Uniqueness for nonlinear Fokker-Planck equations with general diffusion terms and their associated nonlinear Markov processes

Viorel Barbu [*] Yuqi Li [†] Michael Röckner[†] [‡]

July 16, 2026

Abstract

This work is concerned with the uniqueness of distributional solutions to nonlinear Fokker-Planck equations with non-diagonal diffusion terms of type

$$u_t - \sum_{i,j=1}^{d} D_{ij}^2(a_{ij}(x)\beta(x,u)) + \operatorname{div}(b(x,u)u) = 0 \quad \text{in } (0,\infty)\times\mathbb{R}^d,$$

with initial condition $u(0,x) \equiv u_0(x)$, where a_{ij}, β, and b are suitable functions. Under suitable assumptions, this equation generates a continuous contraction semigroup $S(t) : L^1(\mathbb{R}^d) \to L^1(\mathbb{R}^d)$, and $u(t) = S(t)u_0$ is a mild solution to the equation. Our main contribution is to prove that this mild solution is unique in the much larger class of distributional solutions. This extends previous uniqueness results for the diagonal (also called isotropic) diffusion case $a_{ij} \equiv \delta_{ij}$. Another key analytical result of this paper is the uniqueness for distributional solutions of the associated linearized equation. As a main application, we prove weak uniqueness for the corresponding McKean-Vlasov SDEs. Furthermore, we establish a new L^∞ estimate for mild solutions starting from data in $L^1 \cap L^\infty$ and this estimate is used in the construction of nonlinear Markov processes. Finally, we prove that the path laws of the solutions to the McKean-Vlasov SDEs form a nonlinear Markov process in the sense of McKean.

Keywords: nonlinear Fokker-Planck equation, McKean-Vlasov stochastic differential equation, distributional solution, linearized uniqueness, weak uniqueness, nonlinear Markov processes.

---

[*]Al.I. Cuza University and Octav Mayer Institute of Mathematics of Romanian Academy, Iaşi, Romania

[†]Faculty of Mathematics, Bielefeld University, D-33501 Bielefeld, Germany

[‡]Academy for Mathematics and Systems Science, CAS, Beijing, China and School of Data Science, The Chinese University of Hong Kong, Shenzhen (CUHK-Shenzhen), China

1 Introduction

We study nonlinear Fokker-Planck equations of the form

$$\begin{aligned} &\frac{\partial}{\partial t}u(t,x) - \sum_{i,j=1}^{d} D_{ij}^2(a_{ij}(x,u(t,x))u(t,x)) + \operatorname{div}(b(x,u(t,x))u(t,x)) = 0, \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad t>0,\ x\in\mathbb{R}^d, \\ &u(0,x) = u_0(x),\ x\in\mathbb{R}^d, \end{aligned} \tag{1}$$

for diffusion terms of the type $a_{ij}(x,u) = a_{ij}(x)\beta(x,u)/u$. Thus the equation considered in this paper is

$$\begin{aligned} &\frac{\partial}{\partial t}u(t,x) - \sum_{i,j=1}^{d} D_{ij}^2(a_{ij}(x)\beta(x,u(t,x))) + \operatorname{div}(b(x,u(t,x))u(t,x)) = 0, \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad t>0,\ x\in\mathbb{R}^d, \\ &u(0,x) = u_0(x),\ x\in\mathbb{R}^d, \end{aligned} \tag{2}$$

where $a_{ij} : \mathbb{R}^d \to \mathbb{R}$, $\beta : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$, $b : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$ are given functions. We study the existence of mild solutions (see Definition 2.1 below) and, more importantly, the uniqueness of distributional solutions (see Definition 2.2 below) for (2) under the following hypotheses on the coefficients a_{ij}, β, and b.
Hypothesis (I)
(***i***) $\beta \in C^1(\mathbb{R}^d \times \mathbb{R})$, $\beta_r(x,r) \geqslant 0$ for all $x \in \mathbb{R}^d$ and $r \in \mathbb{R}$; $\beta(x,0) \equiv 0$ for all $x \in \mathbb{R}^d$; $\beta_r \in L^\infty(\mathbb{R}^d \times (-N,N))$ for every $N > 0$.
(***ii***) $a_{ij} \in C^\infty(\mathbb{R}^d) \cap C_b(\mathbb{R}^d)$, $(a_{ij})_x \in C_b(\mathbb{R}^d;\mathbb{R}^d)$, $a_{ij} = a_{ji}$, and $\exists\gamma > 0$ such that

$$\sum_{i,j=1}^{d} a_{ij}(x)\xi_i\xi_j \geqslant \gamma\,|\xi|^2\,, \quad \forall \xi \in \mathbb{R}^d, x \in \mathbb{R}^d.$$

(***iii***) $b \in C_b(\mathbb{R}^{d+1};\mathbb{R}^d)$ and $b(x,\cdot) \in C^1(\mathbb{R};\mathbb{R}^d)$ for all $x \in \mathbb{R}^d$.
(***iv***) For each compact $K \subseteq \mathbb{R}$, $\exists \alpha_K \in (0,\infty)$ such that

$$|b(x,r_1)r_1 - b(x,r_2)r_2| \leqslant \alpha_K\,|\beta(x,r_1) - \beta(x,r_2)|\,, \quad \forall r_1,r_2 \in K, \forall x \in \mathbb{R}^d.$$

Here, $b(x,u) = \{b^i(x,u)\}_{i=1}^d$, $(a_{ij})_x = \nabla_x a_{ij}$.

Remark 1.1 If $\sup\limits_{x\in\mathbb{R}^d, r\in K} |b_r(x,r)| < \infty$, $\inf\limits_{x\in\mathbb{R}^d, r\in K} \beta_r(x,r) > 0$ for every compact $K \subseteq \mathbb{R}$, where $b_r(x,r) = \dfrac{\partial}{\partial r}b(x,r)$ and $\beta_r(x,r) = \dfrac{\partial}{\partial r}\beta(x,r)$, then condition (***iv***) can be omitted, because it then follows from (***iii***).

Nonlinear Fokker-Planck equations (abbreviated NFPEs) arise naturally in mean-field game theory and statistical mechanics and have also found numerous applications in areas such as information theory, graph theory, and economics (see, for instance, [17], [19], [21] and the references therein). Beyond these applications, nonlinear Fokker-Planck equations provide a powerful framework for the analysis of nonlinear diffusion phenomena and stochastic processes. A particularly important characteristic of nonlinear Fokker-Planck equations is their connection with McKean-Vlasov stochastic differential equations (also called distribution-dependent stochastic differential equations, abbreviated DDSDEs or McKean-Vlasov SDEs), whose coefficients depend not only on the solution process, but also on the one-dimensional time-marginal law of the solution. More precisely, (2) can be used to solve the following McKean-Vlasov stochastic differential equation

$$\begin{aligned} &dX(t) = b(X(t), u(t, X(t)))dt + \sqrt{\frac{2\beta(X(t), u(t, X(t)))}{u(t, X(t))}}\sigma(X(t))dW(t), \forall t > 0, \\ &\mathcal{L}_{X(0)}(dx) = u_0(x)dx, \quad \mathcal{L}_{X(t)}(dx) = u(t, x)dx, \quad t > 0, \end{aligned} \tag{3}$$

on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with normal filtration $(\mathscr{F}_t)_{t\geqslant 0}$ and an $(\mathscr{F}_t)$-Brownian motion $W(t)$ on it with values in $\mathbb{R}^d$. Here $\sigma = (\sigma_{ij})_{1\leqslant i,j\leqslant d}$ with $(\sigma\sigma^T)_{ij} = a_{ij}$, and $\mathcal{L}_{X_t}$ denotes the law of X_t under $\mathbb{P}$, and $u_0 \in \mathcal{P}$ is given.

Equation (2) extends the isotropic diffusion model, i.e. where $a_{ij} \equiv \delta_{ij}$, $\beta(x, u) \equiv \beta(u)$, and $b(x, u) = D(x)b(u)$, namely

$$\begin{aligned} &\frac{\partial}{\partial t}u(t, x) - \Delta\beta(u(t, x)) + \operatorname{div}(D(x)b(u(t, x))u(t, x)) = 0, \quad t > 0, x \in \mathbb{R}^d, \\ &u(0, x) = u_0(x), \quad x \in \mathbb{R}^d. \end{aligned} \tag{4}$$

This isotropic case can also be viewed as an extension of the classical Smoluchowski equation and has been studied extensively by V. Barbu and M. Röckner. In [8], they proved the existence of mild solutions to (4) and of weak solutions to the associated McKean-Vlasov SDEs. They also proved existence for measures of bounded variation as initial data and established a smoothing effect for the initial data. Building on this existence theory, they further developed the uniqueness theory, in particular, they proved the existence and uniqueness of mild solutions and the uniqueness of distributional solutions in [7, 8, 12], as well as the uniqueness of weak solutions to the corresponding McKean-Vlasov SDEs. Additionally, they considered equation (4) with the fractional Laplacian $(-\Delta)^s$, $0 < s < 1$, in the diffusion term, namely $(-\Delta)^s\beta(u(t, x))$, and developed the existence and uniqueness theory for nonlinear Fokker-Planck equations and McKean-Vlasov SDEs with Lévy noise [13]. Recently, in [3], V. Barbu extended the well-posedness results for (4) (and for the associated McKean-Vlasov SDEs) to a more general class of nonlinear Fokker-Planck equations with singular integral drifts $\operatorname{div}((D(x)b(u) + K * u)u))$, where $K : \mathbb{R}^d \to \mathbb{R}^d$ is a kernel and $*$ denotes convolution. For nonlinear Fokker-Planck equations with time-dependent coefficients, to the best of our knowledge, there is only one related reference, [10]. More specifically, V. Barbu and M. Röckner studied (4) with coefficients $\beta \equiv a(t, x, u)u$ and $Db \equiv b(t, x, u)$ and obtained existence results in appropriate Hilbert spaces. Finally, it should be mentioned that, for porous media equations (that is, $D \equiv 0$ in (4)), uniqueness of distributional solutions was first proved by H. Brezis and M.G. Crandall [18] and also by M. Pierre [26], similar results were established in [6], [7] and [16].

For the general diffusion terms $a_{ij}(x, u)$, the existence of mild solutions to (2) has been obtained in [5] via nonlinear semigroup theory in $L^1(\mathbb{R}^d)$, while the uniqueness of distributional solutions to (2) remained open. Thus, one of the main contributions of this paper is to prove uniqueness of distributional solutions to NFPE (2), see Theorem 3.3 and Corollary 3.1 below. These results are not covered by previous work. In addition, we would like to point out that we also treat the general drift term $b(x, u)$ (i.e., the case with explicit x-dependence). Providing a supplementary proof based on [5], we prove an L^∞ estimate for mild solutions to (2) starting from initial data in $L^1 \cap L^\infty$. This estimate is essential for the solution class used subsequently, especially in Section 5, where nonlinear Markov processes are constructed, see Theorem 3.2 and Corollary 3.2 for details. An additional major analytical result is the proof of "linearized uniqueness" for distributional solutions to (2), see Theorem 4.1 and Corollary 4.1.

Furthermore, we prove uniqueness of probabilistically weak solutions to the corresponding McKean-Vlasov SDE (3) (see Theorem 4.2 below), which is another main contribution of this paper. For related work on weak uniqueness of McKean-Vlasov SDEs, we refer to [6], [7], [15], [25] and [29], but none of these papers covers the results in Theorem 4.2. As another important consequence, we also prove that the path laws of the solutions to McKean-Vlasov SDEs form a nonlinear Markov process in the sense of McKean (see [23]). This concept of a nonlinear Markov process was recently elaborated by M. Rehmeier and M. Röckner in [28]. The results on nonlinear Markov processes in our case are presented and proved in Section 5.

Notation. By $C^k(\mathbb{R}^d)$ ($C^k(\mathbb{R})$, $C^k(\mathbb{R}^d \times \mathbb{R})$, respectively) we denote the space of k-times continuously differentiable functions on $\mathbb{R}^d$ ($\mathbb{R}$, $\mathbb{R}^d \times \mathbb{R}$, respectively); by $C_b(\mathbb{R}^d)$ ($C_b(\mathbb{R})$, $C_b(\mathbb{R}^d \times \mathbb{R})$, respectively) we denote the space of continuous and bounded functions on $\mathbb{R}^d$ ($\mathbb{R}$, $\mathbb{R}^d \times \mathbb{R}$, respectively); and $C([0, T])$ denotes the space of continuous real-valued functions on $[0, T]$. We shall simply write C^k for $C^k(\mathbb{R}^d)$ and C_b for $C_b(\mathbb{R}^d)$.

$L^p(\mathbb{R}^d) = L^p$ for $1 \leqslant p \leqslant \infty$ denotes the space of all real-valued, p-integrable functions on $\mathbb{R}^d$, with standard norm $|\cdot|_p$. By L^p_{loc} we denote the corresponding local space. $(\cdot, \cdot)$ shall denote the scalar product in L^2. Let $H^k(\mathbb{R}^d) = H^k$, $k = 1, 2$, denote the standard Sobolev space on $\mathbb{R}^d$ and by $H^{-k}(\mathbb{R}^d) = H^{-k}$ we denote the dual space of $H^k(\mathbb{R}^d)$.

$C_0^\infty(\mathbb{R}^d)$ is the space of infinitely differentiable real-valued functions with compact support in $\mathbb{R}^d$. $C_0^\infty([0, \infty) \times \mathbb{R}^d)$ denotes the space of all $\varphi \in C_0^\infty([0, \infty) \times \mathbb{R}^d)$ with compact support in $[0, \infty) \times \mathbb{R}^d$. By $\mathcal{D}'(\mathbb{R}^d)$ ($\mathcal{D}'((0, \infty) \times \mathbb{R}^d)$, respectively) we denote the space of Schwartz distributions on $\mathbb{R}^d$ ($(0, \infty) \times \mathbb{R}^d$, respectively).

$\mathcal{P}$ denotes the set of all probability densities on $\mathbb{R}^d$, that is

$$\mathcal{P} = \left\{ \rho \in L^1(\mathbb{R}^d);\ \rho \geqslant 0, a.e. \text{in } \mathbb{R}^d,\ \int_{\mathbb{R}^d} \rho(x) dx = 1 \right\}.$$

$D_i u$ denotes the partial derivative of a function $u = u(x_1, \cdots, x_d)$ with respect to x_i, $1 \leqslant i \leqslant d$, and by $D^2_{ij} u$ we denote the second order derivatives $\frac{\partial^2 u}{\partial x_i \partial x_j}$. We also set, for each $u \in C^1(\mathbb{R}^d \times \mathbb{R})$,

$$u_r = \frac{\partial}{\partial r} u(x, r), \quad u_x = \nabla_x u(x, r).$$

The function $t \to (u(t, \cdot)dx)_{[0,T]}$ is said to be narrowly continuous if, for every $\eta \in C_b(\mathbb{R}^d)$,

$$\lim_{t \to s} \int_{\mathbb{R}^d} \eta(x) u(t, x) dx = \int_{\mathbb{R}^d} \eta(x) u(s, x) dx, \quad s \geqslant 0.$$

The remainder of this paper is structured as follows. Section 2 contains the preliminary definitions and results used throughout the paper. Section 3 contains the existence of mild solutions and the uniqueness of distributional solutions for NFPE (2). In Section 4, we prove weak uniqueness for the corresponding McKean-Vlasov SDE (3). Section 5 contains the main results on nonlinear Markov processes.

2 Preliminaries

We now define mild and distributional solutions to NFPE (2), which are the basic notions used in this work.

Definition 2.1 (mild solution) A continuous function $u : [0, \infty) \to L^1$ is called a mild solution to NFPE (2) if, for each $T > 0$,

$$u(t) = \lim_{h \to 0} u_h(t) \text{ in } L^1, \quad \text{uniformly on } [0, T],$$

where $u_h : [0, T] \to L^1$ is the step function

$$u_h(t) = u_h^k, \ \forall t \in [kh, (k+1)h), k = 0, 1, \cdots, N_h = [\frac{T}{h}],$$
$$u_h^{k+1} + hA(u_h^{k+1}) = u_h^k, \ \forall k = 0, 1, \cdots, N_h; \ u_h^0 = u_0.$$

Here, A is the operator defined as follows

$$A(u) = -\sum_{i,j=1}^{d} D_{ij}^2(a_{ij}(x)\beta(x, u)) + \operatorname{div}(b(x, u)u), \ \forall u \in D(A),$$
$$D(A) = \left\{u \in L^1; A(u) \in L^1\right\},$$

where D_{ij}^2 and div are taken in the sense of Schwartz distributions on $\mathbb{R}^d$.

Definition 2.2 (1) (distributional solution) Let $d \geqslant 1$ and $T > 0$. A function $u \in L^1_{\text{loc}}([0, T) \times \mathbb{R}^d)$ is said to be a distributional solution to NFPE (2) with initial condition u_0 if $a_{ij}(x)\beta(x, u) \in L^1_{\text{loc}}([0, T) \times \mathbb{R}^d)$, $b(x, u)u \in L^1_{\text{loc}}([0, T) \times \mathbb{R}^d; \mathbb{R}^d)$, and, for all $\varphi \in C_0^\infty([0, T) \times \mathbb{R}^d)$ and $u_0 \in L^1(\mathbb{R}^d)$,

$$\begin{aligned}\int_0^T \int_{\mathbb{R}^d} \Bigg(& u(t, x)\frac{\partial \varphi}{\partial t}(t, x) + \sum_{i,j=1}^{d} a_{ij}(x)\beta(x, u(t, x))D_{ij}^2\varphi(t, x) \\ & + u(t, x)b(x, u(t, x)) \cdot \nabla\varphi(t, x)\Bigg) dxdt + \int_{\mathbb{R}^d} u_0(x)\varphi(0, x)dx = 0.\end{aligned} \tag{5}$$

(2) A distributional solution to NFPE (2) is called a probability solution to NFPE (2) if $u_0 \in \mathcal{P}(\mathbb{R}^d)$, and $u(t, \cdot) \in \mathcal{P}(\mathbb{R}^d)$ for a.e. $t \in (0.T)$.

(3) Let $\tilde{\mathcal{P}}(\mathbb{R}^d)$ denote the space of all Borel probability measure on $\mathbb{R}^d$. Suppose $\mathcal{P}_0 \subset \tilde{\mathcal{P}}(\mathbb{R}^d)$ such that for each $(s, \zeta) \in [0, \infty) \times \mathcal{P}_0$, there exists a weakly continuous probability solution $[s, \infty) \ni t \mapsto \mu_t^{s,\zeta} \in \mathcal{P}_0$ to NFPE (2) with starting time $s \in [0, \infty)$ and initial condition u_0 renamed as ζ, and $\mu_t^{s,\zeta} := u^{s,\zeta}(t, x)dx$, such that flow property

$$\mu_t^{s,\zeta} = \mu_t^{r, \mu_r^{s,\zeta}}, \ \ \forall\, 0 \leqslant s \leqslant r \leqslant t, \ \ \zeta \in \mathcal{P}_0$$

holds. Then $(\mu^{s,\zeta})_{(s,\zeta) \in [0,\infty) \times \mathcal{P}_0}$ is called a solution flow of NFPE (2) in $\mathcal{P}_0$.

Remark 2.1 For an initial condition $u_0 \in L^1$, in general, equation (2) does not have a classical strong solution and the best expectation is to obtain mild solutions in the sense of the above definition (which turn out to be also distributional solutions to (2)). Another class of solutions are entropy solutions which is a convenient concept of solutions in statistical physics (see, e.g. [14], Chapter 1). But entropy solutions are too special for our purpose, since time marginal densities of solutions to McKean-Vlasov SDEs are in general not entropy solutions. Therefore, distributional solutions constitute the correct class for stochastic analysis.

Finally, for later use, we introduce the operator $(L, D(L))$

$$L : D(L) \to \mathcal{D}'(\mathbb{R}^d), \quad Lu := \sum_{i,j=1}^{d} D_i(a_{ij}(x)D_j u), \quad D(L) = L^1_{\text{loc}}(\mathbb{R}^d).$$

$(L, D(L))$ is well-defined because $a_{ij} \in C^\infty(\mathbb{R}^d), i, j = 1, 2, \ldots d$. Moreover, for $u : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ and $\beta : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$, we use the shorthand

$$L\beta(x, u) := L[\beta(\cdot, u(t, \cdot))](x), \ \text{ if } \beta(\cdot, u(t, \cdot)) \in D(L).$$

Here and in the sequel, $\beta(x, u)$ stands for $\beta(x, u(t, x))$.

3 Uniqueness of distributional solutions to NFPEs

3.1 Existence of mild solutions to NFPEs

The existence of mild solutions to NFPE (2) was proved in [14], Section 2.1 (see also [5]). To check Hypotheses (H1)-(H3) in there, we need to introduce the following conditions.

(***i***') $\beta \in C^1(\mathbb{R}^d \times \mathbb{R})$, $\beta_r(x, r) > \delta, \forall x \in \mathbb{R}^d, \forall r \in \mathbb{R}$ and some $\delta > 0$; $\beta(x, 0) \equiv 0$, $\beta(x, r) = -\beta(x, -r), \forall x \in \mathbb{R}^d, \forall r \in \mathbb{R}$, and for the function $\Phi(x, r) := \dfrac{\beta(x, r)}{r}, \forall x \in \mathbb{R}^d, \forall r \in \mathbb{R}$, where we set $\Phi(x, 0) := \lim\limits_{\varepsilon \to 0} \dfrac{\beta(x, \varepsilon)}{\varepsilon} = \beta_r(x, 0), x \in \mathbb{R}^d$, we have $\Phi \in C^2(\mathbb{R}^d \times \mathbb{R}) \cap C_b(\mathbb{R}^d \times \mathbb{R})$, $\Phi_x \in C_b(\mathbb{R}^d \times \mathbb{R}; \mathbb{R}^d)$,
(***iii***') $b \in C_b(\mathbb{R}^{d+1}; \mathbb{R}^d) \cap C^1(\mathbb{R}^{d+1}; \mathbb{R}^d)$.

Theorem 3.1 *[14, Thm. 2.1] Assume that (**i**'), (**ii**) and (**iii**') hold. Then there exists a continuous contraction semigroup $S(t) : [0, \infty) \to L^1$ such that, for each $u_0 \in L^1(\mathbb{R}^d)$, $u = u(t, u_0) = S(t)u_0$ is a mild solution to (2) and*

$$|u(t, u_0) - u(t, \tilde{u}_0)|_1 \leqslant |u_0 - \tilde{u}_0|_1, \quad \forall u_0, \tilde{u}_0 \in L^1, \ t \geqslant 0$$

$$u \geqslant 0, \textit{a.e. on } (0, \infty) \times \mathbb{R}^d, \quad \textit{if } u_0 \geqslant 0, \textit{a.e. on } \mathbb{R}^d,$$

$$\int_{\mathbb{R}^d} u(t, x)dx, = \int_{\mathbb{R}^d} u_0(x)dx, \quad \forall u_0 \in L^1, \ t \geqslant 0.$$

Moreover, u is a distributional solution to (2), that is, for all $\varphi \in C_0^\infty([0,T) \times \mathbb{R}^d)$,

$$\int_0^T \int_{\mathbb{R}^d} \bigg(u(t,x) \frac{\partial \varphi}{\partial t}(t,x) + \sum_{i,j=1}^d a_{ij}(x)\beta(x,u(t,x)) D_{ij}^2 \varphi(t,x)$$

$$+ u(t,x) b(x,u(t,x)) \cdot \nabla \varphi(t,x) \bigg) dxdt + \int_{\mathbb{R}^d} u_0(x) \varphi(0,x) dx = 0.$$

Proof. Set $\tilde{a}_{ij}(x,u) = \dfrac{a_{ij}(x)\beta(x,u)}{u} = a_{ij}(x)\Phi(x,u)$. Then, by (***i***') and (***ii***), one readily checks that a_{ij} satisfies the following condition

(H1) $\tilde{a}_{ij} \in C^2(\mathbb{R}^d \times \mathbb{R}) \cap C_b(\mathbb{R}^d \times \mathbb{R})$, $(\tilde{a}_{ij})_x \in C_b(\mathbb{R}^d \times \mathbb{R}; \mathbb{R}^d)$, $\tilde{a}_{ij} = \tilde{a}_{ji}$, $\tilde{a}_{ij}(x,u) = \tilde{a}_{ij}(x,|u|), x \in \mathbb{R}^d, u \in \mathbb{R}, i,j = 1,2,\dots,d$.

Since $\tilde{a}_{ij}(x,u) + (\tilde{a}_{ij})_u(x,u)u = a_{ij}(x)(\Phi(x,u) + u\Phi_u(x,u)) = a_{ij}(x)\beta_r(x,u)$, condition (***ii***) implies that $\tilde{a}_{ij}$ also satisfies

(H2) $\sum\limits_{i,j=1}^d \big(\tilde{a}_{ij}(x,u) + (\tilde{a}_{ij})_u(x,u)u\big)\xi_i\xi_j \geqslant \gamma\delta\, |\xi|^2, \quad \forall \xi \in \mathbb{R}^d, x \in \mathbb{R}^d, u \in \mathbb{R}$, for γ as in (***ii***) and δ as in (***i***').

By (***iii***'), b_i satisfies

(H3) $b_i \in C_b(\mathbb{R}^d \times \mathbb{R}) \cap C^1(\mathbb{R}^d \times \mathbb{R}), i = 1, 2 \dots, d$.

Hence, by Theorem 2.1 in [14], the assertions follow. □

However, the results in [14] do not include the additional property that a mild solution starting from data in $L^1 \cap L^\infty$ belongs to $L^\infty((0,T) \times \mathbb{R}^d)$. The following theorem whose proof is given in the Appendix, provides estimate (8) below for the corresponding general diffusion terms $a_{ij}(x,u)$ as in [14]. It applies, in particular, to the present setting and is needed for the nonlinear Markov property proved in Section 5.

Theorem 3.2 *Assume that (**i**'), (**ii**) and (**iii**') hold and, in addition,*

$$\delta_1(r) := \sup \left\{ \left| b_x^i(x,r) \right| ;\ i = 1,\dots,d,\ x \in \mathbb{R}^d \right\}, \tag{6}$$

$$\delta_2(r) := \sup \left\{ \left| (\tilde{a}_{ij})_{x_i x_j}(x,r) \right| ;\ i,j = 1,\dots,d,\ x \in \mathbb{R}^d \right\}, \tag{7}$$

where $\tilde{a}_{ij}(x,u) = \dfrac{a_{ij}(x)\beta(x,u)}{u} = a_{ij}(x)\Phi(x,u)$.

*(**v**) $\delta_1,\ \delta_2 \in C_b(\mathbb{R})$.*

Then, for each $u_0 \in L^1 \cap L^\infty$, a mild solution to (2) given in Theorem 3.1 also satisfies, for some constant C

$$|u(t)|_\infty \leqslant exp(Ct)\, |u_0|_\infty, \quad \forall t \in [0,T]. \tag{8}$$

3.2 Uniqueness of distributional solutions to NFPEs

Theorem 3.3 *Assume that **Hypothesis (I)** holds. Let $u_1, u_2 \in L^1((0,T) \times \mathbb{R}^d) \cap L^\infty((0,T) \times \mathbb{R}^d)$ be two distributional solutions to (2) in $(0,T) \times \mathbb{R}^d$ such that $u_1 - u_2 \in L^\infty((0,T); H^{-1})$ and*

$$\lim_{t \to 0} \operatorname*{ess\,sup}_{s \in (0,t)} |(u_1(s) - u_2(s), \varphi)_2| = 0, \quad \forall \varphi \in C_0^\infty(\mathbb{R}^d). \tag{9}$$

Then, $u_1 \equiv u_2$.

Proof. We adapt the technique developed in [12], though it has to be modified to apply to our substantially more general situation. We first note that for $i = 1, 2$, $\beta(x, u_i) \in L^1_{\text{loc}} = D(L)$. Hence

$$\begin{aligned}\sum_{i,j=1}^{d} D^2_{ij}(a_{ij}(x)\beta(x, u_i)) &= \sum_{i,j=1}^{d} D_i(a_{ij}(x)D_j\beta(x, u_i)) + \text{div}(A(x)\beta(x, u_i)) \\ &= L\beta(x, u_i) + \text{div}(A(x)\beta(x, u_i)),\end{aligned}$$

where $A(x) = \{A_i(x)\}_{i=1}^d$ and $A_i(x) = \sum_{j=1}^{d} D_j a_{ij}(x)$. Below we also set $b^*(x, u) := b(x, u)u$, and define $N := \max\{|u_1|_\infty, |u_2|_\infty\}$.
We set

$$z := u_1 - u_2,\ w := \beta(x, u_1) - \beta(x, u_2),\ \zeta := b^*(x, u_1) - b^*(x, u_2), \forall x \in \mathbb{R}^d.$$

Then, by Hypothesis (I) on $[0, T] \times \mathbb{R}^d$,

$$|\zeta| \leqslant \alpha_2 |w|, \tag{10}$$

$$wz \geqslant \alpha_1 |w|^2, \tag{11}$$

where $\alpha_2 = \alpha_{[-N,N]}$ and $\alpha_1^{-1} = |\beta_r|_{L^\infty(\mathbb{R}^d\times[-N,N])}$. Furthermore, our conditions imply that $z, w, \zeta \in L^1((0, T) \times \mathbb{R}^d) \cap L^\infty((0, T) \times \mathbb{R}^d)$ and

$$z_t - Lw + \text{div}(\zeta - Aw) = 0, \quad \text{in } \mathcal{D}'((0, T) \times \mathbb{R}^d). \tag{12}$$

We define

$$z_\delta := z * \theta_\delta, w_\delta := w * \theta_\delta, \zeta_\delta := \zeta * \theta_\delta, (Aw)_\delta := (Aw) * \theta_\delta, (Lw)_\delta := (Lw) * \theta_\delta,$$

where $\theta \in C_0^\infty(\mathbb{R}^d)$, $\theta_\delta(x) \equiv \frac{1}{\delta^d}\theta(\frac{x}{\delta})$, $\delta > 0$, is a standard mollifier. Since $z_\delta, w_\delta \in L^2(0, T; H^2)$, and ζ_δ, $(Aw)_\delta$, $(Lw)_\delta$, divζ_δ, div$(Aw)_\delta \in L^2(0, T; L^2)$, we have

$$(z_\delta)_t - (Lw)_\delta + \text{div}(\zeta_\delta - (Aw)_\delta) = 0, \quad \text{in } \mathcal{D}'((0, T); L^2). \tag{13}$$

We recall that by our assumptions on a_{ij} and standard elliptic regularity theory (see, e.g. [24, Theorem 1 and Theorem 2.2]) we have that

$$\epsilon I - L_{|H^2} : H^2 \longrightarrow L^2$$

is a homeomorphism for every $\epsilon > 0$, where I denotes the identity operator. We define its inverse operator Ψ_ε as follows: for any $\varepsilon > 0$,

$$\Psi_\varepsilon : L^2 \to H^2,$$

$$\Psi_\varepsilon(u) = (\varepsilon I - L_{|H^2})^{-1}u, \quad \forall u \in L^2.$$

Then $\Psi_\varepsilon(z_\delta)$, $\Psi_\varepsilon(w_\delta) \in L^2(0,T;H^2)$; $\Psi_\varepsilon(\text{div}\zeta_\delta)$, $\Psi_\varepsilon(\text{div}(Aw)_\delta) \in L^2(0,T;L^2)$. Therefore, applying Ψ_ϵ to (13) yields

$$\begin{aligned}(\Psi_\varepsilon(z_\delta))_t =&\Psi_\varepsilon(Lw_\delta) + \Psi_\varepsilon(\text{div}((Aw)_\delta - \zeta_\delta)) + \Psi_\varepsilon(R_\delta)\\ =&L\Psi_\varepsilon(w_\delta) + \Psi_\varepsilon(\text{div}(Aw)_\delta) - \Psi_\varepsilon(\text{div}\zeta_\delta) + \Psi_\varepsilon(R_\delta), \quad \text{in } \mathcal{D}'((0,T);L^2),\end{aligned} \tag{14}$$

where $R_\delta = (Lw)_\delta - Lw_\delta$. We note that, since $w(t), w_\delta(t) \in L^2$, we have

$$(Lw(t))_\delta \longrightarrow Lw(t) \quad \text{in } \ H^{-2},$$

$$Lw_\delta(t) \longrightarrow Lw(t) \quad \text{in } \ H^{-2}.$$

Since $R_\delta = (Lw)_\delta - Lw_\delta = (I-L)w_\delta - ((I-L)w)_\delta$, for some constant $C>0$

$$|R_\delta(t)|_{-2} \leqslant C\,|w(t)|_2 \quad \text{a.e. } \ t \in [0,T].$$

Since $w \in L^2(0,T;L^2)$, we altogether obtain

$$\lim_{\delta\to 0} |R_\delta|_{L^2(0,T;H^{-2})} = 0. \tag{15}$$

We also note that

$$z_\delta(t) = \varepsilon\Psi_\varepsilon(z_\delta(t)) - L\Psi_\varepsilon(z_\delta(t)). \tag{16}$$

We set

$$h_{\varepsilon,\delta}(t) = (\Psi_\varepsilon(z_\delta(t)), z_\delta(t))_2, \quad \text{a.e. } t \in (0,T).$$

Below we shall prove two claims.

Claim 3.1 $\lim\limits_{t\to 0} \operatorname*{ess\,sup}\limits_{s\in(0,t)} h_{\varepsilon,\delta}(s) = 0.$

Proof. By (14) we have $\frac{d}{dt}\Psi_\varepsilon(z_\delta) \in L^2(0,T;L^2)$, where $\frac{d}{dt}$ is taken in the sense of vector-valued distributions in L^2 on $(0,T)$. Thus, z_δ, $\Psi_\varepsilon(z_\delta) \in H^1(0,T;L^2)$. On the other hand, we have $\Psi_\varepsilon(z_\delta) \in L^2(0,T;H^2)$.
Therefore,

$$\Psi_\varepsilon(z_\delta) \in C([0,T];H^1),$$

since the spaces L^2 and H^2 are in duality with the pivot space $H^1 = (H^2, L^2)_{\frac{1}{2}}$, after modifying $t \to \Psi_\varepsilon(z_\delta(t))$ on a subset of measure zero (see, e.g. [2] and [12]). Since $t \to \Psi_\varepsilon(z_\delta(t))$ has an H^1-continuous version on $[0,T]$, there exists $f \in H^1$ such that

$$\lim_{t\to 0} \Psi_\varepsilon(z_\delta(t)) = f \quad \text{in } H^1. \tag{17}$$

By **(ii)** and (16), we observe that

$$\begin{aligned} h_{\varepsilon,\delta}(t) =&(\Psi_\varepsilon(z_\delta(t)), z_\delta(t))_2 = \int_{\mathbb{R}^d} \Psi_\varepsilon(z_\delta(t))(\varepsilon\Psi_\varepsilon(z_\delta(t) - L\Psi_\varepsilon(z_\delta(t))dx\\ =&\varepsilon\,|\Psi_\varepsilon(z_\delta(t)|_2^2 - \int_{\mathbb{R}^d} \Psi_\varepsilon(z_\delta(t)L\Psi_\varepsilon(z_\delta(t)dx\\ =&\varepsilon\,|\Psi_\varepsilon(z_\delta(t)|_2^2 + \int_{\mathbb{R}^d}\sum_{i,j=1}^d a_{ij}(x)D_j\Psi_\varepsilon(z_\delta(t)D_i\Psi_\varepsilon(z_\delta(t)dx\\ \geqslant&\varepsilon\,|\Psi_\varepsilon(z_\delta(t)|_2^2 + \gamma\,|\nabla\Psi_\varepsilon(z_\delta(t)|_2^2 \geqslant 0. \end{aligned} \tag{18}$$

Furthermore,

$$0 \leqslant h_{\varepsilon,\delta}(s) = (\Psi_\varepsilon(z_\delta(s)), z_\delta(s))_2 \leqslant |\Psi_\varepsilon(z_\delta(s) - f|_{H^1} \cdot |z_\delta(s)|_{H^{-1}} \\ + |f-\varphi|_{H^1} \cdot |z_\delta(s)|_{H^{-1}} + |(\varphi * \theta_\delta, z(s))_2|, \quad \forall \varphi \in C_0^\infty(\mathbb{R}^d),\ s \in (0,T).$$

Hence, by (9), (17),

$$\begin{aligned} 0 &\leqslant \lim_{t\to 0} \operatorname*{ess\,sup}_{s\in(0,t)} h_{\varepsilon,\delta}(s) \\ &\leqslant \left(\lim_{t\to 0} |\Psi_\varepsilon(z_\delta(t) - f|_{H^1} + |f-\varphi|_{H^1}\right) |z_\delta|_{L^\infty(0,T;H^{-1})} + \lim_{t\to 0} \operatorname*{ess\,sup}_{s\in(0,t)} |(\varphi * \theta_\delta, z(s))_2| \\ &= |f-\varphi|_{H^1} |z_\delta|_{L^\infty(0,T;H^{-1})}. \end{aligned}$$

Therefore, Claim 3.1 follows since $C_0^\infty(\mathbb{R}^d)$ is dense in $H^1(\mathbb{R}^d)$. $\square$

Claim 3.2 *Each $h_{\varepsilon,\delta}$ has an absolutely continuous version and for these versions, there exists a choice of the mollification parameter $\delta = \delta(\varepsilon) \to 0$ such that $\lim\limits_{\varepsilon\to 0} h_{\varepsilon,\delta(\varepsilon)}(t) = 0$ uniformly on $t \in [0,T]$, i.e. $\lim\limits_{\varepsilon\to 0} h_{\varepsilon,\delta(\varepsilon)} = 0$ in $C([0,T])$.*

Proof. By (13) we have $\frac{d}{dt} z_\delta \in L^2(0,T;L^2)$, hence $z_\delta \in H^1(0,T;L^2)$ and thus $[0,T] \ni t \mapsto z_\delta(t) \in L^2$ and, as seen below, also $[0,T] \ni t \mapsto \Psi_\varepsilon(z_\delta(t)) \in L^2$ are absolutely continuous (see [2], p. 23, Theorem 1.17). Then, by the definition of $h_{\varepsilon,\delta}(t)$, (16) and (14), we have

$$\begin{aligned} h'_{\varepsilon,\delta}(t) =& 2((\Psi_\varepsilon(z_\delta(t)))_t, z_\delta(t))_2 \\ =& 2\left(L\Psi_\varepsilon(w_\delta(t)) + \Psi_\varepsilon(\operatorname{div}((Aw)_\delta(t))) - \Psi_\varepsilon(\operatorname{div}(\zeta_\delta(t)), z_\delta(t)\right)_2 + 2(R_\delta(t), \Psi_\varepsilon(z_\delta(t)))_2 \\ =& 2\varepsilon(\Psi_\varepsilon(z_\delta(t)), w_\delta(t))_2 - 2(w_\delta(t), z_\delta(t))_2 + 2(R_\delta(t), \Psi_\varepsilon(z_\delta(t)))_2 \\ & - 2(\nabla\Psi_\varepsilon(z_\delta(t)), (Aw)_\delta(t))_2 + 2(\nabla\Psi_\varepsilon(z_\delta(t)), \zeta_\delta(t))_2, \quad \text{a.e. } t \in (0,T). \end{aligned} \tag{19}$$

Next, we derive several estimates for $h'_{\varepsilon,\delta}(t)$.
a) First by (11) we have

$$(w_\delta(t), z_\delta(t))_2 \geqslant \alpha_1 |w(t)|_2^2 + \gamma_\delta(t),$$

where $\gamma_\delta(t) = (w_\delta(t), z_\delta(t))_2 - (w(t), z(t))_2$.
b) By (10) we have

$$\begin{aligned} (\nabla\Psi_\varepsilon(z_\delta(t)), \zeta_\delta(t))_2 - (\nabla\Psi_\varepsilon(z_\delta(t)), (Aw)_\delta(t))_2 &\leqslant |\nabla\Psi_\varepsilon(z_\delta(t)|_2 \cdot (|\zeta(t)|_2 + |(Aw)(t)|_2) \\ &\leqslant |\nabla\Psi_\varepsilon(z_\delta(t)|_2 \cdot (\alpha_2 |w(t)|_2 + |A|_\infty |w(t)|_2) \\ &\leqslant C\, |\nabla\Psi_\varepsilon(z_\delta(t)|_2 |w(t)|_2, \end{aligned}$$

where $C = \alpha_2 + |A|_\infty$.
c) For all $\varepsilon, \delta \in (0,1)$, we have

$$\varepsilon(|\Psi_\varepsilon(z_\delta(t))|, |w_\delta(t)|)_2 \leqslant \varepsilon\, |\Psi_\varepsilon(z_\delta(t))|_\infty |w_\delta(t)|_1 \leqslant c\, |z(t)|_\infty |w(t)|_1, \quad \text{a.e. } t \in (0,T). \tag{20}$$

$$\lim_{\varepsilon\to 0}\sup_{\delta\in(0,1)}|\varepsilon(\Psi_\varepsilon(z_\delta(s))|_\infty=0,\quad \text{a.e. } s\in(0,T) \tag{21}$$

To prove the estimates (20) and (21), let K_ε be the integral kernel of $(\varepsilon I-L)^{-1}$. By Theorem 7 in [1], there exists $c\in(1,\infty)$ such that $c^{-1}K_\varepsilon^\Delta(x-\xi)\leqslant K_\varepsilon(x,\xi)\leqslant cK_\varepsilon^\Delta(x-\xi)$, $\forall x,\xi\in\mathbb{R}^d$, where K_ε^Δ is the integral kernel of $(\varepsilon I-\Delta)^{-1}$.
Therefore, for a.e. $t\in[0,T]$, we have

$$\begin{aligned}|\varepsilon\Psi_\varepsilon(z_\delta(t))(x)|\leqslant&\varepsilon\int_{\mathbb{R}^d}K_\varepsilon(x,\xi)\,|z_\delta(t)(\xi)|\,d\xi\\ \leqslant&\varepsilon\int_{\mathbb{R}^d}cK_\varepsilon^\Delta(x-\xi)\,|z_\delta(t)(\xi)|\,d\xi=c\varepsilon(\varepsilon I-\Delta)^{-1}\,|z_\delta(t)|\,(x),\end{aligned}$$

It is well known that

$$\varepsilon\left|(\varepsilon I-\Delta)^{-1}y\right|_p\leqslant|y|_p\,,\quad \forall y\in L^p,\ p\in[1,\infty], \tag{22}$$

Hence, $\varepsilon\,|\Psi_\varepsilon(z_\delta(t))|_\infty\leqslant c\,|z_\delta(t)|_\infty$, and therefore (20) holds.
To prove (21) we proceed similarly as in [13, Proof of Thorem 3.1] (see also [18]) by writing for a.e. $t\in(0,T)$,

$$\varepsilon(\varepsilon I-\Delta)^{-1}\,|z_\delta(t)|\,(x)=\varepsilon^{\frac{d}{2}}\int_{\mathbb{R}^d}K_1^\Delta(\sqrt{\varepsilon}(x-\xi))\,|z_\delta(t)(\xi)|\,d\xi,\quad x\in\mathbb{R}^d,$$

where K_1^Δ is the kernel associated with the operator $(I-\Delta)^{-1}$ (see [30]), i.e.

$$(I-\Delta)^{-1}(y)(x)=\int_{\mathbb{R}^d}K_1^\Delta(x-\xi)y(\xi)d\xi,\quad x\in\mathbb{R}^d.$$

This yields for a.e. $x\in\mathbb{R}^d$,

$$\left|\varepsilon(\varepsilon I-\Delta)^{-1}\,|z_\delta(t)|\,(x)\right|\leqslant c_r\varepsilon^{\frac{d}{2}}\,|z(t)|_1+\varepsilon^{\frac{d}{2}}\,|z(t)|_\infty\int_{\sqrt{\varepsilon}|x-\xi|\leqslant r}K_1^\Delta(\sqrt{\varepsilon}(x-\xi))d\xi,\quad r>0,$$

where $c_r=\sup\{K(x);|x|\geqslant r\}<\infty$, $\forall r>0$. Then, for a.e. $t\in(0,T)$, we get

$$\limsup_{\varepsilon\downarrow 0}\sup_{\delta\in(0,1)}\left|\varepsilon(\varepsilon I-\Delta)^{-1}\,|z_\delta(t)|\right|_\infty\leqslant|z(t)|_\infty\int_{|\xi|\leqslant r}K_1^\Delta(\xi)d\xi,\quad \forall r>0,$$

letting $r\to 0$, since K_1^Δ as a probability kernel is locally integrable, (21) follows.
By (20), (21) and the dominated convergence theorem, it follows that

$$\lim_{\varepsilon\to 0}\varepsilon\int_0^T\sup_{\delta\in(0,1)}(|\Psi_\varepsilon(z_\delta(s))|\,,|w_\delta(s)|)_2ds=0. \tag{23}$$

d) For every $\varepsilon\in(0,1)$, there exists $\delta(\varepsilon)\in(0,\varepsilon]$ such that

$$\int_0^T\left|(R_{\delta(\varepsilon)}(s),\Psi_\varepsilon(z_{\delta(\varepsilon)}(s)))_2\right|ds\leqslant\varepsilon\,|z|_{L^2(0,T;L^2)}\,. \tag{24}$$

It easily follows by (22) that the operator norm of $\varepsilon\Psi_\varepsilon:L^2\longrightarrow H^2$ is less than one. Hence for a.e. $s\in[0,T]$

$$|(R_\delta(s),\Psi_\varepsilon(z_\delta(s)))_2|\leqslant|R_\delta(s)|_{-2}\cdot|\Psi_\varepsilon(z_\delta(s))|_{H^2}\leqslant(1/\varepsilon)\,|R_\delta(s)|_{-2}\cdot|z(s)|_2\,.$$

By (15) we can choose $\delta(\varepsilon) \in (0, \varepsilon]$ such that $\left|R_{\delta(\varepsilon)}\right|_{L^2(0,T;H^{-2})} \leqslant \varepsilon^2$. Hence, (24) holds. Then by Claim 3.1, (18), (19), a), b), d), we have

$$\begin{aligned}0 \leqslant h_{\varepsilon,\delta(\varepsilon)}(t) \leqslant & 2\varepsilon \int_0^t (\left|\Psi_\varepsilon(z_{\delta(\varepsilon)}(s))\right|, \left|w_{\delta(\varepsilon)}(s)\right|)_2 ds - 2\alpha_1 \int_0^t |w(s)|_2^2 \, ds + 2\int_0^t \left|\gamma_{\delta(\varepsilon)}(s)\right| ds \\ & + 2\varepsilon \, |z|_{L^2(0,T;L^2)} + \frac{C^2}{\alpha_1} \int_0^t \left|\nabla \Psi_\varepsilon(z_{\delta(\varepsilon)}(s)\right|_2^2 ds + \alpha_1 \int_0^t |w(s)|_2^2 \, ds \\ \leqslant & \tilde{\zeta}_\varepsilon + \frac{C^2}{\alpha_1 \gamma} \int_0^t h_{\varepsilon,\delta(\varepsilon)}(s) ds\end{aligned} \tag{25}$$

where we used (18) again in the last step and set

$$\tilde{\zeta}_\varepsilon := 2\varepsilon \int_0^T (\left|\Psi_\varepsilon(z_{\delta(\varepsilon)}(s))\right|, \left|w_{\delta(\varepsilon)}(s)\right|)_2 ds + 2\varepsilon \, |z|_{L^2(0,T;L^2)} + 2\int_0^T \left|\gamma_{\delta(\varepsilon)}(s)\right| ds.$$

Clearly by (23) we have $\lim_{\varepsilon \to 0} \tilde{\zeta}_\varepsilon = 0$. Hence by Gronwall's inequality, we obtain

$$0 \leqslant h_{\varepsilon,\delta(\varepsilon)}(t) \leqslant \tilde{\zeta}_\varepsilon \exp(\frac{C^2}{\alpha_1 \gamma} t), \quad \forall t \in [0, T]. \tag{26}$$

This implies that $h_{\varepsilon,\delta(\varepsilon)}(t) \to 0$ as $\varepsilon \to 0$ uniformly on $[0, T]$ and hence Claim 3.2 is proved. $\square$

Proof of Theorem 3.3 (continued): We observe from (18) that, $h_{\varepsilon,\delta(\varepsilon)}(t)$ can be rewritten as $h_{\varepsilon,\delta(\varepsilon)}(t) = \varepsilon \left|\Psi_\varepsilon(z_{\delta(\varepsilon)}(t))\right|_2^2 + \left|(-L)^{\frac{1}{2}} \Psi_\varepsilon(z_{\delta(\varepsilon)}(t))\right|_2^2$ (see, e.g. Theorem 1.3.1 in [20] or p. 27 in [22]). Hence, by (21) and Claim 3.2, the right-hand side of (16) converges to 0 in $\mathcal{D}'$. Therefore, $0 = \lim_{\varepsilon \to 0} z_{\delta(\varepsilon)}(t) = z(t)$ in $\mathcal{D}'$ for a.e. $t \in (0, T)$, which implies $u_1 \equiv u_2$.

$\square$

Corollary 3.1 *Assume that Hypothesis (I) holds. Let $T > 0$ and $u_0 \in \mathcal{P} \cap L^\infty$, and let $u_1, u_2 \in L^1((0,T) \times \mathbb{R}^d) \cap L^\infty((0,T) \times \mathbb{R}^d)$ be two nonnegative distributional solutions to (5). Then $u_1 \equiv u_2$.*

Proof. Because of Hypothesis (I) (**i**)-(**iii**), it is elementary to check that by (5)

$$\int_{\mathbb{R}^d} u_i(t,x) dx = \int_{\mathbb{R}^d} u_0(x) dx \;\text{ for a.e. } t \in (0,T),\; i = 1, 2. \tag{27}$$

Hence, it follows by Lemma 2.3 in [27], there exist $dt \otimes dx$-versions $\bar{u}_1$ of u_1 and $\bar{u}_2$ of u_2 such that for $\bar{u}_1(t)(dx) := u_1(t,x)dx$, $\bar{u}_2(t)(dx) := u_2(t,x)dx$, $\bar{u}_1(0,x)dx = \bar{u}_2(0,x)dx = u_0(x)dx$, the map $[0,T] \ni t \mapsto \bar{u}_i(t,x)dx$ is narrowly continuous for $i = 1, 2$. In particular, (27) holds for every $t \in [0,T]$, and hence $u_1, u_2 \in L^\infty((0,T); L^1 \cap L^\infty) \subset L^\infty(0,T;L^2) \subset L^\infty(0,T;H^{-1})$. Therefore, for every $\varphi \in C_0^\infty(\mathbb{R}^d)$,

$$\lim_{t \to 0} \operatorname*{ess\,sup}_{s \in (0,t)} |(u_1(s) - u_2(s), \varphi)_2| = \lim_{t \to 0} \left| \int_{\mathbb{R}^d} (\bar{u}_1(t,x) - \bar{u}_2(t,x)) \, \varphi(x) dx \right| = 0.$$

So, (9) holds and Theorem 3.3 implies the assertion. $\square$

By Theorem 3.2 and Corollary 3.1, we obtain the following existence and uniqueness result for NFPE (2).

Corollary 3.2 *Under Hypotheses **(i')**,**(ii)**, **(iii')**, **(iv)**-**(v)**, for each $u_0 \in \mathcal{P} \cap L^\infty$, NFPE (2) has a unique distributional solution*

$$u \in L^1((0,T); L^1) \cap L^\infty((0,T) \times \mathbb{R}^d),$$

which is in fact a weakly continuous probability solution to (2).

4 Weak uniqueness of the corresponding McKean-Vlasov SDEs

In this section, we prove weak uniqueness for the McKean-Vlasov SDE (3)

$$\begin{aligned} &dX(t) = b(X(t), u(t, X(t)))dt + \sqrt{\frac{2\beta(X(t), u(t, X(t)))}{u(t, X(t))}}\sigma(X(t))dW(t), \forall t > 0, \\ &\mathcal{L}_{X(0)}(dx) = u_0(x)dx, \quad \mathcal{L}_{X(t)}(dx) = u(t,x)dx, \quad t > 0, \end{aligned}$$

corresponding to the nonlinear Fokker-Planck equation (2), where u is a distributional solution to it with $u_0 \in \mathcal{P} \cap L^\infty$ under Hypothesis (I). By the superposition principle, first applied to the linearized equation (see (28) below) and subsequently derived for NFPE (2), it follows (see, e.g. [5], [14] and [31]) that there exists a probabilistically weak solution $X(t)$ to the McKean-Vlasov equation (3) with law density $u(t)$. More precisely, there is a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with normal filtration $(\mathscr{F}_t)_{t\geqslant 0}$ and an $(\mathscr{F}_t)$-Brownian motion $W(t)$ on it with values in $\mathbb{R}^d$ and a (with respect to t) continuous $(\mathscr{F}_t)$-progressively measurable map $X : [0, \infty) \times \Omega \to \mathbb{R}^d$, which is a solution to (3), and $u(t,x)dx = \mathbb{P} \circ (X(t))^{-1}(dx)$, $u_0(x)dx = \mathbb{P} \circ (X_0)^{-1}(dx)$.

To prove weak uniqueness of (3), we first establish the so-called "linearized uniqueness" for the "frozen" linearized version of NFPE (2), namely

$$\begin{aligned} &\nu_t - \sum_{i,j=1}^d D_{ij}^2 \left(\frac{a_{ij}(x)\beta(x,u)}{u} \nu \right) + \operatorname{div}(b(x,u)\nu) = 0, \quad \text{in } \mathcal{D}'((0,T) \times \mathbb{R}^d), \\ &\nu(0,x) = \nu_0(x), \quad x \in \mathbb{R}^d, \end{aligned} \tag{28}$$

where $u \in L^1((0,T) \times \mathbb{R}^d) \cap L^\infty((0,T) \times \mathbb{R}^d)$ is a distributional solution to NFPE (2), i.e., $\forall \varphi \in C_0^\infty([0,T) \times \mathbb{R}^d)$, $\nu_0 \in L^1(\mathbb{R}^d)$,

$$\begin{aligned} \int_0^T \int_{\mathbb{R}^d} \bigg(&\frac{\partial \varphi}{\partial t}(t,x) + \sum_{i,j=1}^d a_{ij}(x) \frac{\beta(x, u(t,x))}{u(t,x)} D_{ij}^2 \varphi(t,x) \\ &+ b(x, u(t,x)) \cdot \nabla \varphi(t,x) \bigg) \nu(t,x) dx dt + \int_{\mathbb{R}^d} \nu_0(x) \varphi(0,x) dx = 0. \end{aligned} \tag{29}$$

Theorem 4.1 *(Linearized Uniqueness) Assume that Hypothesis (I) holds. Let $T > 0$, let $u \in L^1((0,T) \times \mathbb{R}^d) \cap L^\infty((0,T) \times \mathbb{R}^d)$ be a distributional solution to NFPE (2), and let $\nu_1, \nu_2 \in L^1((0,T) \times \mathbb{R}^d) \cap L^\infty((0,T) \times \mathbb{R}^d)$ be two distributional solutions to (29) for $\nu_0 \in \mathcal{P} \cap L^\infty$ such that $\nu_1 - \nu_2 \in L^\infty((0,T); H^{-1})$. If (9) holds with ν_i replacing u_i, $i = 1, 2$, then $\nu_1 \equiv \nu_2$.*

Proof. Since the proof is almost the same as that for Theorem 3.1, we only give a sketch of it below.
We first rewrite the original linearized equation using the operator L as follows:

$$\frac{\partial \nu}{\partial t} - L(\frac{\beta(x,u)}{u}\nu) + \mathrm{div}((b(x,u) - A\frac{\beta(x,u)}{u})\nu) = 0 \quad \text{in } \mathcal{D}'((0,T)\times\mathbb{R}^d),$$
$$\nu(0,x) = \nu_0(x), \quad x\in\mathbb{R}^d.$$

We set

$$z := \nu_1 - \nu_2,\ w := \frac{\beta(x,u)}{u}(\nu_1-\nu_2),$$

and define a mollification analogous to that in the proof of Theorem 3.1

$$z_\delta := z*\theta_\delta, w_\delta := w*\theta_\delta, \zeta_\delta := (b(x,u)(\nu_1-\nu_2))*\theta_\delta, (Aw)_\delta := (Aw)*\theta_\delta.$$

We note that our assumptions imply

$$\frac{\beta(x,u)}{u}, b(x,u) \in L^\infty((0,T)\times\mathbb{R}^d), \quad \frac{\beta(x,u)}{u} \geqslant 0, \forall x\in\mathbb{R}^d,$$

we can then define the same Ψ_ε and $h_{\varepsilon,\delta}(t)$ as in the proof of Theorem 3.1. Repeating the proof of Claim 3.1, we obtain $\lim\limits_{t\to 0}\operatorname*{ess\,sup}\limits_{s\in(0,t)} h_{\varepsilon,\delta}(s) = 0$. Repeating the proof of Claim 3.2, we have the following estimates.
a)

$$(w_\delta(t), z_\delta(t))_2 \geqslant \alpha_3 \left|w(t)\right|_2^2 + \gamma_\delta(t),$$

where $\gamma_\delta(t) = (w_\delta(t), z_\delta(t))_2 - (w(t), z(t))_2$, $\alpha_3^{-1} = \left|\frac{\beta(x,u)}{u}\right|_{L^\infty} + 1$.
b)

$$\begin{aligned}(\nabla\Psi_\varepsilon(z_\delta(t)), \zeta_\delta(t))_2 - (\nabla\Psi_\varepsilon(z_\delta(t)), (Aw)_\delta(t))_2 &\leqslant \left|\nabla\Psi_\varepsilon(z_\delta(t)\right|_2 \cdot (\left|\zeta(t)\right|_2 + \left|(Aw)_\delta(t)\right|_2)\\ &\leqslant \left|\nabla\Psi_\varepsilon(z_\delta(t)\right|_2 \cdot (\alpha_4\left|w(t)\right|_2 + \left|A\right|_\infty \left|w(t)\right|_2)\\ &\leqslant C\left|\nabla\Psi_\varepsilon(z_\delta(t)\right|_2 \left|w(t)\right|_2,\end{aligned}$$

where $\alpha_4 = \alpha_{[-|u|_\infty, |u|_\infty]}$, $C = \alpha_4 + |A|_\infty$.
c) For all $\varepsilon,\delta\in(0,1)$, we have

$$\varepsilon(\left|\Psi_\varepsilon(z_\delta(t))\right|, \left|w_\delta(t)\right|)_2 \leqslant \varepsilon\left|\Psi_\varepsilon(z_\delta(t))\right|_\infty \left|w_\delta(t)\right|_1 \leqslant c\left|z(t)\right|_\infty \left|w(t)\right|_1, \quad \text{a.e. } t\in(0,T).$$

$$\lim_{\varepsilon\to 0}\sup_{\delta\in(0,1)} \left|\varepsilon(\Psi_\varepsilon(z_\delta(s))\right|_\infty = 0, \quad \text{a.e. } s\in(0,T).$$

d) For $\delta(\varepsilon)$ as in Claim 3.2, we have

$$\int_0^T \left|(R_{\delta(\varepsilon)}(s), \Psi_\varepsilon(z_{\delta(\varepsilon)}(s)))_2\right| ds \leqslant \varepsilon \left|z\right|_{L^2(0,T;L^2)}.$$

Therefore, we have

$$0 \leqslant h_{\varepsilon,\delta(\varepsilon)}(t) \leqslant \tilde{\zeta}_\varepsilon \exp(\frac{C^2}{\alpha_3\gamma}t), \quad \forall t\in[0,T],$$

where $\tilde{\zeta}_\varepsilon$ is as in the proof of Claim 3.2. This implies that $h_{\varepsilon,\delta(\varepsilon)} \to 0$ as $\varepsilon\to 0$ uniformly on $[0,T]$, and hence $\nu_1 \equiv \nu_2$, as claimed.

$\square$

Corollary 4.1 *Assume that Hypothesis (I) holds. Let $T > 0$ and $\nu_0 \in \mathcal{P} \cap L^\infty$ and let $u \in L^1((0,T) \times \mathbb{R}^d) \cap L^\infty((0,T) \times \mathbb{R}^d)$ be a distributional solution to (2). If $\nu_1, \nu_2 \in L^1((0,T) \times \mathbb{R}^d) \cap L^\infty((0,T) \times \mathbb{R}^d)$ are two nonnegative distributional solutions to (29), then $\nu_1 \equiv \nu_2$.*

Proof. This follows immediately from Theorem 4.1 by the same method as in the proof of Corollary 3.1. □

The following result then standardly follows (see [5, Thm. 4.1]). For the readers' convenience we include the proof.

Theorem 4.2 *Assume that Hypothesis (I) holds. Let $T > 0$, and let $X^1(t)$ and $X^2(t)$ be two probabilistically weak solutions to (3) on the stochastic bases $(\Omega_1, \mathscr{F}_1, (\mathscr{F}^1_t)_{t\geqslant 0}, \mathbb{P}_1)$, $(\Omega_2, \mathscr{F}_2, (\mathscr{F}^2_t)_{t\geqslant 0}, \mathbb{P}_2)$ such that $\mathcal{L}_{X^1(t)}, \mathcal{L}_{X^2(t)}$ are absolutely continuous w.r.t Lebesgue measure dx and for*

$$u_1(t,\cdot) := \frac{d\mathcal{L}_{X^1(t)}}{dx}, \quad u_2(t,\cdot) := \frac{d\mathcal{L}_{X^2(t)}}{dx},$$

we have

$$u_1, u_2 \in L^1((0,T) \times \mathbb{R}^d) \cap L^\infty((0,T) \times \mathbb{R}^d).$$

Then, X^1 and X^2 have the same law, i.e. $\mathbb{P}_1 \circ (X^1)^{-1} = \mathbb{P}_2 \circ (X^2)^{-1}$.

Proof. By narrow continuity, we have

$$u(0,x)dx = u_0(dx) = (\mathbb{P}_1 \circ X^1(0)^{-1})(dx) = (\mathbb{P}_2 \circ X^2(0)^{-1})(dx),$$

$$u_1(t,\cdot), u_2(t,\cdot) \in L^\infty(\mathbb{R}^d), \text{ for all } t \in (0,T].$$

By Itô's formula, u_1 and u_2 are probability solutions to the nonlinear Fokker-Planck equation (5), hence, by Corollary 3.1, $u_1 \equiv u_2$. Again by Itô's formula, both marginal laws $\mathbb{P}_1 \circ (X^1)^{-1}$ and $\mathbb{P}_2 \circ (X^2)^{-1}$ satisfy the martingale problem with initial condition u_0 for the linear Kolmogorov operator

$$L_u := \sum_{i,j=1}^d \frac{a_{ij}(x)\beta(x,u)}{u} \frac{\partial^2}{\partial x_i \partial x_j} + b(x,u) \cdot \nabla.$$

Hence, by Corollary 4.1, the assertion follows from Lemma 2.12 in [31]. More specifically, for $s \in [0,t]$, choose the set $\mathcal{R}_{[s,T]}$ appearing in Lemma 2.12 to be the set of all weakly continuous probability solutions ν of (29) such that, for each $t \in [s,T]$, $t > 0$, $\nu(t,\cdot) \in L^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ and $\nu \in L^1((0,T) \times \mathbb{R}^d) \cap L^\infty((0,T) \times \mathbb{R}^d)$. Therefore, (2.9) and (2.14) in [31] hold. By Corollary 4.1, condition i) in Lemma 2.12 of [31] holds, and hence condition ii) in Lemma 2.12 of [31] follows; that is, X^1 and X^2 have the same law.

□

5 Nonlinear Markov processes

The results in this section standardly follow from our distributional uniqueness results in Section 3 (see e.g., [14] or [28]). For the readers' convenience we recall the necessary definitions and present the precise arguments necessarily to derive them from the main result in [28].

We first review the basic notions from [28]. Let $\Omega_s := C([s,\infty),\mathbb{R}^d)$ denote the space of continuous paths in $\mathbb{R}^d$ starting at time s, equipped with the Borel σ-algebra $\mathcal{B}(\Omega_s)$. For $\tau \geqslant s$,

$$\pi^s_\tau : \Omega_s \to \mathbb{R}^d, \ \ \pi^s_\tau(\omega) := \omega(\tau), \omega \in \Omega_s, \quad \mathcal{F}_{s,r} := \sigma(\pi^s_\tau | s \leqslant \tau \leqslant r).$$

Let $\tilde{\mathcal{P}}(\mathbb{R}^d)$ denote the set of all Borel probability measures on $\mathbb{R}^d$, and

$$\mathcal{P}_a = \left\{\mu \in \tilde{\mathcal{P}} | \mu \ll dx\right\}, \ \ \mathcal{P}_a^\infty = \left\{\mu \in \mathcal{P}_a | \frac{d\mu}{dx} \in L^\infty(\mathbb{R}^d)\right\} = \mathcal{P} \cap L^\infty.$$

As in Section 4, for an initial condition $(s,\zeta) \in \mathbb{R}^+ \times \tilde{\mathcal{P}}(\mathbb{R}^d)$, we denote the probabilistically weak solution of (3) by $(X^{s,\zeta}_t)_{t \geqslant s}$, defined on a stochastic basis $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant s}, \mathbb{P})$, and set $\mathbb{P}_{s,\zeta} := \mathbb{P} \circ (X^{s,\zeta}_\cdot)^{-1}$. Nonlinear Markov processes are given by such families of laws $\mathbb{P}_{s,\zeta}$, $s \geqslant 0$, on path space, where ζ ranges over all admissible initial probability measures on $\mathbb{R}^d$. The following is the rigorous definition.

Definition 5.1 *(Nonlinear Markov process [28]) Let $\mathcal{P}_0 \subseteq \tilde{\mathcal{P}}(\mathbb{R}^d)$. A nonlinear Markov process is a family $(\mathbb{P}_{s,\zeta})_{(s,\zeta)\in\mathbb{R}^+\times\mathcal{P}_0}$ of probability measures $\mathbb{P}_{s,\zeta}$ on $\mathcal{B}(\Omega_s)$ such that*
(i) *The marginals $\mathbb{P}_{s,\zeta} \circ (\pi^s_t)^{-1} =: \mu^{s,\zeta}_t$ belong to $\mathcal{P}_0$ for all $0 \leqslant s \leqslant t$, $\zeta \in \mathcal{P}_0$.*
(ii) *The nonlinear Markov property holds, i.e., for all $0 \leqslant s \leqslant r \leqslant t$, $\zeta \in \mathcal{P}_0$,*

$$\mathbb{P}_{s,\zeta}(\pi^s_t \in A | \mathcal{F}_{s,r})(\cdot) = p_{(s,\zeta),(r,\pi^s_r(\cdot))}(\pi^r_t \in A) \quad \mathbb{P}_{s,\zeta} - a.s. \quad \text{for all } A \in \mathcal{B}(\mathbb{R}^d),$$

where $p_{(s,\zeta),(r,y)}$, $y \in \mathbb{R}^d$, is a regular conditional probability kernel from $\mathbb{R}^d$ to $\mathcal{B}(\Omega_r)$ of $\mathbb{P}_{r,\mu^{s,\zeta}_r}[\cdot|\pi^r_r = y]$, $y \in \mathbb{R}^d$ (i.e., in particular $p_{(s,\zeta),(r,y)} \in \tilde{\mathcal{P}}(\Omega_r)$ and $p_{(s,\zeta),(r,y)}(\pi^r_r = y) = 1$).

Remark 5.1 The one-dimensional time-marginals $\mu^{s,\zeta}_t$ of a nonlinear Markov process satisfy the flow property, i.e.,

$$\mu^{s,\zeta}_t = \mu^{r,\mu^{s,\zeta}_r}_t, \ \ \forall\, 0 \leqslant s \leqslant r \leqslant t, \ \ \zeta \in \tilde{\mathcal{P}}(\mathbb{R}^d),$$

which corresponds to the Chapman-Kolmogorov equations for the linear Markov process.

Theorem 5.1 *Consider NFPE (2), assume that Hypotheses **(i')**,**(ii)**, **(iii')**, **(iv)**-**(v)** hold. Then there exists a nonlinear Markov process $(\mathbb{P}_{s,\zeta})_{(s,\zeta)\in[0,T]\times\mathcal{P}_a^\infty}$ such that, for each $(s,\zeta) \in [0,T] \times \mathcal{P}_a^\infty$, the measure $\mathbb{P}_{s,\zeta}$ is the path law of the unique weak solution to the McKean-Vlasov SDE (3) with initial condition (s,ζ).*

Proof. We identify a measure that is absolutely continuous with respect to Lebesgue measure dx with its density, that is, we identify $u^{s,\zeta}$ with $(\mu^{s,\zeta}_t)_{t\geqslant s}$, where $\mu^{s,\zeta}_t := u^{s,\zeta}(t,x)dx$. We first note that all previous theorems continue to hold when the nonlinear Fokker-Planck equation (2) is considered on $[s,\infty) \times \mathbb{R}^d$ for any $s \geqslant 0$, with an initial condition u_0 renamed as $\zeta \in \mathcal{P}_a^\infty$. Then, by Corollary 3.2, there exists a weakly continuous probability solution $u^{s,\zeta}(t,x)$ to (2) from (s,ζ) such that

$$u^{s,\zeta}(t,x) \in \bigcap_{T>s} L^\infty((s,T)\times\mathbb{R}^d), \quad u^{s,\zeta}(t,\cdot) \in \mathcal{P}_a^\infty, \text{ for all } t \geqslant s,$$

and $\left\{u^{s,\zeta}(t,x)\right\}_{(s,\zeta)\in\mathbb{R}^+\times\mathcal{P}_a^\infty}$ is a solution flow in $\mathcal{P}_a^\infty$. By the superposition principle, there exists a probabilistically weak solution $X_t^{s,\zeta}$ to (3) with initial condition $(s,\zeta) \in [0,T]\times\mathcal{P}_a^\infty$.

Moreover, by Corollary 4.1, the linearized equation (28) has a unique weakly continuous probability solution in $\bigcap_{T>s} L^\infty((s,T)\times\mathbb{R}^d)$. Therefore, by Corollary 3.9 in [28], there exists a nonlinear Markov process $(\mathbb{P}_{s,\zeta})_{(s,\zeta)\in[0,T]\times\mathcal{P}_a^\infty}$ with one-dimensional time-marginals $\mathbb{P}_{s,\zeta}\circ(\pi_t^s)^{-1} = u^{s,\zeta}(t,x)dx$. Finally, for $(s,\zeta)\in[0,T]\times\mathcal{P}_a^\infty$, by Theorem 4.2, $\mathbb{P}_{s,\zeta}$ is the path law of the unique probabilistically weak solution $X_t^{s,\zeta}$ to the McKean-Vlasov SDE (3).

□

Appendix

Proof of Theorem 3.2:

Here, we even prove the theorem for general diffusion terms $a_{ij}(x,u)$ satisfying the conditions (H1)-(H3) in [14], Section 2.1, and not only for our special case $\tilde{a}_{ij}(x,u) = \dfrac{a_{ij}(x)\beta(x,u)}{u}$. We first recall the main steps of the proof of the existence of mild solutions in [14] (also in [5]). We define in the space L^1 the operator $A_0 : D(A_0) \subset L^1 \to L^1$,

$$\begin{aligned} &A_0(u) = -\sum_{i,j=1}^d D_{ij}^2(a_{ij}(x,u)u) + \operatorname{div}(b(x,u)u), \quad \forall u \in D(A_0),\\ &D(A_0) = \left\{u \in L^1;\ A_0(u) \in L^1\right\}. \end{aligned}$$

Once the operator $A \subset A_0$ is shown to be m-accretive in L^1, the operator A generates a continuous contraction semigroup $S(t)$ in L^1 given by the exponential formula

$$S(t)u_0 = \lim_{n\to\infty}\left(I + \frac{t}{n}A\right)^{-n} u_0, \quad u_0 \in L^1,\ \forall t \geqslant 0. \tag{30}$$

Formula (30) means that $u(t) = S(t)u_0$ is a mild solution to the Cauchy problem

$$\begin{aligned} &\frac{du}{dt} + Au = 0 \quad \text{on } (0,\infty),\\ &u(0) = u_0, \end{aligned}$$

and hence a mild solution to NFPE (1). Consider the following regularized functions

$$a_{ij}^\epsilon(x,u) := \phi_\epsilon(x,u)a_{ij}(x,u) + (1-\phi_\epsilon(x,u))\max(\gamma,\tilde{\gamma})\delta_{ij},$$

$$b_i^\epsilon(x,u) := \phi_\epsilon(x,u)b_i(x,u),$$

for $1 \leqslant i,j \leqslant d$, $x \in \mathbb{R}^d$, $u \in \mathbb{R}$, $\tilde{\gamma} := d \sup_{1\leqslant i,j\leqslant d} |a_{ij}|_\infty$ such that $\sum_{i,j=1}^{d} a_{ij}(x,u)\xi_i\xi_j \leqslant \tilde{\gamma}\,|\xi|^2$ for any $u \in \mathbb{R}$, $\xi \in \mathbb{R}^d$. $\phi_\epsilon(x,u) := \phi(\epsilon\,|u|)\phi(\epsilon\,|x|^2)$ with $\phi \in C^\infty(\mathbb{R})$, $\mathbb{1}_{(-\infty,1]} \leqslant \phi \leqslant 1 - \mathbb{1}_{[2,\infty)}$, $-2 \leqslant \phi' \leqslant 0$. a^ϵ_{ij} and b^ϵ_i also satisfy the following conditions

$$\begin{aligned}(a^{\epsilon,*}_{ij}(x,u))_u \in & C_b(\mathbb{R}^d\times\mathbb{R}), \quad \text{and for some } \tilde{C}\in(0,\infty), g\in L^2, 1\leqslant i,j\leqslant d,\\ & \left|(a^{\epsilon,*}_{ij})_u(x,u) - (a^{\epsilon,*}_{ij})_u(x,\bar{u})\right| \leqslant \tilde{C}\,|u-\bar{u}|\,,\\ & \left|\tilde{b}^{\epsilon,*}(x,u) - \tilde{b}^{\epsilon,*}(x,\bar{u})\right| \leqslant g(x)\,|u-\bar{u}|\,, \quad \forall u,\bar{u}\in\mathbb{R},\ x\in\mathbb{R}^d\end{aligned} \tag{31}$$

where $a^{\epsilon,*}_{ij}(x,u) = a^\epsilon_{ij}(x,u)u$, $b^{\epsilon,*}(x,u) = b^\epsilon(x,u)u$, $\tilde{b}^{\epsilon,*}_i(x,u) := b^{\epsilon,*}_i - \sum_{j=1}^{d}(a^{\epsilon,*}_{ij})_{x_j}$, $\tilde{b}^{\epsilon,*} := (\tilde{b}^{\epsilon,*}_1,\dots,\tilde{b}^{\epsilon,*}_d)$. Then for $\lambda^*_0 \in (0, \frac{1}{2}\gamma(b_\infty + c_\infty)^{-2})$, where

$$b_\infty := \sup\left\{|b_i(x,u)|\,; (x,u)\in\mathbb{R}^d\times\mathbb{R}, i=1,\dots,d\right\},$$

$$\begin{aligned}c_\infty := \sup & \left\{\left|(a_{ij})_{x_j}(x,u)\right|; (x,u)\in\mathbb{R}^d\times\mathbb{R}, i,j=1,\dots,d\right\}\\ & + 8(\max_{i,j}|a_{ij}|_\infty \cdot \max_{i,j}\left|(a_{ij})_{x_j}\right|_\infty + \max(\gamma,\tilde{\gamma})),\end{aligned}$$

the equation

$$u - \lambda\sum_{i,j=1}^{d} D^2_{ij}a^{\epsilon,*}_{ij}(x,u) + \lambda\mathrm{div}(b^\epsilon(x,u)u) = f, \tag{32}$$

has a solution $u_\epsilon = u_\epsilon(\lambda,f)$ for each $\lambda\in(0,\lambda_0]$ and $\forall f\in L^1\cap L^2$ (also $\forall f \in L^1$). We also have, for $\lambda\in(0,\lambda^*_0]$, $f\in L^1$, along a subsequence $\{\epsilon\}\to 0$,

$$\begin{aligned}u_\epsilon(\lambda,f)\to u = u(\lambda,f) \quad & \text{weakly in } H^1,\\ & \text{strongly in } L^2_{\mathrm{loc}},\\ & \text{weak-star in } L^\infty.\end{aligned} \tag{33}$$

Defining $u(\lambda,f) := J_\lambda(f)$ and $u_\epsilon(\lambda,f) := J^\epsilon_\lambda(f)$ for $f\in L^1$, where $J^\epsilon_\lambda : L^1 \to D(A^\epsilon_0) \subset L^1$ and $(A^\epsilon_0, D(A^\epsilon_0))$ is defined as $(A_0, D(A_0))$ with a^ϵ_{ij} and b^ϵ_i replacing a_{ij} and b_i, respectively, we have

$$J^\epsilon_\lambda(f)\to J_\lambda(f) \quad \text{strongly in } L^1. \tag{34}$$

We now continue the proof. Take $M = \dfrac{\lambda}{\lambda_0}|f|_\infty$, where $\lambda_0 = \min\left\{\dfrac{\gamma}{2(b_\infty+c_\infty)}, \dfrac{1}{2C_0}\right\}$, $\lambda\in(0,\lambda_0)$, $C_0 = 2c_d\,|\delta|_\infty$, $\delta(r) = \max(\delta_1(r),\delta_2(r))$, where δ_1 is as defined in (6) and δ_2 is as defined in (7) with a_{ij} replacing $\tilde{a}_{ij}$, and c_d is a constant depending only on the dimension d. By (32), we have

$$\begin{aligned}u_\epsilon - |f|_\infty - M - \lambda & \sum_{i,j=1}^{d} D^2_{ij}(a^{\epsilon,*}_{ij}(x,u_\epsilon) - a^{\epsilon,*}_{ij}(x,|f|_\infty + M))\\ & + \lambda\mathrm{div}\,(b^\epsilon(x,u_\epsilon)u_\epsilon - b^\epsilon(x,|f|_\infty+M)(|f|_\infty+M)) = f - |f|_\infty - M\\ & - \lambda(|f|_\infty+M)[\mathrm{div}\,(b^\epsilon(x,|f|_\infty+M)) - \sum_{i,j=1}^{d} D^2_{ij}(a^{\epsilon,*}_{ij}(x,|f|_\infty+M))].\end{aligned} \tag{35}$$

Since

$$
\begin{aligned}
-\lambda(|f|_\infty + M)[\mathrm{div}\,(b^\epsilon(x, |f|_\infty + M)) - \sum_{i,j=1}^{d} D_{ij}^2(a_{ij}^{\epsilon,*}(x, |f|_\infty + M))] \\
&\leqslant c_d(\delta_1(|f|_\infty + M) + \delta_2(|f|_\infty + M)) \cdot \frac{\gamma(|f|_\infty + M)}{2(b_\infty + c_\infty)^2} \\
&\leqslant \lambda C_0(|f|_\infty + M) \\
&\leqslant M,
\end{aligned} \tag{36}
$$

the right-hand side of (35) is non-positive. Thus, multiplying (35) by $\mathcal{X}_\delta((u_\epsilon - |f|_\infty - M)^+)$ and integrating over $\mathbb{R}^d$, by (31), we obtain

$$
\begin{aligned}
&\int_{\mathbb{R}^d} (u_\epsilon - |f|_\infty - M)\mathcal{X}_\delta((u_\epsilon - |f|_\infty - M)^+)dx \\
&\leqslant \int_{\mathbb{R}^d} \lambda \sum_{i,j=1}^{d} D_{ij}^2(a_{ij}^{\epsilon,*}(x, u_\epsilon) - a_{ij}^{\epsilon,*}(x, |f|_\infty + M)) \cdot \mathcal{X}_\delta((u_\epsilon - |f|_\infty - M)^+)dx \\
&- \int_{\mathbb{R}^d} \lambda \mathrm{div}\,(b^\epsilon(x, u_\epsilon)u_\epsilon - b^\epsilon(x, |f|_\infty + M)(|f|_\infty + M)) \cdot \mathcal{X}_\delta((u_\epsilon - |f|_\infty - M)^+)dx \\
&= -\int_{\mathbb{R}^d} \lambda \sum_{i,j=1}^{d} D_i(a_{ij}^{\epsilon,*}(x, u_\epsilon) - a_{ij}^{\epsilon,*}(x, |f|_\infty + M)) \cdot D_j\mathcal{X}_\delta((u_\epsilon - |f|_\infty - M)^+)dx \\
&+ \int_{\mathbb{R}^d} \lambda\,(b^\epsilon(x, u_\epsilon)u_\epsilon - b^\epsilon(x, |f|_\infty + M)(|f|_\infty + M)) \cdot \nabla\mathcal{X}_\delta((u_\epsilon - |f|_\infty - M)^+)dx \\
&= \int_{\mathbb{R}^d} \lambda \left(\tilde{b}^{\epsilon,*}(x, u_\epsilon) - \tilde{b}^{\epsilon,*}(x, |f|_\infty + M)\right) \cdot \nabla\mathcal{X}_\delta((u_\epsilon - |f|_\infty - M)^+)dx \\
&= \frac{\lambda}{\delta} \int_{[(u_\epsilon - |f|_\infty - M)^+ \leqslant \delta]} \left(\tilde{b}^{\epsilon,*}(x, u_\epsilon) - \tilde{b}^{\epsilon,*}(x, |f|_\infty + M)\right) \cdot \nabla u_\epsilon dx.
\end{aligned}
$$

Since

$$
\begin{aligned}
\lim_{\delta\to 0} \frac{\lambda}{\delta} \int_{[(u_\epsilon - |f|_\infty - M)^+ \leqslant \delta]} \left| \left(\tilde{b}^{\epsilon,*}(x, u_\epsilon) - \tilde{b}^{\epsilon,*}(x, |f|_\infty + M)\right) \cdot \nabla u_\epsilon \right| dx \\
\leqslant \lambda\, |g|_2 \lim_{\delta\to 0} \left(\int_{[(u_\epsilon - |f|_\infty - M)^+ \leqslant \delta]} |\nabla u_\epsilon|^2\, dx \right)^{\frac{1}{2}} = 0,
\end{aligned}
$$

we have

$$
|u_\epsilon|_\infty \leqslant (1 + \frac{\lambda}{\lambda_0})\, |f|_\infty .
$$

Thus, by (33) and (34), we obtain

$$
|J_\lambda(u)|_\infty \leqslant (1 + \frac{\lambda}{\lambda_0})\, |u|_\infty, \quad \forall u \in L^1 \cap L^\infty,
$$

take $C = (1/\lambda_0)$, by (30), the proof is complete. $\square$

Acknowledgements. This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project-ID 317210226 - SFB 1283. The second named author also gratefully acknowledges financial support by the DFG - IRTG 2235 - Project-ID 282638148.

References

[1] Aronson, D.G., Non-negative solutions of linear parabolic equations, Ann. Scuola Norm. Sup. Pisa. Cl. Sci., 22 (1968. Addendum 25:221-228, 1971), pp. 607-694.

[2] Barbu, V., Nonlinear Differential Equations of Monotone Type in Banach Spaces, Springer, Berlin. Heidelberg. New York, 2010.

[3] Barbu, V., Nonlinear Fokker–Planck equations with singular integral drifts and McKean–Vlasov SDEs, (2024), arXiv:2410.09822.

[4] Barbu, V., Röckner, M., Probabilistic representation for solutions to nonlinear FP equations, SIAM J. Math. Anal., 50 (2018), 2588-2607.

[5] Barbu, V., Röckner, M., From Fokker–Planck equations to solutions of distribution dependent SDE, Annals of Probability, 48 (2020), 1902- 1920.

[6] Barbu, V., Röckner, M., Uniqueness for nonlinear Fokker–Planck equations and weak uniqueness for McKean–Vlasov SDEs, Stoch. PDE Anal. Comput., 9 (2021), 702-713.

[7] Barbu, V., Röckner, M., Stochastic PDEs, Correction to: Uniqueness for nonlinear Fokker–Planck equations and weak uniqueness for McKean–Vlasov SDEs, Stoch. PDE: Anal. Comp. https://doi.org/ 10.1007/S40072-020-0018-8.

[8] Barbu, V., Röckner, M., Solutions for nonlinear Fokker–Planck equations with measures as initial data and McKean–Vlasov equations, J. Functional Anal., 280 (7) (2021), 1-35.

[9] Barbu, V., Röckner, M., The invariance principle for nonlinear Fokker–Planck equations, J. Differential Equations, 315 (2022), 200-221.

[10] Barbu, V., Röckner, M., Nonlinear Fokker–Planck equations with time-dependent coefficients, SIAM J. Math. Anal., 55 (1) (2023), 1-18.

[11] Barbu, V., Röckner, M., The evolution to equilibrium of solutions to nonlinear Fokker–Planck equations, Indiana University Math. Journal, 72 (1) (2023), 89-131.

[12] Barbu, V., Röckner, M., Uniqueness for nonlinear Fokker–Planck equations and for McKean–Vlasov SDEs: The degenerate case, J. Functional Anal., 285 (2023) 109980.

[13] Barbu, V., Röckner, M., Nonlinear Fokker–Planck equations with fractional Laplacian and McKean–Vlasov SDEs with Lévy-Noise, Prob. Th. Rel. Fields 189 (2024), no. 3-4, 849-878.

[14] Barbu, V., Röckner, M., Nonlinear Fokker–Planck Flows and their Probabilistic Counterparts, Lecture Notes in Mathematics, vol. 2353, Springer, 2024.

[15] Barbu, V., Röckner, M., Zhang, D., Uniqueness of distributional solutions to the 2D vorticity Navier–Stokes equation and its associated Markov process, J. European Math. Soc. (to appear).

[16] Belaribi, N., Russo, F., Uniqueness for Fokker–Planck equations with measurable coefficients and applications to the fast diffusion equations, Electron. J. Probab., 17 (2012), 1-28.

[17] Bogachev, V.I., Krylov, N.V., Röckner, M., Stanislav V. Shaposhnikov, S.V., Fokker–Planck-Kolmogorov equations, Mathematical Surveys and Monographs, vol. 207, American Mathematical Society, Providence, RI, 2015.

[18] Brezis, H., Crandall, M.G., Uniqueness of solutions of the initial-value problem for $u_t - \Delta\beta(u) = 0$, J. Math. Pures et Appl., 58 (1979), 153-163.

[19] Franck, T.D., Nonlinear Fokker–Planck Equations, Springer, Berlin. Heidelberg. New York, 2005.

[20] Fukushima, M., Oshima, Y., Takeda, M., Dirichlet Forms and Symmetric Markov Processes, De Gruyter, Berlin, New York, 1994.

[21] Lasry, J.M., Lions, P.L., Mean field games, Japan J. Math., 2 (2007), 229-260.

[22] Ma, Z., Röckner, M., Introduction to the theory of (nonsymmetric) Dirichlet forms, Universitext. Springer Verlag, Berlin, 1992.

[23] McKean, H.P., A class of Markov processes associated with nonlinear parabolic equations, Proc. Natl. Acad. Sci. USA 56(6), 1907–1911 (1966).

[24] Metafune, G., Pallara, D., Vespri, V., L^p-estimates for a class of elliptic operators with unbounded coefficients in R^N, Houston Journal of Mathematics, 31 (2005), no. 2, 605–620.

[25] Mishura, Y., Veretennikov, A., Existence and uniqueness theorems for solutions of McKean–Vlasov stochastic equations. Theory of Probability and Mathematical Statistics, 2020. MR4421344.

[26] Pierre, M., Uniqueness of the solutions of $u_t - \Delta\varphi(u) = 0$ with initial data measure, Nonlinear Anal. Theory Methods Appl., 6 (2) (1982), 175-187.

[27] Rehmeier, M., Flow selections for (nonlinear) Fokker–Planck–Kolmogorov equations, J. Differential Equations, 328 (2022), 105-132.

[28] Rehmeier, M., Röckner, M., On Nonlinear Markov Processes in the Sense of McKean, J. Theoret. Prob. 38 (2025), no. 3, 60.

[29] Röckner, M., Zhang, X., Weak uniqueness of Fokker–Planck equations with degenerate and bounded coefficients. C. R. Math. Acad. Sci. Paris, 348(7-8):435–438, 2010.

[30] Stein, E., Weiss, G., Introduction to Fourier Analysis on Euclidean Spaces, Princeton University Press, Princeton, N.J., 1971.

[31] Trevisan, D., Well-posedness of multidimensional diffusion processes with weakly differentiable coefficients, Electron J. Probab., 21 (2016), 22-41.