant C_\beta\big(\|u_0-\widetilde{u}_0\|_{\mathbb{Z}_{\beta_0}} + \|b-\widetilde{b}\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')}\big), \tag{3.56}$$

where $C_\beta = C_\beta\big(T_0,\widetilde{T}_0,\Theta,\|b\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')},\|\widetilde{b}\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')}\big) > 0$.

Proof. (i) Since for $q_b \in [1,\infty)$,

$$\lim_{T\downarrow 0}\|b\mathbb{1}_{[0,T]}\|_{\mathbb{L}_{\gamma_b}^{q_b}(\mathbb{X}_0')} = 0,$$

for any $u_0 \in \mathbb{Z}_{\beta_0}$, one can choose T_0 small enough so that for the constant C_0 in Theorem 3.11,

$$\|u_0\|_{\mathbb{Z}_{\beta_0}} \leqslant C_0/\|b\mathbb{1}_{[0,T_0]}\|_{\mathbb{L}_{\gamma_b}^{q_b}(\mathbb{X}_0')}.$$

Thus, we conclude the proof by Theorem 3.11.

(ii) Suppose $u_0 \in \widetilde{\mathbb{Z}}_{\beta_0}$. By the definition of $\mathbb{S}_0$ and (3.55), we have

$$\lim_{T\downarrow 0}\|\mathbb{1}_{(0,T]}(P_.u_0)\|_{\mathbb{S}_0} = \lim_{T\downarrow 0}\sup_{t\in(0,T]}\left(t^{\frac{\mathscr{A}_0-\beta_0-\beta_b}{\alpha}}\|P_tu_0\|_{\mathbb{X}_0} + t^{-\frac{\beta_0}{\alpha}}\|P_tu_0\|_{\mathbb{Y}_0}\right) = 0.$$

From the proof of Theorem 3.11, one can choose T_0 small enough so that $\mathcal{U}$ defined by (3.34) is a contraction mapping when we limit on the time interval $[0,T_0]$. Thus we can repeat the proof of Theorem 3.11 to conclude the proof. $\square$

Remark 3.18. If the condition (3.5) is replaced with

$$0 < \mathscr{A}_0 - \beta_0 - \beta_b < \mathscr{Q} - 1 \quad \text{and} \quad \mathscr{A}_0 - 2\beta_0 - \beta_b < \mathscr{Q},$$

which is equivalent to

$$0 < \mathscr{A}_0 - \beta_0 - \beta_b < \mathscr{Q} + \beta_0 \wedge (-1), \tag{3.57}$$

then the cases (i) and (ii) in Theorem 3.17 can be combined with only one case that $b \in \mathbb{L}_0^{q_b}(\mathbb{X}_0')$ for some $q_b \in [1,\infty]$ and $u_0 \in \mathbb{Z}_{\beta_0}$ since $\mathbb{Z}_{\beta_0} \subset \cap_{\varepsilon>0}\widetilde{\mathbb{Z}}_{\beta_0+\varepsilon}$. We would like to point out that the above (i) and (ii) allows us to consider some critical cases.

3.4. **Global well-posedness of large initial value.** In this subsection we shall consider two cases that the smallness condition (3.29) in Theorem 3.11 can be dropped.

Our first result of global well-posedness with large initial value is

Theorem 3.19. *Let* $\alpha \in (1,2]$, $q_b \in (\frac{\alpha}{\alpha-1},\infty]$ *and* $\boldsymbol{\varrho}_0, \boldsymbol{\varrho}_1 \in [1,\infty]^2$ *with* $\boldsymbol{\varrho}_0 \leqslant \boldsymbol{\varrho}_1$. *Let* $\beta_b \in \mathbb{R}$ *and* $\beta_0 \in (\frac{\alpha}{q_b}-\alpha,0)$. *Suppose that* $\beta_b \neq \boldsymbol{a}\cdot\frac{d}{\boldsymbol{\varrho}_i}-\beta_0$, $i=0,1$ *and*

$$0 < \boldsymbol{a}\cdot\frac{d}{\boldsymbol{\varrho}_0}-\beta_0-\beta_b \leqslant \alpha-\frac{\alpha}{q_b}-1, \quad \boldsymbol{a}\cdot\frac{d}{\boldsymbol{\varrho}_0}-2\beta_0-\beta_b < \alpha-\frac{\alpha}{q_b}, \tag{3.58}$$

and one of the following two conditions holds:

(i) $b \in \mathbb{L}_0^{q_b}(\mathbb{X}_0')$ *for some* $q_b \in [1,\infty)$ *and* $u_0 \in \mathbf{B}_{1;\boldsymbol{a}}^{\beta_0;\infty}$;

(ii) $b \in \mathbb{L}_0^\infty(\mathbb{X}_0')$ *and* $u_0 \in \cup_{q \in [1,\infty)} \mathbf{B}_{1;\boldsymbol{a}}^{\beta_0,q} \subset \mathbf{B}_{1;\boldsymbol{a}}^{\beta_0,\infty}$.

Then there is a unique smooth solution u *to* (3.1) *on* $\mathbb{R}_+$ *so that for any* $T, \beta > 0$,

$$\sup_{t \in (0,T]} \left(t^{(\beta-\beta_0)/\alpha} \|u(t)\|_{\mathbf{B}_{1;\boldsymbol{a}}^{\beta,1}} \right) < \infty.$$

Moreover, for another $(\widetilde{b}, \widetilde{u}_0)$ *satisfying the above (i) or (ii), let* $\widetilde{u}$ *be the unique solution corresponding to* $(\widetilde{b}, \widetilde{u}_0)$. *For any* $T, \beta > 0$, *there is a constant* $C_\beta = C_\beta(T, \Theta, \|b\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')}, \|\widetilde{b}\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')}) > 0$ *such that*

$$\sup_{t \in (0,T]} \left(t^{(\beta-\beta_0)/\alpha} \|u(t) - \widetilde{u}(t)\|_{\mathbf{B}_{1;\boldsymbol{a}}^{\beta,1}} \right) \leqslant C_\beta \left(\|u_0 - \widetilde{u}_0\|_{\mathbf{B}_{1;\boldsymbol{a}}^{\beta_0}} + \|b - \widetilde{b}\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')} \right).$$

Proof. If we take $\boldsymbol{p}_0 = 1$ in $(\mathbf{H}_0)$, then (3.5) reduces to (3.58). By Theorem 3.17, let $T_0 > 0$ be the short time of the existence of a smooth solution. For any fixed $t_0 \in (0, T_0]$, we have

$$u \in \cap_{\beta>0} C([t_0, T_0]; \mathbf{B}_{1;\boldsymbol{a}}^{\beta})$$

and

$$\partial_t u = \Delta_v^{\alpha/2} u - v \cdot \nabla_x u - \operatorname{div}_v((b * u)u), \ t \in [t_0, T_0].$$

Now, starting from the time t_0, since $\beta_0 < 0$ and $\|u(t_0)\|_{\mathbf{B}_{1;\boldsymbol{a}}^{\beta_0,q}} \leqslant C \|u(t_0)\|_{\mathbb{L}^1}$ for any $q \in [1,\infty]$, by Theorem 3.17 again, there is a time length $\delta > 0$ only depending on Θ, $\|b\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')}$ and the bound of $\|u(t_0)\|_{\mathbb{L}^1}$ so that there is a unique smooth solution on the time interval $[t_0, t_0 + \delta]$. On the other hand, by Lemma C.1 in appendix we have

$$\|u(t_0 + \delta/2)\|_{\mathbb{L}^1} \leqslant \|u(t_0)\|_{\mathbb{L}^1}.$$

Next starting from the time $t_0 + \delta/2$, we can proceed to find a unique smooth solution to (3.1) on the time interval $[t_0 + \delta/2, t_0 + 3\delta/2]$ and so on. Thus we obtain a global solution. The regularity and stability estimates are direct consequences of Theorem 3.17. $\qquad\square$

Our second global well-posedness result of large initial value is

Theorem 3.20. *Suppose that* $(\mathbf{H}_0)$ *holds with* $\boldsymbol{p}_0 = (p_0, p_0)$ *and* $\boldsymbol{\varrho}_1 = (\varrho_1, \varrho_1)$, *and one of the following two conditions holds:*

(i) $b \in \mathbb{L}_0^{q_b}(\mathbb{X}_0')$ *for some* $q_b \in [1,\infty)$, $\operatorname{div}_v b = 0$ *and* $u_0 \in \mathbb{Z}_{\beta_0} = \mathbf{B}_{\boldsymbol{p}_0;\boldsymbol{a}}^{\beta_0,\infty} \cap \mathbf{B}_{\boldsymbol{p}_1;\boldsymbol{a}}^{\beta_0,\infty}$;

(ii) $b \in \mathbb{L}_0^\infty(\mathbb{X}_0')$, $\operatorname{div}_v b = 0$ *and* $u_0 \in \widetilde{\mathbb{Z}}_{\beta_0}$, *where* $\widetilde{\mathbb{Z}}_{\beta_0} := \cup_{q \in [1,\infty)} (\mathbf{B}_{\boldsymbol{p}_0;\boldsymbol{a}}^{\beta_0,q} \cap \mathbf{B}_{\boldsymbol{p}_1;\boldsymbol{a}}^{\beta_0,q}) \subset \mathbb{Z}_{\beta_0}$.

Then there is a unique smooth solution u *to* (3.1) *on* $\mathbb{R}_+$ *so that for any* $T, \beta > 0$,

$$\|u \mathbb{1}_{[0,T]}\|_{\mathbb{S}_\beta} < \infty,$$

where $\mathbb{S}_\beta$ *is defined by* (3.21). *Moreover, for another* $(\widetilde{b}, \widetilde{u}_0)$ *satisfying the above (i) or (ii), let* $\widetilde{u}$ *be the unique solution on* $\mathbb{R}_+$ *corresponding to* $(\widetilde{b}, \widetilde{u}_0)$. *Then for any* $T, \beta > 0$, *there is a constant* $C_\beta = C_\beta(T, \Theta, \|b\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')}, \|\widetilde{b}\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')}) > 0$ *such that*

$$\|(u - \widetilde{u}) \mathbb{1}_{[0,T]}\|_{\mathbb{S}_\beta} \leqslant C_\beta \left(\|u_0 - \widetilde{u}_0\|_{\mathbb{Z}_{\beta_0}} + \|b - \widetilde{b}\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')} \right).$$

Proof. Let $T_0 > 0$ be the short time of the existence of a smooth solution in Theorem 3.17. For any fixed $t_0 \in (0, T_0)$, we have

$$u \in C([t_0, T_0]; C_b^\infty(\mathbb{R}^{2d})), \quad H := b * u \in L^1([t_0, T_0]; C_b^1(\mathbb{R}^{2d})),$$

and

$$\partial_t u = \Delta_v^{\alpha/2} u - v \cdot \nabla_x u - \operatorname{div}_v(Hu), \ t \in [t_0, T_0].$$

Since $\operatorname{div}_v b = 0$, we have $\operatorname{div}_v H = 0$. Thus by Lemma C.1 in appendix, for $t \in (t_0, T_0)$,

$$\|u(t)\|_{\mathbb{L}^1} \leqslant \|u(t_0)\|_{\mathbb{L}^1}, \quad \|u(t)\|_{\mathbb{L}^\infty} \leqslant \|u(t_0)\|_{\mathbb{L}^\infty}. \tag{3.59}$$

Noting that for $p \in [1, \infty]$,

$$\|f\|_{\mathbb{L}^p} \leqslant \|f\|_{\mathbb{L}^\infty}^{1-\frac{1}{p}} \|f\|_{\mathbb{L}^1}^{\frac{1}{p}},$$

by $\boldsymbol{p}_0 = (p_0, p_0)$ and $\boldsymbol{\varrho}_1 = (\varrho_1, \varrho_1)$, recalling $p_1 = p_0 \wedge \frac{\varrho_1}{\varrho_1 - 1}$, we have for any $q \in [1, \infty]$,

$$\sum_{i=0,1} \|u(t_0)\|_{\mathbf{B}_{\boldsymbol{p}_i;\boldsymbol{a}}^{\beta_0;q}} \overset{(2.12)}{\lesssim} \sum_{i=0,1} \|u(t_0)\|_{\mathbb{L}^{p_i}} \leqslant \sum_{i=0,1} \|u(t_0)\|_{\mathbb{L}^\infty}^{1-\frac{1}{p_i}} \|u(t_0)\|_{\mathbb{L}^1}^{\frac{1}{p_i}} =: \flat(t_0).$$

Now, starting from t_0, by Theorem 3.17 again, there is a time length $\delta > 0$ depending only on Θ, $\|b\|_{\mathbb{L}_0^{q_b}(\mathbb{X}_0')}$ and the bound of $\flat(t_0)$ so that there is a unique solution on the time interval $[t_0, t_0 + \delta]$. On the other hand, by (3.59) we have

$$\flat(t_0 + \delta/2) \leqslant \flat(t_0).$$

Next starting from the time $t_0 + \delta/2$, we can proceed to find a unique smooth solution to (3.1) on the time interval $[t_0 + \delta/2, t_0 + 3\delta/2]$ and so on. Thus we obtain a global solution. The regularity and stability estimates are direct consequences of Theorem 3.17. $\qquad\square$

Remark 3.21. If u is probability measure, then the condition $\mathrm{div}_v b = 0$ can be replaced by $(\mathrm{div}_v b)^- = 0$ since $(\mathrm{div}_v H)^- = (\mathrm{div}_v b)^- * u = 0$, and by Lemma C.1 we still have (3.59).

As a special case, we also consider the non-degenerate fractional Fokker-Planck equation:

$$\partial_t u = \Delta^{\alpha/2} u + \mathrm{div}((b * u)u), \quad u(0) = u_0. \tag{3.60}$$

By Duhamel's formula, we have

$$u(t) = P_t u_0 + \int_0^t P_{t-s} \mathrm{div}((b * u)u) \mathrm{d}s = P_t u_0 + \mathscr{H}_t^b(u, u),$$

where P_t is the semigroup associated with $\Delta^{\alpha/2}$. We assume

$(\widetilde{\mathbf{H}}_0)$ Let $\alpha \in (1, 2]$, $q_b \in (\frac{\alpha}{\alpha-1}, \infty]$ and $p_0, \varrho_0, \varrho_1 \in [1, \infty]$ with $\varrho_0 \leqslant \varrho_1$ and $1 \leqslant \frac{1}{p_0} + \frac{1}{\varrho_0}$. Let $\bar{\varrho}_0, \bar{\varrho}_1 \in [1, \infty]$ be defined by $\frac{1}{\bar{\varrho}_0} + \frac{1}{\varrho_0} = 1$ and $\frac{1}{\bar{\varrho}_1} + \frac{1}{\varrho_1} = 1$. Let $p_1 := p_0 \wedge \bar{\varrho}_1$ and

$$\mathscr{Q} := \alpha - \frac{\alpha}{q_b}, \quad \mathscr{A}_i := \frac{d}{p_i} + \frac{d}{\varrho_i} - d = \frac{d}{p_i} - \frac{d}{\bar{\varrho}_i}, \quad i = 0, 1.$$

Let $\beta_0 \in (-\mathscr{Q}, 0)$ and $\mathbb{R} \ni \beta_b \neq \mathscr{A}_i - \beta_0, i = 0, 1$. Suppose that

$$0 < \mathscr{A}_0 - \beta_0 - \beta_b \leqslant \mathscr{Q} - 1 \quad \text{and} \quad \mathscr{A}_0 - 2\beta_0 - \beta_b < \mathscr{Q}. \tag{3.61}$$

As in (3.8), we introduce three isotropic Besov spaces: for $\beta \geqslant 0$,

$$\bar{\mathbb{X}}_\beta := \bigcap_{i=0,1} \mathbf{B}_{\bar{\varrho}_i}^{\beta-\beta_b,1}, \quad \bar{\mathbb{Y}}_\beta := \bigcap_{i=0,1} \mathbf{B}_{p_i}^{\beta,1}, \quad \bar{\mathbb{Z}}_\beta := \bigcap_{i=0,1} \mathbf{B}_{p_i}^{\beta}. \tag{3.62}$$

The following result is a combination of Theorems 3.11, 3.14, 3.19 and 3.20. Since it can be considered as a special case and the proofs are completely the same, we omit the details.

Theorem 3.22. Suppose $(\widetilde{\mathbf{H}}_0)$ and $b \in \mathbb{L}_0^{q_b}(\bar{\mathbb{X}}_0')$, where $\bar{\mathbb{X}}_0' = \mathbf{B}_{\varrho_0}^{\beta_0} + \mathbf{B}_{\varrho_1}^{\beta_0}$ is the dual space of $\bar{\mathbb{X}}_0$. For any $\gamma_b > (\mathscr{Q} - 1 - \mathscr{A}_1) \vee 0$, there is a constant $C_0 = C_0(\Theta) > 0$ such that for any fixed $T \geqslant 1$ and $u_0 \in \bar{\mathbb{Z}}_{\beta_0}$ satisyfying

$$\|u_0\|_{\bar{\mathbb{Z}}_{\beta_0}} \|b\|_{\mathbb{L}_0^{q_b}(\bar{\mathbb{X}}_0')} \leqslant C_0 T^{-\gamma_b/\alpha}, \tag{3.63}$$

there is a unique weak solution u to FPE (3.60) on $[0, T]$ such that

$$\sup_{t \in (0,T]} \left(t^{-\frac{\beta_0}{\alpha}} \|u(t)\|_{\bar{\mathbb{Y}}_0} + t^{\frac{\mathscr{A}_0 - \beta_0 - \beta_b}{\alpha}} \|u(t)\|_{\bar{\mathbb{X}}_0} \right) < \infty.$$

Moreover, we have the following conclusions:

(i) **(Smoothness)** For any $\beta \geqslant 0$, it holds that

$$\sup_{t \in (0,T]} \left(t^{\frac{\beta-\beta_0}{\alpha}} \|u(t)\|_{\bar{\mathbb{Y}}_\beta} + t^{\frac{\beta+\mathscr{A}_0 - \beta_0 - \beta_b}{\alpha}} \|u(t)\|_{\bar{\mathbb{X}}_\beta} \right) < \infty.$$

(ii) **(Stability)** *For another $\widetilde{b} \in \mathbb{L}_0^{q_b}(\bar{\mathbb{X}}_0')$ and $\widetilde{u}_0 \in \bar{\mathbb{Z}}_{\beta_0}$ satisfying the same condition* (3.63), *let $\widetilde{u}$ be the unique solution on $[0,T]$ corresponding to $\widetilde{b}$ and $\widetilde{u}_0$. For any $\beta \geqslant 0$, there is a constant $C_\beta = C_\beta(T,\Theta,\|b\|_{\mathbb{L}_0^{q_b}(\bar{\mathbb{X}}_0')}, \|\widetilde{b}\|_{\mathbb{L}_0^{q_b}(\bar{\mathbb{X}}_0')}) > 0$ such that*

$$\sup_{t\in(0,T]} \left(t^{\frac{\beta-\beta_0}{\alpha}} \|u(t) - \widetilde{u}(t)\|_{\bar{\mathbb{Y}}_\beta} \right) \leqslant C_\beta \big(\|u_0 - \widetilde{u}_0\|_{\bar{\mathbb{Z}}_{\beta_0}} + \|b - \widetilde{b}\|_{\mathbb{L}_0^{q_b}(\bar{\mathbb{X}}_0')} \big).$$

(iii) **(Large time decay estimate)** *If $\frac{d}{\varrho_1} > \mathcal{Q} - 1$, $p_0 \neq 1$ and $u_0 \in \bar{\mathbb{Z}}_{\beta_0} \cap \mathbf{B}_1^{\beta_0}$, then we can take $\gamma_b = 0$ in* (3.63), *and the constant C_0 may depend on $\|u_0\|_{\mathbf{B}_1^{\beta_0}}$, and there is a global solution on $\mathbb{R}_+$ so that for any $\beta \geqslant 0$ and $i = 0, 1$,*

$$\sup_{t\geqslant 1} \left(t^{\frac{d-d/p_i}{\alpha}} \|u(t)\|_{\mathbf{B}_{p_i}^{\beta,1}} \right) < \infty.$$

(iv) **(Global solution)** *In addition, we suppose that one of the following four conditions holds:*
   (a) $q_b \in [1,\infty)$ *and* $u_0 \in \mathbf{B}_1^{\beta_0}$.
   (b) $q_b \in [1,\infty)$, $\mathrm{div} b = 0$ *and* $u_0 \in \mathbf{B}_{p_0}^{\beta_0} \cap \mathbf{B}_{p_1}^{\beta_0}$.
   (c) $q_b = \infty$ *and* $u_0 \in \cup_{q\in[1,\infty)} \mathbf{B}_1^{\beta_0,q}$.
   (d) $q_b = \infty$, $\mathrm{div} b = 0$ *and* $u_0 \in \cup_{q\in[1,\infty)}(\mathbf{B}_{p_0}^{\beta_0,q} \cap \mathbf{B}_{p_1}^{\beta_0,q})$.
   Then we can drop the smallness assumption (3.63) *and have a global solution.*

4. Well-posedness of mean-field kinetic SDEs

In this section we use previous results to show the weak and strong well-posedness for mean-field kinetic SDEs. Although the associated nonlinear Fokker-Planck equation has a smooth solution for any positive time, we have to carefully deal with the time singularity at initial time.

4.1. Well-posedness of linear kinetic SDEs.

In this subsection, we first study the weak/strong well-posedness to the following linear kinetic SDE: for $\alpha \in (1,2]$,

$$\begin{cases} X_t = X_0 + \displaystyle\int_0^t V_s \mathrm{d}s, \\ V_t = V_0 + \displaystyle\int_0^t H_s(X_s, V_s)\mathrm{d}s + L_t^{(\alpha)}, \end{cases} \tag{4.1}$$

where $H : \mathbb{R}_+ \times \mathbb{R}^{2d} \to \mathbb{R}^d$ is a measurable vector field, and $(L_t^{(\alpha)})_{t\geqslant 0}$ is the α-stable process as in the introduction. Moreover, if we let $Z_t := (X_t, V_t)$, $\mathbf{H}_t(z) = \mathbf{H}_t(x,v) = (v, H_t(x,v))^* \in \mathbb{R}^{2d}$ and $\sigma := (0, \mathbb{I})^* \in \mathbb{R}^{2d} \otimes \mathbb{R}^d$, then SDE (4.1) can be written as

$$Z_t = Z_0 + \int_0^t \mathbf{H}_s(Z_s)\mathrm{d}s + \sigma L_t^{(\alpha)}, \ Z_0 \sim \mu_0 \in \mathscr{P}(\mathbb{R}^{2d}). \tag{4.2}$$

For $q \in [1,\infty]$, $T > 0$ and a Banach space $\mathbb{B}$, we write

$$\mathbb{L}_T^q(\mathbb{B}) := L^q([0,T];\mathbb{B}), \ \mathbb{L}_T^\infty := L^\infty([0,T] \times \mathbb{R}^{2d}).$$

We have the following well-posedness result for SDE (4.2).

Theorem 4.1. *Let $\alpha \in (1,2]$, $q \in (\frac{\alpha}{\alpha-1}, \infty]$ and $T, \beta > 0$.*

(i) **(Weak well-posedness)** *Suppose $H \in \mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)$. For any initial distribution μ_0, there is a unique weak solution Z to kinetic SDE* (4.2).

(ii) **(Strong well-posedness)** *Suppose $H \in \mathbb{L}_T^q(\mathbf{C}_x^{\frac{\alpha+\beta}{1+\alpha}} \cap \mathbf{C}_{\boldsymbol{a}}^\beta)$, where $\beta > 1 - \frac{\alpha}{2}$. For any initial value Z_0, there is a unique strong solution Z to kinetic SDE* (4.2).

(iii) **(Strong stability)** Let $\widetilde{H} \in \mathbb{L}_T^q(\mathbf{C}_x^{\frac{\alpha+\beta}{1+\alpha}} \cap \mathbf{C}_{\boldsymbol{a}}^\beta)$ be another vector field with the same $\beta > 1 - \frac{\alpha}{2}$ and $\widetilde{Z}_0$ another initial random variable on the same probability space. Let $\widetilde{Z}$ be the strong solution of (4.2) corresponding to $\widetilde{H}$ and $\widetilde{Z}_0$. Then we have

$$\mathbb{E}\left(\sup_{t\in[0,T]}|Z_t - \widetilde{Z}_t|^2\right) \lesssim_C \mathbb{E}|Z_0 - \widetilde{Z}_0|^2 + \|H - \widetilde{H}\|_{\mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)}^2. \tag{4.3}$$

To show this result, we need to study the following linear kinetic equation with transport drift:

$$\partial_t u = \Delta_v^{\alpha/2} u + v \cdot \nabla_x u - \lambda u + H \cdot \nabla_v u + f, \quad u_0 = 0, \tag{4.4}$$

where $\lambda > 0$, $H : \mathbb{R}_+ \times \mathbb{R}^{2d} \to \mathbb{R}^d$ and $f : \mathbb{R}_+ \times \mathbb{R}^{2d} \to \mathbb{R}$ are two Borel measurable functions.

We first show the following simple result.

Theorem 4.2. Let $T > 0$, $\alpha \in (1, 2]$, $\beta, \gamma > 0$ and $q \in (\frac{\alpha}{\alpha-1}, \infty]$.

(i) For any $H, f \in \mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)$, there is a unique weak solution u to equation (4.4) with the regularity that for any $q' \in [q, \infty]$, $\beta' \in [0, \alpha + \frac{\alpha}{q'} - \frac{\alpha}{q})$ and all $\lambda \geqslant 1$,

$$\|u\|_{\mathbb{L}_T^{q'}(\mathbf{C}_{\boldsymbol{a}}^{\beta'+\beta})} \lesssim_C \lambda^{\frac{1}{q}+\frac{\beta'}{\alpha}-1-\frac{1}{q'}} \|f\|_{\mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)}, \tag{4.5}$$

where $C > 0$ only depends on $T, d, \beta, \alpha, q, q', \beta'$ and $\|H\|_{\mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)}$.

(ii) Let $H, f, \widetilde{H}, \widetilde{f} \in \mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)$ and $u, \widetilde{u}$ be the solutions of equation (4.4) corresponding to (H, f) and $(\widetilde{H}, \widetilde{f})$, respectively. For any $q' \in [q, \infty]$ and $\beta' \in [0, \alpha + \frac{\alpha}{q'} - \frac{\alpha}{q})$, there is a constant $C > 0$ depending on $T, d, \beta, \alpha, q, q', \beta'$ and $\|H\|_{\mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)}$, $\|\widetilde{H}\|_{\mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)}$ such that for all $\lambda \geqslant 1$,

$$\|u - \widetilde{u}\|_{\mathbb{L}_T^{q'}(\mathbf{C}_{\boldsymbol{a}}^{\beta'+\beta})} \lesssim_C \lambda^{\frac{1}{q}+\frac{\beta'}{\alpha}-1-\frac{1}{q'}} \left(\|H - \widetilde{H}\|_{\mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)} + \|f - \widetilde{f}\|_{\mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)}\right). \tag{4.6}$$

(iii) Let $H, f \in \mathbb{L}_T^q(\mathbf{C}_{x,\boldsymbol{a}}^{\gamma,\beta})$ and u be the solution of (4.4). For any $q' \in [q, \infty]$ and $\beta' \in [0, \alpha + \frac{\alpha}{q'} - \frac{\alpha}{q})$, there is a constant $C > 0$ depending on $T, d, \gamma, \beta, \alpha, q, q', \beta'$ and $\|H\|_{\mathbb{L}_T^q(\mathbf{C}_{x,\boldsymbol{a}}^{\gamma,\beta})}$ such that for all $\lambda \geqslant 1$,

$$\|u\|_{\mathbb{L}_T^{q'}(\mathbf{C}_{x,\boldsymbol{a}}^{\gamma,\beta'+\beta})} \lesssim_C \lambda^{\frac{1}{q}+\frac{\beta'}{\alpha}-1-\frac{1}{q'}} \|f\|_{\mathbb{L}_T^q(\mathbf{C}_{x,\boldsymbol{a}}^{\gamma,\beta})}, \tag{4.7}$$

Proof. (i) We only prove the a priori estimate (4.5). By Duhamel's formula, we have

$$u_t = \int_0^t P_{t-s}^\lambda(H_s \cdot \nabla_v u_s + f_s)\mathrm{d}s,$$

where

$$P_t^\lambda f(z) = \mathrm{e}^{-\lambda t} P_t f(z), \quad z \in \mathbb{R}^{2d}, \ t > 0.$$

By (2.33) with $\boldsymbol{p}_0 = \boldsymbol{p}' = \infty$, for any $\beta' \in [0, \alpha)$, we have

$$\|u_t\|_{\mathbf{C}_{\boldsymbol{a}}^{\beta'+\beta}} \lesssim \int_0^t \mathrm{e}^{-\lambda(t-s)}(t-s)^{-\frac{\beta'}{\alpha}} \left(\|H_s \cdot \nabla_v u_s\|_{\mathbf{C}_{\boldsymbol{a}}^\beta} + \|f_s\|_{\mathbf{C}_{\boldsymbol{a}}^\beta}\right)\mathrm{d}s$$

$$\lesssim \int_0^t \mathrm{e}^{-\lambda(t-s)}(t-s)^{-\frac{\beta'}{\alpha}} \left(\|H_s\|_{\mathbf{C}_{\boldsymbol{a}}^\beta}\|u_s\|_{\mathbf{C}_{\boldsymbol{a}}^{1+\beta}} + \|f_s\|_{\mathbf{C}_{\boldsymbol{a}}^\beta}\right)\mathrm{d}s. \tag{4.8}$$

In particular, taking $\beta' = 1$, by Hölder's inequality, we have for $r = \frac{q}{q-1}$,

$$\|u_t\|_{\mathbf{C}_{\boldsymbol{a}}^{1+\beta}}^r \lesssim \|H\|_{\mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)}^r \int_0^t \mathrm{e}^{-r\lambda(t-s)}(t-s)^{-\frac{r}{\alpha}} \|u_s\|_{\mathbf{C}_{\boldsymbol{a}}^{1+\beta}}^r \mathrm{d}s + \|f\|_{\mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)}^r \int_0^t \mathrm{e}^{-r\lambda s} s^{-\frac{r}{\alpha}}\mathrm{d}s.$$

Since $\frac{1}{\alpha} < \frac{1}{r} = 1 - \frac{1}{q}$, by Gronwall's inequality, we get

$$\|u\|_{\mathbb{L}_T^\infty(\mathbf{C}_{\boldsymbol{a}}^{1+\beta})} \lesssim \|f\|_{\mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)} \lambda^{\frac{1}{\alpha}-\frac{1}{r}}, \ \lambda \geqslant 1.$$

Let $1 + \frac{1}{q'} = \frac{1}{r} + \frac{1}{q}$. By (4.8) and Young's inequality we have for $\lambda \geqslant 1$,

$$\|u\|_{\mathbb{L}_T^{q'}(\mathbf{C}_{\boldsymbol{a}}^{\beta'+\beta})} \lesssim \left(\|u\|_{\mathbb{L}_T^\infty(\mathbf{C}_{\boldsymbol{a}}^{1+\beta})}\|H\|_{\mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)} + \|f\|_{\mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)} \right) \left(\int_0^T \mathrm{e}^{-r\lambda s} s^{-\frac{\beta' r}{\alpha}} \mathrm{d}s \right)^{1/r}$$

$$\lesssim \|f\|_{\mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)} \lambda^{\frac{\beta'}{\alpha}-\frac{1}{r}} \left(\int_0^\infty \mathrm{e}^{-rs} s^{-\frac{\beta' r}{\alpha}} \mathrm{d}s \right)^{1/r} \lesssim \|f\|_{\mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)} \lambda^{\frac{1}{q}+\frac{\beta'}{\alpha}-1-\frac{1}{q'}}.$$

Thus we get (4.5).

(ii) Let $U = u - \widetilde{u}$. Then U solves the following PDE:

$$\partial_t U = \Delta_v^{\alpha/2} U + v \cdot \nabla_x U - \lambda U + H \cdot \nabla_v U + G, \quad U_0 = 0,$$

where

$$G := f - \widetilde{f} + (H - \widetilde{H}) \cdot \nabla_v \widetilde{u}.$$

By (4.5), we have

$$\|U\|_{\mathbb{L}_T^{q'}(\mathbf{C}_{\boldsymbol{a}}^{\beta'+\beta})} \lesssim \lambda^{\frac{1}{q}+\frac{\beta'}{\alpha}-1-\frac{1}{q'}} \|G\|_{\mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)}. \tag{4.9}$$

Note that

$$\|G\|_{\mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)} \leqslant \|f - \widetilde{f}\|_{\mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)} + \|(H - \widetilde{H}) \cdot \nabla_v \widetilde{u}\|_{\mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)}$$

$$\leqslant \|f - \widetilde{f}\|_{\mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)} + \|H - \widetilde{H}\|_{\mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)} \|\widetilde{u}\|_{\mathbb{L}_T^\infty(\mathbf{C}_{\boldsymbol{a}}^{1+\beta})}.$$

Substituting it into (4.9) and by (4.5) with $q' = \infty$ and $\beta' = 1$, we obtain (4.6).

(iii) Noting that for any $j \geqslant 0$,

$$\|\mathscr{R}_j^x P_t f\|_{\mathbf{C}_{\boldsymbol{a}}^{\beta'+\beta}} \lesssim (1 \wedge t)^{-\frac{\beta'}{\alpha}} \|\mathscr{R}_j^x f\|_{\mathbf{C}_{\boldsymbol{a}}^\beta},$$

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

$$\|P_t f\|_{\mathbf{C}_{x,\boldsymbol{a}}^{\gamma,\beta'+\beta}} \lesssim (1 \wedge t)^{-\frac{\beta'}{\alpha}} \|f\|_{\mathbf{C}_{x,\boldsymbol{a}}^{\gamma,\beta}}.$$

By this estimate, the estimate (4.7) is completely the same as (4.5). $\square$

For $\alpha \in (1,2)$, let $N(\mathrm{d}w, \mathrm{d}t)$ be the Poisson random measure associated with α-stable process $L_t^{(\alpha)}$. More precisely, for any Borel set $E \in \mathscr{B}(\mathbb{R}^d)$ and $t > 0$,

$$N(E \times [0,t]) := \sum_{s \in (0,t]} \mathbb{1}_E(L_s^{(\alpha)} - L_{s-}^{(\alpha)}).$$

Let $\widetilde{N}(\mathrm{d}w, [0,t]) := N(\mathrm{d}w, [0,t]) - \nu^{(\alpha)}(\mathrm{d}w)t$ be the compensated Poisson random measure, where $\nu^{(\alpha)}(\mathrm{d}w) = c_{d,\alpha}\mathrm{d}w/|w|^{d+\alpha}$ is the Lévy measure of α-stable process $L_t^{(\alpha)}$, and $c_{d,\alpha}$ is a normalized constant only depending on d, α. In particular,

$$L_t^{(\alpha)} = \int_{\mathbb{R}^d} w \widetilde{N}(\mathrm{d}w, [0,t]).$$

The following lemma is crucial for weak uniqueness and pathwise uniqueness.

Lemma 4.3. *Let $T > 0$, $\alpha \in (1,2]$, $\beta > 0$ and $q \in (\frac{\alpha}{\alpha-1}, \infty]$. Suppose that $H, f \in \mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)$. Let u be the unique weak solution of the following backward kinetic PDE (see Theorem 4.2):*

$$\partial_t u + \Delta_v^{\alpha/2} u + v \cdot \nabla_x u + H \cdot \nabla_v u = f, \quad u(T) = 0, \tag{4.10}$$

and Z be a solution of kinetic SDE (4.2). If $\alpha \in (1,2)$, then for any $t \in [0,T]$,

$$u(t, Z_t) = u(0, Z_0) + \int_0^t f(s, Z_s)\mathrm{d}s + \int_0^t \int_{\mathbb{R}^d} (u(s, Z_{s-} + \sigma w) - u(s, Z_{s-})) \widetilde{N}(\mathrm{d}w, \mathrm{d}s); \tag{4.11}$$

if $\alpha = 2$, *then for any* $t \in [0, T]$,

$$u(t, Z_t) = u(0, Z_0) + \int_0^t f(s, Z_s)\mathrm{d}s + \int_0^t \nabla_v u(s, Z_s)\mathrm{d}W_s. \tag{4.12}$$

Proof. We only prove (4.11) since (4.12) is similar and easier. Let $\Gamma \in C_c^\infty(\mathbb{R}^{2d})$ be a smooth density function with support in $B_1^{\boldsymbol{a}}$. For each $n \in \mathbb{N}$, define

$$\Gamma_n(z) = \Gamma_n(x, v) := n^{(2+\alpha)d}\Gamma(nx, n^{1+\alpha}v) \tag{4.13}$$

and

$$u^n(t, z) := (u(t, \cdot) * \Gamma_n)(z).$$

Taking convolutions for both sides of (4.10) with Γ_n, we get

$$\partial_t u_n + (\Delta_v^{\alpha/2} + v \cdot \nabla_x)u_n + H \cdot \nabla_v u_n = f_n + g_n + h_n,$$

where

$$g_n = v \cdot \nabla_x u_n - (v \cdot \nabla_x u) * \Gamma_n, \quad h_n := H \cdot \nabla_v u_n - (H \cdot \nabla_v u) * \Gamma_n.$$

By Itô's formula, we have

$$u_n(t, Z_t) = u_n(0, Z_0) + \int_0^t (f_n + h_n + g_n)(s, Z_s)\mathrm{d}s$$

$$+ \int_0^t \int_{\mathbb{R}^d} (u_n(s, Z_{s-} + \sigma w) - u_n(s, Z_{s-})) \, \widetilde{N}(\mathrm{d}w, \mathrm{d}s).$$

Noting that by (4.5) with $q' = \infty$ and $\beta' \in (0, \alpha - \frac{\alpha}{q})$,

$$\|u_n - u\|_{\mathbb{L}_T^\infty} \leqslant \sup_{t \in [0,T]} \sup_{x,v} \int_{\mathbb{R}^{2d}} |u(t, x - y, v - w) - u(t, x, v)|\Gamma_n(y, w)\mathrm{d}y\mathrm{d}w$$

$$\leqslant \|u\|_{\mathbb{L}_T^\infty(\mathbf{C}_{\boldsymbol{a}}^{\beta'})} \int_{\mathbb{R}^{2d}} (|y|^{\frac{1}{1+\alpha}} + |w|)\Gamma_n(y, w)\mathrm{d}y\mathrm{d}w \lesssim n^{-\frac{\beta'}{1+\alpha}}, \tag{4.14}$$

we have

$$(u_n(t, Z_t), u_n(0, Z_0)) \to (u(t, Z_t), u(0, Z_0)) \quad a.s., \ n \to \infty.$$

Let

$$A_n := \int_0^t \int_{\mathbb{R}^d} (u_n(s, Z_{s-} + \sigma w) - u_n(s, Z_{s-})) \, \widetilde{N}(\mathrm{d}w, \mathrm{d}s)$$

and

$$A := \int_0^t \int_{\mathbb{R}^d} (u(s, Z_{s-} + \sigma w) - u(s, Z_{s-})) \, \widetilde{N}(\mathrm{d}w, \mathrm{d}s).$$

Noting that

$$|(u_n(s, z + \sigma w) - u_n(s, z)) - (u(s, z + \sigma w) - u(s, z))|$$

$$\leqslant \|u_n(s) - u(s)\|_\infty \wedge (\|\nabla_v u(s)\|_\infty |\sigma w|),$$

by the isometry formula of stochastic integrals, we have

$$\mathbb{E}|A_n - A|^2 \lesssim \int_0^t \int_{\mathbb{R}^d} \left(\|u_n(s) - u(s)\|_\infty^2 \wedge \left(\|\nabla_v u(s)\|_\infty^2 |\sigma w|^2 \right) \right) \nu^{(\alpha)}(\mathrm{d}w)\mathrm{d}s.$$

Recalling $\nu^{(\alpha)}(\mathrm{d}w) = c_{d,\alpha}|w|^{-d-\alpha}\mathrm{d}w$, by (4.14) and the dominated convergence theorem, we get

$$\lim_{n \to \infty} \mathbb{E}|A_n - A|^2 = 0.$$

Finally, noting that

$$g_n(t, z) = \int_{\mathbb{R}^{2d}} u(t, x - y, v - w)w \cdot \nabla_y \Gamma_n(y, w)\mathrm{d}y\mathrm{d}w,$$

since $\Gamma_n(y,w) = n^{(2+\alpha)d}\Gamma(ny, n^{1+\alpha}w)$, we have

$$\|g_n(t)\|_\infty \lesssim n^{-\alpha}\|u(t)\|_\infty$$

and

$$\|h_n(t)\|_\infty \lesssim n^{-\beta/(1+\alpha)}\|H_t\|_{\mathbf{C}_{\boldsymbol{a}}^\beta}\|\nabla_v u(t)\|_\infty.$$

Hence, by (4.5) with $(q', \beta') = (\infty, 1)$ and the dominated convergence theorem,

$$\int_0^t (|f_n - f| + |h_n| + |g_n|)(s, Z_s)\mathrm{d}s \to 0, \quad a.s., \ n \to \infty.$$

Thus we complete the proof of (4.11). $\qquad\square$

Remark 4.4. The stochastic integrals in (4.11) and (4.12) are martingales by the regularity estimates of u in (4.5).

Now, we can give the

Proof of Theorem 4.1: (**Existence of weak solutions**) Since $H \in \mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)$ is continuous in the spatial variable, the existence of a weak solution is standard by weak convergence method. We referred to [35] and [28].

(**Weak uniqueness**) Let Z^1 and Z^2 be two weak solutions of SDE (4.2) with the same initial distribution. For any $f \in \mathbf{C}_{\boldsymbol{a}}^\beta$, by Lemma 4.3 with $T = t$, we have

$$\mathbb{E}\int_0^t f(Z_s^1)\mathrm{d}s = -\mathbb{E}u(0, Z_0^1) = -\mathbb{E}u(0, Z_0^2) = \mathbb{E}\int_0^t f(Z_s^2)\mathrm{d}s.$$

Hence,

$$\mathbb{E}f(Z_t^1) = \mathbb{E}f(Z_t^2).$$

In particular, Z^1 and Z^2 have the same one dimensional marginal distribution. Thus, as in [35, Theorem 4.4.3], by a standard method, we obtain the uniqueness.

(**Strong uniqueness and stability**) We use the well-known Zvonkin transformation to prove the strong stability, which automatically imply the pathwise uniqueness. The strong uniqueness then follows by Yamada-Watanabe's theorem. Below we only prove it for $\alpha \in (1, 2)$. Recall that

$$H, \widetilde{H} \in \mathbb{L}_T^q(\mathbf{C}_x^{\frac{\gamma}{1+\alpha}} \cap \mathbf{C}_{\boldsymbol{a}}^\beta), \ \ \beta > 1 - \tfrac{\alpha}{2}, \ \ \gamma = \alpha + \beta.$$

By the interpolation inequality (2.20), we have for any $\theta \in [0, 1]$,

$$H, \widetilde{H} \in \mathbb{L}_T^q(\mathbf{C}_{x,\boldsymbol{a}}^{\theta\gamma, (1-\theta)\beta}).$$

Fix $T > 0$ and $\lambda \geqslant 1$. Let $\mathbf{u} := (u_1, \cdots, u_d)$ solve the following backward PDE:

$$\partial_t \mathbf{u} + (\Delta_v^{\alpha/2} + v \cdot \nabla_x - \lambda)\mathbf{u} + H \cdot \nabla_v \mathbf{u} + H = 0, \quad \mathbf{u}(T) = 0.$$

By (4.7), we have for any $\beta' \in [0, \alpha + \tfrac{\alpha}{q'} - \tfrac{\alpha}{q})$ and $\theta \in [0, 1]$,

$$\|\mathbf{u}\|_{\mathbb{L}_T^{q'}(\mathbf{C}_{x,\boldsymbol{a}}^{\theta\gamma,\beta'+(1-\theta)\beta})} \lesssim_C \lambda^{\frac{1}{q}+\frac{\beta'}{\alpha}-1-\frac{1}{q'}}\|H\|_{\mathbb{L}_T^q(\mathbf{C}_{x,\boldsymbol{a}}^{\theta\gamma,(1-\theta)\beta})}. \tag{4.15}$$

Similarly, we define $\widetilde{\mathbf{u}}$ by $\widetilde{H}$. Then (4.15) also holds for $\widetilde{H}$ and $\widetilde{u}$. Moreover, by (4.6) we have

$$\|\mathbf{u} - \widetilde{\mathbf{u}}\|_{\mathbb{L}_T^{q'}(\mathbf{C}_{\boldsymbol{a}}^{\beta'+\beta})} \lesssim_C \lambda^{\frac{1}{q}+\frac{\beta'}{\alpha}-1-\frac{1}{q'}}\|H - \widetilde{H}\|_{\mathbb{L}_T^q(\mathbf{C}_{\boldsymbol{a}}^\beta)}. \tag{4.16}$$

In particular, letting $q' = \infty$, $\theta = 1$ in (4.15) and $\beta' \in (1, \alpha - \tfrac{\alpha}{q})$, by (2.18), we obtain

$$\|\nabla\mathbf{u}\|_{\mathbb{L}_T^\infty} \vee \|\nabla\widetilde{\mathbf{u}}\|_{\mathbb{L}_T^\infty} \lesssim \|\mathbf{u}\|_{\mathbb{L}_T^\infty(\mathbf{C}_{x,\boldsymbol{a}}^{\gamma,\beta'})} \vee \|\widetilde{\mathbf{u}}\|_{\mathbb{L}_T^\infty(\mathbf{C}_{x,\boldsymbol{a}}^{\gamma,\beta'})} \lesssim \lambda^{\frac{1}{q}+\frac{\beta'}{\alpha}-1},$$

where the implicit constant is independent of $\lambda \geqslant 1$. Since $\frac{1}{q} + \frac{\beta'}{\alpha} < 1$, one can choose λ large enough so that

$$\|\nabla\mathbf{u}\|_{\mathbb{L}_T^\infty} \vee \|\nabla\widetilde{\mathbf{u}}\|_{\mathbb{L}_T^\infty} \leqslant \tfrac{1}{2}.$$

Now if we define
$$\Phi(t,z) = \Phi(t,x,v) := (x, v + \mathbf{u}(t,x,v)),$$
then for $z = (x,v)$ and $z' = (x', v')$,
$$|z - z'| \leqslant |\Phi(t,z) - \Phi(t,z')| + |\mathbf{u}(t,z) - \mathbf{u}(t,z')| \leqslant |\Phi(t,z) - \Phi(t,z')| + \tfrac{1}{2}|z - z'|,$$
which implies
$$|z - z'| \leqslant 2|\Phi(t,z) - \Phi(t,z')|. \tag{4.17}$$

Similarly, we define $\widetilde{\Phi}$ in terms of $\widetilde{\mathbf{u}}$. On the other hand, since $\beta > 1 - \frac{\alpha}{2}$, by (2.18) and (4.15) with $q' = q$ and $\theta = 0, 1$, it is easy to see that for some $\varepsilon_0 > 0$,
$$\|\nabla \mathbf{u}\|_{\mathbb{L}_T^q(\mathbf{C}_v^{\alpha/2+\varepsilon_0})} \vee \|\nabla \widetilde{\mathbf{u}}\|_{\mathbb{L}_T^q(\mathbf{C}_v^{\alpha/2+\varepsilon_0})} < \infty. \tag{4.18}$$

By (4.11) with $f = -H$, one finds that
$$\Phi(t, Z_t) = \Phi(0, Z_0) + \int_0^t (V_s, 0) \mathrm{d}s + \int_0^t \int_{\mathbb{R}^d} g(s, Z_{s-}, w) \widetilde{N}(\mathrm{d}w, \mathrm{d}s), \tag{4.19}$$
where
$$g(s, z, w) := \Phi(s, z + \sigma w) - \Phi(s, z) = (0, w + \mathbf{u}(s, z + \sigma w) - \mathbf{u}(s, z)).$$
Similarly, we define
$$\widetilde{g}(s, z, w) := \widetilde{\Phi}(s, z + \sigma w) - \widetilde{\Phi}(s, z) = (0, w + \widetilde{\mathbf{u}}(s, z + \sigma w) - \widetilde{\mathbf{u}}(s, z)).$$
Let $\widetilde{Z}$ be any solution of SDE (4.12) corresponding to $\widetilde{H}$ and initial value $\widetilde{Z}_0$. Define
$$U_t := Z_t - \widetilde{Z}_t, \quad \Lambda_t := \Phi(t, Z_t) - \widetilde{\Phi}(t, \widetilde{Z}_t), \quad G_s(w) := g(s, Z_{s-}, w) - \widetilde{g}(s, \widetilde{Z}_{s-}, w).$$
By (4.19), we have
$$\Lambda_t = \Lambda_0 + \int_0^t (V_s - \widetilde{V}_s, 0) \mathrm{d}s + \int_0^t \int_{\mathbb{R}^d} G_s(w) \widetilde{N}(\mathrm{d}w, \mathrm{d}s).$$
By BDG's inequality, we have
$$\mathbb{E}\left(\sup_{s \in [0,t]} |\Lambda_s|^2\right) \lesssim \mathbb{E}|\Lambda_0|^2 + \int_0^t \mathbb{E}|V_s - \widetilde{V}_s|^2 \mathrm{d}s + \mathbb{E}\left(\sup_{s' \in [0,t]} \left|\int_0^{s'} \int_{\mathbb{R}^d} G_s(w) \widetilde{N}(\mathrm{d}w, \mathrm{d}s)\right|^2\right)$$
$$\lesssim \mathbb{E}|\Lambda_0|^2 + \int_0^T \mathbb{E}|U_s|^2 \mathrm{d}s + \mathbb{E}\left(\int_0^t \int_{\mathbb{R}^d} |G_s(w)|^2 \nu^{(\alpha)}(\mathrm{d}w) \mathrm{d}s\right). \tag{4.20}$$

Noting that
$$g(s, z, w) - g(s, \widetilde{z}, w) = (z - \widetilde{z}) \cdot \int_0^1 \nabla_z g(s, \theta z + (1 - \theta)\widetilde{z}, w) \mathrm{d}\theta,$$
and
$$\nabla_z g(s, z, w) = \begin{pmatrix} 0, & \nabla_x \mathbf{u}(s, z + \sigma w) - \nabla_x \mathbf{u}(s, z) \\ 0, & \nabla_v \mathbf{u}(s, z + \sigma w) - \nabla_v \mathbf{u}(s, z) \end{pmatrix},$$
by (4.18), we have for $\delta = \frac{\alpha}{2} + \varepsilon_0$,
$$|g(s, z, w) - g(s, \widetilde{z}, w)| \lesssim |z - \widetilde{z}| \left(\left(\|\nabla \mathbf{u}(s)\|_{\mathbf{C}_v^\delta} |w|^\delta\right) \wedge \|\nabla \mathbf{u}(s)\|_\infty \right).$$
Similarly, we have
$$|g(s, z, w) - \widetilde{g}(s, z, w)| \lesssim \left(\|\mathbf{u}(s) - \widetilde{\mathbf{u}}(s)\|_{\mathbf{C}_v^\delta} |w|^\delta\right) \wedge \|\mathbf{u}(s) - \widetilde{\mathbf{u}}(s)\|_\infty.$$
Hence,
$$|G_s(w)| \lesssim |U_s| \left(\left(\|\nabla \mathbf{u}(s)\|_{\mathbf{C}_v^\delta} |w|^\delta\right) \wedge \|\nabla \mathbf{u}(s)\|_\infty \right)$$
$$+ \left(\|\mathbf{u}(s) - \widetilde{\mathbf{u}}(s)\|_{\mathbf{C}_v^\delta} |w|^\delta\right) \wedge \|\mathbf{u}(s) - \widetilde{\mathbf{u}}(s)\|_\infty.$$

Recalling $\nu^{(\alpha)}(\mathrm{d}w) = c_{d,\alpha}|w|^{-d-\alpha}\mathrm{d}w$, for $\delta > \frac{\alpha}{2}$ and any $a, b > 0$, we have

$$\int_{\mathbb{R}^d} (a|w|^\delta \wedge b)^2 \nu^{(\alpha)}(\mathrm{d}w) \leqslant C(a^2 + b^2).$$

Moreover, by (4.17) we also have

$$|U_s| \leqslant 2|\Lambda_s| + 2\|\mathbf{u}(s) - \widetilde{\mathbf{u}}(s)\|_\infty.$$

Combining the above calculations and by (4.20), we get for all $t \in [0, T]$,

$$\mathbb{E}\left(\sup_{s\in[0,t]}|U_s|^2\right) \lesssim \mathbb{E}|U_0|^2 + \|\mathbf{u} - \widetilde{\mathbf{u}}\|_{\mathbb{L}_T^\infty}^2 + \int_0^T \|\mathbf{u}(s) - \widetilde{\mathbf{u}}(s)\|_{\mathbf{C}_v^\delta}^2 \mathrm{d}s$$

$$+ \int_0^t \left(1 + \|\nabla\mathbf{u}(s)\|_{\mathbf{C}_v^\delta}^2\right) \mathbb{E}|U_s|^2 \mathrm{d}s.$$

By Gronwall's inequality and (4.18), we obtain

$$\mathbb{E}\left(\sup_{s\in[0,T]}|U_s|^2\right) \lesssim \mathbb{E}|U_0|^2 + \|\mathbf{u} - \widetilde{\mathbf{u}}\|_{\mathbb{L}_T^\infty}^2 + \|\mathbf{u} - \widetilde{\mathbf{u}}\|_{\mathbb{L}_T^2(\mathbf{C}_v^\delta)}^2,$$

which in turn implies (4.3) by (4.16). $\qquad\qquad\square$

4.2. **Proofs of Theorems 1.3 and 1.9.** Suppose that (1.22) and (1.23) hold. Let

$$\gamma_b > (\mathcal{Q} - 1 - \mathscr{A}_1) \vee 0 = \mathcal{Q} - 1 - (\mathcal{Q} - 1) \wedge \mathscr{A}_1. \qquad (4.21)$$

Recall the parameter set Θ defined in (3.7). Under (1.22) and (1.23), by Theorem 3.11, there is a constant $C_0 = C_0(\Theta) > 0$ such that for any fixed $T \geqslant 1$ and $\mu_0 \in \mathbb{Z}_{\beta_0} \cap \mathscr{P}(\mathbb{R}^{2d})$ with

$$\|\mu_0\|_{\mathbb{Z}_{\beta_0}} \leqslant C_0 T^{-\gamma_b/\alpha}/\kappa_b, \qquad (4.22)$$

where κ_b is from (1.23), there is a unique smooth solution u_T to FPE (1.5) with $b_T = b\mathbb{1}_{[0,T]}$ and initial value $u_T(0) = \mu_0$ so that for any $\beta \geqslant 0$,

$$\|u_T\|_{\mathbb{S}_\beta} \leqslant C_\beta,$$

where $\mathbb{S}_\beta$ is defined by (3.21). In the following we fix time $T \geqslant 1$ and drop the subscript T in u_T.

Recalling the notations in Section 3, we first show the following L^1-estimate.

Lemma 4.5. *For any $t \in [0, T]$, it holds that*

$$\|u(t)\|_{\mathbb{L}^1} \leqslant 1. \qquad (4.23)$$

Proof. We use the stability in Theorem 3.11 to show (4.23). Let $b^n := b * \Gamma_n$ and $u_0^n := \mu_0 * \Gamma_n$ be the modifying approximation of b and μ_0, where Γ_n is defined in (4.13). By definitions, it is easy to see that

$$b^n \in \mathbb{L}_{\gamma_b}^{q_b}(C_b^\infty(\mathbb{R}^{2d})), \quad u_0^n \in C_b^\infty(\mathbb{R}^{2d}) \cap L^1(\mathbb{R}^{2d}),$$

and

$$\sup_{n\in\mathbb{N}} \|b^n\|_{\mathbb{L}_{\gamma_b}^{q_b}(\mathbb{X}_0')} \leqslant \|b\|_{\mathbb{L}_{\gamma_b}^{q_b}(\mathbb{X}_0')}, \quad \sup_{n\in\mathbb{N}} \|u_0^n\|_{\mathbb{Z}_{\beta_0}} \leqslant \|u_0\|_{\mathbb{Z}_{\beta_0}},$$

and for any $\varepsilon > 0$,

$$\|b^n - b\|_{\mathbb{L}_{\gamma_b}^{q_b}(\mathbb{X}_\varepsilon')} + \|u_0^n - u_0\|_{\mathbb{Z}_{\beta_0-\varepsilon}} \to 0 \quad \text{as } n \to \infty. \qquad (4.24)$$

Let $u^n \in C([0, T]; C_b^2(\mathbb{R}^{2d}))$ be the smooth solution to the following kinetic Fokker-Planck equation

$$\partial_t u^n = (\Delta_v^{\frac{\alpha}{2}} - v \cdot \nabla_x)u^n + \mathrm{div}_v((b^n * u^n)u^n), \quad u^n(0) = u_0^n.$$

By Lemma C.1 in appendix, we have

$$\|u^n(t)\|_{\mathbb{L}^1} \leqslant \|u^n(0)\|_{\mathbb{L}^1} = \|\mu_0 * \Gamma_n\|_{\mathbb{L}^1} = 1. \qquad (4.25)$$

By (1.22), one can fix $\varepsilon > 0$ being small enough so that

$$\mathscr{A}_0 - (\beta_0 - \varepsilon) - (\beta_b - \varepsilon) < \mathscr{Q} + (\beta_0 - \varepsilon) \wedge (-1).$$

Using $(\beta_0 - \varepsilon, \beta_b - \varepsilon)$ to replace (β_0, β_b) in Theorem 3.11 and slightly adjusting the smallness constant in (4.22), by (4.24) and the stability estimate (3.32), for each $t \in (0, T]$, one has

$$\lim_{n \to \infty} \|u^n(t) - u(t)\|_{\boldsymbol{p}_0} = 0.$$

Thus by (4.25) and Fatou's lemma, we conclude (4.23). $\qquad\square$

The following lemma improves the regularity estimate in (3.31).

Lemma 4.6. *For any $\beta > \beta_0$, there is a constant $C = C(\Theta, \beta, T, \kappa_b, \|\mu_0\|_{\mathbb{Z}_{\beta_0}}) > 0$ such that*

$$\|u(t)\|_{\mathbf{B}^{\beta,1}_{\boldsymbol{p}_0;\boldsymbol{a}}} + \|u(t)\|_{\mathbf{B}^{\beta,1}_{\boldsymbol{p}_1;\boldsymbol{a}}} \leqslant Ct^{-\frac{\beta-\beta_0}{\alpha}}, \ t \in (0, T]. \tag{4.26}$$

Proof. For $\theta = \mathscr{A}_0 - 2\beta_0 - \beta_b$, by (1.22) and (4.21), we have

$$\theta \in (0, \mathscr{Q}), \quad \beta_0 \leqslant \mathscr{Q} - 1 - \theta, \quad 0 \gtrless \boldsymbol{\gamma}_0 + \boldsymbol{\ell}_0 + \boldsymbol{\gamma}_b - (\theta, \mathscr{Q} - 1).$$

Thus for $i = 0, 1$, by (3.14) with $(\boldsymbol{p}', \boldsymbol{p}) = (\boldsymbol{p}_i, \boldsymbol{p}_i)$,

$$\|\mathscr{H}^b(u, u)\|_{\mathbb{L}^\infty_0(\mathbf{B}^{\beta_0,1}_{\boldsymbol{p}_i;\boldsymbol{a}})} \overset{(2.31)}{\lesssim} \|\mathscr{H}^b(u, u)\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}'}(\mathbf{B}^{\mathscr{Q}-1-\theta,1}_{\boldsymbol{p}_i;\boldsymbol{a}})} \lesssim \|b_T\|_{\mathbb{L}^{q_b}_{\boldsymbol{\gamma}_b}(\mathbb{X}'_0)} \|u\|_{\mathbb{L}^\infty_{\boldsymbol{\ell}_0}(\mathbb{X}_0)} \|u\|_{\mathbb{L}^\infty_{\boldsymbol{\gamma}_0}(\mathbb{L}^{\boldsymbol{p}_i})}$$

$$\overset{(2.12)}{\lesssim} \|b\|_{\mathbb{L}^{q_b}_0(\mathbb{X}'_0)} T^{\gamma_b/\alpha} \|u\|^2_{\mathbb{S}_0} \overset{(3.30)}{\lesssim} \|b\|_{\mathbb{L}^{q_b}_0(\mathbb{X}'_0)} T^{\gamma_b/\alpha} \|\mu_0\|^2_{\mathbb{Z}_{\beta_0}}.$$

On the other hand, by Lemma 2.16 we have

$$\|P.u_0\|_{\mathbb{L}^\infty_0(\mathbf{B}^{\beta_0}_{\boldsymbol{p}_i;\boldsymbol{a}})} \lesssim \|\mu_0\|_{\mathbf{B}^{\beta_0}_{\boldsymbol{p}_i;\boldsymbol{a}}}, \ i = 0, 1.$$

Combining the above two estimates, we obtain

$$\|u(t)\|_{\mathbf{B}^{\beta_0}_{\boldsymbol{p}_0;\boldsymbol{a}}} + \|u(t)\|_{\mathbf{B}^{\beta_0}_{\boldsymbol{p}_1;\boldsymbol{a}}} \leqslant C, \ t \in (0, T],$$

For general case $\beta > \beta_0$, it follows by (3.30) and the real interpolation theorem (see [8, Theorem 6.4.5, (1)]). $\qquad\square$

Define

$$H(s, z) := (b(s) * u(s))(z). \tag{4.27}$$

To use Theorem 4.1, the following lemma provides necessary estimates of H for linear SDE.

Lemma 4.7. *(i) Under (1.22) and (1.23), there are $q > \frac{\alpha}{\alpha-1}$ and $\beta > 0$ such that*

$$H \in \mathbb{L}^q_T(\mathbf{C}^\beta_{\boldsymbol{a}}).$$

(ii) Under (1.25) and (1.26), there are $q > \frac{\alpha}{\alpha-1}$ and $\beta > 1 - \frac{\alpha}{2}$ such that

$$H \in \mathbb{L}^q_T(\mathbf{C}^\gamma_x \cap \mathbf{C}^\beta_{\boldsymbol{a}}), \quad \gamma := \frac{\beta+\alpha}{1+\alpha}.$$

(iii) Let $\widetilde{b}$ be another kernel satisfying (1.23) and (1.26), and $\widetilde{\mu}_0 \in \mathbb{Z}_{\beta_0}$ satisfy (4.22). Let $\widetilde{u}$ be the unique smooth solution of FPE (1.5) corresponding to $\widetilde{b}$ and $\widetilde{u}_0$. Define $\widetilde{H}_s(z) := (\widetilde{b}(s) * \widetilde{u}(s))(z)$. Then there are $q > \frac{\alpha}{\alpha-1}$, $\beta > 0$ and constant $C = C(T, \Theta, q, \beta) > 0$ such that

$$\|H - \widetilde{H}\|_{\mathbb{L}^q_T(\mathbf{C}^\beta_{\boldsymbol{a}})} \lesssim_C \|\mu_0 - \widetilde{\mu}_0\|_{\mathbb{Z}_{\beta_0}} + \|b - \widetilde{b}\|_{\mathbb{L}^{q_b}_0(\mathbb{X}'_0)}.$$

Proof. For any $\beta \geqslant 0$, since $\mathscr{A}_1 \leqslant \mathscr{A}_0$, we have for any $t \in (0, T]$,

$$\|u(t)\|_{\mathbb{X}_\beta} = \sum_{i=0,1} \|u(t)\|_{\mathbf{B}^{\beta-\beta_b,1}_{\bar{\varrho}_i;a}} \overset{(2.16)}{\lesssim} \sum_{i=0,1} \|u(t)\|_{\mathbf{B}^{\beta-\beta_b+\mathscr{A}_i,1}_{\boldsymbol{P}_i;a}} \overset{(4.26)}{\lesssim} t^{-\frac{\beta+\mathscr{A}_0-\beta_0-\beta_b}{\alpha}}. \tag{4.28}$$

(i) Recall $b = b_0 + b_1$. For $\beta \geqslant 0$, by (2.13) we have

$$\|b(t) * u(t)\|_{\mathbf{C}^\beta_a} = \|b(t) * u(t)\|_{\mathbf{B}^\beta_{\infty;a}} \lesssim \left(\|b_0(t)\|_{\mathbf{B}^{\beta_b}_{\varrho_0;a}} + \|b_1(t)\|_{\mathbf{B}^{\beta_b}_{\varrho_1;a}} \right) \|u(t)\|_{\mathbb{X}_\beta} = \|b(t)\|_{\mathbb{X}'_0} \|u(t)\|_{\mathbb{X}_\beta}.$$

Since $0 < \mathscr{A}_0 - \beta_0 - \beta_b < \alpha - \frac{\alpha}{q_b} - 1$ and $q_b > \frac{\alpha}{\alpha-1}$, one can choose $q \in (\frac{\alpha}{\alpha-1}, q_b)$ and $\beta > 0$ so that

$$0 < \beta + \mathscr{A}_0 - \beta_0 - \beta_b < \frac{\alpha}{q} - \frac{\alpha}{q_b} =: \frac{\alpha}{r}.$$

Thus by Hölder's inequality and (4.28), we have

$$\|H\|_{\mathbb{L}^q_T(\mathbf{C}^\beta_a)} = \|b * u\|_{\mathbb{L}^q_T(\mathbf{C}^\beta_a)} \lesssim \|b\|_{\mathbb{L}^{q_b}_0(\mathbb{X}'_0)} \left(\int_0^T t^{-r(\beta+\beta_0-\beta_b+\mathscr{A}_0)/\alpha} \mathrm{d}t \right)^{1/r} < \infty.$$

(ii) Since $0 < \mathscr{A}_0 - \beta_0 - \beta_b < \frac{3}{2}\alpha - \frac{\alpha}{q_b} - 2$, one can choose $q > \frac{\alpha}{\alpha-1}$ and $\beta > 1 - \frac{\alpha}{2}$ so that

$$0 < \beta + \mathscr{A}_0 - \beta_0 - \beta_b < \frac{\alpha}{q} - \frac{\alpha}{q_b} =: \frac{\alpha}{r}.$$

Let $\delta = \frac{\alpha}{2} - 1$. Fix $i = 0, 1$. By (2.13), we have

$$\|b_i(t) * u(t)\|_{\mathbf{B}^\beta_{\infty;a}} \lesssim \|b_i(t)\|_{\mathbf{B}^{\beta_b}_{\varrho_i;a}} \|u(t)\|_{\mathbf{B}^{\beta-\beta_b,1}_{\bar{\varrho}_i;a}},$$

and for all $j \geqslant 0$,

$$\|\mathscr{R}^x_j b_i(t) * u(t)\|_{\mathbf{B}^{\delta+\beta}_{\infty;a}} \lesssim \|\mathscr{R}^x_j b_i(t)\|_{\mathbf{B}^{\delta+\beta_b}_{\varrho_i;a}} \|u(t)\|_{\mathbf{B}^{\beta-\beta_b,1}_{\bar{\varrho}_i;a}}.$$

In particular, by Definition 2.2 and (4.28), for $i = 0, 1$,

$$\|b_i(t) * u(t)\|_{\mathbf{B}^{\alpha-\delta,\delta+\beta}_{\infty;x,a} \cap \mathbf{B}^\beta_{\infty;a}} \lesssim \|b_i(t)\|_{\mathbf{B}^{\alpha-\delta,\delta+\beta_b}_{\varrho_i;x,a} \cap \mathbf{B}^{\beta_b}_{\varrho_i;a}} t^{-\frac{\beta+\mathscr{A}_0-\beta_0-\beta_b}{\alpha}}.$$

By Hölder's inequality, we get

$$\|b_i * u\|_{\mathbb{L}^q_T(\mathbf{B}^{\alpha-\delta,\delta+\beta}_{\infty;x,a} \cap \mathbf{B}^\beta_{\infty;a})} \lesssim \|b_i\|_{\mathbb{L}^{q_b}_T(\mathbf{B}^{\alpha-\delta,\delta+\beta_b}_{\varrho_i;x,a} \cap \mathbf{B}^{\beta_b}_{\varrho_i;a})} \left(\int_0^T t^{-\frac{r(\beta+\mathscr{A}_0-\beta_0-\beta_b)\vee 0}{\alpha}} \mathrm{d}t \right)^{1/r} < \infty.$$

Since $\delta + \beta > 0$, by (2.19), we have for $\gamma = \frac{\alpha+\beta}{1+\alpha}$,

$$\|b_i * u\|_{\mathbb{L}^q_T(\mathbf{C}^\gamma_x \cap \mathbf{C}^\beta_a)} = \|b_i * u\|_{\mathbb{L}^q_T(\mathbf{B}^\gamma_{\infty;x} \cap \mathbf{B}^\beta_{\infty;a})} \lesssim \|b_i * u\|_{\mathbb{L}^q_T(\mathbf{B}^{\alpha-\delta,\delta+\beta}_{\infty;x,a} \cap \mathbf{B}^\beta_{\infty;a})} < \infty.$$

(iii) It follows by the calculations in (i) and (3.32). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

Remark 4.8. In the situation of Remark 3.6, by Remark 3.13, the above (i) still holds.

Now we are in a position to give

Proof of Theorem 1.3. Let $\alpha \in (1, 2]$. Consider the following linear SDE:

$$\mathrm{d}Z_t = \mathbf{H}_t(Z_t)\mathrm{d}t + \sigma \mathrm{d}L_t^{(\alpha)}, \quad Z_0 \sim \mu_0,$$

where $\mathbf{H}_t(x, v) = (v, H_t(x, v))$ is defined in (4.27). By Lemma 4.7 and Theorem 4.1, there is a unique weak/strong solution to the above SDE under (1.22)-(1.23) and (1.25)-(1.26), respectively. It remains to show that

$$\mu_t(\mathrm{d}z) := \text{Law of } Z_t = u(t, z)\mathrm{d}z. \tag{4.29}$$

Let

$$w(t, \mathrm{d}z) := \mu_t(\mathrm{d}z) - u(t, z)\mathrm{d}z \in \mathcal{M}(\mathbb{R}^{2d}).$$

Clearly, by (4.23) one has

$$\|w(t)\|_m = \sup_{\|f\|_\infty \leqslant 1} \left| \int_{\mathbb{R}^{2d}} f(z) w(t, \mathrm{d}z) \right| \leqslant 2,$$

and w solves the following FPE with initial value $w(0) = 0$ in the distributional sense:

$$\partial_t w = (\Delta_{\bar{v}}^{\frac{\alpha}{2}} - v \cdot \nabla_x) w - \mathrm{div}_v(Hw).$$

By Duhamel's formula, we have

$$w(t, z) + \int_0^t P_{t-s} \mathrm{div}_v(H_s w_s) \mathrm{d}s = 0, \ t \in (0, T].$$

Let P_t^* be the adjoint semigroup of P_t (see (2.22)). By the definition of total variation norm,

$$\|w(t)\|_m = \sup_{\|f\|_\infty \leqslant 1} \left| \int_0^t \int_{\mathbb{R}^{2d}} P_{t-s} \mathrm{div}_v(H_s w_s)(z) f(z) \mathrm{d}z \mathrm{d}s \right|$$

$$= \sup_{\|f\|_\infty \leqslant 1} \left| \int_0^t \int_{\mathbb{R}^{2d}} H_s(z) \cdot \nabla_v P_{t-s}^* f(z) w(s, \mathrm{d}z) \mathrm{d}s \right|$$

$$\leqslant \sup_{\|f\|_\infty \leqslant 1} \int_0^t \|H_s\|_\infty \|\nabla_v P_{t-s}^* f\|_\infty \|w(s)\|_m \mathrm{d}s.$$

Let $q > \frac{\alpha}{\alpha-1}$ be as in (i) of Lemma 4.7 and $\frac{1}{r} + \frac{1}{q} = 1$. Since $\|\nabla_v P_{t-s}^* f\|_\infty \lesssim (t-s)^{-\frac{1}{\alpha}} \|f\|_\infty$ (see Lemma 2.10), by Hölder's inequality we further have

$$\|w(t)\|_m \lesssim \int_0^t (t-s)^{-\frac{1}{\alpha}} \|H_s\|_\infty \|w(s)\|_m \mathrm{d}s \ \leqslant \|H\|_{\mathbb{L}_T^q(\mathbb{L}^\infty)} \left(\int_0^t (t-s)^{-\frac{r}{\alpha}} \|w(s)\|_m^r \mathrm{d}s \right)^{1/r},$$

which, by $\frac{r}{\alpha} < 1$ and Gronwall's inequality, yields $w(t) = 0$. Thus we obtain (4.29).

(i) The smoothness of density and short time singularity estimates follow by (4.26).

(ii) The large time decay estimate follows by Theorem 3.14.

(iii) The stability of the density follows by (3.32).

(iv) The strong stability follows by Theorem 4.1 and (iii) of Lemma 4.7.

(v) The global existence follows by Theorems 3.19, 3.20 and Remark 3.18.   □

Remark 4.9. Under the assumptions of Remark 3.6, by Remark 4.8, it is easy to see that the weak well-posedness still holds.

Next we give the

Proof of Theorem 1.3. We only show the two-sides estimates of the density since the others are completely the same as Theorem 1.3. Let u be the unique solution of (3.60) and define $H_t(x) = b(t) * u(t)(x)$. Consider the following linear SDE:

$$\mathrm{d}X_t = H_t(X_t)\mathrm{d}t + \mathrm{d}L_t^{(\alpha)}.$$

Since $H \in \mathbb{L}_T^q(\mathbf{C}^\beta)$ for some $q > \frac{\alpha}{\alpha-1}$ and $\beta > 0$, it is well known that for each starting point $X_0 = x$, X_t has a density $p(t, x, y)$ enjoying two-sides estimates: for $\alpha \in (1, 2)$ (see [26]),

$$p(t, x, y) \asymp t(t^{1/\alpha} + |x - y|)^{-d-\alpha},$$

and for $\alpha = 2$ (see [26, Theorem 1.1]),

$$C_1^{-1} t^{-d/2} \mathrm{e}^{-\frac{|x-y|^2}{C_2 t}} \leqslant p(t, x, y) \leqslant C_1 t^{-d/2} \mathrm{e}^{-\frac{C_2|x-y|^2}{t}}.$$

For general starting point $X_0 \sim \mu_0$, we have

$$\rho(t, y) = \int_{\mathbb{R}^d} p(t, x, y) \mu_0(\mathrm{d}x).$$

The desired estimates now follows.   □

5. Application to fractional Vlasov-Poisson-Fokker-Planck equation

Let $d \geqslant 3$. In this section, we apply Theorem 3.14 to the following VPFP equation in $\mathbb{R}^{2d}$:

$$\partial_t f = \Delta_v^{\alpha/2} f - v \cdot \nabla_x f + \gamma \nabla U \cdot \nabla_v f, \tag{5.1}$$

where $\alpha \in (1, 2]$ and $\gamma = \pm 1$ stands for the attractive or repulsive force in physics, respectively,

$$U(t, x) = (|\cdot|^{2-d} * \langle f \rangle(t))(x), \quad \langle f \rangle(t, x) := \int_{\mathbb{R}^d} f(t, x, v) \mathrm{d}v.$$

Let

$$K(x) := \nabla |\cdot|^{2-d}(x), \quad \widetilde{K}(x, v) := K(x).$$

With these notations, we have

$$E(x) := \nabla U * \langle f \rangle(x) = K * \langle f \rangle(x) = (\widetilde{K} * f)(x, v).$$

By Lemma A.3 and Proposition A.4, we have

Lemma 5.1. *Let* $\varrho_0, \varrho_1 \in [1, \infty]$ *and* $\boldsymbol{\varrho}_i := (\varrho_i, \infty)$, $i = 0, 1$. *If either* $\varrho_0 > \frac{d}{d-1}$ *and* $\varrho_1 \in [1, \infty]$ *or* $\varrho_0 < \frac{d}{d-1}$ *and* $\varrho_1 \in (\frac{d}{d-1}, \infty]$, *then for any* $\beta_b \leqslant (1 + \alpha)(\frac{d}{\varrho_0} + 1 - d)$,

$$\widetilde{K} \in \mathbf{B}_{\boldsymbol{\varrho}_0; \boldsymbol{a}}^{\beta_b} + \mathbf{B}_{\boldsymbol{\varrho}_1; \boldsymbol{a}}^{\beta_b}.$$

Moreover, for $s_0 = 1 + \frac{\alpha}{2}$ *and* $s_1 = \frac{\alpha}{2} - 1 + \beta_b$,

$$\widetilde{K} \in \mathbf{B}_{\boldsymbol{\varrho}_0; x, \boldsymbol{a}}^{s_0, s_1}.$$

As an application of Theorem 3.14, we have

Theorem 5.2. *Let* $\alpha \in (1, 2]$, $p_0 \in [2, \infty]$, $\boldsymbol{p}_0 := (p_0, 1)$ *and* $\beta_0 \in (-\alpha, 0)$. *Suppose that*

$$\frac{(1+\alpha)d}{p_0} = 2\alpha + \beta_0, \ \beta_0 \in (-1, 0), \quad \frac{(1+\alpha)d}{p_0} < 2\alpha + 1 + 2\beta_0, \ \beta_0 \in (-\alpha, -1]. \tag{5.2}$$

For any $f_0 \in \mathbf{B}_{\boldsymbol{p}_0; \boldsymbol{a}}^{\beta_0} \cap \mathbf{B}_{\mathbf{1}; \boldsymbol{a}}^{\beta_0}$, *there is a constant* $C_0 = C_0(\alpha, \beta_0, p_0, d, \|f_0\|_{\mathbf{B}_{\mathbf{1}; \boldsymbol{a}}^{\beta_0}}) > 0$ *such that if*

$$\|f_0\|_{\mathbf{B}_{\boldsymbol{p}_0; \boldsymbol{a}}^{\beta_0}} \leqslant C_0, \tag{5.3}$$

then there is a unique global smooth solution f *to VPFP* (5.1) *such that for any* $\beta \geqslant 0$,

$$\sup_{t \in (0, 1]} \left(t^{\frac{\beta - \beta_0}{\alpha}} \left(\|f(t)\|_{\mathbf{B}_{\boldsymbol{p}_0; \boldsymbol{a}}^{\beta, 1}} + \|f(t)\|_{\mathbf{B}_{\mathbf{1}; \boldsymbol{a}}^{\beta, 1}} \right) \right) < \infty, \tag{5.4}$$

and for any $p \in [2, \infty]$ *and* $\boldsymbol{p} := (p, 1)$,

$$\sup_{t \geqslant 1} \left(t^{\frac{d(1+\alpha)}{\alpha} \cdot (1 - \frac{1}{p})} \|f(t)\|_{\mathbf{B}_{\boldsymbol{p}; \boldsymbol{a}}^{\beta, 1}} \right) < \infty. \tag{5.5}$$

In particular,

$$\sup_{t \geqslant 1} \left(t^{\frac{(\alpha+1)(d-1)}{\alpha}} \|E(t)\|_\infty \right) < \infty. \tag{5.6}$$

Proof. To apply Theorem 3.14, we need to first choose suitable parameters so that $(\mathbf{H}_0)$ holds. Let

$$\boldsymbol{\varrho}_0 = (1, \infty), \boldsymbol{\varrho}_1 = (2, \infty), q_b = \infty, \ \beta_b = 1 + \alpha, \ \boldsymbol{p}_1 := (2, 1).$$

With these choice of parameters, it is easy to see that

$$\mathscr{A}_0 := \boldsymbol{a} \cdot \left(\frac{d}{\boldsymbol{p}_0} + \frac{d}{\boldsymbol{\varrho}_0} - \boldsymbol{d} \right) = \frac{(1+\alpha)d}{p_0}, \quad \mathscr{A}_1 := \boldsymbol{a} \cdot \left(\frac{d}{\boldsymbol{p}_1} + \frac{d}{\boldsymbol{\varrho}_1} - \boldsymbol{d} \right) = 0.$$

By (5.2), one sees that $\beta_b \neq \mathscr{A}_i - \beta_0$, $i = 0, 1$ and

$$0 < \mathscr{A}_0 - \beta_0 - \beta_b \leqslant \alpha - 1, \quad \mathscr{A}_0 - \beta_0 - \beta_b < \alpha + \beta_0,$$

which means that $(\mathbf{H_0})$ holds. Moreover, $\boldsymbol{a} \cdot \frac{d}{\varrho_1} = \frac{(1+\alpha)d}{2} > \alpha$. Thus by Lemma 5.1 and Theorem 3.14, there is a constant $C_0 > 0$ depending on $\|f_0\|_{\mathbf{B}^{\beta_0}_{1;\boldsymbol{a}}}$ such that if

$$\|f_0\|_{\mathbf{B}^{\beta_0}_{\boldsymbol{p}_0;\boldsymbol{a}}} + \|f_0\|_{\mathbf{B}^{\beta_0}_{\boldsymbol{p}_1;\boldsymbol{a}}} \leqslant C_0, \tag{5.7}$$

then there is a unique smooth solution f so that (5.4) and (5.5) for $\boldsymbol{p} = \boldsymbol{p}_0, \boldsymbol{p}_1$ hold. Let $\theta \in [0,1]$ be defined by $\frac{1-\theta}{p_0} + \theta = \frac{1}{2}$. Note that by the interpolation inequality (2.16),

$$\|f_0\|_{\mathbf{B}^{\beta_0}_{\boldsymbol{p}_1;\boldsymbol{a}}} \lesssim \|f_0\|^{1-\theta}_{\mathbf{B}^{\beta_0}_{\boldsymbol{p}_0;\boldsymbol{a}}} \|f_0\|^{\theta}_{\mathbf{B}^{\beta_0}_{1;\boldsymbol{a}}}.$$

By adjusting the small constant C_0 in (5.3), one sees that (5.3) implies (5.7).

Now we devote to showing (5.5) for all $\boldsymbol{p} = (p,1)$, where $p \in [2,\infty]$. If $p_0 = \infty$, then by the interpolation inequality (2.16), we immediately have

$$\|f(t)\|_{\mathbf{B}^{\beta,1}_{\boldsymbol{p};\boldsymbol{a}}} \leqslant \|f(t)\|^{\frac{2}{p}}_{\mathbf{B}^{\beta,1}_{(2,1);\boldsymbol{a}}} \|f(t)\|^{1-\frac{2}{p}}_{\mathbf{B}^{\beta,1}_{(\infty,1);\boldsymbol{a}}} \lesssim t^{-\frac{d(1+\alpha)}{\alpha}\cdot(1-\frac{1}{p})}, \quad t \geqslant 1. \tag{5.8}$$

If $p_0 < \infty$, then for $\beta > \frac{d}{p_0}$, by the Sobolev embedding (2.17) and (5.5) for $\boldsymbol{p} = \boldsymbol{p}_0$, we have

$$\|f(t)\|_{\mathbf{B}^{\beta}_{(\infty,1);\boldsymbol{a}}} \lesssim \|f(t)\|_{\mathbf{B}^{\beta,1}_{\boldsymbol{p}_0;\boldsymbol{a}}} \lesssim t^{-\frac{d(1+\alpha)}{\alpha}\cdot(1-\frac{1}{p_0})}, \quad t \geqslant 1.$$

In particular, one can choose $t_0 \geqslant 1$ large enough so that

$$\|f(t_0)\|_{\mathbf{B}^{\beta}_{(\infty,1);\boldsymbol{a}}} \leqslant C_0.$$

Starting from the time t_0 and with the choice of $p_0 = \infty$, we then deduce (5.5) by (5.8).

Finally, for (5.6), by (ii) of Lemma A.3 with $\gamma = d-1$ and $p = 1$, we have for $t \geqslant 1$,

$$\|E(t)\|_\infty = \|\nabla U * \langle f \rangle(t)\|_\infty \lesssim \|\langle f \rangle(t)\|^{\frac{1}{d}}_1 \|\langle f \rangle(t)\|^{\frac{d-1}{d}}_\infty$$
$$\lesssim \|f(t)\|^{\frac{1}{d}}_{\mathbb{L}^1} \|f(t)\|^{\frac{d-1}{d}}_{\mathbf{B}^{0,1}_{(\infty,1);\boldsymbol{a}}} \lesssim t^{-\frac{(\alpha+1)(d-1)}{\alpha}}.$$

The proof is complete. $\qquad\square$

Remark 5.3. Theorem 5.2 extends the well-posedness theory for VPFP equations to the distribution initial data (see [77], [10], [14], [64]). Notice that when $\alpha = 2$, in three dimensional case, Carrillo and Soler [14] studied the well-posedness of VPFP equation with initial data being in Morrey spaces. More precisely, under more restricted smallness assumption on initial data,

$$\|f_0\|_m + \|f_0\|_{M\widetilde{L}_{9/4}} \leqslant \varepsilon,$$

where $M\widetilde{L}_p$ is some Morrey space (see [14, (1.4)]), they showed global well-posedess and obtained the following decay estimate for the force field $E(t,x)$,

$$\|E(t)\|_\infty \lesssim t^{-1/2}, \quad t \geqslant 1.$$

Note that one can check that for any $p_0 \in (9/4, 3)$,

$$M\widetilde{L}_{9/4} \cap m \subset \mathbf{B}^0_{(9/4,1);\boldsymbol{a}} \cap \mathbf{B}^0_{1;\boldsymbol{a}} \subset \mathbf{B}^{9/p_0-4}_{(p_0,1);\boldsymbol{a}} \cap \mathbf{B}^{9/p_0-4}_{1;\boldsymbol{a}}.$$

Later, under smallness condition $\|f_0\|_{L^1} + \|f_0\|_{L^1_v(L^\infty_x)} \leqslant \varepsilon$, Ono and Strauss [64] showed the following optimal rate of large time decay

$$\|E(t)\|_\infty \lesssim t^{-3(d-1)/2}, \quad t \geqslant 1,$$

which coincides with (5.6). It is pointed out that requiring small $\|f_0\|_m$ rules out the application for the associated mean-field SDE since we always have $\|f_0\|_m \equiv 1$ for SDEs. In fact, we have paid much efforts to drop the smallness of $\|f_0\|_m$ in Theorem 3.14.

Now we consider the following mean-field SDE:

$$\begin{cases} X_t = X_0 + \displaystyle\int_0^t V_s \mathrm{d}s, \\ V_t = V_0 + \displaystyle\int_0^t (K * \mu_{X_s})(X_s)\mathrm{d}s + L_t^{(\alpha)}. \end{cases} \qquad (5.9)$$

By Theorem 1.3 we have

Theorem 5.4. *Let $\alpha \in (1,2]$, $\beta_0 \in (-\alpha, 0)$, $p_0 \in [2, \infty]$, $\boldsymbol{p}_0 := (p_0, 1)$ and $(X_0, V_0) \sim \mu_0 \in \mathscr{P}(\mathbb{R}^{2d})$.*
(A) *Suppose that*

$$\frac{(1+\alpha)d}{p_0} < 2\alpha + 1 + \beta_0 + \beta_0 \wedge (-1).$$

Then there is a constant $C_0 = C_0(\alpha, \beta_0, p_0, d) > 0$ such that if

$$\|\mu_0\|_{\mathbf{B}_{\boldsymbol{p}_0;\boldsymbol{a}}^{\beta_0}} \leqslant C_0, \qquad (5.10)$$

then there is a unique weak solution $Z_t = (X_t, V_t)$ to SDE (5.9) such that for each $t > 0$, Z_t admits a smooth density $f(t)$ that satisfies (5.4) and (5.5).
(B) *Suppose that*

$$\frac{(1+\alpha)d}{p_0} < \tfrac{5}{2}\alpha - 2 + \beta_0.$$

Then under (5.10), there is a unique strong solution to SDE (5.9).

*Moreover, when $d = 3$ and $\alpha = 2$, and if $\mathbb{E}|V_0|^2 < \infty$, then the above **(A)** and **(B)** still hold without smallness condition (5.10).*

Proof. Note that by (2.9),

$$\|\mu_0\|_{\mathbf{B}_{1;\boldsymbol{a}}^0} \leqslant C(d, \alpha).$$

As the choice of parameters in Theorem 5.2, **(A)** and **(B)** are direct applications of Theorem 1.3. When $d = 3$ and $\alpha = 2$, in order to obtain a global solution without smallness condition (5.10), we shall use a result of [64, Theorem 5.1]. More precisely, it suffices to find a small time $t_0 > 0$ so that

$$\|f(t_0)\|_{L_v^1(L_x^\infty)} := \int_{\mathbb{R}^d} \sup_{x \in \mathbb{R}^d} |f(t_0, x, v)| \mathrm{d}v < \infty$$

and

$$\int_{\mathbb{R}^{2d}} |v|^2 f(t_0, x, v) \mathrm{d}x \mathrm{d}v < \infty.$$

For any $f_0 \in \mathbf{B}_{\boldsymbol{p}_0;\boldsymbol{a}}^{\beta_0}$, by Theorem 3.17, one can find a $t_0 > 0$ small enough so that there is a local solution for SDE (5.9) up to time t_0. Moreover, for any $\beta \geqslant 0$,

$$\sup_{t \in (0,t_0]} \left(t^{\beta/\alpha} \|f(t)\|_{\mathbb{Z}_{\beta+\beta_0}} \right) < \infty.$$

By (5.9) we have

$$(\mathbb{E}|V_t|^2)^{1/2} \leqslant (\mathbb{E}|V_0|^2)^{1/2} + \left(\mathbb{E}\left| \int_0^t (K * \mu_{X_s})(X_s)\mathrm{d}s \right|^2 \right)^{1/2} + \sqrt{2}(\mathbb{E}|W_t|^2)^{1/2}$$

$$\leqslant (\mathbb{E}|V_0|^2)^{1/2} + \int_0^t \|K * \mu_{X_s}\|_\infty \mathrm{d}s + \sqrt{2t},$$

which implies by (i) of Lemma 4.7 that for any $t \in [0, t_0]$,

$$\int_{\mathbb{R}^{2d}} |v|^2 f(t, x, v) \mathrm{d}x \mathrm{d}v = \mathbb{E}|V_t|^2 < \infty.$$

By (2.12) and (2.17), for any $t \in (0, t_0]$ and $\beta > 3(1 + \alpha)(1 - \frac{1}{p_0})$, we have

$$\|f(t)\|_{L^1_v(L^\infty_x)} \leqslant \sum_{j \geqslant 0} \|\mathcal{R}^{\boldsymbol{a}}_j f(t)\|_{L^1_v(L^\infty_x)} = \|f(t)\|_{\mathbf{B}^{0,1}_{(\infty,1);\boldsymbol{a}}} \lesssim \|f(t)\|_{\mathbf{B}^\beta_{\boldsymbol{p}_0;\boldsymbol{a}}} < \infty,$$

and for $\beta > 3(1 + \alpha)(1 - \frac{1}{p_0}) + 3$,

$$\|f(t)\|_{L^\infty_{x,v}} \lesssim \|f(t)\|_{\mathbf{B}^\beta_{\boldsymbol{p}_0;\boldsymbol{a}}} < \infty.$$

Now starting from time $t_0/2$, we can use [64, Theorem 5.1] to find a unique global smooth solution $\widetilde{f}$ for (5.1) with initial value $\widetilde{f}(0) = f(t_0/2)$. Since f is smooth on $[\frac{t_0}{2}, t_0]$, by the uniqueness we have $f(t + t_0/2) = \widetilde{f}(t)$ on $[0, \frac{t_0}{2}]$. Thus we can patch up the solutions and conclude the proof. $\square$

Remark 5.5. For $\alpha \in (1, 2)$, it seems to be an open question to obtain the global solution for fractional VPFP equation since there is no finite second-order moment for α-stable process.

6. Application to fractional Navier-Stokes equations and SQG equations

In this section we apply our results to the 2D and 3D fractional Navier-Stokes equations:

$$\partial_t u = \Delta^{\alpha/2} u + u \cdot \nabla u + \nabla p, \quad \text{div} u = 0, \tag{6.1}$$

where $\alpha \in (1, 2]$, u is the velocity field and p is the pressure. For 2D fractional Navier-Stokes equations, we show the global well-posedness when the initial vorticity is in $\mathbf{B}^{-\frac{1+\alpha}{2}+}_\infty := \cup_{\varepsilon > 0} \mathbf{B}^{\varepsilon - \frac{1+\alpha}{2}}_\infty$. For 3D fractional Navier-Stokes equations, we show the global well-posedness when the initial vorticity is in $\mathbf{B}^{-\frac{1+\alpha}{2}+}_\infty$ and small. To the best of our knowledge, these results are new for fractional Navier-Stokes equations. Moreover, we also consider the SQG equation in Example 2 in the introduction.

We first consider the 2D fractional NSEs with initial vorticity in Besov space $\mathbf{B}^{\beta_0}_{p_0}$. Let $w = \text{curl}(u)$ be the vorticity of u that satisfies the following scalar equation in $\mathbb{R}^2$:

$$\partial_t w = \Delta^{\alpha/2} w + u \cdot \nabla w, \tag{6.2}$$

where the velocity field u can be recovered from w by the Biot-Savart law (cf. [60]):

$$u(x) = K_2 * w(x) = \frac{1}{2\pi} \int_{\mathbb{R}^2} \frac{(y_2 - x_2, x_1 - y_1)}{|x - y|^2} w(y) \mathrm{d}y.$$

By (iv) of Lemma A.3, one has

$$K_2 \in \mathbf{B}^\beta_{\varrho_0} + \mathbf{B}^\beta_{\varrho_1}, \quad \beta = \frac{2}{\varrho_0} - 1, \quad \varrho_0 \in [1, 2), \quad \varrho_1 \in (2, \infty]. \tag{6.3}$$

By Theorem 3.22 we have

Theorem 6.1. Let $\alpha \in (1, 2]$, $p_0 \in [1, \infty]$ and $\beta_0 \in (-\frac{\alpha+1}{2}, 0)$ satisfy

$$\frac{2}{p_0} \leqslant \alpha + \beta_0, \quad \beta_0 \in (-1, 0), \quad \frac{2}{p_0} < \alpha + 1 + 2\beta_0, \quad \beta_0 \in (-\frac{\alpha+1}{2}, -1]. \tag{6.4}$$

For any $w_0 \in \cup_{q \in [1, \infty)} (\mathbf{B}^{\beta_0, q}_{p_0} \cap \mathbf{B}^{\beta_0, q}_1)$, there is a unique global smooth solution w to vorticity equation (6.2) *on $\mathbb{R}_+$ so that for any $T > 0$ and $\beta \geqslant 0$,*

$$\sup_{t \in (0, T]} \left(t^{-(\beta - \beta_0)/\alpha} \left(\|w(t)\|_{\mathbf{B}^{\beta,1}_{p_0}} + \|w(t)\|_{\mathbf{B}^{\beta,1}_1} \right) \right) < \infty.$$

Proof. We only need to check that under (6.4), the assumptions in Theorem 3.22 hold for $b = K_2$ and $q_b = \infty$. Let $\varrho_0 = 1$, $\varrho_1 = \infty$ and $\beta_b = \frac{2}{\varrho_0} - 1 = 1$. By (6.3), one sees that the assumptions of Theorem 3.22 hold. Moreover, since $\text{div} K_2 = 0$, we have the global solution. $\square$

Remark 6.2. If we take $p_0 = \infty$ in (6.4), then for any $w_0 \in \mathbf{B}^{-\frac{1+\alpha}{2}+}_\infty \cap \mathbf{B}^{-\frac{1+\alpha}{2}+}_1$, there is a unique global smooth solution to equation (6.2) on $\mathbb{R}_+$. Since the two dimensional Brownian white noise locally belongs to $\mathbf{B}^{-1-}_\infty$, one can apply Theorem 6.1 to the 2D fractional Navier-Stokes equations with initial vorticity being the Brownian white noise.

Now we consider the following mean-field SDE in $\mathbb{R}^2$ with kernel K_2:

$$\mathrm{d}X_t = (K_2 * \mu_{X_t})(X_t)\mathrm{d}t + \mathrm{d}L_t^{(\alpha)}, \quad X_0 \sim \mu_0 \in \mathscr{P}(\mathbb{R}^2). \qquad (6.5)$$

By Theorem 1.9 with $\varrho_1 = \infty$, we have

Theorem 6.3. *Let* $\alpha \in (1,2]$, $p_0 \in [1,\infty]$ *and* $\beta_0 \in (-\frac{\alpha+1}{2}, 0)$ *satisfy*

$$\begin{cases} \frac{2}{p_0} < \alpha + 1 + \beta_0 + \beta_0 \wedge (-1), & \text{weak solution,} \\ \frac{2}{p_0} < \frac{3\alpha}{2} - 1 + \beta_0, & \text{strong solution.} \end{cases}$$

For any $\mu_0 \in \mathscr{P}(\mathbb{R}^2) \cap \mathbf{B}_{p_0}^{\beta_0}$, *there exists a unique weak/strong solution* X_t *for SDE* (6.5) *so that for each* $t > 0$, X_t *has a smooth density* $\mu_{X_t}(\mathrm{d}x) = \rho_t(x)\mathrm{d}x$ *with regularity that for any* $T, \beta \geqslant 0$,

$$\sup_{t \in (0,T]} \left(t^{\frac{\beta}{\alpha}} \|\rho_t\|_{\mathbf{B}_{p_0}^{\beta+\beta_0}} \right) < \infty.$$

Next we consider the following mean-field SDE in $\mathbb{R}^2$ associated with SQG (1.16):

$$\mathrm{d}X_t = (\mathscr{R} * \mu_{X_t})(X_t)\mathrm{d}t + \mathrm{d}L_t^{(\alpha)}, \quad X_0 \sim \mu_0 \in \mathscr{P}(\mathbb{R}^2), \qquad (6.6)$$

where $\mathscr{R} = (-x_2, x_1)/|x|^3$. Note that by (iii) of Lemma A.3,

$$\mathscr{R} \in \mathbf{B}_{\varrho_0}^{\beta}, \quad \beta = \frac{2}{\varrho_0} - 2, \ \varrho_0 \in (1, \infty].$$

By Theorem 1.9 with $\varrho_1 = \varrho_0$, we similarly have

Theorem 6.4. *Let* $\alpha \in (1,2]$, $p_0 \in [1,\infty]$ *and* $\beta_0 \in (-\alpha, 0)$ *satisfy*

$$\begin{cases} \frac{2}{p_0} < \alpha + \beta_0 + \beta_0 \wedge (-1), & \text{weak solution,} \\ \frac{2}{p_0} < \frac{3\alpha}{2} - 2 + \beta_0, & \text{strong solution.} \end{cases}$$

For any $\mu_0 \in \mathscr{P}(\mathbb{R}^2) \cap \mathbf{B}_{p_0}^{\beta_0}$, *there exists a unique weak/strong solution* X_t *for SDE* (6.6) *so that for each* $t > 0$, X_t *has a smooth density* $\mu_{X_t}(\mathrm{d}x) = \rho_t(x)\mathrm{d}x$ *with regularity that for any* $T, \beta \geqslant 0$,

$$\sup_{t \in (0,T]} \left(t^{\frac{\beta}{\alpha}} \|\rho_t\|_{\mathbf{B}_{p_0}^{\beta+\beta_0}} \right) < \infty.$$

Remark 6.5. From the condition, one sees that for strong solutions, we need to assume $\alpha \in (\frac{4}{3}, 2]$.

Finally we consider the 3D fractional NSEs with initial vorticity in non-homogeneous space $\mathbf{B}_{p_0}^{\beta_0}$. Let $w = \mathrm{curl}(u)$ be the vorticity of u. Then

$$\partial_t w = \Delta^{\frac{\alpha}{2}} w - u \cdot \nabla w + w \cdot \nabla u = \Delta^{\frac{\alpha}{2}} w + \mathrm{div}(u \otimes w - w \otimes u), \qquad (6.7)$$

where for two divergence free vector fields w_1, w_2,

$$\mathrm{div}(w_1 \otimes w_2) := w_1 \cdot \nabla w_2.$$

By the Biot-Savart law, one knows that (cf. [60])

$$u(x) = K_3 * w(x) = \int_{\mathbb{R}^3} K_3(x-y)w(y)\mathrm{d}y,$$

where

$$K_3(x)h = \frac{1}{4\pi} \frac{x \times h}{|x|^3}, \ h \in \mathbb{R}^3.$$

Note that by (iv) of Lemma A.3,

$$K_3 \in \mathbf{B}_{\varrho_0}^{\beta} + \mathbf{B}_{\varrho_1}^{\beta}, \quad \beta = \frac{3}{\varrho_0} - 2, \ \varrho_0 \in [1, \tfrac{3}{2}), \ \varrho_1 \in (\tfrac{3}{2}, \infty]. \qquad (6.8)$$

By Duhamel's formula, we have

$$w(t) = P_t w_0 + \int_0^t P_{t-s}\mathrm{div}((K_3 * w) \otimes w - w \otimes (K_3 * w))\mathrm{d}s.$$

The following result can be proved by using similar argument as in proving Theorem 3.14.

Theorem 6.6. *Let $\alpha \in (1,2]$, $p_0 \in [2,\infty]$ and $\beta_0 \in (-\alpha,0)$ satisfy*

$$\tfrac{3}{p_0} \leqslant \alpha + \beta_0, \ \ \beta_0 \in (-1,0), \ \ \ \tfrac{3}{p_0} < \alpha + 1 + 2\beta_0, \ \ \beta_0 \in (-\tfrac{\alpha+1}{2}, -1]. \tag{6.9}$$

For any initial vorticity $w_0 \in \mathbf{B}_{p_0}^{\beta_0} \cap \mathbf{B}_1^{\beta_0}$, there is a constant $C_0 = C_0(\alpha,\beta_0,p_0, \|w_0\|_{\mathbf{B}_1^{\beta_0}}) > 0$ such that if

$$\|w_0\|_{\mathbf{B}_{p_0}^{\beta_0}} \leqslant C_0, \tag{6.10}$$

then there is a unique global smooth solution w to equation (6.7) on $\mathbb{R}_+$ so that for any $\beta \geqslant 0$,

$$\sup_{t>0} \left(t^{(\beta-\beta_0)/\alpha} \big(\|w(t)\|_{\mathbf{B}_{p_0}^{\beta,1}} + \|w(t)\|_{\mathbf{B}_1^{\beta,1}} \big) \right) < \infty \tag{6.11}$$

and for any $p \in [2,\infty]$ and $\beta \geqslant 0$, there is a constant $C = C(\alpha,p,\beta,\beta_0,p_0, \|w_0\|_{\mathbf{B}_1^{\beta_0}}) > 0$ such that

$$\sup_{t\geqslant 1} \left(t^{-(3-\frac{3}{p})/\alpha} \|w(t)\|_{\mathbf{B}_p^{\beta,1}} \right) < \infty. \tag{6.12}$$

Proof. For given β_0, p_0 satisfying (6.9), by (6.8), one can choose $\varrho_0 = 1$, $\varrho_1 = 2$, $\beta_b = \tfrac{3}{\varrho_0} - 2 = 1$, $q_b = \infty$ so that $b = K_3 \in \mathbf{B}_{\varrho_0}^{\beta_b} + \mathbf{B}_{\varrho_1}^{\beta_b}$, and (3.61) holds and

$$\tfrac{1}{p_0} + \tfrac{1}{\varrho_0} \geqslant 1, \ \ \ \ \tfrac{3}{\varrho_1} > \alpha - 1.$$

Define

$$\mathscr{H}_t(u,w) := \int_0^t P_{t-s} \mathrm{div}((K_3 * u) \otimes w - w \otimes (K_3 * u)) \mathrm{d}s.$$

As in the proof of Theorem 3.14, one can show that there is a small constant $C_0 > 0$ such that under condition (6.10), there is a unique global smooth solution w satisfying

$$w(t) = P_t w_0 + \mathscr{H}_t(w,w),$$

and (6.11) and (6.12) for $p = 2, p_0$ hold.

Now we devote to showing (6.12) for all $p \in [2,\infty]$. Fix $\beta \geqslant 0$. Note that it is completely the same as proving Theorem 3.14 that if $p_0 = \infty$, then

$$\sup_{t\geqslant 1} \left(t^{3/\alpha} \|w(t)\|_{\mathbf{B}_\infty^{\beta,1}} + t^{3/(2\alpha)} \|w(t)\|_{\mathbf{B}_2^{\beta,1}} \right) < \infty,$$

which implies by interpolation inequality (2.16),

$$\|w(t)\|_{\mathbf{B}_p^{\beta,1}} \leqslant \|w(t)\|_{\mathbf{B}_2^{\beta,1}}^{\frac{2}{p}} \|w(t)\|_{\mathbf{B}_\infty^{\beta,1}}^{1-\frac{2}{p}} \lesssim t^{-(3-\frac{3}{p})/\alpha}, \ \ t \geqslant 1. \tag{6.13}$$

If $p_0 < \infty$, by (6.11) with $\beta = \tfrac{3}{p_0}$ and the Sobolev embedding inequality (2.17), we have

$$\|w(t)\|_\infty \lesssim \|w(t)\|_{\mathbf{B}_{p_0}^{3/p_0,1}} \lesssim t^{-(3-3/p_0)/\alpha}, \ \ t \geqslant 1,$$

and

$$\sup_{t\geqslant 1} \|w(t)\|_{\mathbf{B}_1^{0,1}} < \infty.$$

In particular, we can choose t_0 large enough so that for the small constant C_0 in (6.10),

$$\|w(t_0)\|_{\mathbf{B}_{\infty;\mathbf{a}}^{\beta_0}} \leqslant C \|w(t_0)\|_\infty \leqslant C_0.$$

Now starting from t_0, we can use (6.13) to conclude the proof of (6.12). $\qquad\square$

Remark 6.7. When $\alpha = 2$, it is by no means that Theorem 6.6 is new. For fractional Navier-Stokes equation with $\alpha \in (1,2)$, Wu [80, Theorem 6.1] showed the global well-posedness for small initial velocity in homogeneous Besov spaces. As one knows, there are vast literatures about 3D Navier-Stokes equations. Here we only want to emphasize that we present a general framework that contains Navier-Stokes equations as one application.

7. Application to fractional kinetic porous medium equation with viscosity

In this section we consider the following fractional kinetic porous medium equation with viscosity

$$\partial_t f = \Delta_v^{\alpha/2} f - v \cdot \nabla_x f + \mathrm{div}_v((\nabla_v \Delta_v^{-s} f)f), \qquad (7.1)$$

where $s \in (0,1]$ and $d \geqslant 3$. Note that

$$(\nabla_v \Delta_v^{-s} f)(x,v) = K * f(x,v)$$

with

$$K(x,v) := c_{d,s} \delta_0(\mathrm{d}x) \nabla_v |v|^{2s-d},$$

where $c_{d,s} > 0$ is a normalized constant and δ_0 is the Dirac measure at point 0.

The following lemma provides the necessary regularity estimate for K.

Lemma 7.1. *Fix $p_0 \in [1,\infty]$. Let $\boldsymbol{p} = (p_x, p_v)$ with $p_x \in [1,\infty]$ and $p_v \in [p_0, \infty]$. For any $\beta_0 < \frac{d}{p_0} - \frac{d}{p_v}$, there is a constant $C = C(\alpha, d, p_0, \beta_0, \boldsymbol{p}) > 0$ such that for any $U(v) \in \mathbf{B}_{p_0}^{\beta_0}$,*

$$\|\delta_0 \otimes U\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta_0+d-\frac{d}{p_0}-\boldsymbol{a}\cdot(\boldsymbol{d}-\frac{\boldsymbol{d}}{\boldsymbol{p}})}} \lesssim_C \|U\|_{\mathbf{B}_{p_0}^{\beta_0}}. \qquad (7.2)$$

In particular, for $s \in (0,1]$ and $\boldsymbol{\varrho}_i = (1, \varrho_i)$, $i = 0,1$, where $\varrho_0 \in [1,\infty]$ and $\varrho_1 \in (\frac{d}{d-2s}, \infty]$,

$$\delta_0(\mathrm{d}x) \nabla |v|^{2s-d} \in \mathbf{B}_{\boldsymbol{\varrho}_0;\boldsymbol{a}}^{\beta_b} + \mathbf{B}_{\boldsymbol{\varrho}_1;\boldsymbol{a}}^{\beta_b}, \qquad (7.3)$$

where $\beta_b < (2s - 1 - d + \frac{d}{\varrho_0}) \wedge 0$.

Proof. By definition (2.3) we have

$$\mathscr{R}_j^{\boldsymbol{a}}(\delta_0 \otimes U)(x,v) = (\check{\phi}_j^{\boldsymbol{a}} *_v U)(x,v) = \sum_{k \geqslant 0} (\check{\phi}_j^{\boldsymbol{a}} *_v \mathscr{R}_k^v U)(x,v),$$

where $*_v$ stands for the convolution with respect to the variable v and $\mathscr{R}_k^v U = \check{\phi}_k * U$. Noting that

$$\widehat{(\check{\phi}_j^{\boldsymbol{a}} *_v \mathscr{R}_k^v U)}(\xi,\eta) = \phi_j^{\boldsymbol{a}}(\xi,\eta)\phi_k(\eta)\hat{U}(\eta),$$

and by (2.1) and (2.6),

$$\phi_j^{\boldsymbol{a}}(\xi,\eta)\phi_k(\eta) = 0, \quad k > j + 2,$$

by Young's inequality, scaling of $\check{\phi}_j^{\boldsymbol{a}}$ and Bernstein's inequality (2.7), we have for any $j \geqslant 0$,

$$
\begin{aligned}
\|\mathscr{R}_j^{\boldsymbol{a}}(\delta_0 \otimes U)\|_{\boldsymbol{p}} &\leqslant \sum_{k \leqslant j+2} \|\check{\phi}_j^{\boldsymbol{a}} *_v \mathscr{R}_k^v U\|_{(p_x,p_v)} \leqslant \sum_{k \leqslant j+2} \|\check{\phi}_j^{\boldsymbol{a}}\|_{(p_x,1)} \|\mathscr{R}_k^v U\|_{p_v} \\
&= \sum_{k \leqslant j+2} 2^{j(1+\alpha)(d-\frac{d}{p_x})} \|\check{\phi}_1^{\boldsymbol{a}}\|_{(p_x,1)} \|\mathscr{R}_k^v U\|_{p_v} \\
&\lesssim 2^{j(1+\alpha)(d-\frac{d}{p_x})} \sum_{k \leqslant j+2} 2^{k(\frac{d}{p}-\frac{d}{p_v}-\beta)} \|U\|_{\mathbf{B}_{p_0}^{\beta_0}} \\
&\lesssim 2^{j(\boldsymbol{a}\cdot(\boldsymbol{d}-\frac{\boldsymbol{d}}{\boldsymbol{p}})+\frac{d}{p_0}-d-\beta_0)} \|U\|_{\mathbf{B}_{p_0}^{\beta_0}},
\end{aligned}
$$

where the last step is due to $\beta_0 < \frac{d}{p_0} - \frac{d}{p_v}$. Thus we get (7.2). Now, fix $s \in (0,1]$. Let $U(v) := \nabla_v |v|^{2s-d}$. For any $\varrho_1 \in (\frac{d}{d-2s}, \infty]$, by (iv) of Lemma A.3, we have

$$U \in \mathbf{B}_1^{2s-1} + \mathbf{B}_{\varrho_1}^{2s-1}.$$

Now for any $\varrho_0 \in [1,\infty]$, if we take $\beta_0 < (2s-1) \wedge (d - \frac{d}{\varrho_0})$, $p_0 = 1$ and $\boldsymbol{p} = (1, \varrho_0)$ in (7.2), then for $U_0 \in \mathbf{B}_1^{2s-1} \subset \mathbf{B}_1^{\beta_0}$,

$$\delta_0 \otimes U_0 \in \mathbf{B}_{\boldsymbol{\varrho}_0;\boldsymbol{a}}^{\beta_0+d/\varrho_0-d}.$$

Similarly, for any $\varrho_1 \in (\frac{d}{d-2s}, \infty]$, if we take $\beta_0 = \beta_b < 0$, $p_0 = \varrho_1$ and $\boldsymbol{p} = (1, \varrho_1)$ in (7.2), then for $U_1 \in \mathbf{B}_{\varrho_1}^{2s-1} \subset \mathbf{B}_{\varrho_1}^{\beta_b}$,

$$\delta_0 \otimes U_1 \in \mathbf{B}_{\boldsymbol{\varrho}_1;\boldsymbol{a}}^{\beta_b}.$$

Thus we obtain (7.3) by noting $\beta_0 + \frac{d}{\varrho_0} - d < (2s-1) \wedge (d - \frac{d}{\varrho_0}) + \frac{d}{\varrho_0} - d = (2s - 1 - d + \frac{d}{\varrho_0}) \wedge 0.$ □

The associated density-distribution dependent SDE to (7.1) is taken as the form:

$$
\begin{cases}
X_t = X_0 + \displaystyle\int_0^t V_s \mathrm{d}s, \\[2mm]
V_t = V_0 + \displaystyle\int_0^t (K * f_{Z_s})(X_s, V_s)\mathrm{d}s + L_t^{(\alpha)},
\end{cases}
\tag{7.4}
$$

where f_{Z_s} stands for the distribution density of $Z_s = (X_s, V_s)$. By Theorem 1.3 we have

Theorem 7.2. *Let $s \in (0,1]$, $\beta_0 \in (-\alpha, 0)$, $\boldsymbol{p}_0 = (p_{0x}, p_{0v}) \in [1, \infty]^2$ and $(X_0, V_0) \sim \mu_0 \in \mathscr{P}(\mathbb{R}^{2d})$. Suppose that*

$$
(1 + \alpha)\frac{d}{p_{0x}} < \alpha + (2s - 1 - \frac{d}{p_{0v}}) \wedge 0 + \beta_0 + \beta_0 \wedge (-1).
\tag{7.5}
$$

If $\mu_0 \in \mathbf{B}_{\boldsymbol{p}_0;\boldsymbol{a}}^{\beta_0}$ satisfies the following smallness condition: for some $C_0 = C_0(s, \alpha, d, \boldsymbol{p}_0, \beta_0) > 0$,

$$
\|\mu_0\|_{\mathbf{B}_{\boldsymbol{p}_0;\boldsymbol{a}}^{\beta_0}} \leqslant C_0,
\tag{7.6}
$$

then there is a unique weak solution $Z_t = (X_t, V_t)$ to SDE (7.4) in the class that for each $t > 0$, Z_t admits a smooth density $f(t)$ so that for any $T > 0$ and $\beta \geqslant 0$,

$$
\sup_{t \in (0,T]} \left(t^{(\beta - \beta_0)/\alpha} \|f(t)\|_{\mathbf{B}_{\boldsymbol{p}_0;\boldsymbol{a}}^{\beta,1}} \right) < \infty.
\tag{7.7}
$$

Moreover, if $\boldsymbol{p}_0 = (p_0, p_0)$, then we can drop (7.6).

Proof. Under (7.5), one can choose $\beta_b < (2s - 1 - \frac{d}{p_{0v}}) \wedge 0$ and $\varrho_0 = (1, \frac{p_{0v}}{p_{0v} - 1})$ so that

$$
\boldsymbol{a} \cdot (\tfrac{d}{\boldsymbol{p}_0} + \tfrac{d}{\varrho_0} - \boldsymbol{d}) - \beta_0 - \beta_b < \alpha + \beta_0 \wedge (-1).
$$

Thus by (7.3) and Theorem 1.3, as in the proof of Theorem 5.2, we conclude the proof.

Without (7.6), by Theorem 3.17, there is a local solution to SDE (7.4) on some small time interval $[0, T_0]$, where T_0 depends on the norm $\|\mu_0\|_{\mathbf{B}_{\boldsymbol{p}_0;\boldsymbol{a}}^{\beta_0}}$. Next we want to extend the solution to arbitrary large time when $\boldsymbol{p}_0 = (p_0, p_0)$. Fix $t_0 \in (0, T_0)$. Let us start from time t_0 to make some estimates. Suppose that $\psi : [0, \infty] \to [0, \infty)$ is a C^2-convex function. Since $f \in C([t_0, T_0]; C_b^2(\mathbb{R}^{2d}))$, by (C.6) we have

$$
\int \psi(f(t)) \leqslant \int \psi(f(t_0)) - \int_{t_0}^t \int \nabla_v \Delta_v^{-s} f \cdot \nabla_v f \, \psi''(f) f,
$$

where the integral $\int$ is taken over $\mathbb{R}^{2d}$. Define

$$
G(r) := \int_0^r s \psi''(s)\mathrm{d}s.
$$

It is easy to see that

$$
\int \psi(f(t)) \leqslant \int \psi(f(t_0)) - \int_{t_0}^t \int \nabla_v \Delta_v^{-s} f \cdot \nabla_v G(f) = \int \psi(f(t_0)) + \int_{t_0}^t \int \Delta_v^{1-s} f \, G(f).
$$

By (3.3) and the change of variable, one sees that

$$
\begin{aligned}
\int \Delta_v^{1-s} f \, G(f) &= \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d} \int_{|w| \geqslant \varepsilon} G(f(v)) \frac{f(v+w) - f(v)}{|w|^{d+2(1-s)}} \mathrm{d}w\mathrm{d}v \\
&= \lim_{\varepsilon \downarrow 0} \frac{1}{2} \int_{\mathbb{R}^d} \int_{|w| \geqslant \varepsilon} (G(f(v)) - G(f(v+w))) \frac{f(v+w) - f(v)}{|w|^{d+2(1-s)}} \mathrm{d}w\mathrm{d}v \\
&= -\lim_{\varepsilon \downarrow 0} \frac{1}{2} \int_{\mathbb{R}^d} \int_{|w| \geqslant \varepsilon} \int_0^1 G'(sf(v) + (1-s)f(v+w))\mathrm{d}s \frac{(f(v+w) - f(v))^2}{|w|^{d+2(1-s)}} \mathrm{d}w\mathrm{d}v.
\end{aligned}
$$

Since $G'(r) = r\psi''(r) \geqslant 0$ for $r \geqslant 0$, we get

$$\int \Delta_v^{1-s} f \, G(f) \leqslant 0.$$

Therefore,

$$\int \psi(f(t)) \leqslant \int \psi(f(t_0)).$$

In particular,

$$\|f(t)\|_{p_0} \leqslant \|f(t_0)\|_{p_0}.$$

As in the proof of Theorem 3.19, we can extend the solution to arbitrary time. $\qquad\square$

Appendix A. Several technical lemmas

The following two elementary lemmas are used to show the Schauder estimate in time-weighted spaces for kinetic semigroups.

Lemma A.1. *For any $\gamma > \beta > 0$, there is a constant $C = C(\gamma, \beta) \geqslant 1$ such that for all $\lambda > 0$,*

$$\sum_{j \geqslant 1} 2^{j\beta}((2^{-j\gamma}\lambda) \wedge 1) \asymp_C \lambda^{\frac{\beta}{\gamma}} \int_0^\lambda t^{-\frac{\beta}{\gamma}-1}(t \wedge 1)\mathrm{d}t.$$

Proof. Noting that for $j \in \mathbb{N}$ and $s \in [j-1, j]$,

$$2^{s\beta}((2^{-(s+1)\gamma}\lambda) \wedge 1) \leqslant 2^{j\beta}((2^{-j\gamma}\lambda) \wedge 1) \leqslant 2^{(s+1)\beta}((2^{-s\gamma}\lambda) \wedge 1),$$

we have

$$\mathscr{I}_0 := \int_0^\infty 2^{s\beta}((2^{-(s+1)\gamma}\lambda) \wedge 1)\mathrm{d}s \leqslant \sum_{j \geqslant 1} 2^{j\beta}((2^{-j\gamma}\lambda) \wedge 1) \leqslant \int_0^\infty 2^{(s+1)\beta}((2^{-s\gamma}\lambda) \wedge 1)\mathrm{d}s =: \mathscr{I}_1.$$

By the change of variables, we have

$$\mathscr{I}_0 = \frac{\lambda^{\frac{\beta}{\gamma}}2^{-\beta}}{\gamma \ln 2} \int_0^\lambda t^{-\frac{\beta}{\gamma}-1}(t \wedge 1)\mathrm{d}t, \quad \mathscr{I}_1 = \frac{\lambda^{\frac{\beta}{\gamma}}2^{\beta}}{\gamma \ln 2} \int_0^\lambda t^{-\frac{\beta}{\gamma}-1}(t \wedge 1)\mathrm{d}t.$$

Thus we obtain the desired estimate. $\qquad\square$

Lemma A.2. *For any $\vartheta \in [0, \infty)$ and $\gamma_0, \gamma_1 \in (-\infty, 1)$, there is a constant $C = C(\vartheta, \gamma_0, \gamma_1) > 0$ such that for all $\lambda, t > 0$,*

$$\int_0^t \left[(\lambda(t-s))^{-\vartheta} \wedge 1\right](1 \wedge s)^{-\gamma_0}(1 \vee s)^{-\gamma_1}\mathrm{d}s \leqslant C\ell_\vartheta(\lambda t)(1 \wedge t)^{1-\gamma_0}(1 \vee t)^{1-\gamma_1},$$

where

$$\ell_\vartheta(\lambda) := (\lambda \vee 1)^{-\vartheta \wedge 1}(1 + \mathbb{1}_{\vartheta=1}\ln(1 \vee \lambda)).$$

Proof. We first show the following estimate: for any $\vartheta \in [0, \infty)$ and $\gamma < 1$,

$$\mathscr{I}(\lambda, t) := \int_0^t \left[(\lambda(t-s))^{-\vartheta} \wedge 1\right] s^{-\gamma}\mathrm{d}s \lesssim \ell_\vartheta(\lambda t)t^{1-\gamma}, \quad \forall \lambda, t > 0. \tag{A.1}$$

By the change of variable, we have

$$\mathscr{I}(\lambda, t) = t^{1-\gamma}\mathscr{I}(t\lambda, 1).$$

So, it suffices to prove (A.1) for $t = 1$. Note that

$$\mathscr{I}(\lambda, 1) \leqslant 2^{\gamma \vee 0} \int_0^{\frac{1}{2}} \left[(\lambda s)^{-\vartheta} \wedge 1\right]\mathrm{d}s + \left[(\tfrac{\lambda}{2})^{-\vartheta} \wedge 1\right] \int_{\frac{1}{2}}^1 (1-s)^{-\gamma}\mathrm{d}s$$

$$= 2^{\gamma \vee 0}\lambda^{-1} \int_0^{\frac{\lambda}{2}} \left[s^{-\vartheta} \wedge 1\right]\mathrm{d}s + \left[(\tfrac{\lambda}{2})^{-\vartheta} \wedge 1\right] \int_0^{\frac{1}{2}} s^{-\gamma}\mathrm{d}s.$$

If $0 \leqslant \vartheta \neq 1$, then

$$\mathscr{I}(\lambda, 1) \lesssim \lambda^{-1} \wedge 1 + \lambda^{-\vartheta} \wedge 1 \leqslant 2(\lambda^{-(\vartheta \wedge 1)} \wedge 1) = 2(\lambda \vee 1)^{-\vartheta \wedge 1}.$$

If $\vartheta = 1$, then

$$\mathscr{I}(\lambda, 1) \lesssim (\lambda^{-1} \wedge 1) \ln(1 \vee \lambda) + \lambda^{-1} \wedge 1 \lesssim (\lambda \vee 1)^{-1} \ln(1 \vee \lambda).$$

Hence,

$$\mathscr{I}(\lambda, 1) \lesssim \ell_\vartheta(\lambda).$$

For $t \in (0, 2]$, we have

$$\int_0^t \left[(\lambda(t-s))^{-\vartheta} \wedge 1\right] (1 \wedge s)^{-\gamma_0} (1 \vee s)^{-\gamma_1} \mathrm{d}s$$

$$\lesssim \int_0^t \left[(\lambda(t-s))^{-\vartheta} \wedge 1\right] s^{-\gamma_0} \mathrm{d}s \lesssim_C \ell_\vartheta(\lambda t) t^{1-\gamma_0}.$$

For $t > 2$, we have

$$\int_0^t \left[(\lambda(t-s))^{-\vartheta} \wedge 1\right] (1 \wedge s)^{-\gamma_0} (1 \vee s)^{-\gamma_1} \mathrm{d}s$$

$$\leqslant \int_0^1 \left[(\lambda(t-s))^{-\vartheta} \wedge 1\right] s^{-\gamma_0} \mathrm{d}s + \int_1^t \left[(\lambda(t-s))^{-\vartheta} \wedge 1\right] s^{-\gamma_1} \mathrm{d}s$$

$$\leqslant \left[(\lambda(t-1))^{-\vartheta} \wedge 1\right] \int_0^1 s^{-\gamma_0} \mathrm{d}s + \int_0^t \left[(\lambda(t-s))^{-\vartheta} \wedge 1\right] s^{-\gamma_1} \mathrm{d}s$$

$$\lesssim \left[(\lambda t)^{-\vartheta} \wedge 1\right] + \ell_\vartheta(\lambda t) t^{1-\gamma_1} \lesssim \ell_\vartheta(\lambda t)(1 \vee t)^{1-\gamma_1}.$$

The proof is complete. □

Below we recall some well-known results about singular kernels mentioned in the introduction. For the readers' convenience, we give full proofs.

Lemma A.3. *Let $\gamma \in (0, d)$ and $K : \mathbb{R}^d \to \mathbb{R}^N$ be a measurable function with that for some $C > 0$,*

$$|K(x)| \leqslant C|x|^{\gamma-d}, \quad x \in \mathbb{R}^d.$$

(i) Let $1 < q < p < \infty$ satisfy

$$\tfrac{1}{p} + \tfrac{\gamma}{d} = \tfrac{1}{q}.$$

Then there is a constant $C = C(p, q, d, \gamma) > 0$ such that for any $f \in L^q(\mathbb{R}^d)$,

$$\|K * f\|_p \leqslant C\|f\|_q. \tag{A.2}$$

(ii) Let $p \in [1, d/\gamma)$ and $\theta := p\gamma/d$. Then there is a constant $C = C(p, d, \gamma) > 0$ such that for any $f \in L^p(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$,

$$\|K * f\|_\infty \leqslant C\|f\|_p^\theta \|f\|_\infty^{1-\theta}. \tag{A.3}$$

(iii) For any $p \in (\frac{d}{d-\gamma}, \infty]$, it holds that

$$K \in \mathbf{B}_p^{d/p-d+\gamma,\infty}, \quad \nabla K \in \mathbf{B}_p^{d/p-d+\gamma-1,\infty}. \tag{A.4}$$

(iv) If K is a $(d-\gamma)$-order smooth homogenous function on $\mathbb{R}^d \setminus \{0\}$, i.e., $K(\lambda x) = \lambda^{\gamma-d} K(x)$, $\lambda > 0$, then for any $p_0 \in [1, \frac{d}{d-\gamma})$ and $p_1 \in (\frac{d}{d-\gamma}, \infty]$, it holds that

$$K \in \mathbf{B}_{p_0}^{d/p_0-d+\gamma,\infty} + \mathbf{B}_{p_1}^{d/p_0-d+\gamma,\infty}. \tag{A.5}$$

(v) Let $K(x) = \nabla|x|^{\gamma-d}$ and $K_\varepsilon(x) = \nabla(|x| \vee \varepsilon)^{\gamma-d}$ for $\varepsilon > 0$. Then for any $p \in (\frac{d}{d-\gamma}, \infty]$ and $r \in [1, \frac{d}{d-\gamma})$, there is constant $C = C(r, d, p, \gamma) > 0$ such that for all $\varepsilon > 0$,

$$\|K - K_\varepsilon\|_{\mathbf{B}_p^{\beta,\infty}} \lesssim \varepsilon^{\frac{d}{r}-d+\gamma}, \quad \beta := \tfrac{d}{p} - \tfrac{d}{r}. \tag{A.6}$$

Proof. (i) Let $r = d/(d - \gamma)$. By [1, Theorem 1.5], we have

$$\|K * f\|_p \lesssim \|K\|_{r,\infty}\|f\|_q,$$

where

$$\|K\|_{r,\infty} := \sup_{\lambda > 0} \lambda^r m\{x : |K(x)| > \lambda\} \lesssim \sup_{\lambda > 0} \lambda^r m\{x \in \mathbb{R}^d : C|x|^{\gamma - d} > \lambda\} < \infty.$$

(ii) Let $p' = p/(p-1)$. For any $\lambda > 0$, by Hölder's inequality we have

$$|K * f(x)| \leqslant \int_{|x-y| \lesssim \lambda} \frac{|f(y)|}{|x-y|^{d-\gamma}} \mathrm{d}y + \int_{|x-y| > \lambda} \frac{|f(y)|}{|x-y|^{d-\gamma}} \mathrm{d}y$$

$$\lesssim \|f\|_\infty \lambda^\gamma + \|f\|_p \left(\int_{|x-y| > \lambda} \frac{1}{|x-y|^{(d-\gamma)p'}} \mathrm{d}y \right)^{1/p'}$$

$$\lesssim \|f\|_\infty \lambda^\gamma + \|f\|_p \lambda^{d/p' - d + \gamma},$$

In particular, taking $\lambda = \|f\|_p^{p/d}\|f\|_\infty^{-p/d}$, we get

$$|K * f(x)| \lesssim \|f\|_p^\theta \|f\|_\infty^{1-\theta}.$$

(iii) For $p \in (\frac{d}{d-\gamma}, \infty)$, letting $\frac{1}{p} + \frac{\gamma}{d} = \frac{1}{q}$, by (A.2) we have for any $j \geqslant 1$,

$$\|\mathscr{R}_j K\|_p \lesssim \|\check{\phi}_j\|_q \lesssim 2^{j(d - \frac{d}{q})} \|\check{\phi}_1\|_q = 2^{j(d - \frac{d}{p} - \gamma)} \|\check{\phi}_1\|_q.$$

For $p = \infty$, by (A.3) we have

$$\|\mathscr{R}_j K\|_\infty \leqslant \|\check{\phi}_j\|_1^{\frac{\gamma}{d}} \|\check{\phi}_j\|_\infty^{1 - \frac{\gamma}{d}} = \|\check{\phi}_1\|_1^{\frac{\gamma}{d}} \|\check{\phi}_1\|_\infty^{1 - \frac{\gamma}{d}} 2^{j(d - \gamma)}.$$

Hence, $K \in \mathbf{B}_p^{d/p - d + \gamma, \infty}$. By Bernstein's inequality, we have $\nabla K \in \mathbf{B}_p^{d/p - d + \gamma - 1, \infty}$.

(iv) Now, let χ be a smooth function with $\chi = 1$ on B_1 and $\chi = 0$ on B_2^c. Define

$$K_1(x) := \chi(x)K(x), \quad K_2(x) := (1 - \chi(x))K(x).$$

Fix $p_0 \in [1, \frac{d}{d-\gamma})$ and $p_1 \in (\frac{d}{d-\gamma}, \infty]$. It is easy to see that for any $m \in \mathbb{N}_0$,

$$\nabla^m K_2 \in L^{p_1} \Rightarrow K_2 \in \mathbf{B}_{p_1}^{d/p_0 - d + \gamma, \infty}.$$

Next we show $K_1 \in \mathbf{B}_{p_0}^{d/p_0 - d + \gamma, \infty}$. Note that

$$\|\mathscr{R}_0 K_1\|_{p_0} \leqslant C\|K_1\|_{p_0} < \infty.$$

For $j \geqslant 1$, since $K_1 = K - K_2$, by [1, Proposition 2.21] we have

$$\|\mathscr{R}_j K_1\|_{p_0} \leqslant \|\mathscr{R}_j K\|_{p_0} + \|\mathscr{R}_j K_2\|_{p_0} \leqslant C 2^{j(d - \gamma - \frac{d}{p_0})} + \|\mathscr{R}_j K_2\|_{p_0},$$

and

$$\|\mathscr{R}_j K_2\|_{p_0} = \|\Delta^{-d} \mathscr{R}_j \Delta^d K_2\|_{p_0} \lesssim 2^{-2dj} \|\Delta^d K_2\|_{p_0}$$

$$\lesssim 2^{j(d - \gamma - \frac{d}{p_0})} \left(\|1_{1 \leqslant |x| \leqslant 2} G\|_{p_0} + \|1_{|x| \geqslant 1} \nabla^{2d} K\|_{p_0} \right),$$

where G is a bounded function on $\{x : 1 \leqslant |x| \leqslant 2\}$. Hence,

$$\|K_1\|_{\mathbf{B}_{p_0}^{d/p_0 - d + \gamma, \infty}} = \sup_{j \geqslant 0} 2^{j(\frac{d}{p_0} - d + \gamma)} \|\mathscr{R}_j K_1\|_{p_0} < \infty.$$

(v) For $p \in (\frac{d}{d-\gamma}, \infty)$, letting $1 + \frac{1}{p} = \frac{1}{r} + \frac{1}{q}$, by Young's inequality we have

$$\|\mathscr{R}_j(K - K_\varepsilon)\|_p \leqslant \|\nabla\check{\phi}_j\|_q \||\cdot|^{\gamma - d} - (|\cdot| \vee \varepsilon)^{\gamma - d}\|_r \lesssim 2^{j(d - \frac{d}{q})} \varepsilon^{\frac{d}{r} - d + \gamma}.$$

This completes the proof. $\square$

We also need the following simple lifting proposition.

Proposition A.4. *Suppose* $K \in \mathbf{B}^{s,q}_{\varrho}$ *for some* $s \in \mathbb{R}$ *and* $q, \varrho \in [1, \infty]$. *Let*

$$K_1(x, v) = K(x), \ K_2(x, v) = K(v), \ \boldsymbol{\varrho}_1 = (\varrho, \infty), \ \boldsymbol{\varrho}_2 = (\infty, \varrho).$$

Then we have

$$\|K_1\|_{\mathbf{B}^{(1+\alpha)s,q}_{\boldsymbol{\varrho}_1;\boldsymbol{a}}} \asymp \|K\|_{\mathbf{B}^{s,q}_{\varrho}} \asymp \|K_2\|_{\mathbf{B}^{s,q}_{\boldsymbol{\varrho}_2;\boldsymbol{a}}}.$$

Proof. We only prove $\|K_1\|_{\mathbf{B}^{(1+\alpha)s,q}_{\boldsymbol{\varrho}_1;\boldsymbol{a}}} \asymp \|K\|_{\mathbf{B}^{s,q}_{\varrho}}$. By definition we have for $j \geqslant 0$,

$$\mathscr{R}^{\boldsymbol{a}}_j K_1(x, v) = \int_{\mathbb{R}^{2d}} \check{\phi}^{\boldsymbol{a}}_j(x - x', v - v') K(x') \mathrm{d}x' \mathrm{d}v' = h_j * K(x),$$

where

$$h_j(x) := \int_{\mathbb{R}^d} \check{\phi}^{\boldsymbol{a}}_j(x, v) \mathrm{d}v.$$

Note that for $j \geqslant 1$,

$$\hat{h}_j(\xi) = 2^{j(1+\alpha)d}(2\pi)^{-d} \int_{\mathbb{R}^{2d}} \mathrm{e}^{ix \cdot \xi} \check{\phi}^{\boldsymbol{a}}_1(2^{j(1+\alpha)}x, v) \mathrm{d}x \mathrm{d}v = \phi^{\boldsymbol{a}}_1(2^{-(1+\alpha)j}\xi, 0),$$

where the support of $\phi^{\boldsymbol{a}}_1(2^{-(1+\alpha)j}\cdot, 0)$ is contained in

$$\left\{ 2^{j-1} < |\xi|^{1/(1+\alpha)} < 2^{j+1} \right\} = \left\{ 2^{(1+\alpha)(j-1)} < |\xi| < 2^{(1+\alpha)(j+1)} \right\}.$$

In particular, an equivalent norm of the Besov space $\mathbf{B}^{s,q}_{\varrho}$ is then given by

$$\|f\|_{\mathbf{B}^{s,q}_{\varrho}} \asymp \left(\sum_{j \geqslant 0} \left(2^{j(1+\alpha)s} \|h_j * f\|_{\varrho} \right)^q \right)^{1/q}.$$

Hence,

$$\|K_1\|^q_{\mathbf{B}^{s,q}_{\boldsymbol{\varrho}_1;\boldsymbol{a}}} = \sum_{j \geqslant 0} 2^{j(1+\alpha)sq} \|\mathscr{R}^{\boldsymbol{a}}_j K_1\|^q_{\boldsymbol{\varrho}_1} = \sum_{j \geqslant 0} \left(2^{j(1+\alpha)s} \|h_j * K\|_{\varrho} \right)^q \asymp \|K\|^q_{\mathbf{B}^{s,q}_{\varrho}}.$$

The proof is complete. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

APPENDIX B. BASIC PROPERTIES OF BESOV SPACES

Proof of Lemma 2.6. (i) By definition we have

$$\|f\|_{\boldsymbol{p}} \leqslant \sum_{j \geqslant 0} \|\mathscr{R}^{\boldsymbol{a}}_j f\|_{\boldsymbol{p}} = \|f\|_{\mathbf{B}^{0,1}_{\boldsymbol{p};\boldsymbol{a}}}$$

and

$$\|f\|_{\mathbf{B}^{0,\infty}_{\boldsymbol{p};\boldsymbol{a}}} = \sup_{j \geqslant 0} \|\mathscr{R}^{\boldsymbol{a}}_j f\|_{\boldsymbol{p}} \lesssim \|f\|_{\boldsymbol{p}}.$$

(ii) For $j \geqslant 0$, by (2.5) and Young's inequality, we have

$$\|\mathscr{R}^{\boldsymbol{a}}_j(f * g)\|_{\boldsymbol{p}} = \|(\mathscr{R}^{\boldsymbol{a}}_j f) * (\widetilde{\mathscr{R}}^{\boldsymbol{a}}_j g)\|_{\boldsymbol{p}} \leqslant \|\mathscr{R}^{\boldsymbol{a}}_j f\|_{\boldsymbol{p}_1} \|\widetilde{\mathscr{R}}^{\boldsymbol{a}}_j g\|_{\boldsymbol{p}_2},$$

where $\widetilde{\mathscr{R}}^{\boldsymbol{a}}_j g = \sum_{|i-j| \leqslant 2} \mathscr{R}^{\boldsymbol{a}}_i g$. Hence,

$$\begin{aligned} \|f * g\|_{\mathbf{B}^{\beta,q}_{\boldsymbol{p};\boldsymbol{a}}} &= \left(\sum_j (2^{j\beta} \|\mathscr{R}^{\boldsymbol{a}}_j(f * g)\|_{\boldsymbol{p}})^q \right)^{1/q} \\ &\leqslant \left(\sum_j (2^{j\beta_1} \|\mathscr{R}^{\boldsymbol{a}}_j f\|_{\boldsymbol{p}_1} 2^{j\beta_2} \|\widetilde{\mathscr{R}}^{\boldsymbol{a}}_j g\|_{\boldsymbol{p}_2})^q \right)^{1/q} \\ &\leqslant 5 \|f\|_{\mathbf{B}^{\beta_1,q_1}_{\boldsymbol{p}_1;\boldsymbol{a}}} \|g\|_{\mathbf{B}^{\beta_2,q_2}_{\boldsymbol{p}_2;\boldsymbol{a}}}. \end{aligned}$$

(iii) For $j \geqslant 0$, we define $S_j = \sum_{k \leqslant j-1} \mathscr{R}_j^{\boldsymbol{a}}$ with convention $\mathscr{R}_{-1}^{\boldsymbol{a}} = 0$ and $S_{-1} = 0$. Recall the following Bony decomposition:

$$fg = f \prec g + f \circ g + g \prec f,$$

where

$$f \prec g := \sum_{j \geqslant 0} S_{j-1} f \mathscr{R}_j^{\boldsymbol{a}} g \quad \text{and} \quad f \circ g := \sum_{j \geqslant 0} \mathscr{R}_j^{\boldsymbol{a}} f \widetilde{\mathscr{R}}_j^{\boldsymbol{a}} g.$$

Here $\widetilde{\mathscr{R}}_j^{\boldsymbol{a}}$ is defined in (2.5). Noting that

$$\mathscr{R}_k\big(S_{j-1} f \mathscr{R}_j^{\boldsymbol{a}} g\big) = 0, \quad \forall |k-j| > 3, \tag{B.1}$$

we have for $k \geqslant 0$,

$$\|\mathscr{R}_k^{\boldsymbol{a}}(f \prec g)\|_{\boldsymbol{p}} \leqslant \sum_{j \geqslant 0} \|\mathscr{R}_k\big(S_{j-1} f \mathscr{R}_j^{\boldsymbol{a}} g\big)\|_{\boldsymbol{p}} = \sum_{|j-k| \leqslant 3} \|\mathscr{R}_k^{\boldsymbol{a}}\big(S_{j-1} f \mathscr{R}_j^{\boldsymbol{a}} g\big)\|_{\boldsymbol{p}}$$

$$\lesssim \sum_{|j-k| \leqslant 3} \|S_{j-1} f \mathscr{R}_j^{\boldsymbol{a}} g\|_{\boldsymbol{p}} \leqslant \sum_{|j-k| \leqslant 3} \|S_{j-1} f\|_{\boldsymbol{p}_1} \|\mathscr{R}_j^{\boldsymbol{a}} g\|_{\boldsymbol{p}_2} \lesssim \|f\|_{\boldsymbol{p}_1} \sum_{|j-k| \leqslant 3} \|\mathscr{R}_j^{\boldsymbol{a}} g\|_{\boldsymbol{p}_2}.$$

Hence,

$$\|f \prec g\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{s,q}} \lesssim \|f\|_{\boldsymbol{p}_1} \left(\sum_{k \geqslant -1} 2^{skq} \sum_{|j-k| \leqslant 3} \|\mathscr{R}_j^{\boldsymbol{a}} g\|_{\boldsymbol{p}_2}^q \right)^{1/q} \lesssim \|f\|_{\boldsymbol{p}_1} \|g\|_{\mathbf{B}_{\boldsymbol{p}_2;\boldsymbol{a}}^{s,q}}.$$

By symmetry, we also have

$$\|f \prec g\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{s}} \lesssim \|f\|_{\mathbf{B}_{\boldsymbol{p}_1;\boldsymbol{a}}^{s,q}} \|g\|_{\boldsymbol{p}_2}.$$

Moreover, noting that

$$\mathscr{R}_k^{\boldsymbol{a}}(\mathscr{R}_j^{\boldsymbol{a}} f \widetilde{\mathscr{R}}_j^{\boldsymbol{a}} g) = 0, \quad k > j+5,$$

for $k > 5$ we have

$$\|\mathscr{R}_k^{\boldsymbol{a}}(f \circ g)\|_{\boldsymbol{p}} \leqslant \sum_{j \geqslant k-5} \|\mathscr{R}_k^{\boldsymbol{a}}(\mathscr{R}_j^{\boldsymbol{a}} f \widetilde{\mathscr{R}}_j^{\boldsymbol{a}} g)\|_{\boldsymbol{p}} \lesssim \sum_{j \geqslant k-5} \|\mathscr{R}_j^{\boldsymbol{a}} f \widetilde{\mathscr{R}}_j^{\boldsymbol{a}} g\|_{\boldsymbol{p}}$$

$$\leqslant \sum_{j \geqslant k-5} \|\mathscr{R}_j^{\boldsymbol{a}} f\|_{\boldsymbol{p}_1} \|\widetilde{\mathscr{R}}_j^{\boldsymbol{a}} g\|_{\boldsymbol{p}_2} \lesssim \sum_{j \geqslant k-5} \|\mathscr{R}_j^{\boldsymbol{a}} f\|_{\boldsymbol{p}_1} \|g\|_{\boldsymbol{p}_2}.$$

Since $s > 0$, by Young's inequality for discrete summation, we get

$$\|f \circ g\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{s,q}} \lesssim \left(\sum_{k \geqslant 0} 2^{skq} \left(\sum_{j \geqslant k-5} \|\mathscr{R}_j^{\boldsymbol{a}} f\|_{\boldsymbol{p}_1} \right)^q \right)^{1/q} \|g\|_{\boldsymbol{p}_2} \lesssim \|f\|_{\mathbf{B}_{\boldsymbol{p}_1;\boldsymbol{a}}^{s,q}} \|g\|_{\boldsymbol{p}_2}.$$

Combining the above estimates, we complete the proof of (iii).

(iv) For $\theta = 0, 1$, (2.16) is direct by (2.7). Below we assume $\theta \in (0,1)$. Let $\boldsymbol{p}_2 \leqslant \boldsymbol{w} \in [1, \infty]^N$ and $s_2 \in \mathbb{R}$ be defined by

$$\frac{1}{\boldsymbol{p}} = \frac{1-\theta}{\boldsymbol{p}_1} + \frac{\theta}{\boldsymbol{w}}, \quad s_1(1-\theta) + s_3\theta = s. \tag{B.2}$$

By Hölder's inequality, we have

$$\|\mathscr{R}_j^{\boldsymbol{a}} f\|_{\boldsymbol{p}} = \|(\mathscr{R}_j^{\boldsymbol{a}} f)^{1-\theta} (\mathscr{R}_j^{\boldsymbol{a}} f)^\theta\|_{\boldsymbol{p}} \leqslant \|\mathscr{R}_j^{\boldsymbol{a}} f\|_{\boldsymbol{p}_1}^{1-\theta} \|\mathscr{R}_j^{\boldsymbol{a}} f\|_{\boldsymbol{w}}^\theta.$$

Hence,

$$2^{js} \|\mathscr{R}_j^{\boldsymbol{a}} f\|_{\boldsymbol{p}} \leqslant \left(2^{js_1} \|\mathscr{R}_j^{\boldsymbol{a}} f\|_{\boldsymbol{p}_1} \right)^{1-\theta} \left(2^{js_3} \|\mathscr{R}_j^{\boldsymbol{a}} f\|_{\boldsymbol{w}} \right)^\theta. \tag{B.3}$$

Note that by (2.15) and (B.2),

$$s_2 = s_3 + \boldsymbol{a} \cdot \left(\frac{d}{\boldsymbol{p}_2} - \frac{d}{\boldsymbol{w}} \right),$$

and by (2.7),

$$\|\mathscr{R}_j^{\boldsymbol{a}} f\|_{\boldsymbol{w}} \lesssim 2^{j\boldsymbol{a}\cdot(\frac{d}{\boldsymbol{p}_2}-\frac{d}{\boldsymbol{w}})}\|\mathscr{R}_j^{\boldsymbol{a}} f\|_{\boldsymbol{p}_2}, \ j \geqslant 0.$$

Substituting it into (B.3), we obtain

$$2^{js}\|\mathscr{R}_j^{\boldsymbol{a}} f\|_{\boldsymbol{p}} \lesssim \left(2^{js_1}\|\mathscr{R}_j^{\boldsymbol{a}} f\|_{\boldsymbol{q}}\right)^{1-\theta}\left(2^{js_2}\|\mathscr{R}_j^{\boldsymbol{a}} f\|_{\boldsymbol{q}}\right)^{\theta}.$$

Thus we obtain (iv) by Hölder's inequality with respect to j.

(v) Note that for $\xi = (\xi_1, \xi_2)$,

$$\widehat{\mathscr{R}_k^x \mathscr{R}_j^{\boldsymbol{a}}} f(\xi) = \phi_k(\xi_1)\phi_j^{\boldsymbol{a}}(\xi)\hat{f}(\xi).$$

It is easy to see that

$$\phi_k(\xi_1)\phi_j^{\boldsymbol{a}}(\xi) = 0, \quad k > (1+\alpha)(j+2).$$

Hence,

$$\sup_{k,j\geqslant 1} 2^{\frac{ks_0}{1+\alpha}}2^{js_1}\|\mathscr{R}_k^x\mathscr{R}_j^{\boldsymbol{a}} f\|_{\boldsymbol{p}} = \sup_{k\leqslant(1+\alpha)(j+2)} 2^{\frac{ks_0}{1+\alpha}}2^{js_1}\|\mathscr{R}_k^x\mathscr{R}_j^{\boldsymbol{a}} f\|_{\boldsymbol{p}}$$

$$\leqslant \sup_{k\leqslant(1+\alpha)(j+2)} 2^{\frac{k(s_0-\varepsilon)}{1+\alpha}}2^{j(s_1+\varepsilon)}\|\mathscr{R}_k^x\mathscr{R}_j^{\boldsymbol{a}} f\|_{\boldsymbol{p}}.$$

(vi) By definition and Bernstein's inequality we have

$$\|f\|_{\mathbf{B}_{\boldsymbol{p};x,\boldsymbol{a}}^{\theta\gamma,(1-\theta)\beta}} = \sup_{k,j\geqslant 0}\left(2^{\frac{k\gamma}{1+\alpha}}\|\mathscr{R}_k^x\mathscr{R}_j^{\boldsymbol{a}} f\|_{\boldsymbol{p}}\right)^{\theta}\left(2^{j\beta}\|\mathscr{R}_k^x\mathscr{R}_j^{\boldsymbol{a}} f\|_{\boldsymbol{p}}\right)^{1-\theta}$$

$$\lesssim \sup_{k,j\geqslant 0}\left(2^{\frac{k\gamma}{1+\alpha}}\|\mathscr{R}_k^x f\|_{\boldsymbol{p}}\right)^{\theta}\left(2^{j\beta}\|\mathscr{R}_j^{\boldsymbol{a}} f\|_{\boldsymbol{p}}\right)^{1-\theta} = \|f\|_{\mathbf{B}_{\boldsymbol{p};x}^{\gamma/(1+\alpha)}}^{\theta}\|f\|_{\mathbf{B}_{\boldsymbol{p};\boldsymbol{a}}^{\beta}}^{1-\theta}.$$

The proof is complete. $\qquad\square$

APPENDIX C. A PRIORI L^1-ESTIMATES FOR KINETIC EQUATION

In this appendix we use the probabilistic representation to show some a priori estimates for linear kinetic Fokker-Planck equation with smooth coefficients, which has independent interest.

Lemma C.1. *Let* $T > 0$ *and* $\alpha \in (0, 2]$. *Let* $H \in L^1([0,T]; C_b^1(\mathbb{R}^{2d}))$ *and* $u \in C([0,T]; C_b^2(\mathbb{R}^{2d}))$ *be a smooth solution of the linear Fokker-Planck equation*

$$\partial_t u = \Delta_v^{\alpha/2} u - v \cdot \nabla_x u - \operatorname{div}_v(Hu).$$

For any $t \in [0,T]$, *it holds that*

$$\|u(t)\|_{\mathbb{L}^\infty} \leqslant \|u(0)\|_{\mathbb{L}^\infty} e^{\int_0^t \|(\operatorname{div}_v H_s)^-\|_{\mathbb{L}^\infty} ds} \tag{C.1}$$

and

$$\|u(t)\|_{\mathbb{L}^1} \leqslant \|u(0)\|_{\mathbb{L}^1}. \tag{C.2}$$

Proof. (i) Fix $t_0 \in (0, T)$. Let $\mathbf{H}_t(z) := \mathbf{H}_t(x,v) := (-v, -H_t(x,v))$ and $\sigma := (0, \mathbb{I})^*$. Let $Z_t(z) = Z_t$ solve the following SDE

$$dZ_t = \mathbf{H}_{t_0-t}(Z_t)dt + \sigma dL_t^{(\alpha)}, \quad Z_0 = z, \quad t \in [0, t_0].$$

Since we are considering the additive noise and $H \in L^1([0,T]; C_b^1(\mathbb{R}^{2d}))$, it is easy to see that

$$z \mapsto Z_t(z) \text{ is a } C^1\text{-diffeomorphism},$$

and

$$\nabla Z_t(z) = \mathbb{I}_{2d\times 2d} + \int_0^t \nabla \mathbf{H}_{t_0-s}(Z_s(z))\nabla Z_s(z)ds.$$

Let G_t be the inverse of matrix $\nabla Z_t(z)$, which solves the following ODE

$$G_t(z) = \mathbb{I}_{2d\times 2d} - \int_0^t G_s(z) \cdot \nabla \mathbf{H}_{t_0-s}(Z_s(z))ds. \tag{C.3}$$

Note that
$$\partial_t u = \Delta_v^{\alpha/2} u - v \cdot \nabla_x u - H \cdot \nabla_v u - \operatorname{div}_v Hu.$$

By applying Itô's formula to the function $(s, z) \mapsto u(t_0 - s, z)$, we have

$$u(t_0 - t, Z_t(z)) = u(t_0, z) - \int_0^t (\operatorname{div}_v Hu)(t_0 - s, Z_s(z))\mathrm{d}s + \text{a martingale}.$$

Using product Itô's formula we get

$$\mathbb{E}\left(u(t_0 - t, Z_t(z))\mathrm{e}^{-\int_0^t \operatorname{div}_v H_{t_0-s}(Z_s(z))\mathrm{d}s}\right) = u(t_0, z).$$

Hence,

$$|u(t_0, z)| \leqslant \|u(0)\|_{\mathbb{L}^\infty} \mathrm{e}^{\int_0^{t_0} \|(\operatorname{div}_v H_{t_0-s})^-\|_{\mathbb{L}^\infty}\mathrm{d}s}.$$

Replacing t_0 by t, we obtain (C.1). Moreover, by the change of variables, we also have

$$
\begin{aligned}
\|u(t_0)\|_{\mathbb{L}^1} &\leqslant \mathrm{e}^{\int_0^{t_0} \|(\operatorname{div}_v H_{t_0-s})^-\|_{\mathbb{L}^\infty}\mathrm{d}s} \mathbb{E}\left(\int_{\mathbb{R}^{2d}} |u(0, Z_{t_0}(z))|\mathrm{d}z\right) \\
&\leqslant \mathrm{e}^{\int_0^{t_0} \|(\operatorname{div}_v H_{t_0-s})^-\|_{\mathbb{L}^\infty}\mathrm{d}s} \mathbb{E}\left(\int_{\mathbb{R}^{2d}} |u(0, z)| \det(\nabla Z_{t_0}^{-1}(z))\mathrm{d}z\right) \\
&\leqslant \mathrm{e}^{\int_0^{t_0} \|(\operatorname{div}_v H_{t_0-s})^-\|_{\mathbb{L}^\infty}\mathrm{d}s} \sup_z \mathbb{E}(\det G_{t_0}(z)) \|u(0)\|_{\mathbb{L}^1} \\
&\leqslant \mathrm{e}^{\int_0^{t_0}(\|(\operatorname{div}_v H_{t_0-s})^-\|_{\mathbb{L}^\infty} + \|(\operatorname{div}_v H_{t_0-s})^+\|_{\mathbb{L}^\infty})\mathrm{d}s} \|u(0)\|_{\mathbb{L}^1} \\
&= \mathrm{e}^{\int_0^{t_0} \|\operatorname{div}_v H_s\|_{\mathbb{L}^\infty}\mathrm{d}s} \|u(0)\|_{\mathbb{L}^1}, \hspace{4cm} \text{(C.4)}
\end{aligned}
$$

where we have used that $\nabla Z_{t_0}^{-1}(z) = G_{t_0} \circ Z_{t_0}^{-1}(z)$ in the third inequality, and the last inequality is due to (cf. [60])

$$\det G_t(z) = 1 - \int_0^t \det G_s(z) \cdot \operatorname{div}_v H_{t_0-s}(Z_s(z))\mathrm{d}s.$$

(ii) Let $\psi : \mathbb{R} \to \mathbb{R}$ be a smooth convex function with

$$|\psi(r)| + |\psi'(r)| \leqslant Cr, \quad |\psi''(r)| \leqslant C. \hspace{3cm} \text{(C.5)}$$

Let $\chi \in C_c^\infty(\mathbb{R}^d)$ be a cutoff function in $\mathbb{R}^d$ with $\chi(v) = 1$ for $|v| \leqslant 1$ and $\chi(v) = 0$ for $|v| > 2$. For $R > 0$, define

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

By the chain rule we have

$$\partial_t \psi(u)\chi_R = \psi'(u)\Delta_v^{\alpha/2} u \chi_R - v \cdot \nabla_x \psi(u)\chi_R - \operatorname{div}_v(Hu)\psi'(u)\chi_R.$$

Integrating both sides over $[0, t] \times \mathbb{R}^{2d}$, we get

$$\int \psi(u(t))\chi_R = \int \psi(u(0))\chi_R + \int_0^t \int \Delta_v^{\alpha/2} u \psi'(u)\chi_R - \int_0^t \int \operatorname{div}_v(Hu)\psi'(u)\chi_R,$$

where we have used that

$$\int v \cdot \nabla_x \psi(u)\chi_R = \int \operatorname{div}_x((v\chi_R)\psi(u)) = 0.$$

By (C.4) and (C.5), letting $R \to \infty$ and by the dominated convergence theorem, we get

$$\int \psi(u(t)) = \int \psi(u(0)) + \int_0^t \int \Delta_v^{\alpha/2} u \psi'(u) - \int_0^t \int \operatorname{div}_v(Hu)\psi'(u).$$

We consider the second term in the right hand side. For $\alpha = 2$, we directly have

$$\int \Delta_v u \psi'(u) = -\int |\nabla_v u|^2 \psi''(u) \leqslant 0.$$

For $\alpha \in (0,2)$, by (3.3) and symmetry we have

$$\int_{\mathbb{R}^d} \Delta_v^{\alpha/2} u(v) \psi'(u(v)) \mathrm{d}v = \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^d} \int_{|w| \geqslant \varepsilon} \psi'(u(v)) \frac{u(v+w) - u(v)}{|w|^{d+\alpha}} \mathrm{d}w \mathrm{d}v$$

$$= \lim_{\varepsilon \downarrow 0} \frac{1}{2} \int_{\mathbb{R}^d} \int_{|w| \geqslant \varepsilon} (\psi'(u(v)) - \psi'(u(v+w))) \frac{u(v+w) - u(v)}{|w|^{d+\alpha}} \mathrm{d}w \mathrm{d}v$$

$$= -\lim_{\varepsilon \downarrow 0} \frac{1}{2} \int_{\mathbb{R}^d} \int_{|w| \geqslant \varepsilon} \int_0^1 \psi''(su(v) + (1-s)u(v+w)) \mathrm{d}s \frac{(u(v+w) - u(v))^2}{|w|^{d+\alpha}} \mathrm{d}w \mathrm{d}v \leqslant 0.$$

Hence,

$$\int \psi(u(t)) \leqslant \int \psi(u(0)) + \int_0^t \int (Hu \cdot \nabla_v u) \psi''(u). \tag{C.6}$$

Now we take

$$\psi(r) = \psi_\delta(r) := \sqrt{r^2 + \delta} - \sqrt{\delta}, \quad \delta > 0.$$

Clearly,

$$\lim_{\delta \downarrow 0} \psi_\delta(r) = |r|, \ |\psi_\delta(r)| \leqslant |r|, \ \psi'_\delta(r) = \frac{r}{\sqrt{r^2 + \delta}}, \ \psi''_\delta(r) = \frac{\delta}{(r^2 + \delta)^{3/2}}.$$

By (C.6), we have

$$\int \psi_\delta(u(t)) \leqslant \int \psi_\delta(u(0)) + \int_0^t \int (Hu \cdot \nabla_v u) \psi''_\delta(u).$$

Noting that

$$|\psi''_\delta(u)| \cdot |H| \cdot |u| \cdot |\nabla_v u| \leqslant |H| \cdot |\nabla_v u|,$$

by taking limits $\delta \to 0$ and the dominated convergence theorem, we get

$$\int |u(t)| \leqslant \int |u(0)| + \lim_{\delta \to 0} \int_0^t \int |\psi''_\delta(u)| \cdot |H| \cdot |u| \cdot |\nabla_v u| = \int |u(0)|.$$

The proof is complete. $\square$

REFERENCES

[1] Bahouri H., Chemin J.-Y. and Danchin R. (2011). Fourier analysis and nonlinear partial differential equations. Grundlehren der Mathematischen Wissenschaften, vol. 343, Springer, Heidelberg.

[2] Barbu V. and Röckner M. (2018). Probabilistic representation for solutions to nonlinear Fokker-Planck equations. *SIAM J.Math.Anal.* **50** 4246-4260.

[3] Barbu V. and Röckner M. (2020). From nonlinear Fokker-Planck equations to solutions of distribution dependent SDE. *Ann. Probab.* **48** 1902-1920.

[4] Barbu V. and Röckner M. (2021). Solutions for nonlinear Fokker-Planck equations with measures as initial data and McKean-Vlasov equations. *J. Functional Anal.* **280** 1-35.

[5] Barbu V. and Röckner M. (2023). The evolution to equilibrium of solutions to nonlinear Fokker-Planck equations. Available in arXiv: 1904.08291. *Indiana University Journal* (to appear).

[6] Barbu V. and Röckner M. (2022). Nonlinear Fokker-Planck equations with fractional laplacian and McKean-Vlasov SDEs with Lévy noise. Available in arXiv: 2210.05612.

[7] Benedek A. and Panzone R. (1961). The spaces L^p with mixed norm. *Duke Math. J.* **28** 301-324.

[8] Bergh J. and Löfström J. (1976). Interpolation spaces: an introduction, Springer-Verlag, Berlin Heidelberg New York.

[9] Biler P., Imbert C. and Karch G. (2015). The nonlocal porous medium equation: Barenblatt profiles and other weak solutions. *Arch. Rational Mech. Anal.* **215** 497-529.

[10] Bouchut F. (1993). Existence and uniqueness of a global smooth solution for the Vlasov-Poisson-Fokker-Planck system in three dimensions. *J. Func. Anal.* **111** 239-258.

[11] Bouchut F. (2002). Hypoelliptic regularity in kinetic equations. *J. Math. Pures Appl.* **81** 1135-1159.

[12] Caffarelli L. A. and Vasseur A. (2010). Drift diffusion equations with fractional diffusion and the quasi-geostrophic equation. *Ann. of Math.* **171** 1903-1930.

[13] Cannone M. and Planchon F. (1996). Self-similar solutions for Navier-Stokes equationsin $\mathbb{R}^3$. *Commun. Partial Differ. Equations* **21** 179-193.

[14] Carrillo J. and Soler J. (1997). On the Vlasov-Poisson-Fokker-Planck equations with measures in Morrey spaces as initial data. *J. Math. Anal. Appl.* **207** 475-495.

[15] Cercignani C. (1969). Mathematical methods in kinetic theory. *Plenum*, New York.

[16] Cercignani C. (1988). The Boltzmann equation and its applications. *Applied Mathematical Sciences*, **67**. Springer, New York.

[17] Chae D. and Lee J. (2003). Global well-posedness in the super-critical dissipative quasi-geostrophic equations. *Comm. Math. Phys.* **233** 297-311.

[18] Chaintron L.P. and Diez A. (2021). Propagation of chaos: a review of models, methods and applications. Available in arXiv:2106.14812.

[19] Chapman J., Rubinstein J. and Schatzman M. (1996). A mean-field model for superconducting vortices. *Eur. J. Appl. Math.* **7** 97-111.

[20] 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.

[21] Chaudru de Raynal P.-E. and Menozzi, S. (2022). Regularization effects of a noise propagating through a chain of differential equations: an almost sharp result. *Trans. Amer. Math. Soc.* **375**, no. 1, 1-45.

[22] Chaudru de Raynal P.E., Jabir J.-F. and Menozzi S. (2022). Multidimensional Stable driven McKean-Vlasov SDEs with distributional interaction kernel – a regularization by noise perspective. Available in arXiv:2205.11866.

[23] Chaudru de Raynal P.E., Jabir J.-F. and Menozzi S. (2023). McKean-Vlasov SDEs with distributional interaction kernel and stable additive: Critical thresholds and related models. Available in arXiv:2302.09900.

[24] Chaudru de Raynal P.E., Menozzi S., Pesce A. and Zhang X. (2023). Heat kernel and gradient estimates for kinetic SDEs with low regularity coefficients. *Bull. Sci. Math.*, 103229.

[25] Chen L., Holzinger A., Jüngel A. and Zamponi N. (2022) Analysis and mean-field derivation of a porous-medium equation with fractional diffusion. *Comm. Partial Differential Equations* 1-53.

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

[27] Chen Z.-Q. and Zhang X. (2018). L^p-maximal hypoelliptic regularity of nonlocal kinetic Fokker-Planck operators. *J. Math. Pures Appl.* **116** 52-87.

[28] Chen Z.-Q., Zhang X. and Zhao G. (2021). Supercritical SDEs driven by multiplicative stable-like Lévy processes. *Trans. Amer. Math. Soc.* **374** 7621-7655.

[29] Constantin P., Majda A. and Tabak E. (1994). Formation of strong fronts in the 2-D quasi-geostrophic thermal active scalar. *Nonlinearity* **7** 1495-1533.

[30] Constantin P. and Wu J. (1999). Behavior of solutions to 2D quasi-geostrophic equations. *SIAM J. Math. Anal.* **30** 937-948.

[31] Cottet G.-H. (1986). Equations de Navier-Stokes dans le plan avec tourbillon initial mesure. *C. R. Acad. Sci. Paris Serʹ. I Math.* **303** 105-108.

[32] Delarue F. and Menozzi S. (2010). Density estimates for a random noise propagating through a chain of differential equations. *J. Func. Anal.* **259** 1577-1630.

[33] Dong H. and Yastrzhembskiy T. (2022). Global L^p estimates for kinetic Kolmogorov-Fokker-Planck equations in nondivergence form. *Arch. Ration. Mech. Anal.* **245**, no. 1, 501-564.

[34] E W. (1994). Dynamics of vortex liquids in Ginzburg-Landau theories with applications in superconductivity. *Phys. Rev. B* **50** 1126-1135.

[35] Ethier N. and Kurtz G. (1986). Markov Processes: Characterization and Convergence. Wiley series in probability and mathematical statistics.

[36] Gallagher I. and Gallay T. (2005). Uniqueness for the two-dimensional Navier-Stokes equation with a measure as initial vorticity. *Math. Ann.* **332** 287-327.

[37] Gallay T. and Wayne C.E. (2005). Global stability of vortex solutions of the two-dimensional Navier-Stokes equation. *Comm. Math. Phys.* **255** 97-129.

[38] Giga Y., Miyakawa T. and Osada H. (1988). Two-dimensional Navier-Stokes flow with measuresas initial vorticity. *Arch. Rational Mech. Anal.* **104** 223-250.

[39] Giga Y. and Miyakawa T. (1989). Navier-Stokes flow in $\mathbb{R}^3$ with measures as initial vorticity and Morrey spaces. *Comm. Partial Differential Equations* **14** 577-618.

[40] Golse F. (2016). On the dynamics of large particle systems in the mean field limit. Macroscopic and large scale phenomena: coarse graining, mean field limits and ergodicity, 1-144, Lect. Notes Appl. Math. Mech., 3, Springer.

[41] Hammersley W.R.P., Šiška D. and Szpruch L. (2021). Weak existence and uniqueness for McKan-Vlasov SDEs with common noise. *Ann. Probab.* **49** 527–555.

[42] Han Y. (2022). Smoothness of the density for McKean-Vlasov SDEs with measurable kernel. Available in arXiv:2208.02771.

[43] Hao Z., Röckner M. and Zhang X. (2022). Strong convergence of propagation of chaos for McKean-Vlasov SDEs with singular interactions. Available in arXiv:2204.07952.

[44] Hao Z., Wu M. and Zhang X. (2020). Schauder's estimate for nonlocal kinetic equations and applications. *J. Math. Pures Appl.* **140** 139-184.

[45] Hao Z., Zhang X., Zhu R. and Zhu X. (2021). Singular kinetic equations. Available in arXiv:2108.05042.

[46] Held I., Pierrehumbert R., Garner S. and Swanson K. (1995). Surface quasi-geostrophic dynamics. *J. Fluid Mech.* **282** 1-20.

[47] Huang H., Liu J.-G. and Pickl P. (2020). On the mean-field limit for the Vlasov-Poisson-Fokker-Planck System. *J. Stat. Phys.* **181** 1915-1965.

[48] Huang L., Menozzi S and Priola E. (2019). L^p estimates for degenerate non-local Kolmogorov operators. *J. Math. Pures Appl.* **121** 162-215.

[49] Imbert C. and Silvestre L. (2021). The Schauder estimate for kinetic integral equations. *Analysis and PDE* **14** 717-204.

[50] Jabin P.-E. (2014). A review of the mean field limits for Vlasov equations. *Kinet. Relat. Models* **7** 661-711.

[51] 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.

[52] 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.

[53] Karatzas I. and Shreve S. E. (1988). Brownian Motion and Stochastic Calculus. New York: Springer.

[54] Kato T. (1984). Strong L^p-solutions of the Navier-Stokes equation in Rm, with applications to weak solutions. *Math. Z.* **187** 471-480.

[55] Kato T. (1994). The Navier-Stokes equation for an incompressible fluid in $\mathbb{R}^2$ with a measure as the initial vorticity. *Differential Integral Equations* **7** 949-966.

[56] Koch H. and Tataru D. (2001). Well-posedness for the Navier-Stokes equations. *Adv. Math.* **157** 22-35.

[57] Lacker D. (2021). Hierarchies, entropy, and quantitative propagation of chaos for mean field diffusions. Available in arXiv:2105.02983.

[58] Ling C. and Xie L. (2021). Strong solutions of stochastic differential equations with coefficients in mixed-norm spaces. *Potential Anal.* 1-15.

[59] Lions P.-L. and Perthame B. (1991). Propagation of moments and regularity for the 3-dimensional Vlasov-Poisson system. *Invent. Math.* **105** 415-430.

[60] Majda A.J. and Bertozzi A.L. (2002) *Vorticity and Incompressible Flow.* Cambridge University Press.

[61] 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.

[62] Mishura Y. S. and Veretennikov A. Yu. (2020). Existence and uniqueness theorems for solutions of McKean-Vlasov stochastic equations. *Theory Probab. Math. Statist.* **103** 59-101.

[63] Mouhot C. and Villani C. (2011). On Landau damping. *Acta Math.* **207** 29-201.

[64] Ono K. and Strauss W. A. (2000). Regular solutions of the Vlasov-Poisson-Fokker-Planck system. *Discrete Contin. Dyn. Syst.* **6** 751-772.

[65] Osada H. (1987). Diffusion processes with generators of generalized divergence form. *J. Math. Kyoto Univ.* **27** 597-619.

[66] Pfaffelmoser K. (1992). Global classical solutions of the Vlasov-Poisson system in three dimensions for general initial data. *J. Differential Equations* **95** 281-303.

[67] 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.

[68] Röckner M. and Zhang X. (2021). Well-posedness of distribution dependent SDEs with singular drifts. *Bernoulli* **27** 1131-1158.

[69] Sato K. (1999). Lévy Processes and Infinitely Divisible Processes. Cambridge University Press.

[70] Schaeffer J. (1991). Global existence of smooth solutions to the Vlasov-Poisson system in three dimensions. *Comm. Partial Differential Equations* **16** 1313-1335.

[71] Serfaty S. and Vázquez J. L. (2014). A mean field equation as limit of nonlinear diffusions with fractional Laplacian operators. *Calc. Var. Partial Differential Equations* **49** 1091-1120.

[72] Serfaty S. (2020). Mean field limit for coulomb-type flows. *Duke Math. J.* **169** 2887-2935

[73] Stein E. M. (1970). Singular integrals and differentiability property of functions. Princeton Mathematical Series, No. 30 Princeton University Press, Princeton, N.J.

[74] 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.

[75] Triebel H. (2006). Theory of function spaces, III. Basel, Birkhäuser.

[76] Vazquez J. L. (2017). The mathematical theories of diffusion. Nonlinear and fractional diffusion. In: Bonforte, M., Grillo, G., eds. Nonlocal and Nonlinear Diffusions and Interactions: New Methods and Directions. Lect. Notes Math. 2186. Cham: Springer, pp. 205-278.

[77] Victory H. D. and O'Dwyer B. P. (1990). On classical solutions of Vlasov-Poisson Fokker-Planck systems. *Indiana University mathematics journal* **39** 105-156.

[78] Wang F.-Y. (2018). Distribution dependent SDEs for Landau type equations. *Stoch. Proc. Appl.* **128** 595-621.

[79] 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.

[80] Wu J. (2005). Global solutions of the 2D dissipative quasi-geostrophic equation in Besov spaces. *SIAM J. Math. Anal.* **36** 1014-1030.

[81] Wu J. (2006). Lower bounds for an integral involving fractional Laplacians and the generalized Navier-Stokes equations in Besov spaces. *Comm. Math. Phys.* **263** (2006), no. 3, 803-831.

[82] Xie L. and Zhang. (2020). Ergodicity of stochastic differential equations with jumps and singular coefficients. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** 175-229.

[83] Zhang X. (2018). Stochastic Hamiltonian flows with singular coefficients. *Science China Mathematics* **61** 1353-1384.

[84] Zhang X. (2021). Second order McKean-Vlasov SDEs and kinetic Fokker-Planck-Kolmogorov equations. Available in arXiv:2109.01273.

[85] Zhang X. and Zhang X. (2021). Cauchy problem of stochastic kinetic equations. Available in arXiv:2103.02267.

[86] Zheng Y.X. and Majda, A. (1994). Existence of global weak solutions to one-component Vlasov-Poisson and Fokker-Planck-Poisson systems in one space dimension with measures as initial data. (English summary) *Comm. Pure Appl. Math.* **47** 1365-1401.

Zimo Hao: Fakultät für Mathematik, Universität Bielefeld, 33615, Bielefeld, Germany, and School of Mathematics and Statistics, Wuhan University, Wuhan, Hubei 430072, P.R.China, Email: zhao@math.uni-bielefeld.de

Michael Röckner: Fakultät für Mathematik, Universität Bielefeld, 33615, Bielefeld, Germany, and Academy of Mathematics and Systems Science, CAS, Beijing, China, Email: roeckner@math.uni-bielefeld.de

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