Strong solutions to McKean–Vlasov SDEs associated to a class of degenerate Fokker–Planck equations with coefficients of Nemytskii-type

Sebastian Grube*

May 14, 2024

Abstract

While the nondegenerate case is well-known, there are only few results on the existence of strong solutions to McKean–Vlasov SDEs with coefficients of Nemytskii-type in the degenerate case. We consider a broad class of degenerate nonlinear Fokker–Planck(–Kolmogorov) equations with coefficients of Nemytskii-type. This includes, in particular, the classical porous medium equation perturbed by a first-order term with initial datum in a subset of probability densities, which is dense with respect to the topology inherited from L^1, and, in the one-dimensional setting, the classical porous medium equation with initial datum in an arbitrary point $x_0 \in \mathbb{R}$. For these kind of equations the existence of a Schwartz-distributional solution u is well-known. We show that there exists a unique strong solution to the associated degenerate McKean–Vlasov SDE with time marginal law densities u. In particular, every weak solution to this equation with time marginal law densities u can be written as a functional of the driving Brownian motion. Moreover, plugging any Brownian motion into this very functional yields a weak solution with time marginal law densities u.

1 Introduction

We consider the following McKean–Vlasov stochastic differential equation (abbreviated by McKean–Vlasov SDE or MVSDE) in $\mathbb{R}^d$, $d \in \mathbb{N}$, with coefficients of Nemytskii-type, which in our case is of the form

$$\mathrm{d}X(t) = E(X(t))b\left(\frac{\mathrm{d}\mathcal{L}_{X(t)}}{\mathrm{d}x}(X(t))\right)\mathrm{d}t + \sqrt{2\frac{\beta\left(\frac{\mathrm{d}\mathcal{L}_{X(t)}}{\mathrm{d}x}(X(t))\right)}{\frac{\mathrm{d}\mathcal{L}_{X(t)}}{\mathrm{d}x}(X(t))}}\mathbb{1}_{d\times d}\,\mathrm{d}W(t),$$

$$X(0) = X_0, \tag{MVSDE}$$

where $t \in [0, T]$, $T \in (0, \infty)$, $\mathbb{1}_{d\times d}$ is the d-dimensional unit matrix, $(W(t))_{t\in[0,T]}$ is a standard d-dimensional $(\mathcal{F}_t)$-Brownian motion and X_0 an $\mathcal{F}_0$-measurable function on some stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t\in[0,T]})$, i.e. a complete, filtered probability space, where $(\mathcal{F}_t)_{t\in[0,T]}$ is a normal filtration, and $\mathcal{L}_{X(t)} := \mathbb{P} \circ (X(t))^{-1}, t \in [0, T]$, denote the one-dimensional time marginal laws of the solution process X. Here, we assume that the drift and diffusion coefficient are given through locally bounded Borel measurable functions

$$E : \mathbb{R}^d \to \mathbb{R}^d,\ b : \mathbb{R} \to \mathbb{R},\ \beta : \mathbb{R} \to \mathbb{R},$$

---

*Faculty of Mathematics, Bielefeld University, 33615 Bielefeld, Germany. E-Mail: sgrube@math.uni-bielefeld.de

which satisfy conditions (H1), (H2) below with the convention $\beta(0)/0 := \beta'(0)$. The McKean–Vlasov SDE (MVSDE) arises from the study of the following type of nonlinear Fokker–Planck(–Kolmogorov) equation (FPE)

$$\partial_t u(t,x) + \operatorname{div}(E(x)b(u(t,x))u(t,x)) - \Delta(\beta(u(t,x))) = 0, \quad (t,x) \in (0,T) \times \mathbb{R}^d,$$
$$u|_{t=0} = \nu \in M_b(\mathbb{R}^d), \tag{FPE}$$

(here $M_b(\mathbb{R}^d)$ denotes the set of bounded (signed) Borel measures on $\mathbb{R}^d$). In [BR20], the authors consider also the case of a general non-diagonal diffusion matrix with explicit x-dependence. For simplicity we consider nonlinear Fokker–Planck–Kolmogorov equations of the above type. As usual, (FPE) is considered in the Schwartz-distributional sense. We will say that $u : (0,T) \to L^1(\mathbb{R}^d)$, in short form $(u_t)_{t \in (0,T)}$, is a Schwartz-distributional solution to (FPE) with $u|_{t=0} = \nu \in M_b(\mathbb{R}^d)$ if $u, \beta(u) \in L^1_{loc}([0,T) \times \mathbb{R}^d)$, $[0,T) \ni t \mapsto \int \varphi \left(\mathbb{1}_{(0,T)}(t)u(t,x)\mathrm{d}x + \mathbb{1}_{\{0\}}(t)\mathrm{d}\nu \right)$ is continuous for all $\varphi \in C_b(\mathbb{R}^d)$ and

$$0 = \int_0^T \int_{\mathbb{R}^d} \langle \partial_t \varphi(s,x) + E(x)b(u(s,x)), \nabla\varphi(s,x) \rangle_{\mathbb{R}^d} u(s,x) + \Delta\varphi(s,x)\beta(u(s,x))\mathrm{d}x\mathrm{d}s$$
$$+ \int_{\mathbb{R}^d} \varphi(0,x)\nu(\mathrm{d}x), \tag{1}$$

for all $\varphi \in C_c^\infty([0,T) \times \mathbb{R}^d)$. If, additionally, $\nu \in \mathcal{P}(\mathbb{R}^d)$ and $u(t,\cdot) \in \mathcal{P}(\mathbb{R}^d)$ for all $t \in (0,T)$, then u is simply called a *probability solution* to (FPE) (with $u|_{t=0} = \nu$). Note that a probability solution u can be uniquely extended to a continuous function $u : [0,T] \to \mathcal{P}(\mathbb{R}^d)$, where we consider the topology of narrow convergence on $\mathcal{P}(\mathbb{R}^d)$, see [Reh22, Lemma 2.3].

In applications, (FPE) is used to describe particle transport in disordered media. A prominent particular example for this kind of equation is the classical porous medium equation, in which case $\beta(r) = |r|^{m-1}r, r \in \mathbb{R}, m > 1, E \equiv b \equiv 0$. It is most known for the usage as a model of diffusion of an ideal gas in a homogenous porous medium. For a deep account on the porous medium equation and equations of the general form (FPE), we refer to [Váz07], [Fra05] and the references therein, respectively.

In [BR20] (see also [BR18]), Barbu and Röckner introduced a general approach to solve (MVSDE) weakly in the probabilistic sense by first proving the existence of a Schwartz-distributional solution u to (FPE) and then, via a superposition principle procedure based on Trevisan's superposition principle for SDEs ([Tre16]; see also the recent generalisation [BRS21] and a superposition principle in the nonlocal case [RXZ20]), constructing a probabilistically weak solution (X,W) to (MVSDE); this is done in such a way that the time marginal law densities are given by u, i.e. $\mathcal{L}_{X(t)} = u_t(x)\mathrm{d}x$. This procedure was applied under various conditions on the drift and diffusion coefficients, see [BR21a; BR23b; BR23a]. For related uniqueness results we refer to [BR21b; BR23c].

The above sketched approach constitutes a vital part to McKean's idea. In [McK66], McKean envisioned a connection between certain classes of nonlinear PDEs, including the viscous Burgers' equation or the classical porous medium equations (in dimension one), and nonlinear Markov processes in such a way that the latter's transition probabilities solve the PDE. Recently, Rehmeier and Röckner created a rigorous notion for a nonlinear Markov process in line with McKean's idea, see [RR23]. The authors proved the existence of nonlinear Markov processes for a large class of (MVSDE) including, in particular, those which are associated to the viscous Burgers' equation and the classical porous medium equation even in the multi-dimensional case. Furthermore, their results encompass the 2D vorticity Navier–Stokes equation and a large class of non-local Fokker–Planck equations.

The aim of this paper is to show that the probabilistically weak solutions for a wide class of *degenerate* (i.e. $\beta(r)/r$ is allowed to vanish) McKean–Vlasov SDEs of type (MVSDE) provided by [BR21a] are actually probabilistically strong solutions and pathwise unique among all weak solutions with time marginal law densities u, where u denotes the probability solution constructed in [BR21a]. The class of McKean–Vlasov SDEs includes those associated to the classical porous medium equation perturbed by a nonlinear hyperbolic term of tensor type. In order to be exact, we impose the following conditions on the coefficients of (MVSDE).

Assume there exist $m \in (1, \infty)$ and $\zeta \in [0, 1]$ with $\frac{2\zeta}{m} < 1$ such that the following assumptions hold.

(H1) $\beta \in C^2(\mathbb{R})$; β strictly increasing; $\beta(0) = 0$;
$\forall K > 0 \; \exists a_K > 0, C_K > 0 \; \forall r \in [0, K] : a_K r^{m-1} \leq \beta'(r)$ and $\beta(r) \leq C_K r^m$.

(H2) (a) $E \in L^\infty(\mathbb{R}^d; \mathbb{R}^d), \operatorname{div} E \in L^2_{loc}(\mathbb{R}^d), (\operatorname{div} E)^- \in L^\infty(\mathbb{R}^d)$, and for all $R > 0$ there exists $\iota_R \in L^1(B_R(0))$ such that for a.e. $x, y \in B_R(0)$

$$\langle E(x) - E(y), x - y \rangle_{\mathbb{R}^d} \leq (\iota_R(x) + \iota_R(y)) \, |x - y|^2_{\mathbb{R}^d}. \tag{2}$$

(b) (i) $b \in C_b(\mathbb{R}) \cap C^1(\mathbb{R})$; $b \geq 0$;
(ii) $b \circ (G \circ \beta^{\frac{1}{2}})^{-1} \in \operatorname{Lip}_{loc}([0, \infty))$, where $G(r) := \int_0^r ((\beta^{-1})(s^2))^{-\zeta} \mathrm{d}s, r \geq 0$.

Additionally, one of the following conditions is assumed:

(iii) $\forall K > 0 \; \exists C_K > 0 \; \forall r \in [0, K] : \beta'(r) r \leq C_K \beta(r)$;

or

(iv) $E \in L^2(\mathbb{R}^d; \mathbb{R}^d)$, $\operatorname{div} E \in L^2(\mathbb{R}^d) + L^\infty(\mathbb{R}^d)$, $\operatorname{div} E \geq 0$.

Remark 1.1. *(i) Assume for a moment that β is only assumed to be locally bounded. Then, the condition*

$$\forall K > 0 \; \exists C_K > 0 \; \forall r \in [0, K] : \beta(r) \leq C_K r^m,$$

is clearly equivalent to

$$\exists \varepsilon > 0, C > 0 \; \forall r \in [0, \varepsilon] : \beta(r) \leq C r^m.$$

(ii) (H1) implies that

$$\frac{\beta(r)}{r} \geq 0 \; \text{with} \; \frac{\beta(0)}{0} = \beta'(0) = 0,$$

and, therefore, that (MVSDE) is degenerate. Indeed, for some constant $C > 0$,

$$\left| \frac{\beta(r)}{r} \right| \leq C |r|^{m-1} \to 0, \; \text{as } r \downarrow 0.$$

(iii) If $E \in H^1_{loc}(\mathbb{R}^d; \mathbb{R}^d)$, then clearly $\operatorname{div} E \in L^2_{loc}(\mathbb{R}^d)$ and, setting $\iota_R := \mathrm{M}_{2R} |\nabla E| \; (\in L^2_{loc}(\mathbb{R}^d))$, where M_{2R} is the local maximal function (see Definition A.1), E also fulfills (2) by Lemma A.2 and Lemma A.3 from the Appendix.

(iv) Note that, under condition (H1), G, as introduced in (H2) (b) (ii), is a real-valued function.

(v) Assume that (H1) holds. Then, $G^{-1} \in C^1([0, \infty))$. Hence, additionally assuming (H2) (b) (i), $b \circ (\beta^{\frac{1}{2}})^{-1} \in \operatorname{Lip}_{loc}(\mathbb{R})$ implies (H2) (b) (ii). Indeed, the first claim ist true: Since G is bijective from $[0, \infty)$ to $[0, \infty)$, G^{-1} exists. Furthermore, since $G' > 0$ on $(0, \infty)$, G^{-1} is continuously differentiable on $(0, \infty)$ with $(G^{-1})'(r) = (\beta^{-1}((G^{-1}(r))^2))^\zeta, r \in (0, \infty)$.

(vi) Consider the case when $\beta(r) = |r|^{m-1} r, r \in \mathbb{R}$. Assume that b satisfies (H2) (b) (i) and that $b(r) = r^l, r \in [0, \varepsilon]$ for some $\varepsilon > 0$ and $l \geq 1$. Then, (H2) (b) (ii) is satisfied if and only if $l \geq \frac{m}{2} - \zeta$.

In the following, $\mathcal{P}_0(\mathbb{R}^d)$ denotes the set of all probability densities with respect to the n-dimensional Lebesgue measure. Also, we set

$$\mathcal{Q} := \{ \Phi \in C^2(\mathbb{R}^d) : \Phi \geq 1; \Phi^{-\alpha} \in L^1(\mathbb{R}^d) \text{ for some } \alpha \in [2, \infty); |\nabla \Phi|, \Delta \Phi \in L^\infty(\mathbb{R}^d) \}$$

Note that $\mathcal{Q} \neq \emptyset$, since $\Phi \in \mathcal{Q}$ with $\Phi(x) := (1 + |x|^2)^\gamma, x \in \mathbb{R}^d$, with $\gamma \in (0, \frac{1}{2}]$, see [BR21a, Proof of Lemma 3.3]. Furthermore, we define for nonnegative Borel measurable $\Phi : \mathbb{R}^d \to \mathbb{R}$

$$L^1_\Phi(\mathbb{R}^d) := \left\{ f : \mathbb{R}^d \to [-\infty, \infty] \ \mathcal{B}(\mathbb{R}^d)\text{-measurable} : ||f||_{L^1_\Phi} := \int_{\mathbb{R}^d} |f(x)|\Phi(x)\mathrm{d}x < \infty \right\}.$$

The first main result of this work says that under the conditions above, there exists a probabilistically strong solution to (MVSDE), which is pathwise unique among weak solutions with time marginal law densities u provided by Theorem 3.3. The corresponding precise statement in form of Theorem 1.2 is formulated in the terminology introduced on page 6.

Theorem 1.2 (main result 1). *Let $d \neq 2$. Assume that (H1) and (H2) hold. Let $\nu \in \mathcal{P}_0(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ for which there exists $\Phi \in \mathcal{Q}$ such that $\Phi \in L^1_\Phi(\mathbb{R}^d)$. Then, for every $T > 0$, there exists a unique $P_\nu^{(u_t)}$-strong solution to (MVSDE) up to time T, where u is the probability solution to (FPE) with $u|_{t=0} = \nu$ provided by Theorem 3.3.*

Our second main result says that, in the one-dimensional case, for every $x_0 \in \mathbb{R}$, there exists a strong solution to (MVSDE) with $X_0 := x_0$, which is pathwise unique among all weak solutions with time marginal law densities given by the Barenblatt solution to the classical porous medium equation.

Theorem 1.3 (main result 2). *Let $d = 1$, $m > 1$, $x_0 \in \mathbb{R}$. Furthermore, let $\beta(r) = |r|^{m-1}r, r \in \mathbb{R}$, and $E \equiv b \equiv 0$. Let u denote the Barenblatt solution to (FPE) with $u|_{t=0} = \delta_{x_0}$ as in (BB) below. Then, for every $T > 0$, there exists a unique $P_{\delta_{x_0}}^{(u_t)}$-strong solution to (MVSDE) up to time T.*

In contrast to the well-known nondegenerate case (see [JM98; BJ19; Gru23b; Gru24]), there are only few results on the existence of strong solutions to McKean–Vlasov SDEs with coefficients of Nemytskii-type in the degenerate case. In [Ben+96], the authors constructed a strong solution to (MVSDE) associated to the one-dimensional classical porous medium equation with initial datum $\nu \in \mathcal{P}_0(\mathbb{R}) \cap C_b(\mathbb{R})$ such that $\int_\mathbb{R} \nu(x)|x|^n dx < \infty$, where $n > 2$, and $-\partial_x^2(\nu^{m-1}) \leq C$ in the sense of Schwartz-distributions. In the case $m > 3$, the conditions on the initial data ν can be relaxed to $\nu \in \mathcal{P}_0(\mathbb{R}) \cap L^\infty(\mathbb{R})$ as was shown in [Gru23b] (see also [Gru23a, Section 4.5.2]). Now, with Theorem 1.2, we are able to treat the multidimensional case and, furthermore, even allow for first-order perturbations. To the best of the authors' knowledge, both of these aspects have not been considered in the literature so far. Also, the existence of probabilistically strong solutions to McKean–Vlasov SDEs of the type (MVSDE) seems to not have been considered in the literature so far in the case when X_0 is distributed with a singular part with respect to Lebesgue measure. Therefore, to the best of the authors' knowledge, the result formulated in Theorem 1.3 is new as well.

The proofs of both our main results relies on the restricted Yamada–Watanabe theorem, Theorem 2.1, from [Gru23b] (see also [Gru23a]). Since the existence of weak solutions has already been proved in [BR21a], the task of this paper is to prove suitable pathwise uniqueness results. The most important ingredient for the pathwise uniqueness result related to Theorem 1.2 is Theorem 3.7. There we prove that the in [BR21a] constructed Schwartz-distributional solution u to (FPE) with $u|_{t=0} = \nu \in L^\infty(\mathbb{R}^d) \cap L^1_\Phi(\mathbb{R}^d)$, where $\nu \geq 0$ a.e. and $\Phi \in \mathcal{Q}$, satisfies the following estimate: For all $T > 0$ there exists $C_T \geq 0$ such that for all $t \in (0, T)$

$$\int_{\mathbb{R}^d} \Psi(u_t(x))\mathrm{d}x + \int_0^t \int_{\mathbb{R}^d} |\nabla_x \beta^{\frac{1}{2}}(u(s,x))|^2 \mathrm{d}x\mathrm{d}s \leq \int_{\mathbb{R}^d} \Psi(\nu(x))\mathrm{d}x + C_T, \tag{3}$$

where $\Psi(r) := \int_0^r \ln(\beta(s))\mathrm{d}s, r \geq 0$. This result is strongly motivated by [FG23]. Fehrmann and Gess developed this kind of an estimate in order to prove, in particular, new existence and regularity theorems for analytically weak solutions to a large class of L^1 energy critical FPEs of the form (FPE) on the d-dimensional torus with drift coefficients $(t, x, r) \mapsto E(t, x)b(r)$, where $E \in L^2$, $b(r) = \frac{\beta^{\frac{1}{2}}(r)}{r}$, and β is taken from a large class of functions, which includes those of monomial type, namely $\beta(r) = |r|^{m-1}r$ for $m \geq 1$. For (3), we use a similar idea of proof with differences owed to the transfer of their result to the mild solution framework in $L^1(\mathbb{R}^d)$ (within the Crandall–Ligget theory), in which the Schwartz-distributional solution u from above was constructed in [BR21a].

The combination of the so obtained regularity of $\beta^{\frac{1}{2}}(u)$ and an estimate based on the mean-value theorem motived by [FG23, (76)-(77)], see e.g. (48), are key to our pathwise uniqueness result, Theorem 3.9. The latter is a proper modification of the restricted pathwise uniqueness result in the proof of [RZ10, Theorem 1.1]. The pathwise uniqueness result related to our second main result, Theorem 1.3, is formulated in Theorem 4.3. Here, we already note that the Barenblatt solution u does not satisfy (3), since $u^{\frac{m}{2}} \notin L^2((0,T); W^{1,2}(\mathbb{R}^d))$. However, $u^{\frac{m}{2}} \in L^2((0,T); W^{s,2}(\mathbb{R}^d))$, for all $s \in (0,1)$. This turns out to be enough to prove a suitable pathwise uniqueness result in dimension one. The proof is comparable to the one of Theorem 3.9, but additionally, uses the technique of the one-dimensional result on the existence of strong solutions to rough (ordinary) stochastic differential equations [CJ18, Theorem 2.15 (i)] based on a Lipschitz type estimate for functions with fractional Sobolev regularity of order $1/2$.

Within this work (as well as in [Gru23b; Gru24]), we consider (MVSDE) as a McKean–Vlasov SDE of the form

$$dX(t) = \boldsymbol{b}(t, X(t), \mathcal{L}_{X(t)})\mathrm{d}t + \boldsymbol{\sigma}(t, X(t), \mathcal{L}_{X(t)})\mathrm{d}W(t), \quad t \in [0,T],$$
$$X(0) = X_0,$$

with coefficients

$$[0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \ni (t, x, \nu) \mapsto \boldsymbol{b}(t, x, \nu) := E(x)b(v_a(x)),$$

$$[0,T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \ni (t, x, \nu) \mapsto \boldsymbol{\sigma}(t, x, \nu) := \sqrt{2\frac{\beta(v_a(x))}{v_a(x)}} \mathbb{1}_{d \times d},$$

where $\mathcal{P}(\mathbb{R}^d)$ is the set of all Borel probability measures on $\mathbb{R}^d$ equipped with the topology of narrow convergence of probability measures, v_a denotes the version of the density of the absolutely continuous part of ν both with respect to Lebesgue measure which is obtained by setting $v_a(x) = 0$ if $\lim_{R \to 0} \nu(B_R(x))/\lambda^d(B_R(0))$ does not exist in $\mathbb{R}$. By the Besicovitch derivation theorem, this makes the map $\mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \ni (x, \nu) \mapsto v_a(x)$ and therefore also $\boldsymbol{b}$ and $\boldsymbol{\sigma}$ Borel measurable (for details see [Gru23a, Section 4.2]). The dependence of $\boldsymbol{b}$ and $\boldsymbol{\sigma}$ on ν in terms of v_a evaluated at a fixed point x excludes the continuity of $\boldsymbol{b}$ and $\boldsymbol{\sigma}$ in their measure-component with respect to the topology of weak convergence of probability measures, Wasserstein distance or bounded variation norm. These types of continuity assumptions are made in the major part of the literature (see, e.g. [CD18]).

This paper is structured as follows. First, we introduce some necessary notation and recall the solution and uniqueness concepts for (MVSDE) from [Gru23b]. In Secton 2, we recall the restricted Yamada–Watanabe theorem from [Gru23b], which is employed in the proof of the main results Theorem 1.2 and Theorem 1.3. In Section 3, we prove the first main result, Theorem 1.2. In Section 4, we prove the second main result, Theorem 1.3. In Appendix A, we gather the definition of and some facts on the (local) Hardy–Littlewood maximal function. In Appendix B, we recapitulate the mild solution framework according to Crandall and Ligget in a concise way.

Notation

Within this paper we will use the following notation, which is essentially taken from [Gru23b]. For a topological space $(\mathbf{T}, \tau)$, $\mathcal{B}(\mathbf{T})$ shall denote the Borel σ-algebra on $(\mathbf{T}, \tau)$. Throughout this article, d, n, k denote natural numbers.

On $\mathbb{R}^n$, we will always consider the usual n-dimensional Lebesgue measure λ^n if not said any differently. If there is no risk for confusion, we will just say that some property for elements in $\mathbb{R}^n$ holds *almost everywhere* (or *a.e.*) if and only if it holds λ^n-almost everywhere. Furthermore, on $\mathbb{R}^n$, $|\cdot|_{\mathbb{R}^n}$ denotes the Euclidean norm. If there is no risk for confusion, we will just write $|\cdot| = |\cdot|_{\mathbb{R}^n}$. By $B_R(x)$ we will denote the usual open ball with center $x \in \mathbb{R}^n$ and radius $R > 0$. Furthermore, we set $s \wedge t := \min(t, s)$, $s \vee t := \max(t, s)$ for all $s, t \in \mathbb{R}$.

Let $(S, \mathscr{S}, \eta)$ be a measure space. For $1 \leq p \leq \infty$, $(L^p(S; E), \|\cdot\|_{L^p(S;E)})$ symbolises the usual Bochner

space of strongly measurable E-valued functions f on S for which $\|f\|_E^p$ is integrable. If $S = \mathbb{R}^n$ and $E = \mathbb{R}$, we just write $L^p(\mathbb{R}^n; \mathbb{R}) = L^p(\mathbb{R}^n)$. The set of strongly measurable functions on $\mathbb{R}^n$ with values in E which are locally p-integrable in norm on E will be denoted by $L_{loc}^p(\mathbb{R}^n; E)$. Moreover, $(W^{k,p}(\mathbb{R}^n), \|\cdot\|_{W^{1,p}(S;E)})$ denotes the usual Sobolev space, containing all $L^p(\mathbb{R}^n)$-functions, whose Schwartz-distributional derivatives up to k-th order can be represented by elements in $L^p(\mathbb{R}^n)$. Accordingly, E-valued first-order Sobolev functions on $\mathbb{R}^n$ will be denoted by $W^{k,p}(\mathbb{R}^n; E)$. In the special case when $p = 2$, we also write $W^{k,2} = H^k$. For $\eta \in (0, \infty)$ of the form $\eta = s + k$, where $s \in (0,1)$ and $k \in \mathbb{N}_0$, $\dot{W}^{\eta,p}(\mathbb{R}^n)$ denotes the homogeneous Sobolev–Slobodeckij space, which consists of all functions $f \in L_{loc}^1(\mathbb{R}^n)$ such that for all $\alpha \in \mathbb{N}^k$ with $|\alpha| = k$

$$[D^\alpha f]_{\dot{W}^{s,p}(\mathbb{R}^n)} := \left(\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{|D^\alpha f(x) - D^\alpha f(y)|^p}{|x-y|^{sp+d}} \mathrm{d}x \mathrm{d}y \right)^{\frac{1}{p}} < \infty,$$

with norm $\|f\|_{\dot{W}^{\eta,p}} := \sup_{|\alpha|=k}[D^\alpha f]_{\dot{W}^{s,p}(\mathbb{R}^n)}$. Here, D^α denotes the usual (Schwartz-distributional) derivative of order k, where for each $i \in \{1, ..., d\}$, the derivative in direction x_i is taken α_i-times. The nonhomogeneous Sobolev–Slobodeckij space is defined as $W^{\eta,p}(\mathbb{R}^n) := L^p(\mathbb{R}^d) \cap \dot{W}^{\eta,p}(\mathbb{R}^n)$ with norm $\|f\|_{W^{\eta,p}(\mathbb{R}^d)} := \|f\|_{L^p(\mathbb{R}^d)} + \|f\|_{\dot{W}^{\eta,p}(\mathbb{R}^n)}$. Both the homogenous and nonhomogenous Sobolev–Slobodeckij space are Banach spaces. Recall that $W^{s,2}(\mathbb{R}^n) = \{f \in L^2(\mathbb{R}^n) : \mathcal{F}^{-1}(|\xi|^s \mathcal{F}(f)) \in L^2(\mathbb{R}^n)\}$. Here, $\mathcal{F}$ denotes the Fourier transform on the tempered Schwartz-distributions. For $f \in W^{s,2}(\mathbb{R}^n)$ we set $\nabla^s f := \mathcal{F}^{-1}(|\xi|^s \mathcal{F}(f))$, which is also known as the Riesz derivative. Furthermore, for every real-valued function, we set $f^- := -\min(f, 0)$ and $f^+ := \max(f, 0)$.

Let (M, d) be a metric space. Then $\mathcal{P}(M)$ denotes the set of all Borel probability measures on (M, d). We will consider $\mathcal{P}(M)$ as a topological space with respect to the topology of narrow convergence of probability measures. By $\mathcal{P}_0(\mathbb{R}^n)$ we will denote the set of all probability densities with respect to Lebesgue measure, i.e.

$$\mathcal{P}_0(\mathbb{R}^n) = \left\{ \rho \in L^1(\mathbb{R}^n) \mid \rho \geq 0 \text{ a.e.}, \int_{\mathbb{R}^n} \rho(x) \mathrm{d}x = 1 \right\}.$$

By $C_c^\infty(\mathbb{R}^n)$ we denote the set of all infinitely differentiable functions with compact support. Let $(E, \|\cdot\|_E)$ be a Banach space. The set of continuous functions on the interval $[0, T]$ with values in E is denoted by $C([0, T]; E)$ and is considered with respect to the usual supremum's norm. Furthermore, we define

$$C([0, T]; E)_0 := \{w \in C([0, T]; E) : w(0) = 0\}.$$

For $t \in [0, T]$, $\pi_t : C([0, T]; E) \to E$ denotes the canonical evaluation map at time t, i.e. $\pi_t(w) := w(t), w \in C([0, T]; E)$. Further, we set $\mathcal{B}_t(C([0, T]; E)) := \sigma(\pi_s : s \in [0, t])$ and, correspondingly, $\mathcal{B}_t(C([0, T]; E)_0) := \sigma(\pi_s : s \in [0, t]) \cap C([0, T]; E)_0$. Moreover, $\mathbb{P}_W$ denotes the Wiener measure on $(C([0, T]; \mathbb{R}^d)_0, \mathcal{B}(C([0, T]; \mathbb{R}^d)_0))$.

Throughout this article, $C \in (0, \infty)$ will denote a generic constant, which may change from line to line and which may depend on the spacial dimension d and the final (finite) time T. The derivative ∂_z denotes the derivative with respect to a scalar z-coordinate. Furthermore, $\mathrm{div} = \mathrm{div}_x$, $\nabla = \nabla_x$, $\Delta = \Delta_x$ and $D = D_x$ denote the divergence, gradient, Laplacian and Jacobian with respect to the spacial x-coordinate. All the derivatives are supposed to be understood in the sense of Schwartz-distributions.

$P_\nu^{(u_t)}$-solutions to (MVSDE)

Let us briefly recall the solution concepts for (MVSDE) from [Gru23b] (see also [Gru23a]).

Let $\nu \in \mathcal{P}(\mathbb{R}^d)$ and suppose u is a probability solution to (FPE) with $u|_{t=0} = \nu$. We set

$$P_\nu^{(u_t)} := \{Q \in \mathcal{P}(C([0, T]; \mathbb{R}^d)) : Q \circ \pi_0^{-1} = \nu, Q \circ \pi_t^{-1} = u_t(x)\mathrm{d}x \,\forall t \in (0, T]\}.$$

A $P_\nu^{(u_t)}$-**weak** **solution** $(X, W, (\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t \in [0,T]}))$ is a (probabilistically) weak solution $(X, W, (\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t \in [0,T]}))$ to (MVSDE) in the usual sense such that $\mathcal{L}_{X(0)} = \nu$ and $\mathcal{L}_{X(t)} = u_t(x)\mathrm{d}x$, for all $t \in (0, T]$. For the convenience of the reader, we will just write $(X, W) = (X, W, (\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t \in [0,T]}))$ in cases in which we do not need to refer explicitly to the underlying stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t \in [0,T]})$.

We will say that (MVSDE) has a $P_\nu^{(u_t)}$-**strong solution** if there exists a function $F : \mathbb{R}^d \times C([0,T];\mathbb{R}^d)_0 \to C([0,T];\mathbb{R}^d)$, which is $\overline{\mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(C([0,T];\mathbb{R}^d)_0)}^{\nu \otimes \mathbb{P}_W} / \mathcal{B}(C([0,T];\mathbb{R}^d))$-measurable, such that, for ν-a.e. $x \in \mathbb{R}^d$, $F(x,\cdot)$ is $\overline{\mathcal{B}_t(C([0,T];\mathbb{R}^d)_0)}^{\mathbb{P}_W} / \mathcal{B}_t(C([0,T];\mathbb{R}^d))$-measurable for all $t \in [0,T]$ and, whenever X_0 is an $\mathcal{F}_0$-measurable function with $\mathcal{L}_{X_0} = \nu$ and W is a standard d-dimensional $(\mathcal{F}_t)$-Brownian motion on some stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t \in [0,T]})$, $(F(X_0, W), W, (\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t \in [0,T]}))$ is a $P_\nu^{(u_t)}$-weak solution to (MVSDE). Here, $\overline{\mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(C([0,T];\mathbb{R}^d)_0)}^{\nu \otimes \mathbb{P}_W}$ denotes the completion of $\mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(C([0,T];\mathbb{R}^d)_0)$ with respect to the measure $\nu \otimes \mathbb{P}_W$, and $\overline{\mathcal{B}_t(C([0,T];\mathbb{R}^d)_0)}^{\mathbb{P}_W}$ denotes the completion of $\mathcal{B}_t(C([0,T];\mathbb{R}^d)_0)$ with respect to $\mathbb{P}_W$ on $(C([0,T];\mathbb{R}^d)_0, \mathcal{B}(C([0,T];\mathbb{R}^d)_0))$.

Moreover, $P_\nu^{(u_t)}$-**pathwise uniqueness** holds for (MVSDE) if for every two $P_\nu^{(u_t)}$-weak solutions $(X, W, (\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t \in [0,T]}))$, $(Y, W, (\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t \in [0,T]}))$ (with respect to the same Brownian motion on the same stochastic basis) with $X(0) = Y(0)$ $\mathbb{P}$-a.s., one has $\sup_{t \in [0,T]} |X(t) - Y(t)| = 0$ $\mathbb{P}$-a.s.

We say that there exists a **unique $P_\nu^{(u_t)}$-strong solution** to (MVSDE) if there exists a $P_\nu^{(u_t)}$-strong solution to (MVSDE) with functional F as above, and every $P_\nu^{(u_t)}$-weak solution (X, W) is of the form $X = F(X(0), W)$ almost surely with respect to the underlying probability measure.

2 The restricted Yamada–Watanabe theorem

We recall the restricted Yamada–Watanabe theorem from [Gru23b] (see also [Gru23a]), which is a modification of the original Yamada–Watanabe theorem. In particular, it has the advantage that one can conclude the existence of a strong solution to (MVSDE) under a relaxed pathwise uniqueness condition compared to the original Yamada–Watanabe theorem.

The following theorem can essentially be found in [Gru23b, Theorem 3.3].

Theorem 2.1. *Let $\nu \in \mathcal{P}(\mathbb{R}^d)$ and suppose u is a probability solution to* (FPE) *such that $u|_{t=0} = \nu$. The following statements regarding* (MVSDE) *are equivalent.*

(i) There exists a $P_\nu^{(u_t)}$-weak solution and $P_\nu^{(u_t)}$-pathwise uniqueness holds.

(ii) There exists a unique $P_\nu^{(u_t)}$-strong solution to (MVSDE).

3 Proof of Theorem 1.2

In this section, we prove the first main result, Theorem 1.2. This section is split into three subsections. In Section 3.1, we will recall the results on the existence of solutions to (FPE) from [BR21a] and, moreover, prove a regularity result for $\beta^{\frac{1}{2}}(u)$, where u is the solution to (FPE), which was constructed in [BR21a]. Here, the presentation of the techniques and results of [BR21a] is similar to the one in [Gru23a, Appendix D.2]. In Sections 3.2 and 3.3, we then show that the conditions (i) and (ii) of the restricted Yamada–Watanabe, see Theorem 2.1, are fulfilled, respectively.

Proof of Theorem 1.2. The assertion follows directly from Theorem 3.8 and Theorem 3.9 via Theorem 2.1. $\square$

3.1 Existence of a regular Schwartz-distributional solution to (FPE)

In [BR21a], Barbu and Röckner proved the existence of a Schwartz-distributional solution to (FPE) with initial datum $\nu \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ under even more general conditions on E, b, β than (H1), (H2), namely

(C1) $\beta \in C^2(\mathbb{R})$, β is monotonically nondecreasing, $\beta(0) = 0$,

(C2) $E \in L^\infty(\mathbb{R}^d; \mathbb{R}^d)$, $\operatorname{div} E \in L^2_{loc}(\mathbb{R}^d)$, $(\operatorname{div} E)^- \in L^\infty(\mathbb{R}^d)$,
$b \in C_b(\mathbb{R}) \cap C^1(\mathbb{R})$, $b \geq 0; b = \operatorname{const.}$ if β is not strictly increasing.

Therefore, they considered (FPE) (up to time $T = \infty$) as an evolution equation in $L^1(\mathbb{R}^d)$ of the form

$$\partial_t u(t) + A(u(t)) = 0, \quad t \in [0, \infty), \tag{FPE.Op}$$
$$u|_{t=0} = \nu,$$

where $(A, D(A))$ is a nonlinear operator acting on $L^1(\mathbb{R}^d)$, which coincides with the operator $(A_0, D(A_0))$ defined as

$$A_0(u) := \operatorname{div}(Eb(u)u) - \Delta\beta(u), u \in D(A_0),$$
$$D(A_0) := \{u \in L^1(\mathbb{R}^d) : \beta(u) \in L^1_{loc}(\mathbb{R}^d), -\Delta\beta(u) + \operatorname{div}(Eb(u)u) \in L^1(\mathbb{R}^d)\},$$

on a domain $D(A) \subset D(A_0)$ to be specified later on. The following lemma is a special case of [BR21a, Lemma 3.1].

Lemma 3.1 (cf. [BR21a, Lemma 3.1]). *Assume that (C1) and (C2) hold. We have*

$$\operatorname{Range}(I + \lambda A_0) = L^1(\mathbb{R}^d) \quad \forall\lambda > 0,$$

and there exists an operator $J_\lambda : L^1(\mathbb{R}^d) \to L^1(\mathbb{R}^d)$ such that

$$J_\lambda(v) \in (I + \lambda A_0)^{-1}(v) \quad \forall v \in L^1(\mathbb{R}^d),$$
$$\|J_\lambda(v) - J_\lambda(\bar{v})\|_{L^1} \leq \|v - \bar{v}\|_{L^1} \quad \forall v, \bar{v} \in L^1(\mathbb{R}^d), \lambda > 0, \tag{4}$$
$$J_{\lambda_2}(v) = J_{\lambda_1}\left(\frac{\lambda_1}{\lambda_2}v + \left(1 - \frac{\lambda_1}{\lambda_2}\right)J_{\lambda_2}(v)\right) \quad \forall 0 < \lambda_1, \lambda_2 < \infty. \tag{5}$$

Moreover,

$$\|J_\lambda(v)\|_{L^\infty} \leq \left(1 + \left\|(\operatorname{div} E)^- + |E|\right\|_{L^\infty}^{\frac{1}{2}}\right) \|v\|_{L^\infty} \quad \forall v \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d),$$
$$0 < \lambda < \lambda_0 := \left(\left\|(\operatorname{div} E)^- + |E|\right\|_{L^\infty} + \left(\left\|(\operatorname{div} E)^- + |E|\right\|_{L^\infty}^{\frac{1}{2}}\right)\|b\|_{L^\infty}\right)^{-1} \tag{6}$$

and

$$J_\lambda(\mathcal{P}_0(\mathbb{R}^d)) \subset \mathcal{P}_0(\mathbb{R}^d) \quad \forall\lambda > 0. \tag{7}$$

With the results of Lemma 3.1, Barbu and Röckner defined $D(A)$ as follows:

$$A(v) := A_0(v), v \in D(A), \tag{8}$$
$$D(A) := \{J_\lambda(v) : v \in L^1(\mathbb{R}^d)\}, \lambda > 0.$$

Note that $D(A)$ is independent of the parameter $\lambda > 0$, as can be seen from (13), and is therefore well-defined. From Lemma 3.1 the following corollary is immediately deduced. For the terminology, we refer to Appendix B.

Corollary 3.2 ([BR21a, Lemma 3.2]). *Assume that (C1) and (C2) hold. Then,*

   (i) J_λ coincides with the inverse $(I + \lambda A)^{-1}$ of $(I + \lambda A)$.

   (ii) The operator A is m-accretive in $L^1(\mathbb{R}^d)$.

Moreover, (6) and (7) hold with $(I + \lambda A)^{-1}$ instead of J_λ.

By the Crandall–Ligget theorem (cf. Theorem B.4), the following theorem is then proved. It is a special case of [BR21a, Theorem 2.2], which is one of the main results of [BR21a].

8

Theorem 3.3 (c.f. [BR21a, Theorem 2.2]). *Let $d \neq 2$. Assume that (C1) and (C2) hold. Then $\overline{D(A)}^{L^1(\mathbb{R}^d)} = L^1(\mathbb{R}^d)$ and for every $\nu \in L^1(\mathbb{R}^d)$ there exists a unique mild solution u to (FPE.Op) with $u|_{t=0} = \nu$. Furthermore, for every $\nu \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$*

$$\|u(t)\|_{L^\infty(\mathbb{R}^d)} \leq \exp\left(\left\|(\operatorname{div} E)^- + |E|\right\|_{L^\infty(\mathbb{R}^d)}^{\frac{1}{2}} t\right) \|v\|_{L^\infty(\mathbb{R}^d)} \quad \forall t \in [0, \infty). \tag{9}$$

If $\nu \in \mathcal{P}_0(\mathbb{R}^d)$, then

$$u(t) \in \mathcal{P}_0(\mathbb{R}^d) \ \forall t \in [0, \infty). \tag{10}$$

Moreover, $t \mapsto S(t)\nu = u(t, \cdot)$, where u denotes the mild solution starting in ν, is a strongly continuous semigroup of nonlinear contractions from $L^1(\mathbb{R}^d)$ to $L^1(\mathbb{R}^d)$, that is $S(t+s)\nu = S(t)S(s)\nu$ for all $0 < s < t$, and

$$\|S(t)\nu - S(t)\bar{\nu}\|_{L^1} \leq \|\nu - \bar{\nu}\|_{L^1} \ \forall t > 0, \ \nu, \bar{\nu} \in L^1(\mathbb{R}^d).$$

If $\nu \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$, then u is a Schwartz-distributional solution to (FPE) up to $T = \infty$ such that $u|_{t=0} = \nu$.

In view of the proof of Theorem 3.7 below, we need to provide more details on how Lemma 3.1 is proved. The proof of Lemma 3.1 is extensive and is based on the analysis of the operators $(I + \lambda A_\varepsilon)^{-1}, \lambda > 0, \varepsilon > 0$, where $(A_\varepsilon, D(A_\varepsilon))$ is defined as

$$A_\varepsilon v := -\Delta \tilde{\beta}_\varepsilon(v) + \varepsilon \tilde{\beta}_\varepsilon(v) + \operatorname{div}(E_\varepsilon b_\varepsilon(v)v), v \in D(A_\varepsilon),$$
$$D(A_\varepsilon) := \{v \in L^1(\mathbb{R}^d) : -\Delta \tilde{\beta}_\varepsilon(v) + \varepsilon \tilde{\beta}_\varepsilon(v) + \operatorname{div}(E_\varepsilon b_\varepsilon(v)v) \in L^1(\mathbb{R}^d)\},$$

where $\tilde{\beta}_\varepsilon(r) := \beta_\varepsilon(r) + \varepsilon r, r \in \mathbb{R}$, and for $\varepsilon > 0, r \in \mathbb{R}$,

$$\beta_\varepsilon(r) := \frac{1}{\varepsilon}(r - g_\varepsilon(r)) = \beta(g_\varepsilon(r)), \text{ where } g_\varepsilon(r) := (I + \varepsilon\beta)^{-1}(r),$$

$$E_\varepsilon := \begin{cases} E, & \text{if } |E| \in L^2(\mathbb{R}^d) \text{ and } \operatorname{div}(E) \in L^2(\mathbb{R}^d) + L^\infty(\mathbb{R}^d), \\ \eta_\varepsilon E, & \text{else.} \end{cases}$$

$$b_\varepsilon(r) := \begin{cases} b, & \text{if } b \text{ is constant,} \\ \frac{(b*\rho_\varepsilon)(r)}{1+\varepsilon|r|}, & \text{else.} \end{cases}$$

Here $\rho_\varepsilon(r) := \varepsilon^{-1}\rho(\varepsilon^{-1}r), \rho \in C_c^\infty(\mathbb{R}), \rho \geq 0$, is a standard mollifier and by I we denote the identity map on $\mathbb{R}$. Furthermore, $\eta_\varepsilon \in C_c^1(\mathbb{R}^d), 0 \leq \eta_\varepsilon \leq 1, |\nabla \eta_\varepsilon| \leq 1$ and $\eta_\varepsilon(x) = 1$ if $|x| \leq \varepsilon^{-1}$. Then we have $E_\varepsilon \in L^2(\mathbb{R}^d; \mathbb{R}^d) \cap L^\infty(\mathbb{R}^d; \mathbb{R}^d), |E_\varepsilon| \leq |E|, \lim_{\varepsilon \to 0} E_\varepsilon(x) = E(x)$ and

$$(\operatorname{div} E_\varepsilon)^- \leq (\operatorname{div} E)^- + |E| \mathbb{1}_{\{|\cdot| \geq \varepsilon^{-1}\}}. \tag{11}$$

Consider the following equation in the Schwartz-distributional sense

$$u_\varepsilon + \lambda A_\varepsilon(u_\varepsilon) = f, \tag{12}$$

for $f \in L^1(\mathbb{R}^d)$ and $\lambda > 0$. We have the following lemma on the solvability of (12) in the case $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ together with fundamental estimates for a solution u_ε, which is extracted from [BR21a, Proof of Lemma 3.1].

Lemma 3.4. *Let $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ and $\lambda > 0$. Then (12) has a solution $u_\varepsilon = u_\varepsilon(\lambda, f) \in L^1(\mathbb{R}^d) \cap H^1(\mathbb{R}^d)$ in the strong sense in $L^2(\mathbb{R}^d)$. If $f \geq 0$ a.e. then $u_\varepsilon \geq 0$ a.e. For $\bar{f} \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ we have*

$$\left\|u_\varepsilon(\lambda, f) - u_\varepsilon(\lambda, \bar{f})\right\|_{L^1(\mathbb{R}^d)} \leq \left\|f - \bar{f}\right\|_{L^1(\mathbb{R}^d)}, \tag{13}$$

and $u_\varepsilon(\lambda, 0) = 0$. Furthermore,

$$u_\varepsilon(\lambda) \to J_\lambda f \text{ in } L^1(\mathbb{R}^d), \text{ as } \varepsilon \to 0. \tag{14}$$

Let $f \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$. Then for all $\lambda \in (0, \lambda_0)$, where $\lambda_0 > 0$ is defined as in (6), $u_\varepsilon(\lambda, f) \in L^\infty(\mathbb{R}^d)$ such that

$$\|u_\varepsilon(\lambda, f)\|_{L^\infty(\mathbb{R}^d)} \le C_E \|f\|_{L^\infty(\mathbb{R}^d)},$$

where

$$C_E := \left(1 + \left\|(\operatorname{div} E)^- + |E|\right\|_{L^\infty(\mathbb{R}^d)}^{\frac{1}{2}}\right).$$

The result and proof of the following lemma can be found in [BR21a, Proof of Lemma 3.3], see also [BR23b, Lemma 3.2].

Lemma 3.5. Let $f \in L^1_\Phi(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ for some $\Phi \in \mathcal{Q}$. Let $\varepsilon \in (0, 1), \lambda \in (0, \lambda_0)$, where $\lambda_0 > 0$ is as in (6). Let $u_\varepsilon := u_\varepsilon(\lambda, f)$ denote the solution to (12) provided by Lemma 3.4. Then, we have

$$\|u_\varepsilon\|_{L^1_\Phi} \le \|f\|_{L^1_\Phi} + \lambda \int |\tilde\beta_\varepsilon(u_\varepsilon)| \Delta \Phi + |b_\varepsilon(u_\varepsilon) u_\varepsilon| |\langle E_\varepsilon, \nabla \Phi \rangle_{\mathbb{R}^d}| \mathrm{d}x. \tag{15}$$

Remark 3.6. It is easy to see that Lemma 3.5 and (14) imply the inequality

$$\|J_\lambda(f)\|_{L^1_\Phi} \le \|f\|_{L^1_\Phi} + \lambda \int |\beta(J_\lambda(f))| \Delta \Phi + |b(J_\lambda(f)) J_\lambda(f)| |\langle E, \nabla \Phi \rangle_{\mathbb{R}^d}| \mathrm{d}x.$$

The following theorem provides conditions under which it is guaranteed that $\beta^{\frac{1}{2}}(u) \in L^2([0, T]; H^1(\mathbb{R}^d))$. The technique and the estimate (18) we obtain, are similar to those in [FG23]. As pointed out in Section 1, Fehrmann and Gess established an a priori estimate similar to (18) on the d-dimensional torus $\mathbb{T}^d$ in order to prove the existence and regularity of analytically weak solutions to (FPE) including those with drift coefficient $[0, T] \times \mathbb{R}^d \times \mathbb{R} \ni (t, x, r) \mapsto E(t, x) \frac{\beta^{\frac{1}{2}}(r)}{r}$, $E^i \in L^2([0, T] \times \mathbb{T}^d), i \in \{1, ..., d\}$, and a large class of diffusivity functions β, including those of monomial type, i.e. $\beta(r) = |r|^{m-1} r$ for every $m \ge 1$. Here, we transfer their technique to (FPE.Op) within the mild solution framework in $L^1(\mathbb{R}^d)$.

Theorem 3.7. Assume that (C1) and (C2) hold. Assume that

$$E \in L^2(\mathbb{R}^d; \mathbb{R}^d), \text{ div } E \in L^2(\mathbb{R}^d) + L^\infty(\mathbb{R}^d), \text{ div } E \ge 0; \tag{16}$$

or that

$$\forall K > 0 \ \exists C_K > 0 \ \forall r \in [0, K] : \beta'(r) r \le C_K \beta(r). \tag{17}$$

Let $\nu \in L^1_\Phi(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ for some $\Phi \in \mathcal{Q}$ with $\nu \ge 0$ a.e. and let u denote the Schwartz-distributional solution to (FPE) with $u|_{t=0} = \nu$ provided by Theorem 3.3. Then, for all $T > 0$ there exists $C_T > 0$ such that for all $t \in (0, T)$

$$\int_{\mathbb{R}^d} \Psi(u_t(x)) \mathrm{d}x + \int_0^t \int_{\mathbb{R}^d} |\nabla \beta^{\frac{1}{2}}(u_s(x))|^2 \mathrm{d}x \mathrm{d}s \le \int_{\mathbb{R}^d} \Psi(\nu(x)) \mathrm{d}x + C_T, \tag{18}$$

where $\Psi(r) := \int_0^r \ln(\beta(s)) ds, r \ge 0$. If (16) is assumed, then (18) is true for $C_T = 0$.

Proof. By Definition B.3, we have for each $T > 0$

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

where $u_h : [0, T] \to L^1(\mathbb{R}^d)$ is a step function, defined by the finite difference scheme

$$u_h(t) = u_h^{i+1} \ \forall t \in (ih, (i+1)h], i = 0, 1, ..., N-1, u_h(0) = \nu, \tag{19}$$
$$\text{where } u_h^{i+1} + hA(u_h^{i+1}) = u_h^i, \ Nh = T,$$
$$\text{and } u_h^0 = \nu.$$

We prove the assertion of this lemma according to the following three steps.

1. We show that for every $f \in L_\Phi^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ with $f \geq 0$ a.e. and $\lambda \in (0, \lambda_0)$, the corresponding solution $u_\varepsilon = u_\varepsilon(\lambda, f)$ to

$$u_\varepsilon + \lambda A_\varepsilon(u_\varepsilon) = f, \tag{20}$$

provided by Lemma 3.4, satisfies for $\Psi_\varepsilon(r) := \int_0^r \ln(\tilde{\beta}_\varepsilon(r)) dr, r \geq 0$,

$$\int_{\mathbb{R}^d} \Psi_\varepsilon(u_\varepsilon) \mathrm{d}x + \lambda \int_{\mathbb{R}^d} |\nabla \tilde{\beta}_\varepsilon^{\frac{1}{2}}(u_\varepsilon)|^2 \mathrm{d}x \leq \int_{\mathbb{R}^d} \Psi_\varepsilon(f) \mathrm{d}x + \lambda C_\varepsilon, \tag{21}$$

for a constant $C_\varepsilon \in (0, \infty)$, which is independent of $\lambda \in (0, \lambda_0)$.

2. We show that for every $f \in L_\Phi^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ with $f \geq 0$ a.e. and $\lambda \in (0, \lambda_0)$,

$$\int_{\mathbb{R}^d} \Psi(u) \mathrm{d}x + \lambda \int_{\mathbb{R}^d} |\nabla \beta^{\frac{1}{2}}(u)|^2 \mathrm{d}x \leq \int_{\mathbb{R}^d} \Psi(f) \mathrm{d}x + \lambda C, \tag{22}$$

where $u := J_\lambda(f)$ (cf. Lemma 3.1) for a constant $C > 0$, which is independent of ε and $\lambda \in (0, \lambda_0)$.

3. We conclude (18) via the finite difference scheme (19).

Regarding 1. We use a similar ansatz as in the proof of [FG23, Prop 5.6 (arXiv version 3)] and adapt their techniques to the mild solution framework in $L^1(\mathbb{R}^d)$ in our setting. Let $\varepsilon \in (0, 1)$. By Lemma 3.4 we have that $u_\varepsilon \geq 0$ a.e., $u_\varepsilon \in H^1(\mathbb{R}^d) \cap L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$, $\tilde{\beta}_\varepsilon(u_\varepsilon) \in H^2(\mathbb{R}^d)$, and $||u_\varepsilon||_{L^\infty} \leq C_E ||f||_{L^\infty}$, where $C_E = \left(1 + ||(\operatorname{div} E)^- + |E|||_{L^\infty(\mathbb{R}^d)}^{\frac{1}{2}}\right)$. Here, note that

$$\operatorname{div}(E_\varepsilon b_\varepsilon(u_\varepsilon)u_\varepsilon) = \operatorname{div}(E_\varepsilon)b_\varepsilon(u_\varepsilon)u_\varepsilon + (b_\varepsilon'(u_\varepsilon)u_\varepsilon + b_\varepsilon(u_\varepsilon))\langle E_\varepsilon, \nabla u_\varepsilon \rangle_{\mathbb{R}^d} \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d).$$

Let $\kappa_n \in C_c^\infty(\mathbb{R}^d)$ such that $0 \leq \kappa_n \leq 1$, $\kappa_n = 1$ on $B_n(0)$ and $||\nabla \kappa_n||_{L^\infty} + ||\Delta \kappa_n||_{L^\infty} \leq C$ for some constant $C > 0$ independent of n. We define $l_M(r) := ((\ln(r) \vee -M) \wedge M)\mathbb{1}_{(0,\infty)}(r) - M\mathbb{1}_{(-\infty,0]}(r)$ and $l_{M,n}(x, r) := \kappa_n(x)l_M(r), (r, x) \in \mathbb{R} \times \mathbb{R}^d$. Note that $l_M \in \operatorname{Lip}(\mathbb{R})$ with derivative $l_M'(r) = \mathbb{1}_{(\exp -M, \exp M)}(r)\frac{1}{r}$ for all $r \notin \{\exp -M, \exp M\}$. Let us multiply (20) by $l_{M,n}(\cdot, \tilde{\beta}_\varepsilon(u_\varepsilon))(\in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d) \cap H^1(\mathbb{R}^d))$ and integrate over $\mathbb{R}^d$. We obtain

$$\underbrace{\int u_\varepsilon l_{M,n}(\cdot, \tilde{\beta}_\varepsilon(u_\varepsilon)) \mathrm{d}x}_{=:I} + \lambda \underbrace{\int -\Delta \tilde{\beta}_\varepsilon(u_\varepsilon) l_{M,n}(\cdot, \tilde{\beta}_\varepsilon(u_\varepsilon)) \mathrm{d}x}_{=:II} + \lambda \varepsilon \underbrace{\int \tilde{\beta}_\varepsilon(u_\varepsilon) l_{M,n}(\cdot, \tilde{\beta}_\varepsilon(u_\varepsilon)) \mathrm{d}x}_{=:III}$$

$$+ \lambda \underbrace{\int \operatorname{div}(E_\varepsilon b_\varepsilon(u_\varepsilon)u_\varepsilon) l_{M,n}(\cdot, \tilde{\beta}_\varepsilon(u_\varepsilon)) \mathrm{d}x}_{=:IV} = \underbrace{\int f l_{M,n}(\cdot, \tilde{\beta}_\varepsilon(u_\varepsilon)) \mathrm{d}x}_{=:V}.$$

I: By Lebesgue's dominated convergence theorem, it is standard to prove that

$$\lim_{n \to \infty} \int u_\varepsilon l_{M,n}(\cdot, \tilde{\beta}_\varepsilon(u_\varepsilon)) \mathrm{d}x = \int u_\varepsilon l_M(\tilde{\beta}_\varepsilon(u_\varepsilon)) \mathrm{d}x.$$

II: Using integration by parts, we obtain

$$-\int \Delta\tilde{\beta}_\varepsilon(u_\varepsilon) l_{M,n}(\cdot, \tilde{\beta}_\varepsilon(u_\varepsilon)) \mathrm{d}x = \int \langle \nabla\tilde{\beta}_\varepsilon(u_\varepsilon), \nabla l_{M,n}(\cdot, \tilde{\beta}_\varepsilon(u_\varepsilon))\rangle_{\mathbb{R}^d} \mathrm{d}x$$

$$= \int |\nabla\tilde{\beta}_\varepsilon(u_\varepsilon)|^2 l'_M(\tilde{\beta}_\varepsilon(u_\varepsilon))\kappa_n \mathrm{d}x + \int \langle \nabla\tilde{\beta}_\varepsilon(u_\varepsilon), \nabla\kappa_n\rangle_{\mathbb{R}^d} l_M(\tilde{\beta}_\varepsilon(u_\varepsilon)) \mathrm{d}x$$

Due to our choice of the $\kappa_n, n \in \mathbb{N}$, and the fact that $|\nabla\kappa_n|, \Delta\kappa_n \to 0$ on $\mathbb{R}^d$, as $n \to \infty$, we may use integration by parts and Lebesgue's dominated convergence theorem to conclude that

$$\lim_{n\to\infty} \int \langle \nabla\tilde{\beta}_\varepsilon(u_\varepsilon), \nabla\kappa_n\rangle_{\mathbb{R}^d} l_M(\tilde{\beta}_\varepsilon(u_\varepsilon)) \mathrm{d}x = \lim_{n\to\infty} -\int \tilde{\beta}_\varepsilon(u_\varepsilon) \left(l_M(\tilde{\beta}_\varepsilon(u_\varepsilon))\Delta\kappa_n + \langle \nabla\kappa_n, \nabla\tilde{\beta}_\varepsilon(u_\varepsilon)\rangle_{\mathbb{R}^d} l'_M(\tilde{\beta}_\varepsilon(u_\varepsilon)) \right) \mathrm{d}x$$

$$= 0.$$

Also, via Lebesgue's dominated convergence theorem, we have

$$\lim_{n\to\infty} \int |\nabla\tilde{\beta}_\varepsilon(u_\varepsilon)|^2 l'_M(\tilde{\beta}_\varepsilon(u_\varepsilon))\kappa_n \mathrm{d}x = \int |\nabla\tilde{\beta}_\varepsilon(u_\varepsilon)|^2 l'_M(\tilde{\beta}_\varepsilon(u_\varepsilon)) \mathrm{d}x.$$

Using the monotone convergence theorem, we may, thus, conclude that

$$\lim_{M\to\infty} \lim_{n\to\infty} -\int \Delta\tilde{\beta}_\varepsilon(u_\varepsilon) l_{M,n}(\cdot, \tilde{\beta}_\varepsilon(u_\varepsilon)) \mathrm{d}x = \int_{\{\tilde{\beta}_\varepsilon(u_\varepsilon)>0\}} \frac{|\nabla\tilde{\beta}_\varepsilon(u_\varepsilon)|^2}{\tilde{\beta}_\varepsilon(u_\varepsilon)} \mathrm{d}x.$$

III: Similar to [BR23b, (4.5) and below] (and the references therein) we use that

$$\forall \gamma \in (0,1) \ \exists C_\gamma > 0 \ \forall r \in [0,\infty) : r\ln^-(r) \le C_\gamma r^\gamma, \tag{23}$$

and estimate for all $\gamma \in \left[\frac{\alpha}{\alpha+1}, 1\right)$

$$\lim_{n\to\infty} \int \tilde{\beta}_\varepsilon(u_\varepsilon) l_{M,n}(\cdot, \tilde{\beta}_\varepsilon(u_\varepsilon)) \mathrm{d}x = \int \tilde{\beta}_\varepsilon(u_\varepsilon) l_M(\tilde{\beta}_\varepsilon(u_\varepsilon)) \mathrm{d}x$$

$$\ge -\int \tilde{\beta}_\varepsilon(u_\varepsilon) l_M^-(\tilde{\beta}_\varepsilon(u_\varepsilon)) \mathrm{d}x \ge -\int \tilde{\beta}_\varepsilon(u_\varepsilon) \ln^-(\tilde{\beta}_\varepsilon(u_\varepsilon)) \mathrm{d}x \ge -C_\gamma \int \tilde{\beta}_\varepsilon(u_\varepsilon)^\gamma \mathrm{d}x$$

$$\ge -C_\gamma \left(\int |\tilde{\beta}_\varepsilon(u_\varepsilon)|\Phi \mathrm{d}x\right)^\gamma \left(\int \Phi^{-\frac{\gamma}{1-\gamma}} \mathrm{d}x\right)^{1-\gamma} \ge -C_\gamma \left(\int |\tilde{\beta}_\varepsilon(u_\varepsilon)|\Phi \mathrm{d}x\right)^\gamma \left(\int \Phi^{-\alpha} \mathrm{d}x\right)^{1-\gamma}$$

$$\ge -C_\gamma \left(\sup_{|r|\le C_E\|f\|_{L^\infty}} \beta'(r) + 1\right)^\gamma \left(\sup_{\varepsilon\in(0,1)} \|u_\varepsilon\|_{L^1_\Phi}\right)^\gamma \|\Phi^{-\alpha}\|_{L^1}^{1-\gamma}$$

$$\ge -C_{\Phi,\gamma},$$

where we choose a finite nonnegative constant $C_{\Phi,\gamma}$ independent of $\varepsilon \in (0,1)$ and $\lambda \in (0,\lambda_0)$, which is possible due to Lemma 3.5. In the above calculation, we used Lebesgue's dominated convergence theorem in the first equality and Hölder's inequality in the third line.

IV: We set $h_\varepsilon(r) := \frac{b_\varepsilon(r)r\tilde{\beta}'_\varepsilon(r)}{\tilde{\beta}_\varepsilon(r)}$, for $r \in (0,\infty)$. Since $\beta \in C^2(\mathbb{R})$, h_ε can be continuously extended to a function on $[0,\infty)$ by l'Hospital's rule. We denote this extended function also by h_ε. Furthermore, by (17), we obtain that, via the mean-value theorem, for every $K > 0$ and $C_K^1 := (C_K(\sup_{r\in[0,K]} \beta'(r) + 1)) \vee 1$

$$\beta'(g_\varepsilon(r))r \le C_K^1 \beta(g_\varepsilon(r)) \quad \forall r \in [0, K].$$

Hence, for all $r \in (0, K]$

$$(0 \le) \ h_\varepsilon(r) \le b_\varepsilon(r)\frac{\beta'(g_\varepsilon(r))r + \varepsilon r}{\beta(g_\varepsilon(r)) + \varepsilon r} \le C_K^1 \|b\|_{L^\infty} < \infty. \tag{24}$$

We set $H_\varepsilon(r) := \int_0^r h_\varepsilon(s)ds, r \geq 0$. Note that $H_\varepsilon \geq 0$. Using integration by parts, we obtain

$$\int \mathrm{div}(E_\varepsilon b_\varepsilon(u_\varepsilon)u_\varepsilon)l_{M,n}(\cdot, \tilde\beta_\varepsilon(u_\varepsilon))\mathrm{d}x = -\int \langle E_\varepsilon, \nabla \tilde\beta_\varepsilon(u_\varepsilon)\rangle_{\mathbb{R}^d} b_\varepsilon(u_\varepsilon)u_\varepsilon l'_M(\tilde\beta_\varepsilon(u_\varepsilon))\kappa_n \mathrm{d}x$$

$$- \int \langle E_\varepsilon, \nabla \kappa_n\rangle_{\mathbb{R}^d} b_\varepsilon(u_\varepsilon)u_\varepsilon l_M(\tilde\beta_\varepsilon(u_\varepsilon))\mathrm{d}x.$$

By Lebesgue's dominated convergence theorem, the second summand on the right hand-side vanishes, as $n \to \infty$. Taking the limit of the first summand on the right hand-side yields with the aid of Lebegue's dominated convergence theorem

$$\lim_{n\to\infty} -\int \langle E_\varepsilon, \nabla\tilde\beta_\varepsilon(u_\varepsilon)\rangle_{\mathbb{R}^d} b_\varepsilon(u_\varepsilon)u_\varepsilon l'_M(\tilde\beta_\varepsilon(u_\varepsilon))\kappa_n\mathrm{d}x$$

$$= -\int \langle E_\varepsilon, \nabla\tilde\beta_\varepsilon(u_\varepsilon)\rangle_{\mathbb{R}^d} b_\varepsilon(u_\varepsilon)u_\varepsilon l'_M(\tilde\beta_\varepsilon(u_\varepsilon))\mathrm{d}x = -\int \langle E_\varepsilon, \nabla u_\varepsilon\rangle_{\mathbb{R}^d} h_\varepsilon(u_\varepsilon)\mathbb{1}_{(\exp -M, \exp M)}(\tilde\beta_\varepsilon(u_\varepsilon))\mathrm{d}x.$$

Using (24) with $K := C_E||f||_{L^\infty}$, Lebesgue's dominated convergence theorem and (11) we conclude that

$$-\lim_{M\to\infty} \int \langle E_\varepsilon, \nabla u_\varepsilon\rangle_{\mathbb{R}^d} h_\varepsilon(u_\varepsilon)\mathbb{1}_{(\exp -M, \exp M)}(\tilde\beta_\varepsilon(u_\varepsilon))\mathrm{d}x = -\int \langle E_\varepsilon, \nabla u_\varepsilon\rangle_{\mathbb{R}^d} h_\varepsilon(u_\varepsilon)\mathrm{d}x \tag{25}$$

$$= \int (\mathrm{div}\, E_\varepsilon)H_\varepsilon(u_\varepsilon)\mathrm{d}x \geq -\int (\mathrm{div}\, E_\varepsilon)^- H_\varepsilon(u_\varepsilon)\mathrm{d}x \tag{26}$$

$$\geq -\int \left[(\mathrm{div}\, E)^- + |E|\mathbb{1}_{\{|x|\geq\varepsilon^{-1}\}} \right] H_\varepsilon(u_\varepsilon)\mathrm{d}x \tag{27}$$

$$\geq -C_K^1 ||(\mathrm{div}\, E)^- + |E|||_{L^\infty}||b||_{L^\infty(\mathbb{R}^d)}||f||_{L^1(\mathbb{R}^d)} \tag{28}$$

$$=: -C_{\mathrm{div}} > -\infty. \tag{29}$$

Note that in the first equality we used that $\nabla u_\varepsilon = 0$ a.e. on the set $\{\tilde\beta_\varepsilon(u_\varepsilon) = 0\} = \{u_\varepsilon = 0\}$. Assuming (16) instead of (17), one can choose $C_{\mathrm{div}} := 0$, which is legitimated by $E = E_\varepsilon$, (25)-(26), and the nonnegativity of $H_\varepsilon(u_\varepsilon)$.

V: Similar to the argument in I, we conclude

$$\lim_{n\to\infty} \int f l_{M,n}(\cdot, \tilde\beta_\varepsilon(u_\varepsilon))\mathrm{d}x = \int f l_M(\tilde\beta_\varepsilon(u_\varepsilon))\mathrm{d}x.$$

Now we prove (21). Since $l_M(\tilde\beta_\varepsilon)$ is non-decreasing, we have for $\Psi_\varepsilon^M(r) := \int_0^r l_M(\tilde\beta_\varepsilon(s))ds, r \geq 0$,

$$\Psi_\varepsilon^M(r_1) - \Psi_\varepsilon^M(r_2) \leq l_M(\tilde\beta_\varepsilon(r_1))(r_1 - r_2) \ \forall r_1, r_2 \in \mathbb{R}.$$

Hence, combining the limits and estimates we obtained for **I**-**V**, we obtain for $M \to \infty$ (at least formally up to this point)

$$\liminf_{M\to\infty} \int \Psi_\varepsilon^M(u_\varepsilon)\mathrm{d}x + \lambda \int_{\{\tilde\beta_\varepsilon(u_\varepsilon)>0\}} \frac{|\nabla\tilde\beta_\varepsilon(u_\varepsilon)|^2}{\tilde\beta_\varepsilon(u_\varepsilon)}dx - \lambda(\varepsilon C_{\Phi,\gamma} + C_{\mathrm{div}})$$

$$\leq \liminf_{M\to\infty} \int \Psi_\varepsilon^M(f)\mathrm{d}x. \tag{30}$$

Now, let us verify that the limites inferiores in (30) are actually real limits (which a-posteori makes (30) a rigorous statement). Note that

$$|l_M(r)| = (\ln^+(r) \wedge M) + (\ln^-(r) \wedge M) \ \ \forall r \in [0, \infty). \tag{31}$$

13

Hence, $|\Psi_\varepsilon^M| \le \int_0^\cdot |\ln(\tilde\beta_\varepsilon(r))|dr$ on $[0,\infty)$. Let us show that $\int_0^\cdot |\ln(\tilde\beta_\varepsilon(r))|dr \in L^1(\mathbb{R}^d)$; it will be even shown that $(\int_0^\cdot |\ln(\tilde\beta_\varepsilon(r))|dr)_{\varepsilon\in(0,1)}$ is bounded in $L^1(\mathbb{R}^d)$. Since $\ln^+$ is Lipschitz continuous with $\ln^+(0) = 0$ and Lipschitz constant 1, we may estimate

$$\int_0^{u_\varepsilon(x)} \ln^+(\tilde\beta_\varepsilon(r))dr \le \tilde\beta_\varepsilon(u_\varepsilon)u_\varepsilon(x). \tag{32}$$

Hence, recalling that $||u_\varepsilon||_{L^1} \le ||f||_{L^1}, ||u_\varepsilon||_{L^\infty} \le C_E||f||_{L^\infty}$, we estimate

$$\int_{\mathbb{R}^d}\int_0^{u_\varepsilon(x)} \ln^+(\tilde\beta_\varepsilon(r))drdx \le \left(\sup_{|r|\le C_E||f||_{L^\infty}} \beta'(r)+1\right)C_E||f||_{L^\infty(\mathbb{R}^d)}||f||_{L^1(\mathbb{R}^d)} < \infty,$$

Moreover, by the transformation rule for the Lebesgue integral, we obtain

$$\int_0^{u_\varepsilon(x)} \ln^-(\tilde\beta_\varepsilon(r))dr \le \int_0^{u_\varepsilon(x)} \ln^-(\beta(g_\varepsilon(r)))dr = \int_0^{g_\varepsilon(u_\varepsilon(x))} \ln^-(\beta(r))(1+\varepsilon\beta'(r))dr. \tag{33}$$

Note that (H1) implies that $\ln^-(\beta(r)) \le \ln^-\left(\frac{a_1}{m_1}\right) + m_1\ln^-(r)$ for all $r > 0$. Similar to the estimates we obtained for **III**, noting that $g_\varepsilon(r) \le r \ \forall r \ge 0$, we find that for all $\gamma \in \left[\frac{\alpha}{\alpha+1},1\right)$ and $C_{\beta'} := \left(\sup_{|r|\le C_E||f||_{L^\infty}} \beta'(r)+1\right)$

$$\int_{\mathbb{R}^d}\int_0^{g_\varepsilon(u_\varepsilon(x))} \ln^-(\beta(r))(1+\varepsilon\beta'(r))drdx$$

$$\le C_{\beta'}\left[m_1\int_{\mathbb{R}^d}\int_0^{u_\varepsilon(x)} \ln^-(r)drdx + \ln^-\left(\frac{a_1}{m_1}\right)||u_\varepsilon||_{L^1}\right]$$

$$= C_{\beta'}\left[m_1\int_{\mathbb{R}^d} u_\varepsilon(x)\ln^-(u_\varepsilon(x)) + (u_\varepsilon(x)\wedge 1)dx + \ln^-\left(\frac{a_1}{m_1}\right)||u_\varepsilon||_{L^1}\right]$$

$$\le C_{\beta'}\left[m_1 C_\gamma||u_\varepsilon||_{L^1_\Phi}^\gamma||\Phi^{-\alpha}||_{L^1}^{1-\gamma} + \left(\ln^-\left(\frac{a_1}{m_1}\right)+m_1\right)||u_\varepsilon||_{L^1}\right]$$

$$\le C_{\beta'}\left[m_1 C_\gamma \sup_{\varepsilon\in(0,1)}||u_\varepsilon||_{L^1_\Phi}^\gamma||\Phi^{-\alpha}||_{L^1}^{1-\gamma} + \left(\ln^-\left(\frac{a_1}{m_1}\right)+m_1\right)||f||_{L^1}\right] < \infty. \tag{34}$$

Since $\Psi_\varepsilon^M(r) \to \Psi_\varepsilon(r)$ for all $r \in [0,\infty)$, we therefore conclude that, by Lebesgue's dominated convergence theorem,

$$\Psi_\varepsilon^M(u_\varepsilon) \to \Psi_\varepsilon(u_\varepsilon) \text{ in } L^1(\mathbb{R}^d), \text{ as } M \to \infty,$$

An analogous argument yields

$$\Psi_\varepsilon^M(f) \to \Psi_\varepsilon(f) \text{ in } L^1(\mathbb{R}^d), \text{ as } M \to \infty.$$

Consequently, (30) is equivalent to

$$\int \Psi_\varepsilon(u_\varepsilon)dx + \lambda\int_{\{\tilde\beta_\varepsilon(u_\varepsilon)>0\}} \frac{|\nabla\tilde\beta_\varepsilon(u_\varepsilon)|^2}{\tilde\beta_\varepsilon(u_\varepsilon)}dx - \lambda(\varepsilon C_{\Phi,\gamma} + C_{\text{div}})$$

$$\le \int \Psi_\varepsilon(f)dx. \tag{35}$$

Let $\varphi \in C_c^\infty(\mathbb{R}^d)$ and $\delta > 0$ be arbitrary but fixed. Then, integrating by parts, we obtain

$$\int (\tilde{\beta}_\varepsilon(u_\varepsilon) + \delta)^{\frac{1}{2}} \nabla\varphi \mathrm{d}x = -\int \frac{\nabla\tilde{\beta}_\varepsilon(u_\varepsilon)}{2(\tilde{\beta}_\varepsilon(u_\varepsilon) + \delta)^{\frac{1}{2}}} \varphi \mathrm{d}x = -\int_{\{\tilde{\beta}_\varepsilon(u_\varepsilon)>0\}} \frac{\nabla\tilde{\beta}_\varepsilon(u_\varepsilon)}{2(\tilde{\beta}_\varepsilon(u_\varepsilon) + \delta)^{\frac{1}{2}}} \varphi \mathrm{d}x. \tag{36}$$

where the last equality follows from the fact that $\nabla\tilde{\beta}_\varepsilon(u_\varepsilon) = 0$ a.e. on $\{\tilde{\beta}_\varepsilon(u_\varepsilon) = 0\}$. By Lebesgue's dominated convergence theorem, we may take the limit $\delta \to 0$ in (36) and, hence, obtain from (35) that $\tilde{\beta}_\varepsilon^{\frac{1}{2}}(u_\varepsilon) \in H^1(\mathbb{R}^d)$ with

$$\nabla\tilde{\beta}_\varepsilon^{\frac{1}{2}}(u_\varepsilon) = \mathbb{1}_{\{\tilde{\beta}_\varepsilon(u_\varepsilon)>0\}}(2\tilde{\beta}_\varepsilon^{\frac{1}{2}}(u_\varepsilon))^{-1}\nabla\tilde{\beta}_\varepsilon(u_\varepsilon).$$

Consequently, (21) becomes evident.

Regarding 2. Let $f \in L_\Phi^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ such that $f \geq 0$ a.e. Then, by Lemma 3.1, $u_\varepsilon(\lambda, f) \to J_\lambda(f) =: u$ in $L^1(\mathbb{R}^d)$ and therefore also pointwise almost everywhere for possibly a subsequence, which is not denoted differently. By the proof of Step 1, we have that $(\nabla\tilde{\beta}_\varepsilon^{\frac{1}{2}}(u_\varepsilon))_{\varepsilon\in(0,1)}$ is bounded in $L^2(\mathbb{R}^d)$, as each of the families $\left(\int \Psi_\varepsilon(u_\varepsilon)\mathrm{d}x\right)_{\varepsilon\in(0,1)}$, $\left(\int \Psi_\varepsilon(f)\mathrm{d}x\right)_{\varepsilon\in(0,1)}$ is bounded. Note that $(\tilde{\beta}_\varepsilon)_{\varepsilon\in(0,1)}$ is locally equicontinuous, whence $\tilde{\beta}_\varepsilon(u_\varepsilon) \to \beta(u)$ a.e. Since, for all $\varepsilon \in (0,1)$, $\tilde{\beta}_\varepsilon(u_\varepsilon) \leq \left(1 + \sup_{|r|\leq C_E||f||_{L^\infty}} \beta'(r)\right) f \ (\in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d))$, $\tilde{\beta}_\varepsilon(u_\varepsilon) \to \beta(u)$ in $L^p(\mathbb{R}^d)$, as $\varepsilon \to 0$, for all $p \geq 1$. Hence, via the Banach–Alaoglu theorem, it is easy to see that there exists a (non-relabled) subsubsequence such that

$$\nabla\tilde{\beta}_\varepsilon^{\frac{1}{2}}(u_\varepsilon) \to \nabla\beta^{\frac{1}{2}}(u) \text{ weakly in } L^2(\mathbb{R}^d), \text{ as } \varepsilon \to 0. \tag{37}$$

Note that for all $\gamma \in \left(\frac{\alpha}{\alpha+1}, 1\right)$

$$\int u_\varepsilon^\gamma \mathrm{d}x \to \int u^\gamma \mathrm{d}x, \text{ as } \varepsilon \to 0.$$

Indeed, for all $\rho \in (0,1)$ with $\gamma \in \left[\frac{\alpha}{\alpha+(1-\rho)}, 1\right)$ we have

$$\left|\int u_\varepsilon^\gamma - u^\gamma \mathrm{d}x\right| \leq ||u_\varepsilon - u||_{L^1}^{\gamma\rho} ||u_\varepsilon - u||_{L_\Phi^1}^{\gamma(1-\rho)} ||\Phi^{-\alpha}||_{L^1}^{1-\gamma}.$$

Since by (14) and Lemma 3.5, $||u||_{L_\Phi^1} + \sup_{\varepsilon\in(0,1)} ||u_\varepsilon||_{L_\Phi^1} < \infty$ and $u_\varepsilon \to u$ in $L^1(\mathbb{R}^d)$, the assertion is proved. Furthermore, by the generalised Lebesgue's dominated convergence theorem, it is straight-forward to see that $\Psi_\varepsilon(u_\varepsilon) \to \Psi(u)$ and $\Psi_\varepsilon(f) \to \Psi(f)$ a.e. Hence, using (23) and that for all $\gamma \in (0,1)$

$$|\Psi_\varepsilon(u_\varepsilon)| \leq \left(\sup_{|r|\leq C_E||f||_{L^\infty}} \beta'(r) + 1\right) u_\varepsilon^2 + \ln^-\left(\frac{a_1}{m}\right) u_\varepsilon + m\left(\sup_{|r|\leq C_E||f||_{L^\infty}} \beta'(r) + 1\right)[C_\gamma u_\varepsilon^\gamma + u_\varepsilon],$$

we conclude, via the generalised Lebesgue's dominated convergence theorem, that

$$\int \Psi_\varepsilon(u_\varepsilon)\mathrm{d}x \to \int \Psi(u)\mathrm{d}x, \text{ as } \varepsilon \to 0. \tag{38}$$

Similarly,

$$\int \Psi_\varepsilon(f)\mathrm{d}x \to \int \Psi(f)\mathrm{d}x, \text{ as } \varepsilon \to 0.$$

Hence, taking $\liminf_{\varepsilon\to0}$ on both sides in (35), we obtain

$$\int \Psi(u)\mathrm{d}x + 4\lambda\int |\nabla\tilde{\beta}^{\frac{1}{2}}(u)|^2 dx - \lambda C_{\mathrm{div}} \leq \int \Psi(f)\mathrm{d}x. \tag{39}$$

This finishes the proof for (22).

Regarding 3. Now let us consider the finite difference scheme (19). Due to (39), for all $h \in (0, \gamma_0)$ and $i \in \{0, ..., N-1\}$, we obtain the estimate

$$\int \Psi(u_h^{i+1})\mathrm{d}x + 4h \int |\nabla \beta^{\frac{1}{2}}(u_h^{i+1})|^2 \mathrm{d}x - hC_{\mathrm{div}} \leq \int \Psi(u_h^i)\mathrm{d}x. \tag{40}$$

Let $t \in (0, T)$ and note that $t \leq \left[\frac{Nt}{T}\right]h + h < T$. Now, summing from 0 to $k - 1 := \left[\frac{Nt}{T}\right]$, we obtain

$$\int \Psi(u_h^k)\mathrm{d}x + 4h\sum_{i=0}^{k-1} \int |\nabla \beta^{\frac{1}{2}}(u_h^{i+1})|^2 \mathrm{d}x - TC_{\mathrm{div}} \leq \int \Psi(\nu)\mathrm{d}x.$$

Using (19), we may conclude that

$$\int \Psi(u_h(t))\mathrm{d}x + 4 \int_0^{\left[\frac{Nt}{T}\right]h+h} \int |\nabla \beta^{\frac{1}{2}}(u_h(s))|^2 \mathrm{d}x \mathrm{d}s - TC_{\mathrm{div}} \leq \int \Psi(\nu)\mathrm{d}x. \tag{41}$$

By [BR21a, (3.64) and below], there exists $h_0 \in (0, \gamma_0)$ such that for all $0 < h < h_0$

$$\|u_h(t)\|_{L^\infty} \leq \exp\left(\left\|(\mathrm{div}\,E)^- + |E|\right\|_{L^\infty}^{\frac{1}{2}} t\right) \|\nu\|_{L^\infty} + 1 \ \forall t \in [0, \infty). \tag{42}$$

Let us set $C_\nu := \exp\left(\left\|(\mathrm{div}\,E)^- + |E|\right\|_{L^\infty}^{\frac{1}{2}} T\right) \|\nu\|_{L^\infty} + 1$.

Furthermore, by Remark 3.6, [BR21a, below (3.64)] and (19),

$$\sup_{t \in (0,T), h \in (0,h_0)} \|u_h(t)\|_{L_\Phi^1} \leq \|\nu\|_{L_\Phi^1} + T\left[\left(\sup_{|r| \leq C_\nu} \beta'(r)\right) \|\Delta\Phi\|_{L^\infty} + \|E\|_{L^\infty}\|b\|_{L^\infty}\|\nabla\Phi\|_{L^\infty}\right]\|\nu\|_{L^1}.$$

Therefore, a similar reasoning as for (38) yields that (for a non-relabled subsequence)

$$\int \Psi(u_h(t))\mathrm{d}x \to \int \Psi(u(t))\mathrm{d}x, \text{ as } h \to 0,$$

where the convergence holds in $\mathbb{R}$.

Hence, $(\nabla\beta^{\frac{1}{2}}(u_h))_{h \in (0,h_0)}$ is bounded in $L^2((0,t) \times \mathbb{R}^d)$ for all $t \in (0,T)$. It is easy to check that, for all $p \geq 2$, $\beta^{\frac{1}{2}}(u_h) \to \beta^{\frac{1}{2}}(u)$ in $L^p((0,T) \times \mathbb{R}^d)$ for a subsequence $h \to 0$. Thus, it is straight-forward to see that via the Banach–Alaoglu theorem, for a non-relabled subsubsequence,

$$\nabla\beta^{\frac{1}{2}}(u_h) \to \nabla\beta^{\frac{1}{2}}(u) \text{ weakly in } L^2((0,t) \times \mathbb{R}^d), \text{ as } h \to 0.$$

Hence, taking $\liminf_{h \to 0}$ in (41), we obtain for all $t \in (0,T)$

$$\int \Psi(u(t))\mathrm{d}x + 4\int_0^t \int |\nabla\beta^{\frac{1}{2}}(u(s))|^2 \mathrm{d}x \mathrm{d}s - TC_{\mathrm{div}} \leq \int \Psi(\nu)\mathrm{d}x.$$

This completes the proof. □

In order to prove Theorem (1.2), we will use Theorem 2.1 and show (i) in the latter's formulation. This will be done in the subsequent two subsections.

3.2 Ingredient 1: A $P_\nu^{(u_t)}$-weak solution to (MVSDE)

The following result is part of [BR21a, Theorem 6.1 (a)], which is based on the superposition principle procedure from [BR20, Section 2].

Theorem 3.8 ($P_\nu^{(u_t)}$-weak solution)**.** *Assume that (C1) and (C2) hold. Let $\nu \in \mathcal{P}_0(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$. Then for each $T > 0$, there exists a $P_\nu^{(u_t)}$-weak solution (X, W) to (MVSDE) up to time T, where u is the probability solution to (FPE) with $u|_{t=0} = \nu$ provided by Theorem 3.3.*

3.3 Ingredient 2: $P_\nu^{(u_t)}$-pathwise uniqueness for (MVSDE)

Suppose u is a probability solution to (FPE) with initial condition $u|_{t=0} = \nu \in \mathcal{P}(\mathbb{R}^d)$. Recall that, by definition, $P_\nu^{(u_t)}$-weak solutions to (MVSDE) all have the same time marginal laws. Equivalently, they are $P_\nu^{(u_t)}$-weak solutions to the following *ordinary* SDE

$$dX(t) = \boldsymbol{b}^u(t, X(t))\, dt + \boldsymbol{\sigma}^u(t, X(t))\, dW(t), \quad t \in [0, T], \tag{SDE$_u$}$$
$$\mathcal{L}_{X(0)} = \nu,$$

where $\boldsymbol{b}^u(t, x) := E(x) b(u_t(x))$ and $\boldsymbol{\sigma}^u(t, x) := \sqrt{2 \frac{\beta(u_t(x))}{u_t(x)}} \mathbb{1}_{d \times d}$, $(t, x) \in [0, T] \times \mathbb{R}^d$. As already indicated in the introduction, we will fix a precise representative of the u_t, $t \in [0, T]$, in the coefficients of (MVSDE) such that $\boldsymbol{b}^u$, $\boldsymbol{\sigma}^u$ are Borel measurable. In particular, u is pointwise defined on $[0, T] \times \mathbb{R}^d$. In the following theorem, we are going to prove $P_\nu^{(u_t)}$-pathwise uniqueness for (SDE$_u$). The proof is based on a combination of a proper modification of a pathwise uniqueness result for SDEs performed in the proof of [RZ10, Theorem 1.1], Theorem 3.9, and a trick involving the mean-value theorem, which is motivated by [FG23, (76)-(77)].

Theorem 3.9 ($P_\nu^{(u_t)}$-pathwise uniqueness)**.** *Assume that all the conditions in Theorem 3.7 are fulfilled. Furthermore, let $m \in (1, \infty), \zeta \in [0, 1]$ such that $\frac{2\zeta}{m} < 1$ and assume that*

$$\forall K > 0 \; \exists C_K, a_K \in (0, \infty) \; \forall r \in [0, K] : \beta(r) \leq C_K r^m \text{ and } a_K |r|^{m-1} \leq \beta'(r), \tag{43}$$

that for all $R > 0$ there exits $\iota_R \in L^1(B_R(0))$ such that for a.e. $x, y \in B_R(0)$

$$\langle E(x) - E(y), x - y \rangle_{\mathbb{R}^d} \leq (\iota_R(x) + \iota_R(y)) |x - y|_{\mathbb{R}^d}^2, \tag{44}$$

and that $b \circ (G \circ \beta^{\frac{1}{2}})^{-1} \in \mathrm{Lip}_{loc}([0, \infty))$, where $G(r) := \int_0^r (\beta^{-1}(s^2))^{-\zeta} ds$, for all $r \geq 0$.

Let $\nu \in \mathcal{P}_0(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d) \cap L_\Phi^1(\mathbb{R}^d)$ for some $\Phi \in \mathcal{Q}$ and let u denote the corresponding probability solution to (FPE) with $u|_{t=0} = \nu$ provided by Theorem 3.3. Let $T \in (0, \infty)$. Suppose $(X, W), (Y, W)$ are two $P_\nu^{(u_t)}$-weak solutions to (MVSDE) up to time T on a common stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t \in [0, T]})$ with respect to the same standard d-dimensional $(\mathcal{F}_t)$-Brownian motion $(W(t))_{t \in [0, T]}$ such that $X(0) = Y(0)$ $\mathbb{P}$-a.s. Then $\sup_{t \in [0, T]} |X(t) - Y(t)| = 0$ $\mathbb{P}$-a.s.

Proof. Let $Z(t) := X(t) - Y(t), t \in [0, T]$, $\tau_R := \inf\{t \in [0, T] : |X(t)| \vee |Y(t)| \geq R\} \wedge T$ for $R > 0$, and $\delta > 0$ be arbitrary but fixed. As in the proof of [RZ10, Theorem 1.1], we apply Itô's formula to $f_\delta(Z(t \wedge \tau_R))$, where f_δ is defined as (the C^2-function) $f_\delta(x) := \ln\left(\frac{|x|^2}{\delta^2} + 1\right), x \in \mathbb{R}^d$, with $|\nabla f_\delta(x)| \leq \frac{C}{|x| + \delta}, |\Delta f_\delta(x)| \leq \frac{C}{|x|^2 + \delta^2}$, for

all $x \in \mathbb{R}^d$ and some $C \in (0, \infty)$ independent of x. We obtain the following estimates, utilising Fatou's lemma.

$$
\mathbb{E} \ln \left(1 + \frac{|Z(t \wedge \tau_R)|^2}{\delta^2} \right) \tag{45}
$$

$$
\leq C \mathbb{E} \int_0^{t \wedge \tau_R} \frac{\langle E(X(s))b(u_s(X(s))) - E(Y(s))b(u_s(Y(s))), Z(s) \rangle_{\mathbb{R}^d}}{|Z(s)|^2 + \delta^2} \mathrm{d}s
$$

$$
+ C \mathbb{E} \int_0^{t \wedge \tau_R} \frac{\left| \left(\frac{\beta(u_s(X(s)))}{u_s(X(s))} \right)^{\frac{1}{2}} - \left(\frac{\beta(u_s(Y(s)))}{u_s(Y(s))} \right)^{\frac{1}{2}} \right|^2}{|Z(s)|^2 + \delta^2} \mathrm{d}s
$$

$$
\leq C \mathbb{E} \int_0^{t \wedge \tau_R} \frac{\langle E(X(s))b(u_s(X(s))) - E(Y(s))b(u_s(Y(s))), Z(s) \rangle_{\mathbb{R}^d}}{|Z(s)|^2 + \delta^2} \mathrm{d}s
$$

$$
+ C \liminf_{\eta \to 0} \mathbb{E} \int_0^{t \wedge \tau_R} \frac{\left| \left(\frac{\beta(u_s(X(s)))}{u_s(X(s)) + \eta} \right)^{\frac{1}{2}} - \left(\frac{\beta(u_s(Y(s)))}{u_s(Y(s)) + \eta} \right)^{\frac{1}{2}} \right|^2}{|Z(s)|^2 + \delta^2} \mathrm{d}s
$$

$$
=: \mathrm{I} + \liminf_{\eta \to 0} \mathrm{II}_\eta. \tag{46}
$$

Regarding II_η:

Let $\eta > 0$ and define $f_\eta(r) := \frac{r}{(\beta^{-1}(r^2) + \eta)^{\frac{1}{2}}}, r \geq 0$. Note that f_η is continuously differentiable in a sufficiently small open neighbourhood of $[0, \infty)$ in $\mathbb{R}$. A simple computation yields

$$
f_\eta'(r) = \frac{1}{(\beta^{-1}(r^2) + \eta)^{\frac{1}{2}}} - \frac{r^2 (\beta^{-1})'(r^2)}{(\beta^{-1}(r^2) + \eta)^{\frac{3}{2}}} \quad \forall r \geq 0. \tag{47}
$$

Let $\tilde{K} \in (0, \infty)$ and $K := \beta^{-1}(\tilde{K})$. Then by (43) and the inverse function theorem, one can find $C_1 = C_1(\tilde{K}), C_2 = C_2(\tilde{K}) \in (0, \infty)$ such that

$$
C_1 r^{\frac{1}{m}} \leq \beta^{-1}(r) \quad \forall r \in [0, \tilde{K}] \text{ and } (\beta^{-1})'(r) \leq C_2 r^{-\frac{m-1}{m}} \quad \forall r \in (0, \tilde{K}].
$$

Via (47), we obtain for $C_3 := \frac{C_2}{C_1} + 1$

$$
|f_\eta'(r)| \leq C_3 (C_1 r^{\frac{2}{m}} + \eta)^{-\frac{1}{2}} \quad \forall r \in [0, \tilde{K}].
$$

Therefore, choosing $\tilde{K} := \beta^{\frac{1}{2}}(\|u\|_{L^\infty}) \vee \beta(\|u\|_{L^\infty})$, we may estimate via the mean value theorem for all $x, y \in \mathbb{R}^d, s \in [0, T]$

$$
\left| \left(\frac{\beta(u_s(x))}{u_s(x) + \eta} \right)^{\frac{1}{2}} - \left(\frac{\beta(u_s(y))}{u_s(y) + \eta} \right)^{\frac{1}{2}} \right| = |f_\eta(\beta^{\frac{1}{2}}(u_s(x))) - f_\eta(\beta^{\frac{1}{2}}(u_s(y)))|
$$

$$
\leq \int_0^1 \left| f_\eta' \left(\theta \beta^{\frac{1}{2}}(u_s(x)) + (1 - \theta)\beta^{\frac{1}{2}}(u_s(y)) \right) \right| \mathrm{d}\theta \left| \beta^{\frac{1}{2}}(u_s(x)) - \beta^{\frac{1}{2}}(u_s(y)) \right|
$$

$$
\leq h_\eta(s, x) \left| \beta^{\frac{1}{2}}(u_s(x)) - \beta^{\frac{1}{2}}(u_s(y)) \right|, \tag{48}
$$

where $h_\eta(s, x) := C_3 \int_0^1 (C_4 u_s(x) \theta^{\frac{2}{m}} + \eta)^{-\frac{1}{2}} \mathrm{d}\theta, s \in [0, T], x \in \mathbb{R}^d$, and $C_4 := C_1 (\frac{a}{m})^{\frac{1}{m}}$. Note that $\|h_\eta\|_\infty \leq \frac{C_3}{\sqrt{\eta}}$ and

$$
u_s(x) h_\eta^2(s, x) \leq C_3^2 C_4^{-1} \left(\int_0^1 \theta^{-\frac{1}{m}} \mathrm{d}\theta \right)^2 < \infty. \tag{49}
$$

For $f \in L_{loc}^1(\mathbb{R}^d)$, we set $f_\varepsilon(x) := (f * \varphi_\varepsilon)(x)$, for all $x \in \mathbb{R}^d$, where $\varphi_\varepsilon(x) = \varepsilon^{-d} \varphi(\varepsilon^{-1} x)$, for some fixed

$\varphi \in \mathcal{P}_0(\mathbb{R}^d) \cap C_c^\infty(\mathbb{R}^d), \varepsilon \in (0,1)$. Using this notation and (48) we estimate

$$\text{II}_\eta \leq C\mathbb{E} \int_0^{t \wedge \tau_R} (h_\eta(s, X(s)) \wedge h_\eta(s, Y(s)))^2 \frac{\left| (\beta^{\frac{1}{2}}(u_s))_\varepsilon(X(s)) - (\beta^{\frac{1}{2}}(u_s))_\varepsilon(Y(s)) \right|^2}{|Z(s)|^2 + \delta^2} \mathrm{d}s \tag{50}$$

$$+ C\mathbb{E} \int_0^{t \wedge \tau_R} (h_\eta(s, X(s)) \wedge h_\eta(s, Y(s)))^2 \frac{\left| (\beta^{\frac{1}{2}}(u_s))_\varepsilon(X(s)) - \beta^{\frac{1}{2}}(u_s(X(s))) \right|^2}{|Z(s)|^2 + \delta^2} \mathrm{d}s$$

$$+ C\mathbb{E} \int_0^{t \wedge \tau_R} (h_\eta(s, X(s)) \wedge h_\eta(s, Y(s)))^2 \frac{\left| (\beta^{\frac{1}{2}}(u_s))_\varepsilon(Y(s)) - \beta^{\frac{1}{2}}(u_s(Y(s))) \right|^2}{|Z(s)|^2 + \delta^2} \mathrm{d}s$$

$$=: \text{II}_1^{\varepsilon,\eta} + \text{II}_2^{\varepsilon,\eta} + \text{II}_3^{\varepsilon,\eta}.$$

By Lemma A.2, we have for all $x, y \in \mathbb{R}^d$

$$\left| (\beta^{\frac{1}{2}}(u_s))_\varepsilon(x) - (\beta^{\frac{1}{2}}(u_s))_\varepsilon(y) \right| \leq \left(\left(\mathrm{M} |\nabla \beta^{\frac{1}{2}}(u_s)| \right)_\varepsilon(x) + \left(\mathrm{M} |\nabla \beta^{\frac{1}{2}}(u_s)| \right)_\varepsilon(y) \right) |x - y|.$$

Therefore, using Lemma A.3 and (49), we estimate

$$\text{II}_1^{\varepsilon,\eta} \leq C\mathbb{E} \int_0^{t \wedge \tau_R} (h_\eta(s, X(s)) \wedge h_\eta(s, Y(s)))^2 \frac{\left(\left(\mathrm{M} |\nabla \beta^{\frac{1}{2}}(u_s)| \right)_\varepsilon(X(s)) + \left(\mathrm{M} |\nabla \beta^{\frac{1}{2}}(u_s)| \right)_\varepsilon(Y(s)) \right)^2 |Z(s)|^2}{|Z(s)|^2 + \delta^2} \mathrm{d}s$$

$$\leq C \int_0^T \int_{\mathbb{R}^d} \left(\left(\mathrm{M} |\nabla \beta^{\frac{1}{2}}(u_s)| \right)_\varepsilon(x) \right)^2 h_\eta^2(s, x) u_s(x) \mathrm{d}x \mathrm{d}s \leq C \|\nabla \beta^{\frac{1}{2}}(u)\|_{L^2((0,T) \times \mathbb{R}^d; \mathbb{R}^d)}^2 < \infty, \tag{51}$$

where the right hand-side is finite, according to Theorem 3.7. Furthermore, using (49), we obtain

$$\max\{\text{II}_2^{\varepsilon,\eta}, \text{II}_2^{\varepsilon,\eta}\} \leq \frac{C}{\delta^2} \int_0^T \int_{B_R(0)} |(\beta(u_s))_\varepsilon(x) - \beta(u_s(x))| h_\eta(s, x)^2 u_s(x) \mathrm{d}x \mathrm{d}s$$

$$\leq \frac{C}{\delta^2} \|(\beta(u))_\varepsilon - \beta(u)\|_{L^1((0,T) \times B_R(0))} \to 0, \text{ as } \varepsilon \to 0. \tag{52}$$

Regarding I:
Let us define $G_\eta(r) := \int_0^r ((\beta^{-1})(s^2) + \eta)^{-\zeta} \mathrm{d}s$, for all $r \in [0, \infty)$. Then, $G_\eta(r)$ is continuously differentiable in a sufficiently small neighbourhood of $[0, \infty)$ and, by the monotone convergence theorem, $\lim_{\eta \to 0} G_\eta(r) = G(r)$ for all $r \in [0, \infty)$. Choosing K as above, we obtain similar to (48) that for all $x, y \in \mathbb{R}^d$

$$|G_\eta(\beta^{\frac{1}{2}}(u_s(x))) - G_\eta(\beta^{\frac{1}{2}}(u_s(y)))| \leq (\bar{h}_\eta(x) \wedge \bar{h}_\eta(y)) |\beta^{\frac{1}{2}}(u_s(x)) - \beta^{\frac{1}{2}}(u_s(y))|, \tag{53}$$

where $\bar{h}_\eta(s, x) := \int_0^1 (C_4 u_s(x) \theta^{\frac{2}{m}} + \eta)^{-\zeta} \mathrm{d}\theta, s \in [0, T], x \in \mathbb{R}^d$. Note that

$$\bar{h}_\eta(s, x) u_s(x) \leq C_4^{-\zeta} \|(\cdot)^{-\frac{2\zeta}{m}}\|_{L^1((0,1))} \|u\|_{L^\infty}^{1-\zeta} < \infty,$$

since $\frac{2\zeta}{m} \in [0, 1)$. Since $b \circ (G \circ \beta^{\frac{1}{2}})^{-1} \in \mathrm{Lip}_{loc}([0, \infty))$, we obtain, using (53), that

$$\text{I} \leq C\mathbb{E} \int_0^{t \wedge \tau_R} \frac{\|E\|_{L^\infty} |b(u_s(X(s))) - b(u_s(Y(s)))| + \|b\|_{L^\infty} |E(X(s)) - E(Y(s))|}{|Z(s)| + \delta} \mathrm{d}s$$

$$\leq C \|E\|_{L^\infty} \|b \circ (G \circ \beta^{\frac{1}{2}})^{-1}\|_{\mathrm{Lip}((G \circ \beta^{\frac{1}{2}})(K))}$$

$$\cdot \liminf_{\eta \to 0} \mathbb{E} \int_0^{t \wedge \tau_R} (\bar{h}_\eta(s, X(s)) \wedge \bar{h}_\eta(s, Y(s))) \frac{|\beta^{\frac{1}{2}}(u_s(X(s))) - \beta^{\frac{1}{2}}(u_s(Y(s)))|}{|Z(s)| + \delta} \mathrm{d}s$$

$$+ C \|b\|_{L^\infty} \mathbb{E} \int_0^{t \wedge \tau_R} \frac{|E(X(s)) - E(Y(s))|}{|Z(s)| + \delta} \mathrm{d}s \tag{54}$$

Arguing in a similar fashion as in (50)-(52), the first integral on the right hand-side of (54) can be estimated

as

$$\mathbb{E}\int_0^{t\wedge\tau_R}(\bar{h}_\eta(s,X(s))\wedge\bar{h}_\eta(s,Y(s)))\frac{|\beta^{\frac{1}{2}}(u_s(X(s)))-\beta^{\frac{1}{2}}(u_s(Y(s)))|}{|Z(s)|+\delta}\mathrm{d}s\le C||\nabla\beta^{\frac{1}{2}}(u)||_{L^2((0,T)\times\mathbb{R}^d;\mathbb{R}^d)}<\infty.$$

Using an approximation of E by convolution, we may again perform a similar argument as in (50)-(52) for the second integral on the right hand-side of (54). Indeed; setting $\bar{E}_\varepsilon := E*\varphi_\varepsilon$, where the convolution is meant component-wise, we have

$$\mathbb{E}\int_0^{t\wedge\tau_R}\frac{|E(X(s))-E(Y(s))|}{|Z(s)|+\delta}\mathrm{d}s=\mathbb{E}\int_0^{t\wedge\tau_R}\frac{|\bar{E}_\varepsilon(X(s))-\bar{E}_\varepsilon(Y(s))|}{|Z(s)|+\delta}\mathrm{d}s+\mathbb{E}\int_0^{t\wedge\tau_R}\frac{|\bar{E}_\varepsilon(X(s))-E(X(s))|}{|Z(s)|+\delta}\mathrm{d}s$$
$$+\mathbb{E}\int_0^{t\wedge\tau_R}\frac{|\bar{E}_\varepsilon(Y(s))-E(Y(s))|}{|Z(s)|+\delta}\mathrm{d}s. \tag{55}$$

For the first summand on the right hand-side we estimate similarly to (51)

$$\mathbb{E}\int_0^{t\wedge\tau_R}\frac{|\bar{E}_\varepsilon(X(s))-\bar{E}_\varepsilon(Y(s))|}{|Z(s)|+\delta}\mathrm{d}s\le\mathbb{E}\int_0^{t\wedge\tau_R}\frac{(i_{R+1})_\varepsilon(X(s))+(i_{R+1})_\varepsilon(Y(s))}{|Z(s)|+\delta}\mathrm{d}s$$
$$\le 2\int_0^T\int_{B_R(0)}(i_{R+1})_\varepsilon(x)u_s(x)\mathrm{d}x\mathrm{d}s\le 2T||i_{R+1}||_{L^1(B_R(0))}||u||_{L^\infty}<\infty. \tag{56}$$

The last two summands on the right hand-side in (55) vanish with a similar reasoning as in (52).

All in all, we obtain by monotone convergence that

$$\mathbb{E}\ln\left(1+\sup_{\delta>0}\frac{|Z(t\wedge\tau_R)|^2}{\delta^2}\right)=\sup_{\delta>0}\mathbb{E}\ln\left(1+\frac{|Z(t\wedge\tau_R)|^2}{\delta^2}\right)<\infty. \tag{57}$$

Necessarily, we therefore have for all $t\in[0,T]$

$$X(t\wedge\tau_R)=Y(t\wedge\tau_R)\quad\mathbb{P}\text{-a.s.} \tag{58}$$

Note that $\sup_{t\in[0,T]}|X(t)|+|Y(t)|<\infty$ $\mathbb{P}$-a.s. Hence, $\tau_R\to T$ $\mathbb{P}$-a.s., as $R\to\infty$. Therefore, taking the limit $R\to\infty$ in (58) and recalling that $X,Y\in C([0,T];\mathbb{R}^d)$, we obtain that $\mathbb{P}$-a.s.

$$X(t)=Y(t)\quad\forall t\in[0,T].$$

This finishes the proof. $\qquad\square$

4 Proof of Theorem 1.3

In this section, we prove the second main result, Theorem 1.3. Here, the corresponding Fokker–Planck equation is the classical porous medium equation with start in a Dirac measure. More precisely, we consider (FPE) with $\beta(r)=|r|^{m-1}r, r\in\mathbb{R}$, for some $m>1$, $E\equiv b\equiv 0$, and $\nu=\delta_{x_0}$, i.e.

$$\partial_t u-\Delta(|u|^{m-1}u)=0, u|_{t=0}=\delta_{x_0},\quad t\in[0,\infty), \tag{PME}$$

where $x_0\in\mathbb{R}^d$. In this case, (MVSDE) has the particular form

$$dX(t)=\sqrt{2\left|\frac{d\mathcal{L}_{X(t)}}{dx}(X(t))\right|^{m-1}}\mathbb{1}_{d\times d}\ \mathrm{d}W(t),\quad t\in[0,T],$$
$$X(0)=x_0. \tag{MVSDE.PME}$$

This section is split into three subsections. In Section 4.1, we recall the Barenblatt solution to (PME) and recall and prove regularity results for powers of this special solution. In Sections 4.2 and 4.3, we then show

that the conditions (i) and (ii) of the restricted Yamada–Watanabe, see Theorem 2.1, are fulfilled, respectively. While the results in Sections 4.1 and 4.2 are true in the multi-dimensional case, we need to restrict ourselves to the one-dimensional setting in Section 4.3; more details can be found in Remark 4.4.

Proof of Theorem 1.3. The assertion follows directly Theorem 4.2 and Theorem 4.3 via Theorem 2.1. □

4.1 The classical probability solution to (PME)

It is well known that u, as defined by the following formula, is a probability solution to (PME). It is called the *Barenblatt solution* and has the form

$$u(t,x) := t^{-\alpha}(C - k|(x-x_0)t^{-\beta}|^2)_+^{\frac{1}{m-1}}, \quad (t,x) \in (0,\infty) \times \mathbb{R}^d, \tag{BB}$$

where $m > 1, \alpha := \frac{d}{d(m-1)+2}, k := \frac{\alpha(m-1)}{2md}, \beta := \frac{\alpha}{d}$, and $C \in (0,\infty)$ is a normalising constant. In this section, β shall not be confused with the diffusivity function from (MVSDE). By the argument presented in [GST20, Section 3.1.] (and as repeated in the proof of Proposition 4.1 below), it is simply not true that $\nabla u^{\frac{m}{2}} \in L^2_{(loc)}((0,T) \times \mathbb{R}^d)$. Hence, we cannot perform the same proof as for Theorem 3.9 in order to show that pathwise uniqueness holds among all weak solutions to (MVSDE.PME) with one-dimensional time marginal law densities u. However, it is true that $\nabla^s u^{\frac{m}{2}} \in L^2((0,T) \times \mathbb{R}^d)$ for all $s \in (0,1)$. This will turn out to be enough to prove such a pathwise uniqueness statement in the case $d = 1$, see Section 4.3.

The following proposition provides, separately, a necessary and a sufficient condition for fractional Sobolev regularity in space and integrability in time of certain powers of the Barenblatt solution. Proposition 4.1 (i) is taken from [GST20, Section 3.1.]. The proof of Theorem 4.1 (ii) uses and recalls the proof of the latter.

Proposition 4.1. *Let $m > 1$, $p \in (0,m]$ and let $x_0 \in \mathbb{R}^d$. Let u denote the Barenblatt solution to (PME) with $u|_{t=0} = \delta_{x_0}$, i.e. u is given by (BB).*

(i) *([GST20, Section 3.1.]) Assume that for some $s \in (0,\infty)$*

$$u^p \in L^{\frac{m}{p}}((0,T); \dot{W}^{s,\frac{m}{p}}(\mathbb{R}^d)). \tag{59}$$

Then, $s < \frac{2p}{m}$.

(ii) *Assume that m, p fulfill $m(2p - m + 1) > p$ and let $s \in (0,1)$ such that $s < \frac{2p}{m}$. Then,*

$$u^p \in L^{\frac{m}{p}}((0,T); W^{s,\frac{m}{p}}(\mathbb{R}^d)).$$

Proof. Let us recall the reasoning from [GST20, Section 3.1.]. We set $F(x) := (C - k|x|^2)_+^{\frac{p}{m-1}}, x \in \mathbb{R}^d$. Using the transformation rule for the Lebesgue integral, a direct computation reveals that for $s \in (0,1)$

$$||u^p||^{\frac{m}{p}}_{L^{\frac{m}{p}}((0,T); \dot{W}^{s,\frac{m}{p}}(\mathbb{R}^d))} = ||t^{-\alpha(m-1)-\beta\frac{sm}{p}}||_{L^1((0,T))}||F||^{\frac{m}{p}}_{\dot{W}^{s,\frac{m}{p}}(\mathbb{R}^d)}. \tag{60}$$

The first factor on the right hand-side is finite as

$$-\alpha(m-1) - \beta\frac{sm}{p} > -1 \iff s < \frac{2p}{m}.$$

A similar argument can be performed for $s \in (1,2)$ since $\partial_{x_i} u^p(t,x) = t^{-\alpha m - \beta}\partial_{x_i}F(xt^{-\beta})$, for all $t > 0, x \in \mathbb{R}^d$ and $i \in \{1, ..., d\}$. Hence, by induction, the proof of (i) is complete.

For (ii), let m, p such that $m(2p - m + 1) > p$ and let $0 < s < \frac{2p}{m} \wedge 1$. Using the transformation rule for the Lebesgue integral, we have

$$||u^p||^{\frac{m}{p}}_{L^{\frac{m}{p}}((0,T) \times \mathbb{R}^d)} = ||t^{-\alpha(m-1)}||_{L^1((0,T))}||F||^{\frac{m}{p}}_{L^{\frac{m}{p}}(\mathbb{R}^d)}.$$

Since $-\alpha(m-1) > -1$ and $F \in C_c(\mathbb{R}^d)$, we obtain that $u^p \in L^{\frac{m}{p}}((0,T) \times \mathbb{R}^d)$. By (60), it is sufficient to prove that $\nabla F \in L^{\frac{m}{p}}(\mathbb{R}^d)$, as then, via the Brezis–Mironescu interpolation theorem (see [BM18, Corollary 5.1]),

$$||F||_{\dot{W}^{s,\frac{m}{p}}(\mathbb{R}^d)} \leq ||F||_{L^{\frac{m}{p}}(\mathbb{R}^d)}^{(1-s)} ||\nabla F||_{L^{\frac{m}{p}}(\mathbb{R}^d)}^{s} < \infty.$$

So, let us show that $\nabla F \in L^{\frac{m}{p}}(\mathbb{R}^d)$. Clearly, F is weakly differentiable with

$$\nabla F(x) = \frac{p}{m-1}(C - k|x|^2)^{\frac{p-(m-1)}{m-1}}(-2kx)\mathbb{1}_{B_{\sqrt{C/k}}(0)}(x), \text{ for a.e. } x \in \mathbb{R}^d.$$

Setting $\rho_1 := \frac{m(p-(m-1))}{p(m-1)}$, we have

$$||\nabla F||_{L^{\frac{m}{p}}}^{\frac{m}{p}} = \left(\frac{2kp}{m-1}\right)^{\frac{m}{p}} \int_{B_{\sqrt{C/k}}(0)} (C - k|x|^2)^{\rho_1}|x|^{\frac{m}{p}} \,\mathrm{d}x.$$

Integrating over shells, we obtain for $\rho_2 := \frac{m}{p} + (d-1)$

$$\int_{B_{\sqrt{C/k}}(0)} (C - k|x|^2)^{\rho_1}|x|^{\frac{m}{p}} \,\mathrm{d}x = C_d \int_0^{\sqrt{C/k}} (C - kr^2)^{\rho_1} r^{\rho_2} \,\mathrm{d}r,$$

where $C_d \in (0,\infty)$ is some real constant depending only on the dimension d. Using the transformation rule for the Lebesgue integral, we obtain

$$\int_0^{\sqrt{C/k}} (C - kr^2)^{\rho_1} r^{\rho_2} \,\mathrm{d}r = \frac{1}{2k^{\frac{\rho_2}{2}}} \int_0^C r^{\rho_1}(C - r)^{\frac{\rho_2}{2}} \,\mathrm{d}r.$$

The integral on the right hand-side is finite, since $\rho_1 > -1$ and $\rho_2 > -2$ by our assumptions. This completes the proof. □

4.2 Ingredient 1: A $P_{\delta_{x_0}}^{(u_t)}$-weak solution to (MVSDE.PME)

In [BR21a, Chapter 5], Barbu and Röckner solved a huge class of nonlinear Fokker–Planck equations of the type (FPE) with start in a bounded measure, and then, via the superposition principle procedure from [BR20, Section 2], solved the corresponding McKean–Vlasov SDE (MVSDE) based on the constructed solution to (FPE). The following theorem is implied by [BR21a, Theorem 6.1 (b)].

Theorem 4.2 ($P_{\delta_{x_0}}^{(u_t)}$-weak solution). *Let $x_0 \in \mathbb{R}^d$. Then, there exists a $P_{\delta_{x_0}}^{(u_t)}$-weak solution (X,W) to (MVSDE) where u is the Barenblatt solution to (PME) with $u|_{t=0} = \delta_{x_0}$.*

4.3 Ingredient 2: $P_{\delta_{x_0}}^{(u_t)}$-pathwise uniqueness for (MVSDE.PME)

Let u denote the Barenblatt solution, i.e. u is given by (BB). As emphasised already in Section 4.1, we cannot use the same lines of proof as for Theorem 3.9 in order to prove $P_{\delta_{x_0}}^{(u_t)}$-pathwise uniqueness. The reason for this is that u does not satisfy $\nabla u^{\frac{m}{2}} \in L^2_{loc}((0,T) \times \mathbb{R}^d)$, but only $\nabla^s u^{\frac{m}{2}} \in L^2((0,T) \times \mathbb{R}^d)$ for all $s \in (0,1)$.

The proof of the following theorem is similar to the one of Theorem 3.9, however, it additionally relies on the ideas of the proof of [CJ18, Theorem. 2.15 (i)] together with [CJ18, Lemma 3.5]; the latter is a variation of Lemma A.2 for Sobolev functions of fractional order $1/2$, see (63) below.

Theorem 4.3 ($P_{\delta_{x_0}}^{(u_t)}$-pathwise uniqueness). *Let $d = 1$, $m > 1$, and $x_0 \in \mathbb{R}$. Let u denote the Barenblatt solution to the one-dimensional equation (PME) with $u|_{t=0} = \delta_{x_0}$ as in (BB).*

Let $T \in (0,\infty)$. Suppose $(X,W),(Y,W)$ are two $P_{\delta_{x_0}}^{(u_t)}$-weak solutions to (MVSDE.PME) up to time T on a common stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}; (\mathcal{F}_t)_{t \in [0,T]})$ with respect to the same standard d-dimensional $(\mathcal{F}_t)$-Brownian motion $(W(t))_{t \in [0,T]}$ such that $X(0) = Y(0)$ $\mathbb{P}$-a.s. Then, $\sup_{t \in [0,T]} |X(t) - Y(t)| = 0$ $\mathbb{P}$-a.s.

Proof. Let $Z(t) := X(t) - Y(t), t \in [0, T]$, and let $\tau_R := \inf\{t \in [0, T] : |X(t)| \vee |Y(t)| \geq R\} \wedge T$, for $R > 0$. Let $t \in [0, T]$. Fix an arbitrary $\delta > 0$. Let $f_\delta(r) := |r| \ln\left(\frac{|r|^2}{\delta^2} + 1\right), r \in \mathbb{R}$. Then $f_\delta \in C^2(\mathbb{R})$ with $|f'_\delta(r)| \leq C \ln\left(1 + \frac{|r|^2}{\delta^2}\right)$ and $|f''_\delta(r)| \leq \frac{C}{|r|+\delta}$ for all $r \in \mathbb{R}$ and a constant $C \in (0, \infty)$ independent of δ and r. By Itô's formula and Fatou's lemma, we estimate

$$\mathbb{E}|Z(t \wedge \tau_R)| \ln\left(\frac{|Z(t \wedge \tau_R)|^2}{\delta^2} + 1\right)$$

$$\leq C\mathbb{E} \int_0^{t \wedge \tau_R} \frac{\left|u_s^{\frac{m-1}{2}}(X(s)) - u_s^{\frac{m-1}{2}}(Y(s))\right|^2}{|Z(s)| + \delta} \mathrm{d}s \leq C \liminf_{\eta \to 0} \mathbb{E} \int_0^{t \wedge \tau_R} \frac{\left|\sqrt{\frac{u_s^m(x)}{u_s(x)+\eta}} - \sqrt{\frac{u_s^m(y)}{u_s(y)+\eta}}\right|^2}{|Z(s)| + \delta} \mathrm{d}s \quad (61)$$

A similar argument as in the proof of Theorem 3.9 yields that for all $(s, x) \in [0, T] \times \mathbb{R}$

$$\left|\sqrt{\frac{u_s^m(X(s))}{u_s(X(s)) + \eta}} - \sqrt{\frac{u_s^m(Y(s))}{u_s(Y(s)) + \eta}}\right| \leq (h_\eta(s, x) \wedge h_\eta(s, y)) \left|u_s^{\frac{m}{2}}(x) - u_s^{\frac{m}{2}}(y)\right|, \quad (62)$$

where we define $h_\eta(s, x) := (1 + \frac{1}{m}) \int_0^1 (\theta^{\frac{2}{m}} u_s(x) + \eta)^{-\frac{1}{2}} \mathrm{d}\theta$. Note that $h_\eta^2(s, x) u_s(x) \leq C$ for all $(s, x) \in [0, T] \times \mathbb{R}$ for some constant $C \in (0, \infty)$ independent of s, x, η. We recall that by [CJ18, Lemma 3.2] and Proposition 4.1, for all $t > 0$

$$|u_t^{\frac{m}{2}}(x) - u_t^{\frac{m}{2}}(y)| \leq \left(\mathrm{M}|\partial_x^{\frac{1}{2}} u_t^{\frac{m}{2}}|(x) + \mathrm{M}|\partial_x^{\frac{1}{2}} u_t^{\frac{m}{2}}|(y)\right) |x - y|^{\frac{1}{2}} \quad (63)$$

for a.e. $x, y \in \mathbb{R}$. Using (62) and (63) and that $u^m \in L^1((0, T) \times \mathbb{R}^d)$, we estimate the right hand-side of (61) via a convolution argument similar to (50)-(51) in the proof of Theorem 3.9 and obtain that

$$\liminf_{\eta \to 0} \mathbb{E} \int_0^{t \wedge \tau_R} \frac{\left|\sqrt{\frac{u_s^m(x)}{u_s(x)+\eta}} - \sqrt{\frac{u_s^m(y)}{u_s(y)+\eta}}\right|^2}{|Z(s)| + \delta} \mathrm{d}s \leq C||\partial_x^{\frac{1}{2}} u||_{L^2((0,T) \times \mathbb{R})} < \infty.$$

From the above, we may therefore conclude that

$$\sup_{\delta > 0} \mathbb{E}|Z(t \wedge \tau_R)| \ln\left(\frac{|Z(t \wedge \tau_R)|^2}{\delta^2} + 1\right) = \mathbb{E}|Z(t \wedge \tau_R)| \ln\left(\sup_{\delta > 0} \frac{|Z(t \wedge \tau_R)|^2}{\delta^2} + 1\right) < \infty.$$

Necessarily, we therefore have for all $t \in [0, T]$

$$X(t \wedge \tau_R) = Y(t \wedge \tau_R) \quad \mathbb{P}\text{-a.s.} \quad (64)$$

Note that $\sup_{t \in [0,T]} |X(t)| + |Y(t)| < \infty$ $\mathbb{P}$-a.s. Hence, $\tau_R \to T$ $\mathbb{P}$-a.s., as $R \to \infty$. Therefore, taking the limit $R \to \infty$ in (64) and recalling that $X, Y \in C([0, T]; \mathbb{R})$, we obtain that $\mathbb{P}$-a.s.

$$X(t) = Y(t) \quad \forall t \in [0, T].$$

This finishes the proof. $\square$

Remark 4.4. *Note that in the proof of Theorem 4.3 the fundamental estimates on the first- and second-order derivatives of f_δ only work in the one-dimensional setting, see [CJ18, Proof of Theorem 2.15., p.1533].*

Appendix

A   Lipschitz-type estimates

We recall the definition and properties of (local) Hardy–Littlewood maximal functions on the set of Radon measures on $\mathbb{R}^d$, where the latter are denoted as $M_{loc}(\mathbb{R}^d; \mathbb{R}^d)$ in the following. Furthermore, if $\mu \in M_{loc}(\mathbb{R}^d; \mathbb{R}^d)$,

then $|\mu|$ denotes its variation measure.

The upcoming definition and the two lemmata presented afterwards are essentially taken from [CD08].

Definition A.1 ([CD08, Definition A.1]). *Let $\mu \in M_{loc}(\mathbb{R}^d; \mathbb{R}^n)$ and $R \in (0, \infty]$. We define the (local) maximal function as*

$$\mathrm{M}_R\,|\mu|(x) := \sup_{0 < r < R} \frac{1}{|B_R(0)|} \int_{B_r(x)} |\mu|(\mathrm{d}x), x \in \mathbb{R}^d.$$

In the case $R = \infty$ we set $\mathrm{M} := \mathrm{M}_\infty$. Furthermore, if μ is of the form $\mu(dx) = f(x)\mathrm{d}x$, where $f \in L^1_{loc}(\mathbb{R}^d; \mathbb{R}^n)$, then we write $\mathrm{M}_R\,|f| := \mathrm{M}_R\,|\mu|$.

Lemma A.2 ([CD08, Lemma A.3]). *Let $R \in (0, \infty]$ and $f \in W^{1,1}_{loc}(\mathbb{R}^d; \mathbb{R}^n)$. Then there exists a constant $C_d > 0$ depending only on the dimension d, and a set $N \in \mathcal{B}(\mathbb{R}^d)$ with $\lambda^d(N) = 0$ such that for all $x, y \in N^{\complement}$ with $|x - y| \le R$*

$$|f(x) - f(y)| \le C_d\,(\mathrm{M}_R\,|Df|(x) + \mathrm{M}_R\,|Df|(y))\,|x - y|, \tag{65}$$

where $Df = (\partial_{x_j} f^i)_{1 \le i \le d, 1 \le j \le n}$ denotes the matrix of all weak derivatives of f.

Lemma A.3 ([CD08, Lemma A.2]). *Let $\mu \in \mathcal{M}_{loc}(\mathbb{R}^d; \mathbb{R}^{m \times n})$ and $R \in (0, \infty]$. Then for a.e. $x \in \mathbb{R}^d$, $M_R|\mu|(x) < \infty$.*

Let $p \in (1, \infty)$. Then there exists a constant $C_{d,p} > 0$ such that for all $f \in L^p_{loc}(\mathbb{R}^d; \mathbb{R}^n)$ and all $\rho > 0$

$$\int_{B_\rho(0)} (\mathrm{M}_R\,|f|(x))^p \mathrm{d}x \le C_{d,p} \int_{B_{\rho+R}(0)} |f(x)|^p \mathrm{d}x.$$

For $p = 1$, the previous statement does not hold. However, for $p = 1$ one has the following weak estimate: There exists a constant $C_d > 0$ such that for all $f \in L^1_{loc}(\mathbb{R}^d; \mathbb{R}^n)$

$$\lambda^d\left(\{x \in B_\rho(0) : \mathrm{M}_R\,|f|(x) > \alpha\}\right) \le \frac{C_d}{\alpha} \int_{B_{\rho+R}(0)} |f(x)|\mathrm{d}x.$$

B A concise recapitulation on nonlinear evolution equations in Banach spaces and the Crandall–Ligget theorem

This section's content is borrowed from [Bar10, Chapter 4] and presented in a concise form. For the details, we refer to [Bar10].

The problem. Let $(X, ||\cdot||)$ be a Banach space and let $A : D(A) \subset X \to X$ be a (possibly multi-valued) operator. Under which conditions on $\nu \in X$ may one find a (unique) solution to the evolution equation

$$\partial_t u + A(u) = 0, \quad u|_{t=0} = \nu, \tag{EE}$$

and in what sense?

In the following I will denote the identity map on X.

Definition B.1 ((m-)accretivity). *$(A, D(A))$ is called accretive, if for all $i = 1, 2$, $(x_i, y_i) \in \mathrm{Graph}(A)$,*

$$\|x_1 - x_2\| \le \|x_1 - x_2 + \lambda(y_1 - y_2)\| \quad \forall \lambda > 0. \tag{66}$$

If $(A, D(A))$ is accretive and $(I + \lambda A)(D(A)) = X$, then $(A, D(A))$ is called m-accretive.

Remark B.2. *Due to (66), the operator $(I + \lambda A)^{-1}$ is single-valued for all $\lambda > 0$.*

From now on assume that $(A, D(A))$ is m-accretive. The following definition provides a concept of solution to (EE) based on a finite difference scheme.

Definition B.3. *A continuous function $u : [0, \infty) \to X$ is called a mild solution to (EE), if for each $T > 0$*

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

where $u_h : [0, T] \to X$ is a step function, defined by the finite difference scheme

$$u_h(t) = u_h^{i+1} \ \forall t \in (ih, (i+1)h], i = 0, 1, ..., N-1, u_h(0) = \nu,$$
$$\text{where } u_h^{i+1} + hA(u_h^{i+1}) = u_h^i, \ Nh = T,$$
$$\text{and } u_h^0 = \nu.$$

The following theorem is a special case of the *Crandall–Ligget theorem.*

Theorem B.4 (cf. [Bar10, Theorem 4.3])**.** *Let $(A, D(A))$ be m-accretive. Then, for all $\nu \in \overline{D(A)}$ (closure of $D(A)$ in X) there exits a unique mild solution to (EE).*

References

[Bar10]   Barbu, V. *Nonlinear differential equations of monotone types in Banach spaces.* Springer Monographs in Mathematics. Springer, New York, 2010, pp. x+272.

[Ben+96] Benachour, S., Chassaing, P., Roynette, B., and Vallois, P. "Processus associés à l'équation des milieux poreux". *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze* 4e série, 23.4 (1996), pp. 793–832.

[BJ19]   Bossy, M. and Jabir, J.-F. "On the wellposedness of some McKean models with moderated or singular diffusion coefficient". *Frontiers in stochastic analysis—BSDEs, SPDEs and their applications.* Vol. 289. Springer Proc. Math. Stat. Springer, Cham, 2019, pp. 43–87.

[BM18]   Brezis, H. and Mironescu, P. "Gagliardo–Nirenberg inequalities and non-inequalities: the full story". *Ann. Inst. H. Poincaré C Anal. Non Linéaire* 35.5 (2018), pp. 1355–1376.

[BR18]   Barbu, V. and Röckner, M. "Probabilistic representation for solutions to nonlinear Fokker–Planck equations". *SIAM J. Math. Anal.* 50.4 (2018), pp. 4246–4260.

[BR20]   Barbu, V. and Röckner, M. "From nonlinear Fokker–Planck equations to solutions of distribution dependent SDE". *Ann. Probab.* 48.4 (2020), pp. 1902–1920.

[BR21a]   Barbu, V. and Röckner, M. "Solutions for nonlinear Fokker–Planck equations with measures as initial data and McKean–Vlasov equations". *J. Funct. Anal.* 280.7 (2021), p. 108926. arXiv:2005.02311v5.

[BR21b]   Barbu, V. and Röckner, M. "Uniqueness for nonlinear Fokker–Planck equations and weak uniqueness for McKean–Vlasov SDEs". *Stoch. Partial Differ. Equ. Anal. Comput.* 9.3 (2021), pp. 702–713.

[BR23a]   Barbu, V. and Röckner, M. "Nonlinear Fokker–Planck equations with time-dependent coefficients". *SIAM J. Math. Anal.* 55.1 (2023), pp. 1–18.

[BR23b]   Barbu, V. and Röckner, M. "The evolution to equilibrium of solutions to nonlinear Fokker–Planck equation". *Indiana Univ. Math. J.* 72.1 (2023), pp. 89–131.

[BR23c]   Barbu, V. and Röckner, M. "Uniqueness for nonlinear Fokker–Planck equations and for McKean–Vlasov SDEs: the degenerate case". *J. Funct. Anal.* 285.4 (2023), Paper No. 109980, 37.

[BRS21] Bogachev, V. I., Röckner, M., and Shaposhnikov, S. V. "On the Ambrosio–Figalli–Trevisan Superposition Principle for Probability Solutions to Fokker–Planck–Kolmogorov Equations". *J. Dynam. Differential Equations* 33.2 (2021), pp. 715–739.

[CD08] Crippa, G. and De Lellis, C. "Estimates and regularity results for the DiPerna–Lions flow". *J. Reine Angew. Math.* 616 (2008), pp. 15–46.

[CD18] Carmona, R. and Delarue, F. *Probabilistic theory of mean field games with applications. I&II.* Probability Theory and Stochastic Modelling. Springer, Cham, 2018.

[CJ18] Champagnat, N. and Jabin, P.-E. "Strong solutions to stochastic differential equations with rough coefficients". *Ann. Probab.* 46.3 (2018), pp. 1498–1541.

[FG23] Fehrman, B. and Gess, B. "Non-equilibrium large deviations and parabolic-hyperbolic PDE with irregular drift". *Invent. Math.* 234.2 (2023), pp. 573–636.

[Fra05] Frank, T. D. *Nonlinear Fokker–Planck equations.* Springer Series in Synergetics. Fundamentals and applications. Springer-Verlag, Berlin, 2005, pp. xii+404.

[Gru23a] Grube, S. *Strong solutions to McKean–Vlasov SDEs with coefficients of Nemytskii-type.* PhD thesis. Bielefeld University, 2023.

[Gru23b] Grube, S. "Strong solutions to McKean–Vlasov SDEs with coefficients of Nemytskii-type". *Electron. Commun. Probab.* 28 (2023), Paper No. 11, 13.

[Gru24] Grube, S. "Strong solutions to McKean–Vlasov SDEs with coefficients of Nemytskii type: the time-dependent case". *J. Evol. Equ.* 24.2 (2024), p. 37.

[GST20] Gess, B., Sauer, J., and Tadmor, E. "Optimal regularity in time and space for the porous medium equation". *Anal. PDE* 13.8 (2020), pp. 2441–2480.

[JM98] Jourdain, B. and Méléard, S. "Propagation of chaos and fluctuations for a moderate model with smooth initial data". *Ann. Inst. H. Poincaré Probab. Statist.* 34.6 (1998), pp. 727–766.

[McK66] McKean Jr., H. P. "A class of Markov processes associated with nonlinear parabolic equations". *PNAS* 56.6 (1966), pp. 1907–1911.

[Reh22] Rehmeier, M. "Flow selections for (nonlinear) Fokker–Planck–Kolmogorov equations". *J. Differential Equations* 328 (2022), pp. 105–132.

[RR23] Rehmeier, M. and Röckner, M. "On nonlinear Markov processes in the sense of McKean". 2023. arXiv:2212.12424v2.

[RXZ20] Röckner, M., Xie, L., and Zhang, X. "Superposition principle for non-local Fokker–Planck–Kolmogorov operators". *Probab. Theory Related Fields* 178.3-4 (2020), pp. 699–733.

[RZ10] Röckner, M. and Zhang, X. "Weak uniqueness of Fokker–Planck equations with degenerate and bounded coefficients". *C. R. Math. Acad. Sci. Paris* 348.7-8 (2010), pp. 435–438.

[Tre16] Trevisan, D. "Well-posedness of multidimensional diffusion processes with weakly differentiable coefficients". *Electron. J. Probab.* 21 (2016), Paper No. 22, 41.

[Váz07] Vázquez, J. L. *The porous medium equation.* Oxford Mathematical Monographs. Mathematical theory. The Clarendon Press, Oxford University Press, Oxford, 2007, pp. xxii+624.