Probability Theory and Related Fields (2020) 178:699–733
https://doi.org/10.1007/s00440-020-00985-8

Superposition principle for non-local Fokker–Planck–Kolmogorov operators

Michael Röckner[1] · **Longjie Xie**[2] · **Xicheng Zhang**[3]

Received: 24 October 2019 / Revised: 11 May 2020 / Published online: 13 July 2020

Abstract
We prove the superposition principle for probability measure-valued solutions to non-local Fokker–Planck–Kolmogorov equations, which in turn yields the equivalence between martingale problems for stochastic differential equations with jumps and such non-local partial differential equations with rough coefficients. As an application, we obtain a probabilistic representation for weak solutions of fractional porous media equations.

Keywords Non-local Fokker–Planck–Kolmogorov equation · Superposition principle · Martingale problem · Fractional porous media equation

Mathematics Subject Classification 60H10 · 60J75 · 40K05

Contents

This work is supported by NNSF of China (Nos. 11731009, 11931004), NSF of Jiangsu (BK20170226) and the DFG through the CRC 1283 "Taming uncertainty and profiting from randomness and low regularity in analysis, stochastics and their applications".

✉ Xicheng Zhang
XichengZhang@gmail.com

Michael Röckner
roeckner@math.uni-bielefeld.de

Longjie Xie
longjiexie@jsnu.edu.cn

1 Fakultät für Mathematik, Universität Bielefeld, 33615 Bielefeld, Germany

2 School of Mathematics and Statistics, Jiangsu Normal University, Xuzhou 221000, Jiangsu, People's Republic of China

3 School of Mathematics and Statistics, Wuhan University, Wuhan 430072, Hubei, People's Republic of China

1 Introduction

1.1 Background

Let $\mathcal{P}(\mathbb{R}^d)$ be the space of all probability measures on $\mathbb{R}^d$ endowed with the weak convergence topology. Let $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ be a measurable vector field. In [2], Ambrosio studied the connection between the continuity equation

$$\partial_t \mu_t = \mathrm{div}(b\mu_t), \tag{1.1}$$

and the ordinary differential equation (ODE for short)

$$\mathrm{d}\omega_t = b_t(\omega_t)\mathrm{d}t. \tag{1.2}$$

The following superposition principle was proved therein: Suppose that $t \mapsto \mu_t \in \mathcal{P}(\mathbb{R}^d)$ is a solution of (1.1) and satisfies

$$\int_0^T \int_{\mathbb{R}^d} \frac{|b_t(x)|}{1+|x|} \mu_t(\mathrm{d}x)\mathrm{d}t < \infty, \quad \forall T > 0,$$

then there exists a probability measure η on the space $\mathbb{C}$ of continuous functions from $\mathbb{R}_+$ to $\mathbb{R}^d$, which is concentrated on the set of all ω such that ω is an absolutely continuous solution of (1.2), and for every function $f \in C_b(\mathbb{R}^d)$ and all $t \geq 0$,

$$\int_{\mathbb{R}^d} f(x)\mu_t(\mathrm{d}x) = \int_{\mathbb{C}} f\big(\omega_t\big)\eta(\mathrm{d}\omega).$$

In other words, the measure μ_t coincides with the image of η under the evaluation map $\omega \mapsto \omega_t$. Consequently, the well-posedness of ODE (1.2) is equivalent to the existence and uniqueness of solutions for the continuity Eq. (1.1). In particular, the well-posedness of ODE (1.2) with BV drift whose distributional divergence belongs to L^∞ was obtained in a generalized sense. See also [3–5,29] and the references therein for further developments.

The stochastic counterpart of the above superposition principle was established by Figalli [15]. In this situation, the continuity equation becomes the Fokker–Planck–

Kolmogorov equation, while the ODE becomes a stochastic differential equation (SDE for short). More precisely, let X_t solve the following SDE in $\mathbb{R}^d$:

$$\mathrm{d}X_t = b_t(X_t)\mathrm{d}t + \sigma_t(X_t)\mathrm{d}W_t, \tag{1.3}$$

where $b : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ are measurable functions, W_t is a standard Brownian motion defined on some probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Let $\mu_t \in \mathcal{P}(\mathbb{R}^d)$ be the marginal law of X_t. By Itô's formula, μ_t solves the following Fokker–Planck–Kolmogorov equation in the distributional sense

$$\partial_t \mu_t = \big(\mathscr{A}_t + \mathscr{B}_t\big)^* \mu_t, \tag{1.4}$$

where for $f \in C_b^2(\mathbb{R}^d)$,

$$\mathscr{A}_t f(x) := \mathrm{tr}(a_t(x) \cdot \nabla^2 f(x)), \ \ \mathscr{B}_t f(x) := b_t(x) \cdot \nabla f(x) \tag{1.5}$$

with $a_t(x) = \frac{1}{2}(\sigma_t \sigma_t^T)(x)$, and $\mathscr{A}_t^*$ and $\mathscr{B}_t^*$ stand for the adjoint operators of $\mathscr{A}_t$ and $\mathscr{B}_t$, respectively. When the coefficients a and b are *bounded* measurable, the superposition principle for Eq. (1.4) was proved by Figalli [15, Theorem 2.6], which says that every probability measure-valued solution to the Fokker–Planck–Kolmogorov Eq. (1.4) yields a martingale solution for the operator $\mathscr{A}_t + \mathscr{B}_t$ on the path space $\mathbb{C}$ (or equivalently, a weak solution for SDE (1.3)). We would like to mention that Kurtz in [20, Theorem 2.7] has already proven such a principle if a and b are time-independent and bounded measurable (see [20, Remark 2.8(a)]). In [32], Trevisan extended it to the following natural *integrability* assumption:

$$\int_0^T \int_{\mathbb{R}^d} \Big(|b_t(x)| + |a_t(x)|\Big)\mu_t(\mathrm{d}x)\mathrm{d}t < \infty, \ \ \forall T > 0. \tag{1.6}$$

More precisely, for any probability measure-valued solution μ of (1.4), under (1.6), there is a weak solution X to SDE (1.3) so that for each $t > 0$,

$$\mu_t = \text{Law of } X_t. \tag{1.7}$$

It should be noticed that if μ_t does not have finite first moment, then (1.6) may not be satisfied for b and σ with at most linear growth. Recently, in [12], Bogachev, Röckner and Shaposhnikov obtained the superposition principle under the following more natural assumption:

$$\int_0^T \int_{\mathbb{R}^d} \frac{|\langle x, b_t(x)\rangle| + |a_t(x)|}{1 + |x|^2}\mu_t(\mathrm{d}x)\mathrm{d}t < \infty, \ \ \forall T > 0. \tag{1.8}$$

The proofs in [12] depend on quite involved uniqueness results for Fokker–Planck–Kolmogorov equations obtained in [11]. The superposition principle obtained in [15,32] has been used in the study of the uniqueness of FPKEs with rough coefficients (see e.g. [25,36]), probabilistic representations for solutions to non-linear partial differential equations (PDEs for short) [6] as well as distribution dependent SDEs (see [7,26]).

On the other hand, let $(X_t)_{t\geq 0}$ be a Feller process in $\mathbb{R}^d$ with infinitesimal generator $(\mathscr{L}, \mathrm{Dom}(\mathscr{L}))$ (see [24, page 88]). One says that $\mathscr{L}$ satisfies a positive maximum principle if for all $0 \leq f \in \mathrm{Dom}(\mathscr{L})$ reaching a positive maximum at point $x_0 \in \mathbb{R}^d$, then $\mathscr{L} f(x_0) \leq 0$. Suppose that $C_c^\infty(\mathbb{R}^d) \subset \mathrm{Dom}(\mathscr{L})$. The well-known Courrège theorem states that $\mathscr{L}$ satisfies the positive maximum principle if and only if $\mathscr{L}$ takes the following form

$$\mathscr{L} f(x) = \sum_{i,j=1}^d a_{ij}(x)\partial^2_{ij} f(x) + \sum_{i=1}^d b_i(x)\partial_i f(x) + c(x) f(x) + \int_{\mathbb{R}^d} \big(f(x+z) - f(x) - \mathbf{1}_{|z|\leq 1} z\cdot\nabla f(x)\big)\,\nu_x(\mathrm{d}z), \tag{1.9}$$

where $a = (a_{ij})_{1\leq i,j\leq d}$ is a $d\times d$-symmetric positive definite matrix-valued measurable function on $\mathbb{R}^d$, $b : \mathbb{R}^d \to \mathbb{R}^d$, $c : \mathbb{R}^d \to (-\infty, 0]$ are measurable functions and $\nu_x(\mathrm{d}z)$ is a family of Lévy measures (see [28]). In particular, if we let μ_t be the marginal law of X_t, then by Dynkin's formula,

$$\partial_t \mu_t = \mathscr{L}^* \mu_t.$$

We naturally ask that for any probability measure-valued solution μ_t to the above Fokker–Planck–Kolmogorov equation, is it possible to find some process X so that μ_t is just the law of X_t for each $t \geq 0$? In the next subsection, under some growth assumptions on the coefficients, we shall give an affirmative answer.

1.2 Superposition principle for non-local operators

Our aim in this paper is to develop a *non-local* version of the superposition principle. Let $\{\nu_{t,x}\}_{t\geq 0, x\in\mathbb{R}^d}$ be a family of Lévy measures over $\mathbb{R}^d$, that is, for each $t \geq 0$ and $x \in \mathbb{R}^d$,

$$g_t^\nu(x) := \int_{B_\ell} |z|^2 \nu_{t,x}(\mathrm{d}z) < \infty, \quad \nu_{t,x}(B_\ell^c) < \infty, \tag{1.10}$$

where $\ell > 0$ is a fixed number, and $B_\ell := \{z \in \mathbb{R}^d : |z| < \ell\}$. Without loss of generality we may assume

$$\ell \leq 1/\sqrt{2}.$$

We introduce the following Lévy type operator: for any $f \in C_b^2(\mathbb{R}^d)$,

$$\mathscr{N}_t f(x) := \mathscr{N}_t^\nu f(x) := \mathscr{N}^{\nu_{t,x}} f(x) := \int_{\mathbb{R}^d} \Theta_f(x; z)\nu_{t,x}(\mathrm{d}z), \tag{1.11}$$

where

$$\Theta_f(x;z) := f(x+z) - f(x) - \mathbf{1}_{|z|\le\ell} z \cdot \nabla f(x). \tag{1.12}$$

Let us consider the following non-local Fokker–Planck–Kolmogorov equation (FPKE for short):

$$\partial_t \mu_t = \mathscr{L}_t^* \mu_t, \tag{1.13}$$

where $\mathscr{L}_t$ is a general diffusion operator with jumps, i.e.,

$$\mathscr{L}_t := \mathscr{A}_t + \mathscr{B}_t + \mathscr{N}_t$$

with $\mathscr{A}_t$ and $\mathscr{B}_t$ being defined by (1.5) and $\mathscr{N}_t$ being defined by (1.11). We introduce the following definition of weak solution to Eq. (1.13).

Definition 1.1 *(Weak solution)* Let $\mu : \mathbb{R}_+ \to \mathcal{P}(\mathbb{R}^d)$ be a continuous curve. We call $\mu = (\mu_t)_{t\ge 0}$ a weak solution of the non-local FPKE (1.13) if for any $R > 0$ and $t > 0$,

$$\left\{\begin{array}{l} \int_0^t \int_{\mathbb{R}^d} \mathbf{1}_{B_R}(x)\Big(|a_s(x)| + |b_s(x)| + g_s^\nu(x)\Big)\mu_s(\mathrm{d}x)\mathrm{d}s < \infty, \\ \int_0^t \int_{\mathbb{R}^d} \Big(\nu_{s,x}(B^c_{\ell\vee(|x|-R)}) + \mathbf{1}_{B_R}(x)\nu_{s,x}(B^c_\ell)\Big)\mu_s(\mathrm{d}x)\mathrm{d}s < \infty, \end{array}\right\} \tag{1.14}$$

and for all $f \in C_c^2(\mathbb{R}^d)$ and $t \ge 0$,

$$\mu_t(f) = \mu_0(f) + \int_0^t \mu_s(\mathscr{L}_s f)\mathrm{d}s, \tag{1.15}$$

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

We point out that unlike the local case considered in [2,12,15,32], where the local integrability of the coefficients with respect to $\mu_t(\mathrm{d}x)\mathrm{d}t$ implies the well-definedness of the integrals in (1.15), it is even not clear whether the above integral in (1.15) makes sense in the *non-local* case since in general $\mathscr{N}_t^\nu f$ does not have compact support for $f \in C_c^2(\mathbb{R}^d)$. This is the reason why we need the second assumption in (1.14).

Remark 1.2 Under (1.14), one has $\int_0^t \mu_s(|\mathscr{L}_s f|)\mathrm{d}s < \infty$ for any $f \in C_c^2(\mathbb{R}^d)$. Let us only show

$$\int_0^t \mu_s(|\mathscr{N}_s^\nu f|)\mathrm{d}s < \infty.$$

Note that for $x, z \in \mathbb{R}^d$, by Taylor's expansion, there is a $\theta \in [0, 1]$ such that

$$f(x+z) - f(x) - z \cdot \nabla f(x) = \sum_{i,j=1,\dots,d} z_i z_j \partial_i \partial_j f(x+\theta z)/2. \tag{1.16}$$

Suppose that the support of f is contained in a ball B_R. By definition we have

$$|\Theta_f(x;z)| \leq \|f\|_\infty \mathbf{1}_{|z|>\ell}(\mathbf{1}_{|x+z|<R} + \mathbf{1}_{|x|<R}) + \|\nabla^2 f\|_\infty \mathbf{1}_{|z|\leq\ell}|z|^2 \mathbf{1}_{|x|<R+\ell}.$$

Hence,

$$\int_0^t \mu_s(|\mathscr{N}_s^\nu f|)\mathrm{d}s \lesssim \int_0^t \int_{\mathbb{R}^d} \Big[\nu_{s,x}(B^c_{\ell\vee(|x|-R)}) + \mathbf{1}_{B_R}(x)\nu_{s,x}(B^c_\ell)\Big]\mu_s(\mathrm{d}x)\mathrm{d}s + \int_0^t \int_{\mathbb{R}^d} \mathbf{1}_{B_{R+\ell}}(x) g_s^\nu(x)\mu_s(\mathrm{d}x)\mathrm{d}s < \infty.$$

Let $\mathbb{D}$ be the space of all $\mathbb{R}^d$-valued càdlàg functions on $\mathbb{R}_+$, which is endowed with the Skorokhod topology so that $\mathbb{D}$ becomes a Polish space. Let $X_t(\omega) = \omega_t$ be the canonical process. For $t \geq 0$, let $\mathcal{B}_t^0(\mathbb{D})$ denote the natural filtration generated by $(X_s)_{s\in[0,t]}$, and let

$$\mathcal{B}_t := \mathcal{B}_t(\mathbb{D}) := \cap_{s\geq t}\mathcal{B}_t^0(\mathbb{D}), \quad \mathcal{B} := \mathcal{B}(\mathbb{D}) := \mathcal{B}_\infty(\mathbb{D}).$$

Now we recall the notion of martingale solutions associated with $\mathscr{L}_t$ in the sense of Stroock–Varadhan [31].

Definition 1.3 *(Martingale Problem)* Let $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$, $s \geq 0$ and $\tau \geq s$ be a $\mathcal{B}_t$-stopping time. We call a probability measure $\mathbb{P} \in \mathcal{P}(\mathbb{D})$ a martingale solution (resp. a "stopped" martingale solution) of $\mathscr{L}_t$ with initial distribution μ_0 at time s if

(i) $\mathbb{P}(X_t = X_s, t \in [0,s]) = 1$ and $\mathbb{P} \circ X_s^{-1} = \mu_0$.
(ii) For any $f \in C_c^2(\mathbb{R}^d)$, M_t^f (resp. $M^f_{t\wedge\tau}$) is a $\mathcal{B}_t$-martingale under $\mathbb{P}$, where

$$M_t^f := f(X_t) - f(X_s) - \int_s^t \mathscr{L}_r f(X_r)\mathrm{d}r, \ t \geq s. \tag{1.17}$$

All the martingale solutions (resp. "stopped" martingale solutions) associated with $\mathscr{L}_t$ with initial law μ_0 at time s will be denoted by $\mathcal{M}_s^{\mu_0}(\mathscr{L})$ (resp. $\mathcal{M}_{s,\tau}^{\mu_0}(\mathscr{L})$). In particular, if $\mu_0 = \delta_x$ (the Dirac measure concentrated on x), we shall write $\mathcal{M}_s^x(\mathscr{L}) = \mathcal{M}_s^{\delta_x}(\mathscr{L})$ for simplify.

Remark 1.4 Under (1.18) below, (ii) in Definition 1.3 is equivalent to that for any $f \in C^2(\mathbb{R}^d)$ with $|f(x)| \leq C\log(2+|x|)$, M_t^f is a local $\mathcal{B}_t$-martingale under $\mathbb{P}$. Indeed, let $\chi \in C_c^\infty(\mathbb{R}^d)$ be a smooth function with $\chi(x) = 1$ for $|x| < 1$ and $\chi(x) = 0$ for $|x| > 2$. For each $n, m \in \mathbb{N}$, define $f_n(x) := f(x)\chi(x/n)$ and $\tau_m := \inf\{t > s : |X_t| \vee |X_t - X_{t-}| \geq m\}$. By (ii) of Definition 1.3, one knows that $M^{f_n}_{t\wedge\tau_m}$ is a $\mathcal{B}_t$-martingale. Since $|f(x)| \leq C\log(2+|x|)$, by definition (1.11) and (1.18) below, it is easy to see that for each fixed $m \in \mathbb{N}$,

$$\sup_n \sup_{r\in[0,t]} \sup_{|x|\leq m} |\mathscr{L}_r f_n(x)| < \infty.$$

Thus, for each $t > s$, by the dominated convergence theorem, we have

$$\lim_{n\to\infty} \mathbb{E}\left(\int_s^{t\wedge\tau_m} |\mathscr{L}_r f_n(X_r) - \mathscr{L}_r f(X_r)|\mathrm{d}r\right) = 0.$$

Therefore, for each $t > s$,

$$\lim_{n\to\infty} \mathbb{E}|M^{f_n}_{t\wedge\tau_m} - M^{f}_{t\wedge\tau_m}| = 0,$$

which implies that $M^{f}_{t\wedge\tau_m}$ is a $\mathcal{B}_t$-martingale for each $m \in \mathbb{N}$, and also M^f_t is a local $\mathcal{B}_t$-martingale since $\tau_m \to \infty$ as $m \to \infty$.

Throughout this paper, we make the following assumption:

$$\Gamma^{\nu}_{a,b} := \sup_{t,x}\left[\frac{|a_t(x)| + g^{\nu}_t(x)}{1+|x|^2} + \frac{|b_t(x)|}{1+|x|} + \hbar^{\nu}_t(x)\right] < \infty, \tag{1.18}$$

where $g^{\nu}_t(x)$ is defined by (1.10) and

$$\hbar^{\nu}_t(x) := \int_{B^c_\ell} \log\left(1 + \tfrac{|z|}{1+|x|}\right)\nu_{t,x}(\mathrm{d}z), \tag{1.19}$$

and if $\nu_{t,x}$ is symmetric, then we define

$$\hbar^{\nu}_t(x) := \int_{|z|>1+|x|} \log\left(1 + \tfrac{|z|}{1+|x|}\right)\nu_{t,x}(\mathrm{d}z). \tag{1.20}$$

The main result of this paper is as follows.

Theorem 1.5 (Superposition principle) *Under* (1.18)*, for any weak solution* $(\mu_t)_{t\ge 0}$ *of FPKE* (1.13) *in the sense of Definition* 1.1*, there is a martingale solution* $\mathbb{P} \in \mathcal{M}^{\mu_0}_0(\mathscr{L}_t)$ *such that*

$$\mu_t = \mathbb{P}\circ X_t^{-1}, \quad \forall t \ge 0.$$

Remark 1.6 Under (1.18), condition (1.14) holds. In fact, it suffices to check that

$$\sup_{t,x}\left(\nu_{t,x}(B^c_{\ell\vee(|x|-R)}) + \mathbf{1}_{B_R}(x)\nu_{t,x}(B^c_\ell)\right) < \infty, \quad \forall R > 0. \tag{1.21}$$

By definition we have

$$\begin{aligned}\nu_{t,x}(B^c_{\ell\vee(|x|-R)}) &\le \int_{B^c_\ell} \log\left(1+\tfrac{|z|}{1+|x|}\right)\Big/\log\left(1+\tfrac{\ell\vee(|x|-R)}{1+|x|}\right)\nu_{t,x}(\mathrm{d}z)\\ &= \hbar^{\nu}_t(x)/\log\left(1+\tfrac{\ell\vee(|x|-R)}{1+|x|}\right) \le \hbar^{\nu}_t(x)/\log\left(1+\tfrac{\ell}{1+\ell+R}\right),\end{aligned}$$

and

$$\mathbf{1}_{B_R}(x)\nu_{t,x}(B_\ell^c) \le \mathbf{1}_{B_R}(x)\int_{B_\ell^c} \log\left(1+\tfrac{|z|}{1+|x|}\right)/\log\left(1+\tfrac{\ell}{1+|x|}\right)\nu_{t,x}(\mathrm{d}z)$$
$$= \mathbf{1}_{B_R}(x)\hbar_t^\nu(x)/\log\left(1+\tfrac{\ell}{1+|x|}\right) \le \hbar_t^\nu(x)/\log\left(1+\tfrac{\ell}{1+R}\right).$$

Hence, (1.21) follows by (1.18).

Remark 1.7 Note that our result does not cover the one in [12] (see the above (1.8)). The results in [12] allow to treat SDEs with singular and linear growth coefficients, while our assumption (1.18) only allows the coefficients being of linear growth. Here the main issue is that the elegant push-forward method used in [32] seems not valid in the non-local case. Moreover, in our proof, we borrow some technique from [12] to construct the approximation sequence (see Proposition 3.2 below).

Example 1.8 Let $\nu_{t,x}(\mathrm{d}z) = \kappa_t(x,z)\mathrm{d}z/|z|^{d+\alpha}$ with $\alpha \in (0,2)$, that is, $\mathscr{N}_t$ is an α-stable like operator.

(i) If $|\kappa_t(x,z)| \le c(1+|x|)^{\alpha\wedge 1}/(1+\mathbf{1}_{\alpha=1}\log(1+|x|))$, then $\sup_{t,x}\hbar_t^\nu(x) < \infty$. Indeed, by definition we have

$$\hbar_t^\nu(x) \lesssim \frac{(1+|x|)^{\alpha\wedge 1}}{1+\mathbf{1}_{\alpha=1}\log(1+|x|)}\int_{B_\ell^c}\log\left(1+\tfrac{|z|}{1+|x|}\right)\frac{\mathrm{d}z}{|z|^{d+\alpha}}.$$

We calculate the right hand integral which is denoted by $\mathscr{I}$ as follows: using polar coordinates and integration by parts,

$$\begin{aligned}\mathscr{I} &= c\int_\ell^\infty \log\left(1+\tfrac{r}{1+|x|}\right) r^{-1-\alpha}\mathrm{d}r\\ &\lesssim \log\left(1+\tfrac{\ell}{1+|x|}\right) + \int_\ell^\infty r^{-\alpha}\,(1+|x|+r)^{-1}\,\mathrm{d}r\\ &\lesssim (1+|x|)^{-1} + (1+|x|)^{-1}\int_\ell^{1+|x|} r^{-\alpha}\mathrm{d}r + \int_{1+|x|}^\infty r^{-1-\alpha}\mathrm{d}r\\ &\lesssim (1+|x|)^{-1} + (1+|x|)^{-(\alpha\wedge 1)}(1+\mathbf{1}_{\alpha=1}\log(1+|x|)) + (1+|x|)^{-\alpha}\\ &\lesssim (1+|x|)^{-(\alpha\wedge 1)}(1+\mathbf{1}_{\alpha=1}\log(1+|x|)).\end{aligned}$$

Thus, we have $\hbar_t^\nu(x) \le C$.

(ii) If $\kappa_t(x,z)$ is symmetric, that is, $\kappa_t(x,z) = \kappa_t(x,-z)$, and $|\kappa_t(x,z)| \le c(1+|x|)^\alpha$, $\alpha \in (0,2)$. Then $\sup_{t,x}\hbar_t^\nu(x) < \infty$. In fact, by (1.20) we have for any $\beta \in (0,\alpha\wedge 1)$,

$$\begin{aligned}\hbar_t^\nu(x) &\lesssim (1+|x|)^\alpha\int_{|z|>1+|x|}\left(1+\tfrac{|z|}{1+|x|}\right)^\beta\frac{\mathrm{d}z}{|z|^{d+\alpha}}\\ &\lesssim (1+|x|)^{\alpha-\beta}\int_{|z|>1+|x|}\frac{\mathrm{d}z}{|z|^{d+\alpha-\beta}},\end{aligned}$$

which in turn yields $\sup_{t,x} \hbar_t^\nu(x) < \infty$.

As far as we know, there are very few results concerning the superposition principle for non-local operators. In the constant non-local case, the third author of the present paper [36] used the superposition principle to show the uniqueness of non-local FPKEs. Recently, Fournier and Xu [16] proved a non-local version to the superposition principle in a special case, that is,

$$\mathscr{N}_t^\nu f(x) = \int_{\mathbb{R}^d} [f(x+z) - f(x)]\nu_{t,x}(\mathrm{d}z),$$

and $(\mu_t)_{t\geq 0}$ have finite first order moments, i.e.,

$$\int_{\mathbb{R}^d} |x|\mu_t(\mathrm{d}x) < \infty, \quad \forall t \geq 0.$$

These two assumptions rule out the interesting α-stable processes (see Example 1.8 above). To drop these two limitations, we employ some techniques from [12]. It should be emphasized that the elegant push-forward method used in [32] does not seem to work in the non-local case. Here the main obstacles are to show the tightness and taking limits. One important motivation for studying the superposition principle for nonlocal operators is to solve the Boltzman equation as explained in Subsection 1.2 of [16] (see also [17]).

1.3 Equivalence between FPKEs and martingale problems

The following corollary is a direct consequence of Theorem 1.5 and [14, Theorem 4.4.2] (see also [21, Corollary 1.3] and [32, Lemma 2.12]). For the readers' convenience, we provide a detailed proof here.

Corollary 1.9 *Under (1.18), the well-posedness of the Fokker–Planck–Kolmogorov Eq. (1.13) is equivalent to the well-posedness of the martingale problem associated with $\mathscr{L}$. More precisely, we have the following equivalences:*

- **(Existence)** *For any $\nu \in \mathcal{P}(\mathbb{R}^d)$, the non-local FPKE (1.13) admits a solution $(\mu_t)_{t\geq 0}$ with initial value $\mu_0 = \nu$ if and only if $\mathcal{M}_0^\nu(\mathscr{L})$ has at least one element.*
- **(Uniqueness)** *The following two statements are equivalent.*

  (i) *For each $(s,\nu) \in \mathbb{R}_+ \times \mathcal{P}(\mathbb{R}^d)$, the non-local FPKE (1.13) has at most one solution $(\mu_t)_{t\geq s}$ with $\mu_s = \nu$.*
  (ii) *For each $(s,\nu) \in \mathbb{R}_+ \times \mathcal{P}(\mathbb{R}^d)$, $\mathcal{M}_s^\nu(\mathscr{L})$ has at most one element.*

Proof We only prove the uniqueness part. (ii)⇒(i) is easy by Theorem 1.5. We show (i)⇒(ii). For given $(s,\nu) \in \mathbb{R}_+ \times \mathcal{P}(\mathbb{R}^d)$ and let $\mathbb{P}_1, \mathbb{P}_2 \in \mathcal{M}_s^\nu(\mathscr{L})$. To show $\mathbb{P}_1 = \mathbb{P}_2$, it suffices to prove the following claim by induction:

($\mathbf{C}_n$) for given $n \in \mathbb{N}$, and for any $s \leq t_1 < t_2 < t_n$ and strictly positive and bounded measurable functions $f_1, \ldots, f_n$ on $\mathbb{R}^d$,

$$\mathbb{E}^{\mathbb{P}_1}(f_1(X_{t_1}) \cdots f_n(X_{t_n})) = \mathbb{E}^{\mathbb{P}_2}(f_1(X_{t_1}) \cdots f_n(X_{t_n})). \tag{1.22}$$

First of all, by Theorem 1.5 and the assumption, one sees that $(\mathbf{C}_1)$ holds. Next we assume $(\mathbf{C}_n)$ holds for some $n \geq 2$. For simplicity we write

$$\eta := f_1(X_{t_1}) \cdots f_n(X_{t_n}),$$

and for $i = 1, 2$, we define new probability measures

$$\mathrm{d}\tilde{\mathbb{P}}_i := \eta \mathrm{d}\mathbb{P}_i / \int_\Omega \eta \mathrm{d}\mathbb{P}_i \in \mathcal{P}(\mathbb{D}), \quad \tilde{\nu}_i := \tilde{\mathbb{P}}_i \circ X_{t_n}^{-1} \in \mathcal{P}(\mathbb{R}^d).$$

Now we show

$$\tilde{\mathbb{P}}_i \in \mathcal{M}_{t_n}^{\tilde{\nu}_i}(\mathscr{L}), \quad i = 1, 2.$$

Let M_t^f be defined by (1.17). We only need to prove that for any $t' > t \geq t_n$ and bounded $\mathcal{B}_t$-measurable ξ,

$$\mathbb{E}^{\tilde{\mathbb{P}}_i}\left(M_{t'}^f \xi\right) = \mathbb{E}^{\tilde{\mathbb{P}}_i}\left(M_t^f \xi\right) \Leftrightarrow \mathbb{E}^{\mathbb{P}_i}(M_{t'}^f \xi \eta) = \mathbb{E}^{\mathbb{P}_i}(M_t^f \xi \eta),$$

which follows since $\mathbb{P}_i \in \mathcal{M}_s^\nu(\mathscr{L})$. Thus, by induction hypothesis and Theorem 1.5,

$$\tilde{\nu}_1 = \tilde{\nu}_2 \Rightarrow \tilde{\mathbb{P}}_1 \circ X_{t_{n+1}}^{-1} = \tilde{\mathbb{P}}_2 \circ X_{t_{n+1}}^{-1}, \quad \forall t_{n+1} > t_n.$$

which in turn implies that $(\mathbf{C}_{n+1})$ holds. The proof is complete. □

1.4 Fractional porous media equation

Probabilistic representation of solution to PDEs is a powerful tool to study their analytic properties (well-posedness, regularity, etc) since it allows us to use many probabilistic tools (see [7–9]). As an application of the superposition principle obtained in Theorem 1.5, we intend to derive a probabilistic representation for the weak solution of the following fractional porous media equation (FPME for short):

$$\partial_t u = \Delta^{\alpha/2}(|u|^{m-1}u), \quad u(0, x) = \varphi(x), \tag{1.23}$$

where the porous media exponent $m > 1$, $\alpha \in (0, 2)$ and $\Delta^{\alpha/2} := -(-\Delta)^{\alpha/2}$ is the usual fractional Laplacian with, up to a constant, alternative expression

$$\Delta^{\alpha/2} f(x) = \mathrm{P.V.} \int_{\mathbb{R}^d} (f(x+z) - f(x)) \mathrm{d}z / |z|^{d+\alpha}, \tag{1.24}$$

where P.V. stands for the Cauchy principal value. This equation is a typical non-linear, degenerate and non-local parabolic equation, which appears naturally in statistical mechanics and population dynamics in order to describe the hydrodynamic limit of interacting particle systems with jumps or long-range interactions. In the last decade, there are many works devoted to the study of Eq. (1.23) from the PDE point of view, see [23] and the recent survey paper [33], the monograph [34] and the references therein.

Let $\dot{H}^{\alpha/2}(\mathbb{R}^d)$ be the homogeneous fractional Sobolev space defined as the completion of $C_0^\infty(\mathbb{R}^d)$ with respect to

$$\|f\|_{\dot{H}^{\alpha/2}} := \left(\int_{\mathbb{R}^d} |\xi|^\alpha |\hat{f}(\xi)|^2 \mathrm{d}\xi\right)^{1/2} = \|(-\Delta)^{\alpha/4} f\|_2,$$

where $\hat{f}$ is the Fourier transform of f. The following notion about the weak solution of FPME is introduced in [22, Definition 3.1].

Definition 1.10 A function u is called a weak or L^1-energy solution of FPME (1.23) if

- $u \in C([0, \infty); L^1(\mathbb{R}^d))$ and $|u|^{m-1}u \in L^2_{loc}((0, \infty); \dot{H}^{\alpha/2}(\mathbb{R}^d))$;
- for every $f \in C_0^1(\mathbb{R}_+ \times \mathbb{R}^d)$,

$$\int_0^\infty \int_{\mathbb{R}^d} u \cdot \partial_t f \mathrm{d}x\mathrm{d}t = \int_0^\infty \int_{\mathbb{R}^d} (|u|^{m-1}u) \cdot \Delta^{\alpha/2} f \mathrm{d}x\mathrm{d}t;$$

- $u(0, x) = \varphi(x)$ almost everywhere.

The following result was proved in [22, Theorem 2.1, Theorem 2.2].

Theorem 1.11 *Let $\alpha \in (0, 2)$ and $m > 1$. For every $\varphi \in L^1(\mathbb{R}^d)$, there exists a unique weak solution u for Eq. (1.23). Moreover, u enjoys the following properties:*

(i) *if $\varphi \geq 0$, then $u(t, x) > 0$ for all $t > 0$ and $x \in \mathbb{R}^d$;*
(ii) *$\partial_t u \in L^\infty((s, \infty); L^1(\mathbb{R}^d))$ for every $s > 0$;*
(iii) *for all $t \geq 0$, $\int_{\mathbb{R}^d} u(t, x)\mathrm{d}x = \int_{\mathbb{R}^d} \varphi(x)\mathrm{d}x$;*
(iv) *if $\varphi \in L^\infty(\mathbb{R}^d)$, then for every $t > 0$,*

$$\|u(t, \cdot)\|_\infty \leq \|\varphi\|_\infty;$$

(v) *for some $\beta \in (0, 1)$, $u \in C^\beta((0, \infty) \times \mathbb{R}^d)$.*

Our aim in this subsection is to represent the above solution u as the distributional density of the solution to a nonlinear stochastic differential equation driven by the α-stable process L_t with Lévy measure $\mathrm{d}z/|z|^{d+\alpha}$. More precisely, consider the following distribution dependent stochastic differential equation (DDSDE for short) driven by the d-dimensional isotropic α-stable process L_t:

$$\mathrm{d}Y_t = \rho_{Y_t}\big(Y_{t-}\big)^{\frac{m-1}{\alpha}} \mathrm{d}L_t, \quad \rho_{Y_0}(x) = \varphi(x), \tag{1.25}$$

where $\rho_{Y_t}(x) := (\mathrm{d}\mathcal{L}_{Y_t}/\mathrm{d}x)(x)$ denotes the distributional density of Y_t with respect to Lebesgue measure. We introduce the following notion about the above DDSDE (1.25).

Definition 1.12 Let $(\Omega, \mathcal{F}, \mathbf{P}; (\mathcal{F}_t)_{t\geq 0})$ be a stochastic basis and (Y, L) two $\mathcal{F}_t$-adapted càdlàg processes. For $\mu \in \mathcal{P}(\mathbb{R}^d)$, we call $(\Omega, \mathcal{F}, \mathbf{P}; (\mathcal{F}_t)_{t\geq 0}; Y, L)$ a solution of (1.25) with initial law μ if

(i) L is an α-stable process with Lévy measure $\mathrm{d}z/|z|^{d+\alpha}$;
(ii) for each $t \geq 0$, $\mathbf{P} \circ Y_t^{-1}(\mathrm{d}x) = \rho_{Y_t}(x)\mathrm{d}x$;
(iii) Y_t solves the following SDE:

$$Y_t = Y_0 + \int_0^t \rho_{Y_s}\big(Y_{s-}\big)^{\frac{m-1}{\alpha}} \mathrm{d}L_s.$$

The following is the second main result of this paper.

Theorem 1.13 *Let $\varphi \geq 0$ be bounded and satisfy $\int_{\mathbb{R}^d} \varphi(x)\mathrm{d}x = 1$. Let u be the unique weak solution to FPME (1.23) given by Theorem 1.11 with initial value φ. Then there exists a weak solution Y to DDSDE (1.25) such that*

$$\rho_{Y_t}(x) = u(t, x), \quad \forall t \geq 0.$$

Remark 1.14 Here an open question is to show the uniqueness of weak solutions to the nonlinear SDE (1.25), which can not be derived from the uniqueness of FPME (1.23). We will study this in a future work.

We mention that in the 1-dimensional case, such kind of probabilistic representation for the classical porous media equation (i.e., $\alpha = 2$) was obtained in [8], see also [10] and [6,7] and for the generalization to the multi-dimensional case and more general non-linear equations. We also mention that there has been an increasing interest in DDSDEs driven by Brownian motion in the last decade, see [7,26] and in particular, [13] as well as the references therein. As far as we know, even the weak existence result for DDSDE (1.25) driven by Lévy noise in Theorem 1.13 is also new.

This paper is organized as follows: In Sect. 2, we study the Eq. (1.13) with smooth and non-degenerate coefficients. Then we prove Theorems 1.5 and 1.13 in Sects. 3 and 4, respectively. Throughout this paper we shall use the following conventions:

- The letter C denotes a constant, whose value may change in different places.
- We use $A \lesssim B$ to denote $A \leq CB$ for some unimportant constant $C > 0$.
- $\mathbb{N}_0 := \mathbb{N}\cup\{0\}, \mathbb{R}_+ := [0, \infty), a\vee b := \max(a, b), a\wedge b := \min(a, b), a^+ := a\vee 0.$
- $\nabla_x := \partial_x := (\partial_{x_1}, \dots, \partial_{x_d}), \partial_i := \partial_{x_i} := \partial/\partial x_i$.
- $\mathbb{S}^d_+$ is the set of all $d \times d$-symmetric and non-negative definite matrices.

2 Proof of Theorem 1.5: smooth and nondegenerate coefficients

First of all, we show the following well-posedness result about the martingale problem associated with $\mathscr{L}_t$, which extends Stroock's result [30] to unbounded coefficients case, and is probably well-known at least to experts. However, since we can not find it in the literature, we provide a detailed proof here.

Theorem 2.1 *Suppose that the following conditions are satisfied:*

(A) $a_t(x):\mathbb{R}_+\times\mathbb{R}^d\to\mathbb{S}^d_+$ *is continuous and* $a_t(x)$ *is invertible;*
(B) $b_t(x):\mathbb{R}_+\times\mathbb{R}^d\to\mathbb{R}^d$ *is locally bounded and measurable;*
(C) *for any* $A\in\mathcal{B}(\mathbb{R}^d)$, $(t,x)\mapsto\int_A(1\wedge|z|^2)\nu_{t,x}(\mathrm{d}z)$ *is continuous;*
(D) *the following global growth condition holds:*

$$\bar\Gamma^\nu_{a,b}:=\sup_{t,x}\left(\frac{|a_t(x)|+\langle x,b_t(x)\rangle^++g^\nu_t(x)}{1+|x|^2}+2\hbar^\nu_t(x)\right)<\infty,$$

where $g^\nu_t(x)$ *and* $\hbar^\nu_t(x)$ *are defined by* (1.10) *and* (1.19)*, respectively.*

Then for each $(s,x)\in\mathbb{R}_+\times\mathbb{R}^d$*, there is a unique martingale solution* $\mathbb{P}_{s,x}\in\mathcal{M}^x_s(\mathscr{L}_t)$*. Moreover, the following assertions hold:*

(i) *For each* $A\in\mathcal{B}(\mathbb{D})$, $(s,x)\mapsto\mathbb{P}_{s,x}(A)$ *is Borel measurable.*
(ii) *The following strong Markov property holds: for every bounded measurable* f *and any finite stopping time* τ,

$$\mathbb{E}^{\mathbb{P}_{0,x}}(f(\tau+t,X_{\tau+t})|\mathcal{B}_\tau)=\left(\mathbb{E}^{\mathbb{P}_{s,y}}(f(s+t,X_{s+t}))\right)\Big|_{(s,y)=(\tau,X_\tau)}.$$

Remark 2.2 Condition **(D)** ensures the non-explosion of the solution.

To prove this theorem we first show the following Lyapunov type estimate.

Lemma 2.3 *Let* $\psi\in C^2(\mathbb{R};\mathbb{R}_+)$ *with* $\lim_{r\to\infty}\psi(r)=\infty$ *and*

$$0<\psi'\le1,\quad \psi''\le0. \tag{2.1}$$

Fix $y\in\mathbb{R}^d$ *and define a Lyapunov function* $V_y(x):=\psi(\log(1+|x-y|^2))$*. Then for all* $t\ge0$ *and* $x\in\mathbb{R}^d$*, we have*

$$\mathscr{L}_tV_y(x)\le2\left(\frac{|a_t(x)|+\langle x-y,b_t(x)\rangle^++g^\nu_t(x)}{1+|x-y|^2}+2H^\nu_t(x,y)\right), \tag{2.2}$$

where $g^\nu_t(x)$ *is defined by* (1.10)*, and*

$$H^\nu_t(x,y):=\int_{B^c_\ell}\log\left(1+\tfrac{|z|}{1+|x-y|}\right)\nu_{t,x}(\mathrm{d}z). \tag{2.3}$$

Proof By definition, it is easy to see that

$$\nabla V_y(x)=\frac{2(x-y)}{1+|x-y|^2}\psi'(\log(1+|x-y|^2))$$

and

$$\nabla^2V_y(x)=\frac{4(x-y)\otimes(x-y)}{(1+|x-y|^2)^2}(\psi''-\psi')(\log(1+|x-y|^2))$$

$$+\frac{2\mathbb{I}}{1+|x-y|^2}\psi'(\log(1+|x-y|^2)).$$

Thus by (2.1), one gets that

$$\mathscr{A}_t^a V_y(x) \le \frac{2|a_t(x)|}{1+|x-y|^2}, \quad \mathscr{B}_t^b V_y(x) \le \frac{2\langle x-y, b_t(x)\rangle^+}{1+|x-y|^2}.$$

On the other hand, recalling (1.12), we have for $|z| \le \ell \le 1/\sqrt{2}$,

$$\begin{aligned}\Theta_{V_y}(x;z) &= V_y(x+z) - V_y(x) - z\cdot\nabla V_y(x) = z_i z_j \partial_i\partial_j V_y(x+\theta z)/2\\ &= \frac{2\langle z, x-y+\theta z\rangle^2}{(1+|x-y+\theta z|^2)^2}(\psi''-\psi')(\log(1+|x-y+\theta z|^2))\\ &\quad + \frac{|z|^2}{1+|x-y+\theta z|^2}\psi'(\log(1+|x-y+\theta z|^2))\\ &\overset{(2.1)}{\le} \frac{|z|^2}{1+|x-y+\theta z|^2} \le \frac{|z|^2}{1+|x-y|^2/2-|z|^2} \le \frac{2|z|^2}{1+|x-y|^2},\end{aligned}$$

where $\theta \in [0,1]$. Similarly, by the mean value formula, we have

$$\begin{aligned}V_y(x+z) - V_y(x) &= \psi'(\theta_*)\Big[\log\big(1+|x-y+z|^2\big) - \log\big(1+|x-y|^2\big)\Big]\\ &\le \log\left(1+\frac{2|\langle x-y, z\rangle| + |z|^2}{1+|x-y|^2}\right) \le \log\left(1+\frac{|z|}{\sqrt{1+|x-y|^2}}\right)^2\\ &\le \log\left(1+\frac{2|z|}{1+|x-y|}\right)^2 \le \log\left(1+\frac{|z|}{1+|x-y|}\right)^4,\end{aligned}$$

where $\theta_* \in \mathbb{R}$. Hence,

$$\mathscr{N}_t^\nu V_y(x) \le \int_{\mathbb{R}^d} \Theta_{V_y}(x;z)\nu_{t,x}(\mathrm{d}z) \le 2\frac{g_t^\nu(x)}{1+|x-y|^2} + 4H_t^\nu(x,y).$$

Combining the above calculations, we obtain (2.2). □

The following stochastic Gronwall inequality for continuous martingales was proved by Scheutzow [27], and for general discontinuous martingales in [35, Lemma 3.7].

Lemma 2.4 (Stochastic Gronwall inequality) *Let $\xi(t)$ and $\eta(t)$ be two non-negative càdlàg adapted processes, A_t a continuous non-decreasing adapted process with $A_0 = 0$, M_t a local martingale with $M_0 = 0$. Suppose that*

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

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

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

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

The following localization lemma is well known (see e.g. [31, Theorem 1.3.5]). Although it is only proved for the probability measures on the space of continuous functions, by checking the proof therein, one sees that it also works for $\mathbb{D}$.

Lemma 2.5 *Let* $(\mathbb{P}_n)_{n\in\mathbb{N}} \subset \mathcal{P}(\mathbb{D})$ *be a family of probability measures and* $(\tau_n)_{n\in\mathbb{N}}$ *a non-decreasing sequence of stopping times with* $\tau_0 \equiv 0$. *Suppose that for each* $n \in \mathbb{N}$, $\mathbb{P}_n$ *equals* $\mathbb{P}_{n-1}$ *on* $\mathcal{B}_{\tau_{n-1}}(\mathbb{D})$, *and for any* $T \geq 0$,

$$\lim_{n\to\infty} \mathbb{P}_n(\tau_n \leq T) = 0.$$

Then there is a unique probability measure $\mathbb{P} \in \mathcal{P}(\mathbb{D})$ *such that* $\mathbb{P}$ *equals* $\mathbb{P}_n$ *on* $\mathcal{B}_{\tau_n}(\mathbb{D})$ *and* $\mathbb{P}_n$ *weakly converges to* $\mathbb{P}$ *as* $n \to \infty$.

We now use the above localization lemma to give

Proof of Theorem 2.1 Let $\chi \in C_c^\infty(\mathbb{R}^d)$ be a smooth function with

$$\chi(x) = 1, \quad |x| < 1, \quad \chi(x) = 0, \quad |x| > 2.$$

For any $n \in \mathbb{N}$, define

$$\chi_n(x) := \chi(x/n)$$

and

$$a_t^n(x) := a_t(x\chi_n(x)), \quad b_t^n(x) := \chi_n(x)b_t(x), \quad \nu_{t,x}^n(\mathrm{d}z) := \chi_n(x)\nu_{t,x}(\mathrm{d}z).$$

By the assumptions **(A)**–**(C)**, one can check that (a^n, b^n, ν^n) satisfies for any $T > 0$,

(A′) $a_t^n(x) : [0, T] \times \mathbb{R}^d \to \mathbb{S}_+^d$ is bounded continuous and $a_t^n(x)$ is invertible.
(B′) $b_t^n(x) : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ is bounded measurable.
(C′) For any $A \in \mathcal{B}(\mathbb{R}^d)$, $(t, x) \mapsto \int_A (1 \wedge |z|^2)\nu_{t,x}^n(\mathrm{d}z)$ is bounded continuous.

Let $\mathscr{L}_t^n$ be defined in terms of (a^n, b^n, ν^n). For each $n \in \mathbb{N}$ and $(s, x) \in \mathbb{R}_+ \times \mathbb{R}^d$, by [19, Theorem 2.34, p. 159], there is a unique martingale solution $\mathbb{P}_{s,x}^n \in \mathcal{M}_s^x(\mathscr{L}_t^n)$, and the following properties hold:

(i) For each $A \in \mathcal{B}(\mathbb{D})$, $(s, x) \mapsto \mathbb{P}_{s,x}^n(A)$ is Borel measurable.
(ii) The following strong Markov property holds: for any bounded measurable f and finite stopping time τ,

$$\mathbb{E}^{\mathbb{P}_{0,x}^n}(f(\tau + t, X_{\tau+t})|\mathcal{B}_\tau) = \big(\mathbb{E}^{\mathbb{P}_{s,y}}(f(s + t, X_{s+t}))\big)\big|_{(s,y)=(\tau,X_\tau)}.$$

Moreover, if we define

$$\tau_n := \inf\{t \geq s : |X_t| > n\},$$

then by [19, Theorem 2.41, p. 161], for any $m \geq n$, the "stopped" martingale problem $\mathcal{M}^x_{s,\tau_n}(\mathscr{L}^m_t)$ admits a unique solution, that is,

$$\mathbb{P}^m_{s,x}|_{\mathcal{B}_{\tau_n}(\mathbb{D})} = \mathbb{P}^n_{s,x}|_{\mathcal{B}_{\tau_n}(\mathbb{D})}.$$

To show the well-posedness, by Lemma 2.5, it suffices to show that for any $T > 0$,

$$\lim_{n\to\infty} \mathbb{P}^n_{s,x}(\tau_n \leq T) = 0.$$

Let $V(x) := \log(1 + |x|^2)$. By the definition of martingale solution (see Remark 1.4), there is a càdlàg local $\mathbb{P}^n_{s,x}$-martingale M_t such that

$$\begin{aligned} V(X_{t\wedge\tau_n}) &= V(x) + \int_s^{t\wedge\tau_n} \mathscr{L}^n_r V(X_r)\mathrm{d}r + M_t \\ &= V(x) + \int_s^{t\wedge\tau_n} \mathscr{L}_r V(X_r)\mathrm{d}r + M_t \\ &\overset{(2.2)}{\leq} V(x) + 2\bar{\Gamma}^\nu_{a,b} \cdot (t - s) + M_t, \end{aligned}$$

where $\bar{\Gamma}^\nu_{a,b}$ is defined in **(D)**. By Lemma 2.4 and condition **(D)**, we obtain

$$\sup_n \mathbb{E}^{\mathbb{P}^n_{s,x}} \left(\sup_{t\in[s,T\wedge\tau_n]} V^{\frac{1}{2}}(X_t) \right) < +\infty,$$

which in turn implies that

$$\mathbb{P}^n_{s,x}(\tau_n \leq T) = \mathbb{P}^n_{s,x}\left(\sup_{t\in[s,T\wedge\tau_n]} |X_t| > n \right) \leq \frac{1}{V^{\frac{1}{2}}(n)} \mathbb{E}^{\mathbb{P}^n_{s,x}} \left(\sup_{t\in[s,T\wedge\tau_n]} V^{\frac{1}{2}}(X_t) \right) \overset{n\to\infty}{\to} 0.$$

The proof is complete. □

Now we can give the proof of Theorem 1.5 under the assumptions **(A)**–**(D)**.

Theorem 2.6 *Assume that* **(A)**–**(D)** *hold. Then for any $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$, there are a unique solution $(\mu_t)_{t\geq 0}$ to FPKE (1.13) and a unique martingale solution $\mathbb{P}_{0,\mu_0} \in \mathcal{M}^{\mu_0}_0(\mathscr{L})$ so that $\mu_t = \mathbb{P}_{0,\mu_0} \circ X_t^{-1}$.*

Proof Let $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$ and $\mathbb{P}_{0,x} \in \mathcal{M}^x_0(\mathscr{L})$. Clearly,

$$\mathbb{P}_{0,\mu_0} := \int_{\mathbb{R}^d} \mathbb{P}_{0,x}\mu_0(\mathrm{d}x) \in \mathcal{M}^{\mu_0}_0(\mathscr{L}),$$

and $\mu_t := \mathbb{P}_{0,\mu_0} \circ X_t^{-1}$ solves FPKE (1.13). It remains to show the uniqueness for (1.13). Following the same argument as in [16], due to Horowitz and Karandikar [17, Theorem B1], we only need to verify the following five points:

(a) $C_c^2(\mathbb{R}^d)$ is dense in $C_0(\mathbb{R}^d)$ with respect to the uniform convergence.
(b) $(t,x) \to \mathscr{L}_t f(x)$ is measurable for all $f \in C_c^2(\mathbb{R}^d)$.
(c) For each $t \geq 0$, the operator $\mathscr{L}_t$ satisfies the maximum principle.
(d) There exists a countable family $(f_k)_{k\in\mathbb{N}} \subset C_c^2(\mathbb{R}^d)$ such that for all $t \geq 0$,

$$\{\mathscr{L}_t f,\, f \in C_c^2(\mathbb{R}^d)\} \subset \overline{\{\mathscr{L}_t f_k,\, k \in \mathbb{N}\}},$$

where the closure is taken in the uniform norm.
(e) For each $x \in \mathbb{R}^d$, $\mathcal{M}_0^x(\mathscr{L})$ has exactly one element.

Note that **(a)**–**(c)** are obvious and **(e)** is proven in Theorem 2.1. Thus we only need to check **(d)**. Let $(f_k)_{k\in\mathbb{N}}$ be a countable dense subset of $C_c^2(\mathbb{R}^d)$, that is, for any $f \in C_c^2(\mathbb{R}^d)$ with support in B_R, where $R \geq 2$, there is a subsequence f_{k_n} with support in B_{2R} such that

$$\lim_{n\to\infty} \left(\|f_{k_n} - f\|_\infty + \|\nabla f_{k_n} - \nabla f\|_\infty + \|\nabla^2 f_{k_n} - \nabla^2 f\|_\infty\right) = 0.$$

We want to show

$$\lim_{n\to\infty} \|\mathscr{L}_t(f_{k_n} - f)\|_\infty = 0.$$

Without loss of generality, we may assume $f = 0$ and proceed to prove the following limits:

$$\lim_{n\to\infty} \|\mathscr{A}_t f_{k_n}\|_\infty = 0, \ \lim_{n\to\infty} \|\mathscr{B}_t f_{k_n}\|_\infty = 0, \ \lim_{n\to\infty} \|\mathscr{N}_t^\nu f_{k_n}\|_\infty = 0.$$

The first two limits are obvious. Let us focus on the last one. By definition we have

$$\begin{aligned}|\Theta_{f_{k_n}}(x;z)| &= |f_{k_n}(x+z) - f_{k_n}(x) - \mathbf{1}_{|z|\leq\ell} z \cdot \nabla f_{k_n}(x)| \\ &\leq \mathbf{1}_{|z|>\ell}|f_{k_n}(x+z)| + \mathbf{1}_{|z|>\ell}\mathbf{1}_{B_{2R}}(x)\|f_{k_n}\|_\infty \\ &\quad + \mathbf{1}_{|z|\leq\ell}\mathbf{1}_{B_{2R+2\ell}}(x)\|\nabla^2 f_{k_n}\|_\infty |z|^2.\end{aligned}$$

Note that

$$\mathbf{1}_{|z|>\ell}\mathbf{1}_{B_{5R}}(x) \leq \left[\log(1 + \tfrac{\ell}{1+5R})\right]^{-1} \log\left(1 + \tfrac{|z|}{1+|x|}\right),$$

and if $|x| > 5R$, then for $|x+z| \leq 2R$,

$$\tfrac{|z|}{1+|x|} \geq \tfrac{|x|-|x+z|}{1+|x|} \geq \tfrac{|x|-2R}{1+|x|} > \tfrac{1}{2},$$

and thus,

$$\mathbf{1}_{|z|>\ell}\mathbf{1}_{B^c_{5R}}(x)\mathbf{1}_{B^c_{2R}}(x+z) \le \Big[\log(\tfrac{3}{2})\Big]^{-1}\log\Big(1+\tfrac{|z|}{1+|x|}\Big).$$

Therefore,

$$\begin{aligned}|\mathscr{N}^\nu_t f_{k_n}(x)| &\le \int_{\mathbb{R}^d}|\Theta_{f_{k_n}}(x;z)|\nu_{t,x}(\mathrm{d}z) \le \|\nabla^2 f_{k_n}\|_\infty \sup_{x\in B_{2R+2}}\int_{B_1}|z|^2\nu_{t,x}(\mathrm{d}z)\\ &\quad + C\|f_{k_n}\|_\infty\int_{B^c_1}\log\Big(1+\tfrac{|z|}{1+|x|}\Big)\nu_{t,x}(\mathrm{d}z)\\ &= \|\nabla^2 f_{k_n}\|_\infty \sup_{x\in B_{2R+2}} g^\nu_t(x) + C\|f_{k_n}\|_\infty \sup_{x\in\mathbb{R}^d}\hbar^\nu_t(x),\end{aligned}$$

which in turn implies by (1.18) that

$$\lim_{n\to\infty}\|\mathscr{N}^\nu_t f_{k_n}\|_\infty = 0.$$

The proof is compete. □

3 Proof of Theorem 1.5: general case

Let μ_t be a solution of (1.13) in the sense of Definition 1.1. In order to show the existence of a martingale solution $\mathbb{P}\in\mathcal{M}^{\mu_0}_0(\mathscr{L}_t)$ so that

$$\mu_t = \mathbb{P}\circ X^{-1}_t,$$

we shall follow the same lines of argument as in [12,15,32]. Here and below we use the following convention: for $t\le 0$,

$$\mu_t(\mathrm{d}x) := \mu_0(\mathrm{d}x),\quad a_t(x)=0,\quad b_t(x)=0,\quad \nu_{t,x}(\mathrm{d}z)=0.$$

3.1 Regularization

Let $\rho^{\mathrm{t}}\in C^\infty_c([0,1];\mathbb{R}_+)$ with $\int_0^1\rho^{\mathrm{t}}(s)\mathrm{d}s = 1$ and $\rho^{\mathrm{x}}\in C^\infty_c(B_1;\mathbb{R}_+)$ with $\int_{\mathbb{R}^d}\rho^{\mathrm{x}}(x)\mathrm{d}x = 1$. For $\varepsilon>0$, define

$$\rho^{\mathrm{t}}_\varepsilon(t) := \varepsilon^{-1}\rho^{\mathrm{t}}(t/\varepsilon),\quad \rho^{\mathrm{x}}_\varepsilon(x) := \varepsilon^{-d}\rho^{\mathrm{x}}(x/\varepsilon),\quad \rho_\varepsilon(t,x) := \rho^{\mathrm{t}}_\varepsilon(t)\rho^{\mathrm{x}}_\varepsilon(x).$$

Given a locally finite signed measure $\zeta_t(\mathrm{d}x)\mathrm{d}t$ on $\mathbb{R}^{d+1}$, we define

$$\rho_\varepsilon * \zeta(t,x) := \int_{\mathbb{R}^{d+1}}\rho_\varepsilon(t-s,x-y)\zeta_s(\mathrm{d}y)\mathrm{d}s.$$

Throughout this section we shall fix

$$\ell \in (0, 1/\sqrt{2}).$$

We first show the following regularization estimate.

Lemma 3.1 *Let a, b and ν be as in the introduction. For $\varepsilon \in (0, \ell)$, we have*

$$\frac{|\rho_\varepsilon * (a\mu)|(t,x)}{1+|x|^2} \le \sup_{s,y} \frac{2|a_s(y)|}{1+|y|^2} (\rho_\varepsilon * \mu)(t,x),$$
$$\frac{|\rho_\varepsilon * (b\mu)|(t,x)}{1+|x|} \le \sup_{s,y} \frac{2|b_s(y)|}{1+|y|} (\rho_\varepsilon * \mu)(t,x).$$

Moreover, if we let

$$\bar\nu^\varepsilon_{t,x}(\mathrm{d}z) := \int_{\mathbb{R}^{d+1}} \rho_\varepsilon(t-s, x-y)\nu_{s,y}(\mathrm{d}z)\mu_s(\mathrm{d}y)\mathrm{d}s,$$

then we also have

$$\frac{g^{\bar\nu^\varepsilon}_t(x)}{1+|x|^2} \le \sup_{s,y} \frac{2g^\nu_s(y)}{1+|y|^2} (\rho_\varepsilon * \mu)(t,x),$$
$$H^{\bar\nu^\varepsilon}_t(x,y) \le 2 \sup_{s,y'} H^\nu_s(y', y)(\rho_\varepsilon * \mu)(t,x),$$

where $g^\nu_t(x)$ and $H^\nu_t(x,y)$ are defined by (1.10) and (2.3), respectively.

Proof Note that for $|x-y| \le \ell \le 1/\sqrt{2}$,

$$(1+|y|^2)/2 \le 1+|x|^2 \le 2(1+|y|^2). \tag{3.1}$$

Fix $\varepsilon \in (0, \ell)$ below. By definition we have

$$\begin{aligned}\frac{|\rho_\varepsilon * (a\mu)|(t,x)}{1+|x|^2} &\le \int_{\mathbb{R}^{d+1}} \rho_\varepsilon(t-s, x-y)\frac{|a_s(y)|}{1+|x|^2}\mu_s(\mathrm{d}y)\mathrm{d}s \\ &\le 2\int_{\mathbb{R}^{d+1}} \rho_\varepsilon(t-s, x-y)\frac{|a_s(y)|}{1+|y|^2}\mu_s(\mathrm{d}y)\mathrm{d}s,\end{aligned}$$

and

$$\begin{aligned}\frac{|\rho_\varepsilon * (b\mu)|(t,x)}{1+|x|} &\le \int_{\mathbb{R}^{d+1}} \rho_\varepsilon(t-s, x-y)\frac{|b_s(y)|}{1+|x|}\mu_s(\mathrm{d}y)\mathrm{d}s \\ &\le 2\int_{\mathbb{R}^{d+1}} \rho_\varepsilon(t-s, x-y)\frac{|b_s(y)|}{1+|y|}\mu_s(\mathrm{d}y)\mathrm{d}s.\end{aligned}$$

Similarly, by Fubini's theorem and (3.1), we have

$$\begin{aligned}\frac{g_t^{\bar{\nu}^\varepsilon}(x)}{1+|x|^2} &= \int_{\mathbb{R}^{d+1}}\int_{B_\ell}\frac{|z|^2}{1+|x|^2}\rho_\varepsilon(t-s,x-y)\nu_{s,y}(\mathrm{d}z)\mu_s(\mathrm{d}y)\mathrm{d}s\\ &\le 2\int_{\mathbb{R}^{d+1}}\int_{B_\ell}\frac{|z|^2}{1+|y|^2}\rho_\varepsilon(t-s,x-y)\nu_{s,y}(\mathrm{d}z)\mu_s(\mathrm{d}y)\mathrm{d}s\\ &= 2\int_{\mathbb{R}^{d+1}}\frac{g_s^\nu(y)}{1+|y|^2}\rho_\varepsilon(t-s,x-y)\mu_s(\mathrm{d}y)\mathrm{d}s,\end{aligned}$$

and

$$\begin{aligned}H_t^{\bar{\nu}^\varepsilon}(x,y) &= \int_{\mathbb{R}^{d+1}}\int_{B_\ell^c}\log\left(1+\tfrac{|z|}{1+|x-y|}\right)\rho_\varepsilon(t-s,x-y')\nu_{s,y'}(\mathrm{d}z)\mu_s(\mathrm{d}y')\mathrm{d}s\\ &\le \int_{\mathbb{R}^{d+1}}\int_{B_\ell^c}\log\left(1+\tfrac{2|z|}{1+|y'-y|}\right)\rho_\varepsilon(t-s,x-y')\nu_{s,y'}(\mathrm{d}z)\mu_s(\mathrm{d}y')\mathrm{d}s\\ &\le 2\int_{\mathbb{R}^{d+1}}H_s^\nu(y',y)\rho_\varepsilon(t-s,x-y')\mu_s(\mathrm{d}y')\mathrm{d}s.\end{aligned}$$

Combining the above calculations, we obtain the desired estimates. □

Let $\phi(x) := (2\pi)^{-d}\mathrm{e}^{-|x|^2/2}$ be the normal density. For $\varepsilon\in(0,\ell)$, as in [12], we define the approximation sequence $\mu_t^\varepsilon\in\mathcal{P}(\mathbb{R}^d)$ by

$$\mu_t^\varepsilon(x) := (1-\varepsilon)(\rho_\varepsilon*\mu)(t,x)+\varepsilon\phi(x). \tag{3.2}$$

We have the following easy consequence.

Proposition 3.2 *(i) For each $t\ge 0$ and $\varepsilon\in(0,\ell)$, we have*

$$0<\mu_t^\varepsilon(x)\in C^\infty(\mathbb{R}_+;C_b^\infty(\mathbb{R}^d)),\quad \int_{\mathbb{R}^d}\mu_t^\varepsilon(x)\mathrm{d}x=1.$$

(ii) For each $t\ge 0$, μ_t^ε weakly converges to μ_t, that is, for any $f\in C_b(\mathbb{R}^d)$,

$$\lim_{\varepsilon\to 0}\int_{\mathbb{R}^d}f(x)\mu_t^\varepsilon(x)\mathrm{d}x=\int_{\mathbb{R}^d}f(x)\mu_t(\mathrm{d}x).$$

(iii) μ_t^ε solves the following Fokker–Planck–Kolmogorov equation:

$$\partial_t\mu_t^\varepsilon=(\mathscr{A}_t^\varepsilon+\mathscr{B}_t^\varepsilon+\mathscr{N}_t^\varepsilon)^*\mu_t^\varepsilon=:(\mathscr{L}_t^\varepsilon)^*\mu_t^\varepsilon,$$

where $\mathscr{A}_t^\varepsilon$, $\mathscr{B}_t^\varepsilon$ and $\mathscr{N}_t^\varepsilon$ are defined as in the introduction in terms of

$$a_t^\varepsilon(x) := \frac{(1-\varepsilon)[\rho_\varepsilon*(a\mu)](t,x)+\varepsilon\phi(x)\mathbb{I}}{\mu_t^\varepsilon(x)}, \tag{3.3}$$

$$b_t^\varepsilon(x) := \frac{(1-\varepsilon)[\rho_\varepsilon * (b\mu)](t,x) + \varepsilon\phi(x)x}{\mu_t^\varepsilon(x)}, \tag{3.4}$$

and

$$\nu_{t,x}^\varepsilon(\mathrm{d}z) := \frac{1-\varepsilon}{\mu_t^\varepsilon(x)} \int_{\mathbb{R}^{d+1}} \rho_\varepsilon(t-s, x-y)\nu_{s,y}(\mathrm{d}z)\mu_s(\mathrm{d}y)\mathrm{d}s. \tag{3.5}$$

(iv) *The following uniform estimates hold: for any* $\varepsilon \in (0, \ell)$,

$$\sup_{t,x}\left[\frac{|a_t^\varepsilon(x)| + g_t^{\nu^\varepsilon}(x)}{1+|x|^2} + \frac{|b_t^\varepsilon(x)|}{1+|x|}\right] \le 1 + 2\sup_{t,x}\left[\frac{|a_t(x)| + g_t^\nu(x)}{1+|x|^2} + \frac{|b_t(x)|}{1+|x|}\right] \tag{3.6}$$

and

$$\sup_{t,x} H_t^{\nu^\varepsilon}(x,y) \le \sup_{t,x} H_t^\nu(x,y), \quad y \in \mathbb{R}^d. \tag{3.7}$$

Proof The first two assertions are obvious by definition. Let us show (iii). By definition, it suffices to prove that for any $f \in C_c^\infty(\mathbb{R}^d)$ and $t \ge 0$,

$$\mu_t^\varepsilon(f) = \mu_0^\varepsilon(f) + \int_0^t \mu_s^\varepsilon(\mathscr{L}_s^\varepsilon f)\mathrm{d}s, \tag{3.8}$$

where

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

Note that for any $f \in C_c^\infty(\mathbb{R}^d)$,

$$\Delta\phi + \mathrm{div}(x\cdot\phi) \equiv 0 \Rightarrow \int_{\mathbb{R}^d} \phi(x)(\Delta f(x) - x\cdot\nabla f(x))\mathrm{d}x = 0.$$

By Fubini's theorem and a change of variables, it is easy to see that (3.8) holds. Finally, estimate (3.6) follows by Lemma 3.1. □

The following result follows by Theorem 2.6.

Lemma 3.3 *For any* $\varepsilon \in (0, \ell)$ *and* $(s,x) \in \mathbb{R}_+ \times \mathbb{R}^d$, *there is a unique martingale solution* $\mathbb{P}_{s,x}^\varepsilon \in \mathcal{M}_s^x(\mathscr{L}_t^\varepsilon)$. *In particular, there is also a martingale solution* $\mathbb{Q}^\varepsilon \in \mathcal{M}_0^{\mu_0^\varepsilon}(\mathscr{L}_t^\varepsilon)$ *so that for each* $t \ge 0$,

$$\mu_t^\varepsilon(x)\mathrm{d}x = \mathbb{Q}^\varepsilon \circ X_t^{-1}(\mathrm{d}x).$$

Proof By Theorem 2.6, it suffices to check that $(a^\varepsilon, b^\varepsilon, \nu^\varepsilon)$ satisfies conditions **(A)**–**(D)**. First of all, **(A)** and **(B)** are obvious, and **(D)** follows by (3.6). It remains to check **(C)**. We only check that for any $\varepsilon \in (0, \ell)$, $n \in \mathbb{N}$ and $x, x' \in B_n$, $t, t' \in [0, n]$,

$$\int_{\mathbb{R}^d} (1 \wedge |z|^2) |\nu^\varepsilon_{t,x} - \nu^\varepsilon_{t',x'}|(\mathrm{d}z) \le c_{n,\varepsilon}(|t - t'| + |x - x'|). \tag{3.9}$$

Noting that

$$\inf_t \inf_{x \in B_n} \mu^\varepsilon_t(x) \ge \varepsilon \inf_{x \in B_n} \phi(x),$$

we have by definition that for all $x, x' \in B_n$ and $t, t' \in [0, n]$,

$$\begin{aligned} |\nu^\varepsilon_{t,x} - \nu^\varepsilon_{t',x'}|(\mathrm{d}z) &\le \int_{\mathbb{R}^{d+1}} \left| \frac{\rho_\varepsilon(t-s, x-y)}{\mu^\varepsilon_t(x)} - \frac{\rho_\varepsilon(t'-s, x'-y)}{\mu^\varepsilon_{t'}(x')} \right| \nu_{s,y}(\mathrm{d}z)\mu_s(\mathrm{d}y)\mathrm{d}s \\ &\le c_{n,\varepsilon}(|t-t'| + |x-x'|) \int_0^{n+1} \int_{B_{n+1}} \nu_{s,y}(\mathrm{d}z)\mu_s(\mathrm{d}y)\mathrm{d}s. \end{aligned}$$

Estimate (3.9) then follows since $\sup_{s,y \in [0,n+1] \times B_{n+1}} \int_{\mathbb{R}^d} (1 \wedge |z|^2) \nu_{s,y}(\mathrm{d}z) < \infty$. □

3.2 Tightness

We first prepare the following result (cf. [11, Proposition 7.1.8]).

Lemma 3.4 *For $\mu^\varepsilon_0 \in \mathcal{P}(\mathbb{R}^d)$ being defined by (3.2), there exits a function $\psi \in C^2(\mathbb{R}_+)$ with the properties*

$$\psi \ge 0, \quad \psi(0) = 0, \quad 0 < \psi' \le 1, \quad -2 \le \psi'' \le 0, \quad \lim_{r \to \infty} \psi(r) = +\infty,$$

and such that

$$\sup_{\varepsilon \in [0,\ell)} \int_{\mathbb{R}^d} \psi\big(\log(1 + |x|^2)\big) \mu^\varepsilon_0(\mathrm{d}x) < \infty. \tag{3.10}$$

Proof Since μ^ε_0 weakly converges to μ_0 as $\varepsilon \to 0$, we have

$$\lim_{n \to \infty} \sup_{\varepsilon \in [0,\ell)} \mu^\varepsilon_0(B^c_n) = 0.$$

In particular, we can find a subsequence n_k such that for $z_k := \log(1 + n_k^2)$,

$$z_{k+1} - z_k \ge z_k - z_{k-1} \ge 1,$$

and

$$\sup_{\varepsilon\in[0,\ell)}\int_{\mathbb{R}^d} 1_{[z_k,\infty)}(\log(1+|x|^2))\mu_0^\varepsilon(\mathrm{d}x) = \sup_{\varepsilon\in[0,\ell)} \mu_0^\varepsilon(B_{n_k}^c) \le 2^{-k}.$$

Let $z_0 = 0$ and define

$$\psi_0(s) := \sum_{k=0}^{\infty} \mathbf{1}_{[z_k,z_{k+1}]}(s)\left[k-1+\frac{s-z_k}{z_{k+1}-z_k}\right].$$

Clearly, we have

$$\int_{\mathbb{R}^d}\psi_0(\log(1+|x|^2))\mu_0^\varepsilon(\mathrm{d}x) \le \sum_{k=0}^{\infty} k\int_{\mathbb{R}^d} 1_{[z_k,\infty)}(\log(1+|x|^2))\mu_0^\varepsilon(\mathrm{d}x) \le \sum_{k=0}^{\infty}\frac{k}{2^k}.$$

However, ψ_0 does not belong to the class $C^2(\mathbb{R}_+)$. Let us take

$$\psi(t) := \int_0^t g(r)\mathrm{d}r$$

with $g \in C^1(\mathbb{R}_+)$, $0 \le g \le 1$, $-2 \le g' \le 0$, and

$$g(z) = \psi_0'(z) \quad \text{if} \quad z \in (z_k, z_{k+1} - k^{-1}).$$

It is easy to see that such a function g always exists. The proof is complete. □

Lemma 3.5 *Let $H_t^\nu(x, y)$ be defined by (2.3). We have*

$$H_t^\nu(x, y) \le 2(1+|y|)\hbar_t^\nu(x), \quad \forall t \ge 0, x, y \in \mathbb{R}^d. \tag{3.11}$$

Proof Recall that

$$H_t^\nu(x, y) = \int_{B_\ell^c}\log\left(1+\tfrac{|z|}{1+|x-y|}\right)\nu_{t,x}(\mathrm{d}z).$$

If $|x| \le 2|y|$, then

$$\begin{aligned} H_t^\nu(x, y) &\le \int_{B_\ell^c}\log(1+|z|)\,\nu_{t,x}(\mathrm{d}z) \le \int_{B_\ell^c}\log\left(1+\tfrac{(1+2|y|)|z|}{1+|x|}\right)\nu_{t,x}(\mathrm{d}z)\\ &\le \int_{B_\ell^c}\log\left(1+\tfrac{|z|}{1+|x|}\right)^{1+2|y|}\nu_{t,x}(\mathrm{d}z) = (1+2|y|)\hbar_t^\nu(x). \end{aligned}$$

If $|x| > 2|y|$, then $2|x-y| \ge 2|x| - 2|y| \ge |x|$ and

$$H_t^\nu(x, y) \le \int_{B_\ell^c}\log\left(1+\tfrac{2|z|}{2+|x|}\right)\nu_{t,x}(\mathrm{d}z) \le 2\hbar_t^\nu(x).$$

The proof is complete. □

Now, we prove the following tightness result.

Lemma 3.6 *The family of probability measures* $(\mathbb{Q}^\varepsilon)_{\varepsilon\in(0,\ell)}$ *is tight in* $\mathcal{P}(\mathbb{D})$.

Proof By Aldous' criterion (see [1] or [19, p.356]), it suffices to check the following two conditions:

(i) For any $T > 0$, it holds that

$$\lim_{N\to\infty} \sup_{\varepsilon} \mathbb{Q}^\varepsilon \left(\sup_{t\in[0,T]} |X_t| > N \right) = 0.$$

(ii) For any $T, \delta_0 > 0$ and stopping time $\tau < T - \delta_0$, it holds that

$$\lim_{\delta\to 0} \sup_{\varepsilon} \sup_{\tau} \mathbb{Q}^\varepsilon \left(|X_{\tau+\delta} - X_\tau| > \lambda\right) = 0, \quad \forall \lambda > 0.$$

Verification of (i) Let ψ be as in Lemma 3.4 and $V(x) := \psi(\log(1+|x|^2))$. By the definition of martingale solution (see Remark 1.4), (2.2) and (3.6), there is a càdlàg local $\mathbb{Q}^\varepsilon$-martingale M_t^ε and constant C independent of ε such that for all $t \geq 0$,

$$V(X_t) = V(X_0) + \int_0^t \mathscr{L}_r^\varepsilon V(X_r)\mathrm{d}r + M_t^\varepsilon \leq V(X_0) + Ct + M_t^\varepsilon.$$

By Lemma 2.4 , there is a constant $C > 0$ such that for all $T > 0$,

$$\sup_{\varepsilon\in(0,\ell)} \mathbb{E}^{\mathbb{Q}^\varepsilon} \left(\sup_{t\in[0,T]} V^{\frac{1}{2}}(X_t) \right) \leq C \sup_{\varepsilon\in(0,\ell)} (\mathbb{E}^{\mathbb{Q}_\varepsilon} V(X_0))^{\frac{1}{2}} \overset{(3.10)}{<} \infty, \tag{3.12}$$

which in turn implies that (i) is true.
Verification of (ii) Let $\tau \leq T - \delta_0$ be a bounded stopping time. For any $\delta \in (0, \delta_0)$, by the strong Markov property we have

$$\mathbb{Q}^\varepsilon \left(|X_{\tau+\delta} - X_\tau| > \lambda\right) = \mathbb{E}^{\mathbb{Q}^\varepsilon} \left(\mathbb{P}_{s,y}^\varepsilon \left(|X_{s+\delta} - y| > \lambda\right) \big|_{(s,y)=(\tau,X_\tau)} \right). \tag{3.13}$$

Recalling that $V_y(x) := \psi(\log(1+|x-y|^2)$, and by (2.2), (3.6), (3.7), (3.11) and (1.18) we deduce that

$$\begin{aligned}\mathscr{L}_t^\varepsilon V_y(x) &\leq 2\left(\frac{|a_t^\varepsilon(x)| + \langle x-y, b_t^\varepsilon(x)\rangle^+ + g_t^{\nu^\varepsilon}(x)}{1+|x-y|^2} + 2H_t^{\nu^\varepsilon}(x,y) \right) \\ &\leq C\left(\frac{1+|x|^2 + |x-y|(1+|x|)}{1+|x-y|^2} + H_t^\nu(x,y) \right) \leq C(1+|y|^2),\end{aligned}$$

where $C > 0$ is independent of t, x, y and ε. Furthermore, we have

$$V_y(X_t) = V_y(X_s) + \int_s^t \mathscr{L}_r^\varepsilon V_y(X_r)\mathrm{d}r + M_t^\varepsilon$$
$$\leq V_y(X_s) + C(1+|y|^2)(t-s) + M_t^\varepsilon,$$

where $(M_t^\varepsilon)_{t\geq s}$ is a local $\mathbb{P}_{s,y}^\varepsilon$-martingale with $M_s^\varepsilon = 0$. By Lemma 2.4 again and since $V_y(y) = 0$, we obtain

$$\mathbb{E}^{\mathbb{P}_{s,y}^\varepsilon}\left(V_y(X_{s+\delta})^{1/2}\right) \leq C(1+|y|)\delta^{1/2}.$$

Hence,

$$\mathbb{P}_{s,y}^\varepsilon(|X_{s+\delta} - y| > \lambda) = \mathbb{P}_{s,y}^\varepsilon\left(V_y(X_{s+\delta}) > \psi(\log(1+\lambda^2))\right)$$
$$\leq \mathbb{E}^{\mathbb{P}_{s,y}^\varepsilon}\left(V_y(X_{s+\delta})^{1/2}\right)/\psi^{1/2}(\log(1+\lambda^2))$$
$$\leq C(1+|y|)\delta^{1/2}/\psi^{1/2}(\log(1+\lambda^2)),$$

and by (3.13) and (3.12),

$$\mathbb{Q}^\varepsilon(|X_{\tau+\delta} - X_\tau| > \lambda) \leq \mathbb{Q}^\varepsilon(|X_\tau| > R) + C(1+R)\delta^{1/2}/\psi^{1/2}(\log(1+\lambda^2))$$
$$\leq C/\psi^{1/2}(\log(1+R^2)) + C(1+R)\delta^{1/2}/\psi^{1/2}(\log(1+\lambda^2)).$$

Letting $\delta \to 0$ first and then $R \to \infty$, one sees that (ii) is satisfied. □

3.3 Limits

In order to take weak limits, we rewrite

$$\mathscr{B}_t f(x) + \mathscr{N}_t f(x) = \tilde{b}_t(x) \cdot \nabla f(x) + \int_{\mathbb{R}^d} \Theta_f^\pi(x;z)\nu_{t,x}(\mathrm{d}z) =: \widetilde{\mathscr{B}}_t f(x) + \widetilde{\mathscr{N}}_t f(x),$$

where

$$\tilde{b}_t(x) := b_t(x) + \int_{\mathbb{R}^d} \left[\pi(z) - z\mathbf{1}_{|z|\leq\ell}\right]\nu_{t,x}(\mathrm{d}z), \tag{3.14}$$

and

$$\Theta_f^\pi(x;z) := f(x+z) - f(x) - \pi(z)\cdot\nabla f(x). \tag{3.15}$$

Here, $\pi : \mathbb{R}^d \to \mathbb{R}^d$ is a smooth symmetric function satisfying

$$\pi(z) = z, \quad |z| \leq \ell, \quad \pi(z) = 0, \quad |z| > 2\ell.$$

As in (1.11), we shall also write $\widetilde{\mathscr{N}_t} f(x) = \widetilde{\mathscr{N}_t^\nu} f(x) = \widetilde{\mathscr{N}}^{\nu_{t,x}} f(x)$. We have the following result.

Lemma 3.7 *For any $f \in C_c^2(\mathbb{R}^d)$ with support in B_R, there is a constant $C = C(f) > 0$ such that for all $x \in \mathbb{R}^d$ and $z, z' \in \mathbb{R}^d$ with $|z'| \le |z|$,*

$$|\Theta_f^\pi(x; z) - \Theta_f^\pi(x; z')| \le C(|z - z'| \wedge \ell)(\mathbf{1}_{B_{R+\ell}}(x)\mathbf{1}_{|z|\le\ell}|z| + \mathbf{1}_{|z|>\ell\vee(|x|-R)}).$$

Proof Note that

$$\mathscr{Q} := |\Theta_f^\pi(x; z) - \Theta_f^\pi(x; z')| = |f(x + z) - f(x + z') - (\pi(z) - \pi(z')) \cdot \nabla f(x)|.$$

We make the following decomposition:

$$\mathscr{Q} = \mathscr{Q} \cdot \mathbf{1}_{|z|\le\ell} + \mathscr{Q} \cdot \mathbf{1}_{|z|>\ell}\mathbf{1}_{|x|\le R} + \mathscr{Q} \cdot \mathbf{1}_{|z|>\ell}\mathbf{1}_{|x|>R} =: \mathscr{Q}_1 + \mathscr{Q}_2 + \mathscr{Q}_3.$$

For $\mathscr{Q}_1$, since $\mathrm{supp}(f) \subset B_R$ and $|z'| \le |z|$, we have by (1.16) that

$$|\mathscr{Q}_1| \le |z - z'|^2 \|\nabla^2 f\|_\infty \mathbf{1}_{B_{R+\ell}}(x)\mathbf{1}_{|z|\le\ell} \le C(|z - z'| \wedge \ell)|z|\mathbf{1}_{B_{R+\ell}}(x)\mathbf{1}_{|z|\le\ell}.$$

For $\mathscr{Q}_2$, we have

$$\begin{aligned}|\mathscr{Q}_2| &\le \Big(|f(x + z) - f(x + z')| + |\pi(z) - \pi(z')| \cdot \|\nabla f\|_\infty\Big)\mathbf{1}_{|z|>\ell}\mathbf{1}_{|x|\le R} \\ &\le C(|z - z'| \wedge \ell)\mathbf{1}_{|z|>\ell}\mathbf{1}_{|x|\le R}.\end{aligned}$$

As for $\mathscr{Q}_3$, we have

$$|\mathscr{Q}_3| = |f(x + z) - f(x + z')| \cdot \mathbf{1}_{|z|>\ell}\mathbf{1}_{|x|>R} \le C(|z - z'| \wedge \ell)\mathbf{1}_{|z|>\ell\vee(|x|-R)},$$

where we have used that for $|z'| \le |z| \le |x| - R$,

$$f(x + z) = f(x + z') = 0.$$

Combining the above calculations, we obtain the desired estimate. □

The following approximation result will be crucial for taking weak limits.

Lemma 3.8 *For any $\delta \in (0, 1)$ and $R, T > 0$, there is a family of Lévy measures $\eta_{t,x}(\mathrm{d}z)$ such that for any $f \in C_c^2(B_R)$,*

$$\int_0^T \int_{\mathbb{R}^d} \sup_{x\in B_1(y)} |\widetilde{\mathscr{N}}^{\nu_{s,y}} f(x) - \widetilde{\mathscr{N}}^{\eta_{s,y}} f(x)|\mu_s(\mathrm{d}y)\mathrm{d}s \le \delta, \tag{3.16}$$

and

$$\sup_{s,y} \|\widetilde{\mathscr{N}}^{\eta_{s,y}} f\|_\infty < \infty, \quad (s, y, x) \mapsto \widetilde{\mathscr{N}}^{\eta_{s,y}} f(x) \text{ is continuous.}$$

Moreover, there are continuous functions $\bar{a} : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$ and $\bar{b} : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ with compact supports such that

$$\int_0^T \int_{\mathbb{R}^d} \left(\frac{|\bar{a}_s(x) - a_s(x)|}{1 + |x|^2} + \frac{|\bar{b}_s(x) - \tilde{b}_s(x)|}{1 + |x|} \right) \mu_s(\mathrm{d}x)\mathrm{d}s \leq \delta, \tag{3.17}$$

where $\tilde{b}$ is defined by (3.14).

Proof (i) By the randomization of kernel functions (see [18, Lemma 14.50, p.469]), there is a measurable function

$$h_{t,x}(\theta) : [0, T] \times \mathbb{R}^d \times (0, \infty) \to \mathbb{R}^d \cup \{\infty\}$$

such that

$$\nu_{t,x}(A) = \int_0^\infty \mathbf{1}_A(h_{t,x}(\theta))\mathrm{d}\theta, \quad \forall A \in \mathscr{B}(\mathbb{R}^d).$$

In particular, we have

$$\widetilde{\mathscr{N}}^{\nu_{s,y}} f(x) = \int_0^\infty \Theta_f^\pi(x; h_{s,y}(\theta))\mathrm{d}\theta =: \widetilde{\mathscr{N}}^{h_{s,y}} f(x), \tag{3.18}$$

and

$$g_t^\nu(x) = \int_0^\infty \mathbf{1}_{B_\ell}(h_{t,x}(\theta))|h_{t,x}(\theta)|^2\mathrm{d}\theta, \quad \nu_{t,x}(B_\ell^c) = \int_0^\infty \mathbf{1}_{B_\ell^c}(h_{t,x}(\theta))\mathrm{d}\theta.$$

We introduce $\mathbb{X} := [0, T] \times \mathbb{R}^d \times (0, \infty)$ and a locally finite measure γ over $\mathbb{X}$ by

$$\gamma(\mathrm{d}\theta, \mathrm{d}x, \mathrm{d}t) := \varrho_{t,x}(\theta)\mathrm{d}\theta\mu_t(\mathrm{d}x)\mathrm{d}t$$

with $\varrho_{t,x}(\theta) := \mathbf{1}_{B_\ell}(h_{t,x}(\theta))\mathbf{1}_{B_{R+\ell+1}}(x) + \mathbf{1}_{B^c_{\ell\vee(|x|-R-1)}}(h_{t,x}(\theta))$ so that

$$\begin{aligned}\int_{\mathbb{X}} \left(|h_{t,x}(\theta)|^2 \wedge \ell^2\right)\gamma(\mathrm{d}\theta, \mathrm{d}x, \mathrm{d}t) &= \int_0^T \int_{\mathbb{R}^d} g_t^\nu(x)\mathbf{1}_{B_{R+\ell+1}}(x)\mu_t(\mathrm{d}x)\mathrm{d}t \\ &\quad + \ell^2 \int_0^T \int_{\mathbb{R}^d} \nu_{t,x}(B^c_{\ell\vee(|x|-R-1)})\mu_t(\mathrm{d}x)\mathrm{d}t \overset{(1.14)}{<} \infty.\end{aligned} \tag{3.19}$$

□

Claim There is a sequence of measurable functions $\{\bar{h}^n_{t,x}(\theta), n \in \mathbb{N}\}$ so that for each $n \in \mathbb{N}$, $(t, x, \theta) \mapsto \bar{h}^n_{t,x}(\theta)$ is continuous with compact support, and

$$|\bar{h}^n_{t,x}(\theta)| \leq |h_{t,x}(\theta)|, \tag{3.20}$$

and

$$\lim_{n\to\infty}\int_{\mathbb{X}}\left(|\bar{h}^n_{t,x}(\theta)-h_{t,x}(\theta)|^2\wedge\ell^2\right)\gamma(\mathrm{d}\theta,\mathrm{d}x,\mathrm{d}t)=0. \tag{3.21}$$

Proof of Claim Fix $m\in\mathbb{N}$. Since $\mathbf{1}_{(0,m)}(\theta)\gamma(\mathrm{d}\theta,\mathrm{d}x,\mathrm{d}t)$ is a finite measure over $\mathbb{X}$, by Lusin's theorem, there exists a family of continuous functions $\{\bar{h}^\varepsilon_{t,x}(\theta),\varepsilon\in(0,1)\}$ with compact support in (t,x,θ) such that

$$|\bar{h}^\varepsilon_{t,x}(\theta)|\le|h_{t,x}(\theta)|,\ \bar{h}^\varepsilon_{t,x}(\theta)\to h_{t,x}(\theta),\ \varepsilon\to0,\gamma-a.s.$$

Thus by the dominated convergence theorem,

$$\lim_{\varepsilon\to0}\int_{\mathbb{X}}\left(|\bar{h}^\varepsilon_{t,x}(\theta)-h_{t,x}(\theta)|^2\wedge\ell^2\right)\mathbf{1}_{(0,m)}(\theta)\gamma(\mathrm{d}\theta,\mathrm{d}x,\mathrm{d}t)=0.$$

On the other hand, by (3.19) and the monotone convergence theorem, we have

$$\lim_{m\to\infty}\int_{\mathbb{X}}\left(|h_{t,x}(\theta)|^2\wedge\ell^2\right)\mathbf{1}_{[m,\infty)}(\theta)\gamma(\mathrm{d}\theta,\mathrm{d}x,\mathrm{d}t)=0.$$

By a diagonalizaion argument, we obtain the desired approximation sequence. The claim is proven.

(ii) Let $f\in C^2_c(B_R)$. By (3.18), (3.20) and Lemma 3.7, we have for all $x\in B_1(y)$,

$$\begin{aligned}
&|\widetilde{\mathscr{N}}^{h_{s,y}}f(x)-\widetilde{\mathscr{N}}^{\bar{h}^n_{s,y}}f(x)|\le\int_0^\infty|\Theta^\pi_f(x;h_{s,y}(\theta))-\Theta^\pi_f(x;\bar{h}^n_{s,y}(\theta))|\mathrm{d}\theta\\
&\lesssim\int_0^\infty\left(|h_{s,y}(\theta)|\mathbf{1}_{B_\ell}(h_{s,y}(\theta))\mathbf{1}_{B_{R+\ell}}(x)+\mathbf{1}_{B^c_{\ell\vee(|x|-R)}}(h_{s,y}(\theta))\right)\\
&\quad\times\left(|h_{s,y}(\theta)-\bar{h}^n_{s,y}(\theta)|\wedge\ell\right)\mathrm{d}\theta\\
&\le\left(\int_0^\infty\left(|h_{s,y}(\theta)|^2\mathbf{1}_{B_\ell}(h_{s,y}(\theta))\mathbf{1}_{B_{R+\ell+1}}(y)+\mathbf{1}_{B^c_{\ell\vee(|y|-R-1)}}(h_{s,y}(\theta))\right)\mathrm{d}\theta\right)^{\frac12}\\
&\quad\times\left(\int_0^\infty\left(|h_{s,y}(\theta)-\bar{h}^n_{s,y}(\theta)|^2\wedge\ell^2\right)\varrho_{s,y}(\theta)\mathrm{d}\theta\right)^{\frac12}\\
&=\left(\mathbf{1}_{B_{R+\ell+1}}(y)g^\nu_s(y)+\nu_{s,y}(B^c_{\ell\vee(|y|-R-1)})\right)^{\frac12}\\
&\quad\times\left(\int_0^\infty\left(|h_{s,y}(\theta)-\bar{h}^n_{s,y}(\theta)|^2\wedge\ell^2\right)\varrho_{s,y}(\theta)\mathrm{d}\theta\right)^{\frac12}.
\end{aligned}$$

Hence, by (1.18) and (1.21) we further have

$$\int_0^T\int_{\mathbb{R}^d}\sup_{x\in B_1(y)}|\widetilde{\mathscr{N}}^{h_{s,y}}f(x)-\widetilde{\mathscr{N}}^{\bar{h}^n_{s,y}}f(x)|\mu_s(\mathrm{d}y)\mathrm{d}s$$

$$\lesssim \left(\int_0^T \int_{\mathbb{R}^d} \int_0^\infty \left(|h_{s,y}(\theta) - \bar{h}^n_{s,y}(\theta)|^2 \wedge \ell^2 \right) \varrho_{s,y}(\theta) \mathrm{d}\theta \mu_s(\mathrm{d}y) \mathrm{d}s \right)^{\frac{1}{2}}$$
$$= \left(\int_{\mathbb{X}} (|h_{s,y}(\theta) - \bar{h}^n_{s,y}(\theta)|^2 \wedge \ell^2) \gamma(\mathrm{d}\theta, \mathrm{d}y, \mathrm{d}s) \right)^{\frac{1}{2}} \overset{(3.21)}{\to} 0.$$

(iii) For fixed $n \in \mathbb{N}$, since $f \in C^2_c(B_R)$, by the above claim that $(s, y, \theta) \mapsto \bar{h}^n_{s,y}(\theta)$ is continuous and has compact support, and the dominated convergence theorem, we have that

$$(s, y, x) \mapsto \widetilde{\mathscr{N}}^{\bar{h}^n_{s,y}} f(x) = \int_0^\infty \Theta^\pi_f(x; \bar{h}^n_{s,y}(\theta)) \mathrm{d}\theta \text{ is continuous.}$$

Moreover, we have

$$|\widetilde{\mathscr{N}}^{\bar{h}^n_{s,y}} f(x)| \le \int_0^\infty |\Theta^\pi_f(x; \bar{h}^n_{s,y}(\theta))| \mathrm{d}\theta \le C \int_0^\infty \left(|\bar{h}^n_{s,y}(\theta)|^2 \wedge 1 \right) \mathrm{d}\theta.$$

Since $\bar{h}^n_{s,y}(\theta)$ has compact support in (s, y), we have

$$\sup_{s,y} \|\widetilde{\mathscr{N}}^{\bar{h}^n_{s,y}} f\|_\infty < \infty.$$

Finally we only need to take n large enough and define

$$\eta_{t,x}(A) := \int_0^\infty \mathbf{1}_A(\bar{h}^n_{t,x}(\theta)) \mathrm{d}\theta.$$

(iv) Now let us show (3.17). By Lusin's theorem, the set of continuous functions with compact supports is dense in $L^1([0, T] \times \mathbb{R}^d, \mu_t(\mathrm{d}x)\mathrm{d}t)$. Since $\frac{a_t(x)}{1+|x|^2}$ and $\frac{\tilde{b}_t(x)}{1+|x|}$ are bounded by (1.18), the existence of $\bar{a}$ and $\bar{b}$ with property (3.17) follows. □

Now we are in a position to give:

Proof of Theorem 1.5 Let $\mathbb{Q}$ be any accumulation point of $(\mathbb{Q}^\varepsilon)_{\varepsilon \in (0,\ell)}$ (see Lemma 3.6). By taking weak limits for

$$\mu^\varepsilon_t = \mathbb{Q}^\varepsilon \circ X_t^{-1},$$

we obtain

$$\mu_t = \mathbb{Q} \circ X_t^{-1}.$$

It remains to show that $\mathbb{Q} \in \mathcal{M}^{\mu_0}_0(\mathscr{L}_t)$. We need to show that for any $f \in C^2_c(\mathbb{R}^d)$,

$$M_t := f(X_t) - f(X_0) - \int_0^t \mathscr{L}_s f(X_s) \mathrm{d}s$$

is a $\mathcal{B}_t$-martingale under $\mathbb{Q}$. Let $J := \{t \geq 0 : \mathbb{Q}(\Delta X_t \neq 0) > 0\}$, which is a countable subset of $\mathbb{R}_+$. Since $t \mapsto M_t$ is right continuous and bounded, to show that M_t is a $\mathcal{B}_t$-martingale under $\mathbb{Q}$, it suffices to prove that for any $s < t \notin J$ and any bounded $\mathcal{B}_s$-measurable continuous functional g_s on $\mathbb{D}$,

$$\mathbb{E}^{\mathbb{Q}}(M_t g_s) = \mathbb{E}^{\mathbb{Q}}(M_s g_s).$$

Since $\mathbb{Q}^\varepsilon \in \mathscr{M}_0^{\mu_0^\varepsilon}(\mathscr{L}^\varepsilon)$, by the definition of martingale solution, we have

$$\mathbb{E}^{\mathbb{Q}^\varepsilon}(M_t^\varepsilon g_s) = \mathbb{E}^{\mathbb{Q}^\varepsilon}(M_s^\varepsilon g_s),$$

where

$$M_t^\varepsilon := f(X_t) - f(X_0) - \int_0^t \mathscr{L}_s^\varepsilon f(X_s)\mathrm{d}s.$$

Since $\lim_{\varepsilon\to 0} \mathbb{E}^{\mathbb{Q}^\varepsilon}(f(X_t)g_s) = \mathbb{E}^{\mathbb{Q}}(f(X_t)g_s)$ for $t \notin J$ (see [19, Proposition 3.4, page 349]), we only need to show the following three limits:

$$\lim_{\varepsilon\to 0} \mathbb{E}^{\mathbb{Q}^\varepsilon}\left(g_s \int_s^t \mathscr{A}_r^\varepsilon f(X_r)\mathrm{d}r\right) = \mathbb{E}^{\mathbb{Q}}\left(g_s \int_s^t \mathscr{A}_r f(X_r)\mathrm{d}r\right), \tag{3.22}$$

$$\lim_{\varepsilon\to 0} \mathbb{E}^{\mathbb{Q}^\varepsilon}\left(g_s \int_s^t \widetilde{\mathscr{B}}_r^\varepsilon f(X_r)\mathrm{d}r\right) = \mathbb{E}^{\mathbb{Q}}\left(g_s \int_s^t \widetilde{\mathscr{B}}_r f(X_r)\mathrm{d}r\right), \tag{3.23}$$

$$\lim_{\varepsilon\to 0} \mathbb{E}^{\mathbb{Q}^\varepsilon}\left(g_s \int_s^t \widetilde{\mathscr{N}}_r^{\nu^\varepsilon} f(X_r)\mathrm{d}r\right) = \mathbb{E}^{\mathbb{Q}}\left(g_s \int_s^t \widetilde{\mathscr{N}}_r^{\nu} f(X_r)\mathrm{d}r\right). \tag{3.24}$$

Below we assume that the support of f is contained in the ball B_R. Let us first show (3.24). Fix $\delta \in (0, 1)$. Let $\eta_{t,x}(\mathrm{d}z)$ be as given by Lemma 3.8, and recall that ν^ε is defined by (3.5). We write

$$\begin{aligned}
&\left|\mathbb{E}^{\mathbb{Q}^\varepsilon}\left(g_s \int_s^t \widetilde{\mathscr{N}}_r^{\nu^\varepsilon} f(X_r)\mathrm{d}r\right) - \mathbb{E}^{\mathbb{Q}}\left(g_s \int_s^t \widetilde{\mathscr{N}}_r^{\nu} f(X_r)\mathrm{d}r\right)\right| \\
&\leq \left|\mathbb{E}^{\mathbb{Q}^\varepsilon}\left(g_s \int_s^t \widetilde{\mathscr{N}}_r^{\nu^\varepsilon} f(X_r)\mathrm{d}r\right) - \mathbb{E}^{\mathbb{Q}^\varepsilon}\left(g_s \int_s^t \widetilde{\mathscr{N}}_r^{\eta_\varepsilon} f(X_r)\mathrm{d}r\right)\right| \\
&\quad + \left|\mathbb{E}^{\mathbb{Q}^\varepsilon}\left(g_s \int_s^t \widetilde{\mathscr{N}}_r^{\eta^\varepsilon} f(X_r)\mathrm{d}r\right) - \mathbb{E}^{\mathbb{Q}^\varepsilon}\left(g_s \int_s^t \widetilde{\mathscr{N}}_r^{\eta} f(X_r)\mathrm{d}r\right)\right| \\
&\quad + \left|\mathbb{E}^{\mathbb{Q}^\varepsilon}\left(g_s \int_s^t \widetilde{\mathscr{N}}_r^{\eta} f(X_r)\mathrm{d}r\right) - \mathbb{E}^{\mathbb{Q}}\left(g_s \int_s^t \widetilde{\mathscr{N}}_r^{\eta} f(X_r)\mathrm{d}r\right)\right| \\
&\quad + \left|\mathbb{E}^{\mathbb{Q}}\left(g_s \int_s^t \widetilde{\mathscr{N}}_r^{\eta} f(X_r)\mathrm{d}r\right) - \mathbb{E}^{\mathbb{Q}}\left(g_s \int_s^t \widetilde{\mathscr{N}}_r^{\nu} f(X_r)\mathrm{d}r\right)\right| =: \sum_{i=1}^4 I_i(\varepsilon),
\end{aligned}$$

where η^ε is defined similarly as in (3.5) with ν being replaced by η. For $I_1(\varepsilon)$, by definition, we have

$$\begin{aligned}I_1(\varepsilon) &\le \|g_s\|_\infty \mathbb{E}^{\mathbb{Q}^\varepsilon}\left(\int_s^t |\widetilde{\mathscr{N}_r^{\nu^\varepsilon}} f(X_r) - \widetilde{\mathscr{N}_r^{\eta^\varepsilon}} f(X_r)|\mathrm{d}r\right)\\ &= \|g_s\|_\infty \int_s^t \int_{\mathbb{R}^d} |\widetilde{\mathscr{N}_r^{\nu^\varepsilon}} f(x) - \widetilde{\mathscr{N}_r^{\eta^\varepsilon}} f(x)|\mu_r^\varepsilon(x)\mathrm{d}x\mathrm{d}r\\ &= (1-\varepsilon)\|g_s\|_\infty \int_s^t \int_{\mathbb{R}^d} |\widetilde{\mathscr{N}_r^{\bar\nu^\varepsilon}} f(x) - \widetilde{\mathscr{N}_r^{\bar\eta^\varepsilon}} f(x)|\mathrm{d}x\mathrm{d}r\\ &= (1-\varepsilon)\|g_s\|_\infty \int_s^t \int_{\mathbb{R}^d} \left|\int_{\mathbb{R}^d} \Theta_f^\pi(x;z)(\bar\nu_{r,x}^\varepsilon - \bar\eta_{r,x}^\varepsilon)(\mathrm{d}z)\right| \mathrm{d}x\mathrm{d}r,\end{aligned}$$

where $\Theta_f^\pi(x;z)$ is defined by (3.15) and

$$\bar\nu_{r,x}^\varepsilon(\mathrm{d}z) := \int_{\mathbb{R}^{d+1}} \rho_\varepsilon(r-s, x-y)\nu_{s,y}(\mathrm{d}z)\mu_s(\mathrm{d}y)\mathrm{d}s.$$

By Fubini's theorem we further have

$$\begin{aligned}I_1(\varepsilon) &\le \|g_s\|_\infty \int_s^t \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^{d+1}} \rho_\varepsilon(r-s, x-y)\right.\\ &\quad \times \left.\left(\widetilde{\mathscr{N}}^{\nu_{s,y}} f(x) - \widetilde{\mathscr{N}}^{\eta_{s,y}} f(x)\right)\mu_s(\mathrm{d}y)\mathrm{d}s\right|\mathrm{d}x\mathrm{d}r\\ &\le \|g_s\|_\infty \int_0^T \int_{\mathbb{R}^d} \sup_{x\in B_1(y)} |\widetilde{\mathscr{N}}^{\nu_{s,y}} f(x) - \widetilde{\mathscr{N}}^{\eta_{s,y}} f(x)|\mu_s(\mathrm{d}y)\mathrm{d}s \overset{(3.16)}{\le} \|g_s\|_\infty\delta.\end{aligned}$$

For $I_2(\varepsilon)$, recalling (3.2), we have

$$\begin{aligned}I_2(\varepsilon) &\le \|g_s\|_\infty \mathbb{E}^{\mathbb{Q}^\varepsilon}\left(\int_s^t |\widetilde{\mathscr{N}_r^{\eta^\varepsilon}} f(X_r) - \widetilde{\mathscr{N}_r^{\eta}} f(X_r)|\mathrm{d}r\right)\\ &= \|g_s\|_\infty \int_s^t \int_{\mathbb{R}^d} |\widetilde{\mathscr{N}_r^{\eta^\varepsilon}} f(x) - \widetilde{\mathscr{N}_r^{\eta}} f(x)|\mu_r^\varepsilon(x)\mathrm{d}x\mathrm{d}r\\ &= \|g_s\|_\infty \int_s^t \int_{\mathbb{R}^d} |(1-\varepsilon)\widetilde{\mathscr{N}_r^{\bar\eta^\varepsilon}} f(x) - \mu_r^\varepsilon(x)\widetilde{\mathscr{N}_r^{\eta}} f(x)|\mathrm{d}x\mathrm{d}r\\ &\le (1-\varepsilon)\|g_s\|_\infty \int_s^t \int_{\mathbb{R}^d} \int_{\mathbb{R}^{d+1}} \rho_\varepsilon(r-s, x-y)\\ &\quad \times |\widetilde{\mathscr{N}}^{\eta_{s,y}} f(x) - \widetilde{\mathscr{N}}^{\eta_{r,x}} f(x)|\mu_s(\mathrm{d}y)\mathrm{d}s\mathrm{d}x\mathrm{d}r\\ &\quad + \varepsilon\|g_s\|_\infty \int_s^t \int_{\mathbb{R}^d} |\phi(x)\widetilde{\mathscr{N}_r^{\eta}} f(x)|\mathrm{d}x\mathrm{d}r.\end{aligned}$$

Since $(s, y, x) \mapsto \widetilde{\mathscr{N}}^{\eta_{s,y}} f(x)$ is continuous and $\|\widetilde{\mathscr{N}}^{\eta} f\|_\infty < \infty$, by the dominated convergence theorem, we get

$$\lim_{\varepsilon \to 0} I_2(\varepsilon) = 0.$$

Concerning $I_3(\varepsilon)$, it follows by the definition of weak convergence that

$$\lim_{\varepsilon \to 0} I_3(\varepsilon) = 0.$$

For $I_4(\varepsilon)$, we have

$$I_4(\varepsilon) \le \|g_s\|_\infty \int_s^t \int_{\mathbb{R}^d} |\widetilde{\mathscr{N}}_r^{\eta} f(x) - \widetilde{\mathscr{N}}_r^{\nu} f(x)| \mu_r(\mathrm{d}x)\mathrm{d}r \overset{(3.16)}{\le} \|g_s\|_\infty \delta.$$

Since δ is arbitrary, combining the above calculations, we obtain (3.24). The proofs for (3.22) and (3.23), by (3.17), are completely the same as above. The proof is complete. □

4 Proof of Theorem 1.13

Let u be the unique weak solution of FPME (1.23) given by Theorem 1.11 with initial value $\varphi \ge 0$ being bounded and $\int_{\mathbb{R}^d} \varphi(x)\mathrm{d}x = 1$. Let

$$\sigma_t(x) := |u(t,x)|^{\frac{m-1}{\alpha}}, \quad \kappa_t(x) := u(t,x)^{m-1}, \quad \nu_{t,x}(\mathrm{d}z) := \tfrac{\kappa_t(x)\mathrm{d}z}{|z|^{d+\alpha}}.$$

By the change of variable we have

$$\nu_{t,x}(A) = \int_{\mathbb{R}^d} \mathbf{1}_A(\sigma_t(x)z)\frac{\mathrm{d}z}{|z|^{d+\alpha}}, \quad A \in \mathscr{B}(\mathbb{R}^d \setminus \{0\}), \tag{4.1}$$

and

$$\mathscr{N}_t f(x) := \mathrm{P.V.} \int_{\mathbb{R}^d} (f(x + \sigma_t(x)z) - f(x))\frac{\mathrm{d}z}{|z|^{d+\alpha}} = \kappa_t(x)\Delta^{\alpha/2},$$

where the second equality is due to (1.24). By Definition 1.10 it is easy to see that $u(t, x)$ solves the following non-local FPKE:

$$\partial_t u = \mathscr{N}_t^* u, \quad u(0,x) = \varphi(x),$$

that is, for every $t > 0$ and $f \in C_0^2(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^d} f(x)u(t,x)\mathrm{d}x = \int_{\mathbb{R}^d} f(x)\varphi(x)\mathrm{d}x + \int_0^t \int_{\mathbb{R}^d} \kappa_s(x)\Delta^{\alpha/2} f(x)u(s,x)\mathrm{d}x\mathrm{d}s.$$

Note that for each $t > 0$,

$$|\sigma_t(x)| = |u(t,x)|^{\frac{m-1}{\alpha}} \leq \|\varphi\|_\infty^{\frac{m-1}{\alpha}}.$$

Thus, by Example 1.8 with the above $\nu_{t,x}$ and Theorem 1.5 with $\mu_0(\mathrm{d}x) = \varphi(x)\mathrm{d}x$, there is a martingale solution $\mathbb{P} \in \mathscr{M}_0^{\mu_0}(\mathcal{N}_t)$ so that

$$\mathbb{P} \circ X_t^{-1}(\mathrm{d}x) = u(t,x)\mathrm{d}x, \quad t \geq 0.$$

By (4.1) and [19, Theorem 2.26, p.157] (see Remark 4.1 below), there are a stochastic basis $(\Omega, \mathcal{F}, \mathbf{P}; (\mathcal{F}_t)_{t\geq 0})$ and a Poisson random measure N on $\mathbb{R}^d \times [0,\infty)$ with intensity $|z|^{-d-\alpha}\mathrm{d}z\mathrm{d}t$, as well as an $\mathcal{F}_t$-adapted càdlàg process Y_t such that

$$\mathbf{P} \circ Y_t^{-1}(\mathrm{d}x) = \mathbb{P} \circ X_t^{-1}(\mathrm{d}x), \quad t \geq 0,$$

and

$$\mathrm{d}Y_t = \int_{|z|\leq 1} \sigma_t(Y_{t-}) z \tilde{N}(\mathrm{d}z, \mathrm{d}t) + \int_{|z|>1} \sigma_t(Y_{t-}) z N(\mathrm{d}z, \mathrm{d}t),$$

where $\tilde{N}(\mathrm{d}z, \mathrm{d}t) := N(\mathrm{d}z, \mathrm{d}t) - |z|^{-d-\alpha}\mathrm{d}z\mathrm{d}t$. Finally we just need to define

$$L_t := \int_0^t \int_{|z|\leq 1} z \tilde{N}(\mathrm{d}z, \mathrm{d}s) + \int_0^t \int_{|z|>1} z N(\mathrm{d}z, \mathrm{d}s),$$

then L is a d-dimensional isotropic α-stable process with Lévy measure $\mathrm{d}z/|z|^{d+\alpha}$, and

$$\mathrm{d}Y_t = \sigma_t(Y_{t-})\mathrm{d}L_t.$$

The proof is finished. □

Remark 4.1 For a more recent general analysis on the equivalence of stochastic equations and martingale problem, we refer to [21].

Acknowledgements The authors are very grateful to the referees for their quite useful suggestions.

References

1. Aldous, D.: Stopping times and tightness. Ann. Probab. **6**, 335–340 (1978)
2. Ambrosio, L.: Transport equation and Cauchy problem for BV vector fields. Invent. Math. **158**, 227–260 (2004)
3. Ambrosio, L.: Transport equation and Cauchy problem for non-smooth vector fields. Lect. Notes Math. **1927**, 2–41 (2008)
4. Ambrosio, L., Figalli, A., Friesecke, G., Giannoulis, J., Paul, T.: Semiclassical limit of quantum dynamics with rough potentials and well-posedness of transport equations with measure initial data. Commun. Pure Appl. Math. **64**, 1199–1242 (2011)

5. Ambrosio, L., Trevisan, D.: Lecture notes on the DiPerna–Lions theory in abstract measure spaces. Ann. Fac. Sci. Toulouse Math. **6**, 729–766 (2017)
6. Barbu, V., Röckner, M.: Probabilistic representation for solutions to nonlinear Fokker–Planck equations. SIAM J. Math. Anal. **50**, 4246–4260 (2018)
7. Barbu V. and Röckner M.: From nonlinear Fokker–Planck equations to solutions of distribution dependent SDE. Ann. Probab. (To appear) arXiv:1808.10706
8. Barbu, V., Röckner, M., Russo, F.: Probabilistic representation for solutions of an irregular porous media type equation: the degenerate case. Probab. Theory Relat. Fields **15**, 1–43 (2011)
9. Belaribi, N., Russo, F.: Uniqueness for Fokker–Planck equations with measurable coefficients and applications to the fast diffusion equation. Electron. J. Probab. **17**, 1–28 (2012)
10. Blanchard, Ph, Röckner, M., Russo, F.: Probabilistic representation for solutions of an irregular porous media type equation. Ann. Probab. **38**, 1870–1900 (2010)
11. Bogachev, V.I., Krylov, N.V., Röckner, M., Shaposhnikov, S.V.: Fokker–Planck–Kolmogorov Equations. Mathematical Surveys and Monographs, vol. 207. American Mathematical Society, Providence (2015)
12. Bogachev V. I., Röckner M. and Shaposhnikov S. V.: On the Ambrosio-Figalli-Trevisan superposition principle for probability solutions to Fokker–Planck–Kolmogorov equations. arXiv:1903.10834v1
13. Carmona, R., Delarue, F.: Probabilistic Theory of Mean Field Games with Applications. II. Mean Field Games With Common Noise and Master Equations. Probability Theory and Stochastic Modeling, vol. 84. Springer, Berlin (2018)
14. Ethier, S.N., Kurtz, T.G.: Markov Processes, Characterization and Convergence. Wiley, New York (1986)
15. Figalli, A.: Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients. J. Funct. Anal. **254**, 109–153 (2008)
16. Fournier, N., Xu, L.: On the equivalence between some jumping SDEs with rough coefficients and some non-local PDEs. Ann. Inst. Henri Poincaré Probab. Stat. **55**, 1163–1178 (2019)
17. Horowitz, J., Karandikar, R.: Martingale problems associated with the Boltzmann equation. Seminar on Stochastic Processes, 1989 (San Diego, CA, 1989), Volume 18 of Progress in Probability, pp. 75–122. Birkhäuser, Boston (1990)
18. Jacod, J.: Calcul Stochastique et Problèmes de Martingales, Lecture Notes in Mathematics, vol. 714. Springer, Berlin (1979)
19. Jacod, J., Shiryaev, A.: Limit Theorems For Stochastic Processes. Springer, Berlin (1987)
20. Kurtz, T.G.: Martingale problems for conditional distributions of Markov processes. Electron. J. Probab. **3**, 1–29 (1998)
21. Kurtz T. G.: Equivalence of stochastic equations and martingale problems. In: Stochastic Analysis 2010, pp. 113–130. Springer, New York (2011)
22. Pablo, A., Quirós, F., Rodrïguez, A., Vázquez, J.: A general fractional porous medium equation. Commun. Pure Appl. Math. **65**, 1242–1284 (2012)
23. Ren, J., Röckner, M., Wang, F.-Y.: Stochastic generalized porous media and fast diffusion equations. J. Differ. Equ. **238**, 118–152 (2007)
24. Revuz, D., Yor, M.: Continuous Martingales and Brownian Motion, Grundlehren der Mathematischen Wissenschaften, vol. 293, 3rd edn. Springer, New York (2005)
25. Röckner, M., Zhang, X.: Weak uniqueness of Fokker-Planck equations with degenerate and bounded coefficients. C. R. Math. **348**, 435–438 (2010)
26. Röckner M. and X. Zhang: Well-posedness of distribution dependent SDEs with singular drifts. arXiv:1809.02216
27. Scheutzow, M.: A stochastic Gronwall's lemma. Infin. Dimens. Anal., Quantum Probab. Relat. Top **16**, 1350019 (2013)
28. Schilling, R.L.: Conservativeness and extensions of feller semigroups. Positivity **2**, 239–256 (1998)
29. Stepanov, E., Trevisan, D.: Three superposition principles: currents, continuity equations and curves of measures. J. Funct. Anal. **272**, 1044–1103 (2017)
30. Stroock, D.W.: Diffusion processes associated with Lévy generators. Z. Wahr. Verw. Gebiete **32**, 209–244 (1975)
31. Strook, D.W., Varadhan, S.R.S.: Multidimensional Diffusion Processes. Springer, Berlin (2006)
32. Trevisan, D.: Well-posedness of multidimensional diffusion processes with weakly differentiable coefficients. Electron. J. Probab. **21**, 1–41 (2016)

33. Vázquez, J.L.: Recent progress in the theory of nonlinear diffusion with fractional Laplacian operators. Discrete Contin. Dyn. Syst. **7**, 857–885 (2014)
34. Vázquez J. L.: The mathematical theories of diffusion: nonlinear and fractional diffusion. In: Nonlocal and Nonlinear Diffusions and Interactions: New Methods and Directions, Volume 2186 of Lecture Notes in Math., pp. 205–278. Springer, New York (2017)
35. Xie, L., Zhang, X.: Ergodicity of stochastic differential equations with jumps and singular coefficients. Ann. Inst. Henri Poincaré Probab. Stat. **56**(1), 175–229 (2020)
36. Zhang, X.: Degenerate irregular SDEs with jumps and application to integro-differential equations of Fokker–Planck type. Electron. J. Probab. **18**, 1–25 (2013)