

The Leibenson process

Viorel Barbu* Sebastian Grube† Marco Rehmeier‡ Michael Röckner §¶

August 18, 2025

Abstract

Consider the Leibenson equation

$$\partial_t u = \Delta_p u^q,$$

where $\Delta_p f = \operatorname{div}(|\nabla f|^{p-2} \nabla f)$ for $p > 1$ and $q > 0$, which is a simultaneous generalization of the porous media and the p -Laplace equation. In this paper we identify the Leibenson equation as a nonlinear Fokker–Planck equation and prove that it has a nonlinear Markov process in the sense of McKean as its probabilistic counterpart. More precisely, we obtain a probabilistic representation of its Barenblatt solutions as the one-dimensional marginal density curve of the unique solutions to the associated McKean–Vlasov SDE. The latter is of novel type, since its coefficients depend pointwise both on its solution’s time marginal densities and also on their first and second order derivatives. Moreover, we show that these solutions constitute the aforementioned nonlinear Markov process, which we call the *Leibenson process*. A further main result of this work is to prove that despite the strong degeneracy of the diffusion and the irregularity of the drift coefficient (which is merely of bounded variation) of the McKean–Vlasov SDE these solutions are probabilistically strong, i.e., measurable functionals of the driving Brownian motion and the initial condition.

Keywords: Leibenson equation; Barenblatt solution; nonlinear Fokker–Planck equation; Markov process; McKean–Vlasov stochastic differential equation; probabilistically strong solution

2020 MSC: Primary: 35K55, 35Q84, 60J25. Secondary: 35C06, 35K65, 60H30.

Competing Interests. The authors declare that they have no financial or non-financial interests to disclose.

Acknowledgements. The second and fourth named authors are funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project-ID 317210226 - SFB 1283. We would like to thank Alexander Grigor’yan for valuable discussions on his work on the Leibenson equation on Riemannian manifolds.

1 Introduction

This paper should be seen as a substantial step forward in a general program which, motivated by McKean’s vision laid out in [49], we have started (see [3, 4, 5, 6, 7, 8, 9, 10, 11, 53]), namely to develop a new theory whose aim is to construct a probabilistic counterpart for a large class of nonlinear parabolic PDEs in the same way as has been done in the last 60 years for linear parabolic PDEs (see

*Al.I. Cuza University and Octav Mayer Institute of Mathematics of Romanian Academy, Iași, Romania. E-mail: vb41@uaic.ro

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

‡Institute of Mathematics, TU Berlin, 10623 Berlin, Germany. E-mail: rehmeier@tu-berlin.de

§Faculty of Mathematics, Bielefeld University, 33615 Bielefeld, Germany. E-mail: roeckner@math.uni-bielefeld.de

¶Academy of Mathematics and System Sciences, CAS, Beijing

||corresponding author

[14, 15, 22, 23, 24, 28, 29, 39, 40, 46, 50, 55, 56, 58, 59] and the references therein). The purpose is to then transform problems about the PDE to the probabilistic counterpart and vice versa. This probabilistic counterpart is as in the linear case a Markov process, but a nonlinear one in the sense of McKean and in the (so far) most interesting cases they are given by the path laws of solutions to very singular McKean–Vlasov SDEs. The mentioned transfer between analysis and probability was extremely successful in the linear case with a huge literature (see above) until today. We are convinced that there are very good chances that this also will happen in the nonlinear case.

In this paper, we realize the above program in the case where the nonlinear PDE is the Leibenson equation ([45, 44])

$$\partial_t u(t, x) = \Delta_p u(t, x)^q, \quad (t, x) \in (0, \infty) \times \mathbb{R}^d, \quad (1.1)$$

which is well known as the prototype of a degenerate doubly nonlinear parabolic partial differential equation, with its Barenblatt fundamental solutions (see [12, 13]).

Here $q > 0$ and $\Delta_p f = \operatorname{div}(|\nabla f|^{p-2} \nabla f)$ denotes the p -Laplace operator, $p > 1$. The aforementioned probabilistic counterpart is a nonlinear Markov process in the sense of McKean (see [49, 52]) consisting of the unique solutions to the corresponding McKean–Vlasov stochastic differential equation (McKean–Vlasov SDE)

$$\begin{cases} dX(t) &= q^{p-1} \nabla(|\nabla u(t, X(t))|^{p-2} u(t, X(t))^{(q-1)(p-1)}) dt \\ &\quad + \sqrt{2q^{p-1}} (|\nabla u(t, X(t))|^{\frac{p-2}{2}} u(t, X(t))^{\frac{(q-1)(p-1)}{2}}) dW(t), \\ \mathcal{L}_{X(t)}(dx) &= u(t, x) dx, \end{cases} \quad (1.2)$$

with one-dimensional time marginal densities $\frac{d\mathcal{L}_{X(t)}}{dx}$, $t \geq 0$, equal to the Barenblatt solutions to (1.1) (see (4.1) below). Here and throughout, dx denotes Lebesgue measure, W is a standard d -dimensional Brownian motion and \mathcal{L}_X denotes the distribution of a random variable X . We call this nonlinear Markov process the *Leibenson process*, see Definition 5.9 and Remark 5.8. We significantly extend the techniques and results from the previous paper [3] of the first, third and last named author in which the probabilistic counterpart of the slow diffusion p -Laplace special case ($q = 1, p > 2$) was constructed and was called *p-Brownian motion*.

Before we give details on (1.1), (1.2) and our results in Part 1 and 2 below, we summarize the main achievements of the present paper. First, in Part 1, we identify the Leibenson equation (1.1) as a nonlinear Fokker–Planck equation (see (1.7)). This leads to the corresponding singular McKean–Vlasov SDE (1.2) ($p - 2$ or $(p - 1)(q - 1)$ may be strictly negative), for which we then establish weak well-posedness for prescribed one-dimensional solution time marginals. Then we show that the path measures of the solution processes constitute a nonlinear Markov process. This substantially generalizes the approach from [3] to the doubly nonlinear Leibenson equation (1.1) for a large range of pairs $(p, q) \in (1, \infty) \times (0, \infty)$, including (unlike [3]) the fast diffusion case $p < 2$ (with $q > 1$). Furthermore, the corresponding McKean–Vlasov SDE (1.2) is the first of its kind in the literature, since its coefficients depend both on the solutions' time marginal densities as well as on their first and second order derivatives. Second, in Part 2, we finally solve the problem left open in [3] about p -Brownian motion even for the more general and doubly nonlinear Leibenson equation by proving that the Leibenson process is *probabilistically strong*, i.e. the solutions to (1.2) with Barenblatt one-dimensional time marginals are adapted measurable functionals of the driving Brownian motion. This is a quite surprising result in view of the strong degeneracy of the diffusion coefficient in (1.2) (which, as shown below, is compactly supported in space) and the irregular drift coefficient, which is merely BV (and not Sobolev regular). Furthermore, in addition to the analytic proof for the crucial restricted uniqueness result for the linearized Leibenson equation (see Theorem 7.4 below), we provide a genuinely new probabilistic proof for this result (see Subsection 7.2).

Let us now describe our goals and results in more detail. Probabilistic representations of solutions to (nonlinear) partial differential equations (PDEs) as time marginal densities of stochastic processes, more precisely of canonically associated (nonlinear) Markov processes, allow to transfer analytic

questions on well-posedness, asymptotic behavior and quantitative estimates for PDEs to probabilistic ones in stochastic analysis and vice versa (see [7] for an example in the nonlinear case). In the linear case the simplest, yet very instructive example is the relation between Brownian motion and the heat equation

$$\partial_t u(t, x) = \Delta u(t, x), \quad (t, x) \in (0, \infty) \times \mathbb{R}^d.$$

Its fundamental solution, the classical heat kernel $p(t, x, y)$, consists of the one-dimensional time marginal law densities of Brownian motion started at $y \in \mathbb{R}^d$ in the sense that

$$\mathcal{L}_{W^y(t)}(dx) = p(t, x, y)dx, \quad \forall t > 0, y \in \mathbb{R}^d,$$

where we set $W^y(t) := \sqrt{2}W(t) + y$ (the factor $\sqrt{2}$ may be exchanged for a factor $\frac{1}{2}$ on the RHS of the heat equation). Of course, W^y is the unique probabilistically strong solution to the SDE

$$dX(t) = \sqrt{2}dW(t), \quad X(0) = y,$$

and the family $\{P_y\}_{y \in \mathbb{R}^d}$ of path laws of $\{W^y\}_{y \in \mathbb{R}^d}$ constitutes a uniquely determined (linear) Markov process with transition kernels $y \mapsto p(t, x, y)dx$, $t > 0$.

The Leibenson equation. In this paper we develop comparable relations for the Leibenson equation (1.1) and the McKean–Vlasov SDE (1.2). (1.1) is a simultaneous generalization of the porous medium ($p = 2$) and the p -Laplace equation ($q = 1$). In hydrodynamics, (1.1) models filtration of a turbulent compressible fluid in a porous medium. For more physical background, we refer, for instance, to the introduction of [30]. In the linear case $(p, q) = (2, 1)$, (1.1) is the heat equation. To our knowledge, the Leibenson equation was first introduced in [45, 44]. In the past decades the theory of existence, regularity, and qualitative and quantitative properties of its solutions was developed, e.g., in [37, 38, 61, 30, 31, 63] and in further references mentioned therein. In [30] it is stated that particularly interesting physical cases arise for $q \geq 1$ and $\frac{3}{2} \leq p \leq 2$. We give more details on the admissible values for (p, q) in our results below, but mention already here that we cover large parts of this "physical range". We focus on the explicit *Barenblatt solutions* with point source initial datum in the case $q(p-1) > 1$ (see Definition 4.1), which were first obtained in [12] and can be considered as the analogue of the heat kernel for the heat equation (even though the Barenblatt solutions are not fundamental solutions in the usual sense since, due to the nonlinearity of (1.1), one cannot build from them solutions for non-Dirac initial data by simply taking convex combination). The other cases $q(p-1) = 1$ and $q(p-1) < 1$ give rise to completely different regimes: For $q(p-1) > 1$ the Barenblatt solutions have finite speed of propagation, whereas in the other two regimes they have infinite speed of propagation, see Section 4. We postpone the study of these other regimes and of more general solutions to future work.

Goals and results. We split our program into two parts. The first part is the construction of a uniquely determined nonlinear Markov process as the probabilistic counterpart of the Barenblatt solutions to (1.1), consisting of solutions to (1.2), see Sections 2-5 and 7. The second part consists in proving that these solutions are probabilistically strong, see Section 6.

Part 1: Construction of the Leibenson process. The derivation of a McKean–Vlasov SDE associated with a nonlinear PDE follows by identifying the latter as a nonlinear Fokker–Planck equation (FPE). These are parabolic equations for measures of type (here and throughout we use Einstein summation convention)

$$\partial_t \mu_t = \partial_{ij}(a_{ij}(t, x, \mu_t)\mu_t) - \partial_i(b_i(t, x, \mu_t)\mu_t), \quad (t, x) \in (s, \infty) \times \mathbb{R}^d, \quad (1.3)$$

where $s \geq 0$ is the initial time, $a_{ij}, b_i : (0, \infty) \times \mathbb{R}^d \times \mathcal{M} \rightarrow \mathbb{R}$ are measurable coefficients, \mathcal{M} is some subset of the set of signed locally finite Borel measures on \mathbb{R}^d , and solutions are curves $[s, \infty) \ni t \mapsto \mu_t \in \mathcal{M}$ solving (1.3) in the (Schwartz) distributional sense (see Definition C.2 below).

$a = (a_{ij})_{i,j \leq d}$ is assumed to take values in the space of nonnegative definite $d \times d$ -matrices, thus there exists a (not necessarily unique) measurable square root $\sigma = (\sigma_{ij})_{i,j \leq d} : (0, \infty) \times \mathbb{R}^d \times \mathcal{M} \rightarrow \mathbb{R}^{d \times d}$ with $a = \sigma \sigma^T$. If one takes \mathcal{M} to be a set \mathcal{P}_0 of probability measures on \mathbb{R}^d , it is well-known that for any weakly continuous probability measure-valued solution $t \mapsto \mu_t \in \mathcal{P}_0$ to (1.3) satisfying an additional rather mild integrability condition (see Definition C.2), there exists a probabilistically weak solution $X = (X(t))_{t \geq s}$ to the McKean–Vlasov SDE,

$$dX(t) = b(t, X(t), \mathcal{L}_{X(t)})dt + \sqrt{2} \sigma(t, X(t), \mathcal{L}_{X(t)})dW(t), \quad t > s, \quad (1.4)$$

such that $\mathcal{L}_{X(t)} = \mu_t$ for all $t \geq s$, see [1, 26, 64, 4, 5, 17, 59]. This result, called *superposition principle*, provides a probabilistic counterpart to solutions of (1.3) by identifying the latter as the one-dimensional time marginals of the paths of stochastic processes solving (1.4). For this, *no* regularity assumption on the coefficients in $x \in \mathbb{R}^d$ or $\mu \in \mathcal{P}_0$ is needed.

A particularly interesting type of (1.3) is given by the *Nemytskii-case*, where \mathcal{P}_0 is a set of absolutely continuous measures (w.r.t. Lebesgue measure dx) and the dependence of a_{ij} and b_i on $\mu \in \mathcal{P}_0$ is pointwise via its density, i.e.

$$a_{ij}(t, x, \mu) = \bar{a}_{ij}(t, x, u(x)), \quad b_i(t, x, \mu) = \bar{b}_i(t, x, u(x)),$$

for Borel coefficients $\bar{a}_{ij}, \bar{b}_i : (0, \infty) \times \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}$ and where $\mu(dx) = u(x)dx$ (see [10] for a large number of concrete examples). Rewriting (1.3) as an equation for the density $u(t)$ of μ_t , one arrives at a nonlinear PDE

$$\partial_t u(t, x) = \partial_{ij}(\bar{a}_{ij}(t, x, u(t, x))u(t, x)) - \partial_i(\bar{b}_i(t, x, u(t, x))u(t, x)). \quad (1.5)$$

In this case the corresponding McKean–Vlasov SDE is

$$\begin{cases} dX(t) &= \bar{b}(t, X(t), u(t, X(t)))dt + \sqrt{2} \bar{\sigma}(t, X(t), u(t, X(t)))dW(t), \\ \mathcal{L}_{X(t)}(dx) &= u(t, x)dx, \end{cases} \quad (1.6)$$

where $\bar{\sigma} = (\bar{\sigma}_{ij})_{i,j \leq d} : (0, \infty) \times \mathbb{R}^d \times \mathcal{P}_0 \rightarrow \mathbb{R}^{d \times d}$ is a measurable function with $\bar{a} = \bar{\sigma} \bar{\sigma}^T$. Thus, for every weakly continuous probability measure-valued solution u to (1.5) there is a stochastic process solving (1.6). For instance, solutions to the generalized porous medium equation (PME) perturbed by a nonlinear drift

$$\partial_t u(t, x) = \Delta \beta(u(t, x)) - \operatorname{div}(D(x)b(u(t, x))),$$

where $D : \mathbb{R}^d \rightarrow \mathbb{R}^d$, $b : \mathbb{R} \rightarrow \mathbb{R}_+$ and $\beta : \mathbb{R} \rightarrow \mathbb{R}$, are of this type with $\bar{a}_{ij}(t, x, u(x)) = \delta_{ij}\beta(u(x))$ and $\bar{b}_i(t, x, u(x)) = D(x)b(u(x))$, see [6, 8, 10] and the references therein. The classical PME is retrieved with $D = 0$, $\beta(r) = r^m$, $m > 0$ (cf. [65]).

Since the Leibenson equation at first sight is not of type (1.5), the identification of its McKean–Vlasov SDE is not straightforward. Our approach is as follows. First, we show that (1.1) can be rewritten as the nonlinear FPE

$$\partial_t u(t, x) = q^{p-1} \Delta(|\nabla u(t, x)|^{p-2} u(t, x)^{(p-1)(q-1)} u(t, x)) - q^{p-1} \operatorname{div}(\nabla(|\nabla u(t, x)|^{p-2} u(t, x)^{(p-1)(q-1)} u(t, x))). \quad (1.7)$$

This is indeed of type (1.3) for coefficients

$$a_{ij}(t, x, \mu_t) = q^{p-1} \delta_{ij} |\nabla u(t, x)|^{p-2} u(t, x)^{(p-1)(q-1)}, \quad b_i(t, x, \mu_t) = q^{p-1} \partial_i(|\nabla u(t, x)|^{p-2} |u(t, x)|^{(p-1)(q-1)}), \quad (1.8)$$

where $\mu_t(dx) = u(t, x)dx$. Details, in particular a proof of the equivalence of (1.1) and this FPE, are given in Section 2. The coefficients a and b from (1.8) are not of Nemytskii-type, which would require $a_{ij}(t, x, \cdot)$ to depend on $\mu_t(dx) = u(t, x)dx$ only via $u(t, x)$ (and similarly for b_i). Instead, due to the dependence on the gradient of $u(t)$, $a_{ij}(t, x, \cdot)$ depends on μ_t via the values of $u(t)$ in an arbitrary small open neighborhood of x (similarly for b_i). Then the associated McKean–Vlasov SDE is (1.2).

Our first result is the following construction of the probabilistic counterpart of solutions to the Leibenson equation.

Theorem 1 (see Thm. 3.3, Cor. 3.4, Thm. 4.4 for details). Let $d \geq 1$, $p > 1$, $q > 0$. For every solution u to (1.1) consisting of probability densities with suitable regularity and integrability assumptions there is a probabilistically weak solution to (1.2). If additionally $q > \frac{1}{p-1}$ and $p > \frac{1+d}{d}$, this result applies to the Barenblatt solutions $u = w^y$, $y \in \mathbb{R}^d$, (see Definition 4.1) to (1.1).

In the linear case, i.e. when a_{ij} and b_i in (1.3) do not depend on μ , it is well-known that if for every initial datum (s, δ_y) the FPE (1.3) has a unique weakly continuous probability solution $(\mu_t^{s,y})_{t \geq s}$, then the corresponding SDE (1.4) has a unique probabilistically weak solution from any initial datum (s, δ_y) as well, and the family of these solutions forms a Markov process (for the definition of a Markov process see Section 5.2) with transition kernels $(y, A) \mapsto \mu_t^{s,y}(A)$. This is *not* true for nonlinear FPEs and their associated McKean–Vlasov SDEs, since the proof relies on the stability of the associated martingale problem under convex combinations, which is lost in the nonlinear case. To overcome this issue, inspired by McKean [49], in [52] the third and last-named author introduced *nonlinear Markov processes*, and showed that under general assumptions a solution family to a nonlinear FPE gives rise to a uniquely determined nonlinear Markov process, consisting of solution path laws to the associated McKean–Vlasov SDE (see Section 5.2 for details). We stress that this result applies to ill-posed situations, i.e. uniqueness of the nonlinear FPE-solutions is not required, but only a “restricted linearized uniqueness” result, see Theorem 5.2 and Lemma 5.4. In previous work (cf. e.g. [3, 7, 11, 52, 53] and also [10] and the references therein) this was utilized to construct nonlinear Markov processes with one-dimensional time marginals given by solutions to nonlinear PDEs, as e.g. Burgers’ and (generalized) PME’s, as well as 2D vorticity Euler and Navier–Stokes equations. In each of these cases, it was first shown that the PDE under consideration can be written as a Neymetskii-type FPE.

In the present paper we construct a nonlinear Markov process associated with the Barenblatt solutions to (1.1) via [52, Theorem 3.8]. In particular, we have to prove a delicate restricted linearized uniqueness result, which boils down to a highly nontrivial uniqueness result for the degenerate second-order linear PDE (7.1), see Theorem 7.4. The conditions on p, q in Theorem 2 below are essentially determined by the conditions in Theorem 7.4, which split into two different regimes, for which we prove the uniqueness result separately. One part, proven via analytic methods, is a close adaptation of the proof for the corresponding result in the p -Laplace special case. The second part is based on genuinely new purely probabilistic methods. Our result is:

Theorem 2 (see Thm. 5.6 and Def. 5.9 for details). Let $d \geq 2$, $p > \frac{d}{d-1}$, $q > \frac{1}{p-1}$. If $p < 2$ assume additionally $q > \frac{2-p+d}{d(p-1)} (> \frac{1}{p-1})$. The family of McKean–Vlasov SDE-solutions from Theorem 1 with one-dimensional time marginals equal to the Barenblatt solutions to (1.1) constitutes a uniquely determined nonlinear Markov process, which we call *Leibenson process*.

The construction of the Leibenson process is analogous to the construction of Brownian motion as a Markov process from the classical heat kernel. The slow diffusion p -Laplace special case of Theorems 1 and 2, i.e. $p > 2$ and $q = 1$, was proven in [3], where the corresponding nonlinear Markov process was called *p-Brownian motion* (for $p = 2$, one recovers Brownian motion).

Part 2: The Leibenson process is probabilistically strong. By construction, the Leibenson process from Theorem 2 consists of probabilistically weak solutions to (1.2). It has been an imminent and challenging open question whether these solutions are actually probabilistically strong, i.e. measurable adapted functionals of the driving Brownian motion and the initial condition. For details on the definition of probabilistically strong solutions, see Definition 6.1. In this paper, we answer this question affirmatively via the following result. Note that this includes as a new result the p -Brownian motion

case ($p > 2, q = 1$), for which the question of being a strong solution to its corresponding McKean–Vlasov SDE was left open in [3].

Theorem 3 (see Thm. 6.3 for details). Let $d \geq 2, p > \frac{d-1}{d}, q > \frac{|p-2|+d}{d(p-1)} (> \frac{1}{p-1})$. Then, the probabilistically weak solution (X, W) to (1.2) with $u = w^y$ from Theorem 1 are probabilistically strong from any strictly positive time $\delta > 0$ on (i.e. they are measurable adapted functionals of the driving Brownian motion W and $X(\delta)$).

The proof follows from a *restricted Yamada–Watanabe theorem* [33, Theorem 1.3.1], [32], which in turn relies on the existence of a probabilistically weak solution with one-dimensional time marginal law densities $w^y(t + \delta), t \geq 0$, (as provided by Theorem 1) and a restricted pathwise uniqueness result for (1.2), see Theorem 6.13. Here, *restricted* means uniqueness in the class of solutions with one-dimensional time marginals $w^y(t + \delta, x)dx, t \geq 0$. Thus these marginals may be fixed beforehand, which allows to simplify the McKean–Vlasov SDE (1.2) to a non-distribution dependent SDE

$$dX(t) = \tilde{b}(t, X(t))dt + \tilde{\sigma}(t, X(t))dW(t),$$

with coefficients (using the definition of b_i, a_{ij} from (1.8))

$$\tilde{b}_i(t, x) := b_i(t, x, w^y(t + \delta, x)dx), \quad \tilde{\sigma}_{ij}(t, x) := \sqrt{2a_{ij}(t, x, w^y(t + \delta, x)dx)}, \quad (t, x) \in \mathbb{R}_+ \times \mathbb{R}^d.$$

The proof of this pathwise uniqueness result poses one of the main technical challenges of the present paper due to the degeneracy and irregularity of the diffusion coefficient $\tilde{\sigma}$, as well as the low differentiability of the drift coefficient. Indeed, note that $\tilde{\sigma}(t, \cdot)$ is compactly supported and not Lipschitz-continuous. For $p < 2$, $\tilde{\sigma}(t, \cdot)$ is even unbounded and discontinuous. Moreover, \tilde{b} is only a function of bounded variation in its spatial coordinate.

Although the literature on pathwise uniqueness results for SDEs, also in degenerate cases, is vast (see, e.g., [54, 47, 62, 48, 66, 43], and in particular [20] for a result for bounded Sobolev-coefficients with certain drifts of (spatial) bounded variation, see also [34] for a recent generalization), we could not apply any of these results for our desired result.

Our proof heavily relies on a careful analysis of the degeneracy and discontinuity of $\tilde{\sigma}$ and (spatial) jump sets of $\tilde{b}(t, \cdot), t > 0$, which in turn, of course, crucially relies on the explicit form of the Barenblatt solutions w^y . This analysis is based on the generalized Lipschitz-type estimate, sometimes called *Crippa–De Lellis estimate* ([21, 20])

$$|f(x) - f(y)| \lesssim_d (M|\nabla f|(x) + M|\nabla f|(y))|x - y| \quad \text{for } dx\text{-a.e. } x, y \in \mathbb{R}^d, \quad (1.9)$$

(cf. Appendix D), where $f \in BV(\mathbb{R}^d)$, and $(\nabla f)_i$ denotes the (signed) measure-valued representative of the Schwartz distributional derivative in direction x_i of f . Here, M denotes the usual (Hardy–Littlewood) maximal operator. The above estimate applied to $\tilde{b}(t, \cdot)$ is essential for our pathwise uniqueness result. Here, the key is to control the maximal function, which involves the singular part (with respect to Lebesgue measure) of the drift’s gradient, which turns out to be comparable to the $(d - 1)$ -dimensional Hausdorff (or surface) measure restricted to the boundary of the spatial support of the Barenblatt solution considered as a Borel measure on \mathbb{R}^d , see Lemma 6.7. In Lemma 6.11, we control the maximal operator applied to this measure by the estimate (6.11). This estimate is optimal, see Remark 6.9, which can be seen from explicit calculation using elementary geometric arguments in dimension $d = 3$, see Lemma D.7 (ii) in Appendix D. We also employ (1.9) for differences of certain powers of $\tilde{\sigma}$. In contrast to existing results where the diffusion coefficient is in H^1 with respect to its spatial coordinate (cf. e.g. [18, 32, 36]), which yields sufficient control over the maximal function on the right hand-side of the inequality when applied to $f = \tilde{\sigma}(t, \cdot)$ (with power 1), such a control fails in the present case. To overcome this issue, we employ an inequality based on the mean-value theorem inspired by [25], see (6.28). A similar idea was also used in [35]. Furthermore, we use and extend the method from the latter work in order to deal with the degeneracy and low regularity of $\tilde{\sigma}$.

Structure of the paper. In Section 2 we introduce the Leibenson equation and its reformulation as a nonlinear FPE. Section 3 contains the associated McKean–Vlasov SDE and the *superposition principle*, via which we lift solutions to the Leibenson equation to solutions to this SDE. In Section 4, we consider the Barenblatt solutions and apply to them this lift, while in Section 5 we construct the *Leibenson process*, a nonlinear Markov process constructed via the superposition principle from the Barenblatt solutions. A key ingredient for this is a delicate uniqueness result for a corresponding linear PDE, which is stated and proven in Section 7. The proof consists of an analytic and a probabilistic part. Before, we prove in Section 6 that the probabilistically weak solutions of the Leibenson process are probabilistically strong. The appendix contains some material on Fokker–Planck equations, Hardy–Littlewood maximal function and details on a few aspects of the proof of Theorems 6.13 and 7.4.

Notation. We write $\mathbb{R}_+ := [0, \infty)$, $x \cdot y$ or $\langle x, y \rangle_{\mathbb{R}^d}$ for the usual inner product on \mathbb{R}^d , $f_+ = \max(f, 0)$ for a real-valued function f , and dx for Lebesgue measure on \mathbb{R}^d .

Classical function spaces. For $U \subseteq \mathbb{R}^l$, $l \in \mathbb{N}$, we denote by $C^n(U, \mathbb{R}^k)$ the space of n -times continuously differentiable maps $f : U \rightarrow \mathbb{R}^k$, and write $C(U, \mathbb{R}^k) = C^0(U, \mathbb{R}^k)$. We write $C(U, \mathbb{R}^k)_0$ to denote the set of all $f \in C(U, \mathbb{R}^k)$ with $f(0) = 0$. More generally, for a topological space X , we write $C(U, X)$ for the set of continuous maps from U to X . Let $t \in U$. By $\pi_t : C(U, X) \rightarrow X$, we denote the canonical evaluation map at t , i.e. $\pi_t(w) := w(t)$. $C_c^n(U, \mathbb{R}^k)$ and $C_b^n(U, \mathbb{R}^k)$ are the subsets of $C^n(U, \mathbb{R}^k)$ of compactly supported functions (in U) and those functions which together with its partial derivatives up to order n are bounded, respectively. We set $C^\infty(U, \mathbb{R}^k) = \bigcap_{n \in \mathbb{N}} C^n(U, \mathbb{R}^k)$, and similarly for C_c^∞ and C_b^∞ . Partial time and spatial derivatives of a sufficiently smooth function are denoted $\partial_t f = \frac{\partial f}{\partial t}$, $\partial_i f = \frac{\partial f}{\partial x_i}$ and $\partial_{ij} f = \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j}$.

L^p - and Sobolev spaces. For $p \in [1, \infty]$ and $U \subseteq \mathbb{R}^l$, the usual spaces of (equivalence classes of) Borel measurable p -integrable (w.r.t. dx) functions $f : U \rightarrow \mathbb{R}^k$ are denoted $L^p(U, \mathbb{R}^k)$ with norm $|\cdot|_{L^p(U, \mathbb{R}^k)}$, shortly $|\cdot|_{L^p(U)}$ or simply $|\cdot|_{L^p}$ when no confusion can occur. $L_{\text{loc}}^p(U, \mathbb{R}^k)$ are the corresponding spaces of locally integrable functions. When integrability is considered with respect to a measure μ instead of dx , we write $L^p(U, \mathbb{R}^k; \mu)$. $W_{(\text{loc})}^{m,p}(U, \mathbb{R}^k)$, $m \in \mathbb{N}, p \in [1, \infty]$, are the usual Sobolev spaces of (locally) p -integrable m -times weakly differentiable functions $f : U \rightarrow \mathbb{R}^k$ with (locally) p -integrable weak derivatives up to order m ; we write $H_{(\text{loc})}^{m,2}(U, \mathbb{R}^k) = W_{(\text{loc})}^{m,2}(U, \mathbb{R}^k)$, and $H^{-1}(U, \mathbb{R}^k)$ for the corresponding topological dual space. All these spaces are equipped with their usual norms. If $U \subseteq I \times \mathbb{R}^d$, $I \subseteq \mathbb{R}_+$, we write $W_p^{m,n}(U)$ for the space of p -integrable maps with n p -integrable weak derivatives in $t \in I$ and m p -integrable weak derivatives in $x \in \mathbb{R}^d$. For a normed space $(X, \|\cdot\|)$, by $L_{(\text{loc})}^p(I; X)$ we denote the space of measurable functions $\varphi : I(\subseteq \mathbb{R}_+) \rightarrow X$ such that $[t \mapsto \|\varphi(t)\|] \in L_{(\text{loc})}^p(I, \mathbb{R})$. For all function spaces introduced here, we omit the state space \mathbb{R}^k from the notation when $k = 1$.

Measures and spaces of measures. \mathcal{M}_b^+ is the space of nonnegative finite Borel measures on \mathbb{R}^d . The vague and weak topology on it are defined as usual, i.e. as the initial topology of the maps $\mu \mapsto \int_{\mathbb{R}^d} h d\mu$ for all $h \in C_c(\mathbb{R}^d)$ and $h \in C_b(\mathbb{R}^d)$, respectively. For a topological space X , we write $\mathcal{P}(X)$ for the space of probability measures on the Borel σ -algebra $\mathcal{B}(X)$, and simply $\mathcal{P} = \mathcal{P}(\mathbb{R}^d)$. δ_x is the Dirac measure in x . For a random variable X between a measure space (X, \mathcal{X}, μ) and some measurable space, we write $\mathcal{L}_X = \mu \circ X^{-1}$ for its distribution, i.e. the image (or pushforward) measure of X with respect to μ . This includes the case where $X = (X(t))_{t \geq 0}$ is a stochastic process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, considered as a path-valued map $\omega \mapsto X(\omega)$. We then say \mathcal{L}_X is the (path) law of X . Let $M \subset \mathbb{R}^n$ be a submanifold. By S , we denote the standard *surface measure* on $(M, \mathcal{B}(M))$. Here, M is considered with respect to the subspace topology of \mathbb{R}^n . We only consider the case where M is a $(d-1)$ -dimensional sphere.

Functions of bounded variation. $f \in L^1(\mathbb{R}^l, \mathbb{R}^k)$ is said to be of *bounded variation*, in symbol $f \in BV(\mathbb{R}^l; \mathbb{R}^k)$, if all its components' first-order Schwartz distributional derivatives $\partial_i f^j$ are given by a finite signed Borel measure on \mathbb{R}^l , for all $1 \leq i \leq l$, $1 \leq j \leq k$.

(Hardy–Littlewood) *Maximal function/operator and Muckenhoupt weights.* For a brief introduction

to both maximal operator (in symbol 'M') and Muckenhoupt weights (in symbol ' A_p '), we refer to Appendix D.

Probability theory. We say that $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ is a *stochastic basis* if $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space and (\mathcal{F}_t) a right-continuous filtration on Ω augmented by the \mathbb{P} -zero sets.

Miscellaneous. Let $a, b \in \mathbb{R}$. We write $a \lesssim_{(t)} b$ and $a \lesssim b$ if $a \leq C_t b$ or $a \leq Cb$, where $C_t, C > 0$ are constants dependent and independent from a parameter t , respectively. Likewise, we write $a \cong_t b$ and $a \cong b$, if $a = Cb$ for a constant C dependent or independent from t , respectively.

2 The Leibenson equation and its Fokker–Planck reformulation

Let $d \in \mathbb{N}$, $q > 0$, $p > 1$.

Definition 2.1. $u : (0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R}_+$ is a *weak solution* to the Leibenson equation (1.1), if

$$u^q \in L^1_{\text{loc}}((0, \infty); W^{1,1}_{\text{loc}}(\mathbb{R}^d)), \quad u, |\nabla u^q|^{p-1} \in L^1_{\text{loc}}((0, \infty) \times \mathbb{R}^d), \quad (2.1)$$

$t \mapsto u(t, x)dx$ is vaguely continuous, and for all $\psi \in C^1_c((0, \infty) \times \mathbb{R}^d)$

$$\int_{(0, \infty) \times \mathbb{R}^d} -u \partial_t \psi + |\nabla u^q|^{p-2} \nabla u^q \cdot \nabla \psi \, dx dt = 0. \quad (2.2)$$

u has initial condition $\nu \in \mathcal{M}_b^+$, if $u(t, x)dx \xrightarrow{t \rightarrow 0} \nu$ vaguely. u is called *probability solution*, if dt-a.e. $u(t, x)dx$ and its initial datum are probability measures.

The only reason why we restrict attention to nonnegative solutions is that, ultimately, we will only be interested in probability solutions.

The following equivalent formulation via time-independent test functions will be useful. The equivalence can be proven as in the proof of [16, Prop.6.1.2].

Lemma 2.2. *If u satisfies (2.1), $t \mapsto u(t, x)dx$ is vaguely continuous and*

$$|\nabla u^q|^{p-1} \in L^1_{\text{loc}}([0, \infty) \times \mathbb{R}^d), \quad (2.3)$$

then u is a weak solution with initial condition ν if and only if

$$\int_{\mathbb{R}^d} \varphi u(t) \, dx = \int_{\mathbb{R}^d} \varphi \, d\nu - \int_0^t \int_{\mathbb{R}^d} |\nabla u^q|^{p-2} \nabla u^q \cdot \nabla \varphi \, dx dt, \quad \forall t \geq 0, \quad (2.4)$$

for every $\varphi \in C^1_c(\mathbb{R}^d)$.

Even though the solutions we consider are vaguely continuous, the following remark may be useful for extensions of our results to (a priori) discontinuous solutions.

Remark 2.3. *The vague continuity assumption is not restrictive: If u satisfies (2.3) as well as all assumptions of Definition 2.1 except for the continuity assumption and, in addition,*

$$\text{ess sup}_{t>0} |u(t)|_{L^1(\mathbb{R}^d)} < \infty,$$

then $t \mapsto u(t, x)dx$ has a unique vaguely continuous (measure-valued) dt-version on $[0, \infty)$. This follows from [51, Lemma 2.3]. Moreover, if $|\nabla u^q|^{p-1} \in L^1_{\text{loc}}([0, \infty), L^1(\mathbb{R}^d))$, then $[0, \infty) \ni t \mapsto |u(t)|_{L^1(\mathbb{R}^d)}$ is constant. The latter follows by considering (2.4) for an increasing sequence $(\varphi_n)_{n \in \mathbb{N}} \subseteq C^1_c(\mathbb{R}^d)$ such that $0 \leq \varphi_n \leq 1$, $\varphi_n(x) = 1$ for $|x| < n$, $\sup_n |\nabla \varphi_n(x)| \leq C$ for all $x \in \mathbb{R}^d$ for some $C > 0$ not depending on n or x , and by letting $n \rightarrow \infty$.

Associated nonlinear Fokker–Planck equation. Let $\mathcal{F} = \mathcal{F}(p, q, d) \subseteq W_{\text{loc}}^{1,1}(\mathbb{R}^d)$ be the set of those nonnegative functions $u \in W_{\text{loc}}^{1,1}(\mathbb{R}^d)$ such that $|\nabla u|^{p-2} u^{(p-1)(q-1)} \in W_{\text{loc}}^{1,1}(\mathbb{R}^d)$. Note that both $p-2$ and $(q-1)(p-1)$ could be negative and that hence some care is required in writing expressions as $|\nabla u|^{p-2} u^{(q-1)(p-1)}$, since $|\nabla u|$ and u may vanish. For such terms to be well-defined in the sense of Sobolev functions, an implicit assumption is that either functions raised to a negative power are strictly positive dx-a.s. or, for products, that zeros raised to a negative power are multiplied by zeros raised to a positive power. In the latter case we use the convention $0 \times \infty = 0$.

Then let $a_{ij}, b_i : \mathcal{F} \times \mathbb{R}^d \rightarrow \mathbb{R}$, $1 \leq i, j \leq d$, be defined as in (1.8), i.e.

$$a_{ij}(u, x) = q^{p-1} \delta_{ij} |\nabla u(x)|^{p-2} u(x)^{(p-1)(q-1)}, \quad b_i(u, x) = q^{p-1} \partial_i (|\nabla u(x)|^{p-2} |u(x)|^{(p-1)(q-1)})$$

and consider the nonlinear Fokker–Planck equation (here and throughout using Einstein summation convention)

$$\partial_t u(t, x) = \partial_{ij} (a_{ij}(u(t), x) u(t, x)) - \partial_i (b_i(u(t), x) u(t, x)), \quad (t, x) \in (0, \infty) \times \mathbb{R}^d,$$

i.e. (1.7). We stress that these coefficients are not of the usual Nemytskii-type, since they depend pointwise on the gradient of the solution u , hence on the values of u in a(n arbitrary small) neighborhood of x .

Definition 2.4. A nonnegative function $u \in L_{\text{loc}}^1((0, \infty) \times \mathbb{R}^d)$ is a *distributional solution* to (1.7), if $u(t) \in \mathcal{F}$ for all $t > 0$, $t \mapsto u(t, x) dx$ is vaguely continuous,

$$|\nabla u|^{p-2} u^{(q-1)(p-1)+1}, \quad u \partial_i (|\nabla u|^{p-2} u^{(q-1)(p-1)}) \in L_{\text{loc}}^1((0, \infty) \times \mathbb{R}^d), \quad i \in \{1, \dots, d\}, \quad (2.5)$$

and for all $\psi \in C_c^2((0, \infty) \times \mathbb{R}^d)$

$$\int_{(0, \infty) \times \mathbb{R}^d} \left(\partial_t \psi + q^{p-1} |\nabla u|^{p-2} u^{(p-1)(q-1)} \Delta \psi + q^{p-1} \nabla (|\nabla u|^{p-2} u^{(p-1)(q-1)}) \cdot \nabla \psi \right) u \, dx dt = 0. \quad (2.6)$$

u has initial condition $\nu \in \mathcal{M}_b^+$, if $u(t, x) dx \xrightarrow{t \rightarrow 0} \nu$ vaguely. u is called *probability solution*, if ν and dt -a.e. $u(t, x) dx$ are Borel probability measures.

The following lemma can be proven exactly as [16, Prop. 6.1.2].

Lemma 2.5. If u is as in the previous definition and satisfies (2.5) with $[0, \infty)$ replacing $(0, \infty)$, then u is a distributional solution with initial condition ν if and only if for all $t \geq 0$ and $\varphi \in C_c^2(\mathbb{R}^d)$

$$\int_{\mathbb{R}^d} \varphi u(t) \, dx = \int_{\mathbb{R}^d} \varphi \, d\nu + \int_0^t \int_{\mathbb{R}^d} q^{p-1} \left(|\nabla u|^{p-2} u^{(p-1)(q-1)} \Delta \varphi + \nabla (|\nabla u|^{p-2} u^{(p-1)(q-1)}) \cdot \nabla \varphi \right) u \, dx dt. \quad (2.7)$$

An analogue of Remark 2.3 holds in the Fokker–Planck case. We compare Definitions 2.1 and 2.4:

Lemma 2.6. Assume u satisfies (2.1) with

$$\nabla u^q = q u^{q-1} \nabla u, \quad (2.8)$$

$u(t) \in \mathcal{F}$ for all $t > 0$, (2.5), and $t \mapsto u(t, x) dx$ is vaguely continuous. Then u is a weak solution to (1.1) in the sense of Definition 2.1 if and only if it is a distributional solution to (1.7) in the sense of Definition 2.4.

If in addition u satisfies all integrability assumptions in time on $[0, \infty)$ instead of $(0, \infty)$ and has initial condition ν , then u satisfies (2.4) if and only if it satisfies (2.7) (for all $\varphi \in C_c^1(\mathbb{R}^d)$ and $C_c^2(\mathbb{R}^d)$, respectively).

Note that (2.8) does not follow from (2.1), since the latter does not imply local integrability of $u^{q-1} \nabla u$ (in particular, note that $q-1$ need not be positive).

Proof. By assumption all regularity- and integrability properties of Definitions 2.1 and 2.4 are satisfied, so it remains to prove equivalence of (2.2) and (2.6). To this end, let $\psi \in C_c^2((0, \infty) \times \mathbb{R}^d)$. Then

$$\begin{aligned} \int_{(0, \infty) \times \mathbb{R}^d} |\nabla u^q|^{p-2} \nabla u^q \cdot \nabla \psi \, dx dt &= q^{p-1} \int_{(0, \infty) \times \mathbb{R}^d} u^{(p-1)(q-1)} |\nabla u|^{p-2} \nabla u \cdot \nabla \psi \, dx dt \\ &= q^{p-1} \int_{(0, \infty) \times \mathbb{R}^d} \nabla \left(u^{(p-1)(q-1)} |\nabla u|^{p-2} u \right) \cdot \nabla \psi - \nabla \left(u^{(p-1)(q-1)} |\nabla u|^{p-2} \right) u \cdot \nabla \psi \, dx dt \\ &= -q^{p-1} \int_{(0, \infty) \times \mathbb{R}^d} \left(|\nabla u|^{p-2} u^{(p-1)(q-1)} \Delta \psi + \nabla \left(u^{(p-1)(q-1)} |\nabla u|^{p-2} \right) \cdot \nabla \psi \right) u \, dx dt, \end{aligned}$$

where the first equality is due to (2.8), and the second follows from the product rule for Sobolev functions. To extend (2.2) to all $\psi \in C_c^1((0, \infty) \times \mathbb{R}^d)$, one uses a simple approximation argument. This concludes the first part of the proof. The second part follows immediately from the first part and Lemmas 2.2 and 2.5. \square

3 Associated McKean–Vlasov SDE and superposition principle

For any nonlinear Fokker–Planck equation there is an associated McKean–Vlasov SDE with drift- and diffusion-coefficient given by the drift and the square root of twice the diffusion coefficient of the nonlinear Fokker–Planck equation. Thus, the McKean–Vlasov SDE associated with (1.7) is

$$\begin{cases} dX(t) = q^{p-1} \nabla \left(|\nabla u(t, X(t))|^{p-2} u(t, X(t))^{(q-1)(p-1)} \right) dt + \sqrt{2q^{p-1}} \left(|\nabla u(t, X(t))|^{\frac{p-2}{2}} u(t, X(t))^{\frac{(q-1)(p-1)}{2}} \right) dW(t) \\ \mathcal{L}_{X(t)}(dx) = u(t, x) dx, \end{cases} \quad (3.1)$$

where $W = (W(t))_{t \geq 0}$ is a standard d -dimensional Brownian motion. We stress that here u is *not* a priori given, but part of the solution.

Definition 3.1. A *probabilistically weak solution* to (3.1) is an adapted stochastic process $X = (X(t))_{t \geq 0}$ on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ with an (\mathcal{F}_t) -standard Brownian motion W such that $\mathcal{L}_{X(t)}(dx) = u(t, x) dx$ with $u(t, \cdot) \in \mathcal{F}$ for all $t > 0$,

$$\mathbb{E} \left[\int_0^T \left(|\nabla (|\nabla u(t, X(t))|^{p-2} u(t, X(t))^{(q-1)(p-1)})| + |\nabla u(t, X(t))|^{p-2} u(t, X(t))^{(q-1)(p-1)} \right) dt \right] < \infty \quad \forall T > 0, \quad (3.2)$$

and \mathbb{P} -a.s.

$$\begin{aligned} X(t) &= X(0) + q^{p-1} \int_0^t \nabla (|\nabla u(t, X(t))|^{p-2} u(t, X(t))^{(q-1)(p-1)}) dt \\ &\quad + \sqrt{2q^{p-1}} \int_0^t |\nabla u(t, X(t))|^{\frac{p-2}{2}} u(t, X(t))^{\frac{(q-1)(p-1)}{2}} dW(t), \quad \forall t \geq 0. \end{aligned}$$

The *initial condition* of X is the probability measure $\mathcal{L}_{X(0)}$. We call the probability measure P on $\mathcal{B}(C(\mathbb{R}_+, \mathbb{R}^d))$, $P := \mathbb{P} \circ X^{-1}$, where X is considered as the path-valued map $X : \Omega \rightarrow C(\mathbb{R}_+, \mathbb{R}^d)$, $X(\omega) = [t \mapsto X(t)(\omega)]$, a *solution (path) law* to (3.1).

Strictly speaking, a weak solution is a tuple $((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}), X, W)$. However, we usually shortly refer to X as the weak solution or, when it is relevant in the context of strong solutions (see Definition 6.1), to (X, W) .

Superposition principle. The crucial relation between a nonlinear Fokker–Planck equation and its associated McKean–Vlasov SDE is the following: By Itô’s formula, the curve of one-dimensional time marginals of any McKean–Vlasov probabilistically weak solution is a weakly continuous probability solution to the nonlinear Fokker–Planck equation (in the sense of Definition C.2). Conversely, by the *superposition principle*, see [64, 4, 5], for any such distributional solution with sufficient spatial integrability, there exists a probabilistically weak solution to the McKean–Vlasov SDE with one-dimensional time marginals equal to this distributional solution. Now we state these relations precisely for the nonlinear FPE (1.7) and its associated McKean–Vlasov SDE (3.1).

Proposition 3.2. *Let $X = (X(t))_{t \geq 0}$ be a probabilistically weak solution to (3.1). Then u , given by $t \mapsto u(t, x)dx := \mathcal{L}_{X(t)}(dx)$, is a weakly continuous probability solution to (1.7). Thus, it is a weak solution to (1.1), provided u satisfies (2.1) and (2.8).*

Proof. The first assertion follows from a straightforward application of Itô’s formula, while the second part follows from Lemma 2.6. \square

The following result yields probabilistic representations of weak solutions to the Leibenson equation as one-dimensional marginal density curves of solutions to the associated McKean–Vlasov SDE.

Theorem 3.3 (Superposition principle). *Let u be a weakly continuous probability solution to (1.7) with initial condition ν such that*

$$\int_0^T \int_{\mathbb{R}^d} |\nabla u|^{p-2} u^{(q-1)(p-1)+1} dx dt < \infty, \quad \forall T > 0 \quad (3.3)$$

and

$$\int_0^T \int_{\mathbb{R}^d} |\nabla (|\nabla u|^{p-2} u^{(q-1)(p-1)})| u dx dt < \infty, \quad \forall T > 0. \quad (3.4)$$

Then there exists a probabilistically weak solution $(X(t))_{t \geq 0}$ to (3.1) with $\mathcal{L}_{X(t)}(dx) = u(t, x)dx$ for all $t > 0$ and $\mathcal{L}_{X(0)}(dx) = \nu(dx)$.

The following corollary follows immediately from the previous theorem and Lemma 2.6.

Corollary 3.4. *Let u be a weakly continuous weak probability solution to (1.1) with initial condition ν such that $u(t) \in \mathcal{F}$ for all $t > 0$, and (2.8), (3.3)+(3.4) hold. Then there exists a probabilistically weak solution $(X(t))_{t \geq 0}$ to (3.1) with $\mathcal{L}_{X(t)}(dx) = u(t, x)dx$ for all $t > 0$ and $\mathcal{L}_{X(0)}(dx) = \nu(dx)$.*

Proof of Theorem 3.3. Consider u as a solution (in the sense of Definition C.1) to the linear FPE with coefficients \tilde{a} and \tilde{b} obtained by a priori fixing u in a and b from (1.8), i.e.

$$\tilde{a}_{ij}(t, x) := q^{p-1} \delta_{ij} |\nabla u(x)|^{p-2} |u(x)|^{(p-1)(q-1)}, \quad \tilde{b}_i(t, x) := q^{p-1} \partial_i (|\nabla u(x)|^{p-2} |u(x)|^{(p-1)(q-1)}). \quad (3.5)$$

By [64, Thm.2.5] (which applies due to (3.3)+(3.4)) one obtains a probabilistically weak solution $X = (X(t))_{t \geq 0}$ to the corresponding (non-distribution dependent) SDE with coefficients \tilde{b}_i as above and $\tilde{\sigma} = (\tilde{\sigma}_{ij})$ given by $\frac{1}{2} \tilde{\sigma} \tilde{\sigma}^T = \tilde{a}(t, x)$ such that $\mathcal{L}_{X(t)}(dx) = u(t, x)dx$. Clearly, X then solves (3.1). \square

4 Barenblatt solutions

The Leibenson equation admits explicit solutions with point source initial condition, known as *Barenblatt solutions*. Three very different regimes arise, subject to $q(p-1)$ being strictly greater, equal, or strictly less than 1. In this paper, we focus on the first case. This includes the p -Laplace case ($q = 1, p > 2$).

Definition 4.1 (Barenblatt solutions). Let $d \in \mathbb{N}$, $p > 1$, $q > 0$ such that $q(p-1) > 1$. For $y \in \mathbb{R}^d$, the function

$$w^y(t, x) := t^{-\frac{d}{\beta}} \left[C - \kappa \left(t^{-\frac{1}{\beta}} |x - y| \right)^{\frac{p}{p-1}} \right]_+^\gamma, \quad (t, x) \in (0, \infty) \times \mathbb{R}^d, \quad (4.1)$$

where

$$\beta = p + d(q(p-1) - 1), \quad \gamma = \frac{p-1}{q(p-1) - 1}, \quad \kappa = \frac{q(p-1) - 1}{pq} \beta^{-\frac{1}{p-1}},$$

and $C > 0$ is any constant, is called *Barenblatt solution* to (1.1).

A straightforward calculation shows that there is a unique choice $C = C(d, p, q)$ independent of t such that $\int_{\mathbb{R}^d} w^y(t, x) dx = 1$ for all $t > 0$. We fix this choice, so that $t \mapsto w^y(t, x) dx$ is a curve of probability measures. We often abbreviate

$$f(t, x) := C - \kappa \left(t^{-\frac{1}{\beta}} |x| \right)^{\frac{p}{p-1}}.$$

Furthermore, we set

$$R(t) := \left(\frac{C}{\kappa} \right)^{\frac{p-1}{p}} t^{\frac{1}{\beta}}. \quad (4.2)$$

Remark 4.2. Note that $\beta > p$ and $\gamma, \kappa > 0$.

For the definition of Barenblatt solutions in the cases $q(p-1) = 1$ and $0 < q(p-1) < 1$, we refer to [12, 30]. In these cases the solutions behave fundamentally different than those in (4.1). In particular, while w^y above has *finite* speed of propagation, the Barenblatt solutions in the case $q(p-1) = 1$ and $q(p-1) < 1$ have *infinite* speed of propagation. We postpone the study of these case to future work.

From now on, we consider without loss of generality $y = 0$ and write $w = w^0$. All subsequent statements hold for any $y \in \mathbb{R}^d$, with obvious modifications where necessary.

Lemma 4.3. w has the following properties.

- (i) w is a weakly continuous weak probability solution to (1.1) and attains its initial datum δ_0 weakly in the sense of measures, i.e. $w(t, x) dx \rightarrow \delta_0$ weakly as $t \rightarrow 0$.
- (ii) For all $t > 0$, $w(t) \in C_c(\mathbb{R}^d)$ and $\text{supp } w(t) = \overline{B_{R(t)}(y)}$, with $R(t)$ as in (4.2).
- (iii) $w \in C((0, \infty) \times \mathbb{R}^d) \cap \bigcap_{0 < r < R} C_c([r, R] \times \mathbb{R}^d) \cap \bigcap_{r > 0} L^\infty((r, \infty) \times \mathbb{R}^d)$.

Proof. The nonnegativity as well as (ii) and (iii) follow immediately from the definition of w . The weak continuity of $(0, \infty) \ni t \mapsto w(t, x) dx$ follows from the continuity of w in $(t, x) \in (0, \infty) \times \mathbb{R}^d$, and the weak continuity in $t = 0$ follows by a straightforward calculation via the transformation $\Phi(x) = t^{\frac{1}{\beta}} x$ (similarly to the proof that the classical heat kernel converges weakly to δ_0 when $t \rightarrow 0$). That w solves (1.1) in the sense that both sides of the equation are well-defined and equal in the sense of Sobolev spaces is a straightforward calculation. In particular, w is a weak solution to (1.1) in the sense of Definition 2.1. \square

4.1 Probabilistic representation for Barenblatt solutions

We apply Corollary 3.4 to the Barenblatt solution w . More precisely, we have the following result.

Theorem 4.4. Let $d \in \mathbb{N}$, $p > 1$, $q > 0$ such that $q(p-1) > 1$ and $p > \frac{1+d}{d}$. There is a probabilistically weak solution $(X(t))_{t \geq 0}$ to (3.1) such that $\mathcal{L}_{X(t)}(dx) = w(t, x) dx$ for all $t \geq 0$ (here and below with a slight abuse of notation we write $w(0, x) dx = \delta_0(dx)$).

Recall that we write " \lesssim " and " \lesssim_t " whenever we suppress absolute multiplicative constants, independent or dependent of t , respectively, in an inequality.

Proof. We show that w satisfies the assumptions of Corollary 3.4. That w is a weakly continuous probability solution to (1.1) was already stated in Lemma 4.3, so it remains to prove $u(t) \in \mathcal{F}$ for all $t > 0$, (3.3)-(3.4) and (2.8). To this end, we first prove $w(t) \in W^{1,1}(\mathbb{R}^d)$ for all $t > 0$. Note that $w(t) \in L^1(\mathbb{R}^d)$ follows from Lemma 4.3. If $\gamma \geq 1$, $w(t) \in W^{1,1}(\mathbb{R}^d)$ is obvious, and [recall that we set $f(t, x) = C - \kappa(t^{-\frac{1}{\beta}}|x|)^{\frac{p}{p-1}}$]

$$\nabla w(t, x) = \gamma t^{-\frac{d}{\beta}} f(t, x)^{\gamma-1} \nabla f(t, x) \mathbf{1}_{f(t) \geq 0}(x). \quad (4.3)$$

Since f is locally bounded, $\{f(t) \geq 0\}$ is compact and

$$\nabla f(t, x) = -\kappa \frac{p}{p-1} t^{-\frac{p}{\beta(p-1)}} |x|^{\frac{2-p}{p-1}} x, \quad (4.4)$$

we have indeed $\nabla w(t) \in L^1(\mathbb{R}^d, \mathbb{R}^d)$. If $0 < \gamma < 1$, the RHS of (4.3) is still the weak gradient of $w(t)$, provided we can show $f(t)^{\gamma-1} \nabla f(t) \mathbf{1}_{f(t) \geq 0} \in L^1(\mathbb{R}^d, \mathbb{R}^d)$. Since $f(t, x)^{\gamma-1} |\nabla f(t, x)| \lesssim_t |x|^{\frac{1}{p-1}} f(t, x)^{\gamma-1}$, this translates into proving $f(t)^{\gamma-1} \mathbf{1}_{f(t) \geq 0} \in L^1(\mathbb{R}^d)$. Using polar coordinates and the transformation rule, we find

$$\int_{\mathbb{R}^d} f(t, x)^{\gamma-1} \mathbf{1}_{f(t) \geq 0}(x) dx \lesssim_t \int_0^C r^{\gamma-1} \left(\frac{C-r}{\kappa} \right)^{\frac{(d-1)(p-1)-1}{p}} dr. \quad (4.5)$$

Splitting the latter in integrals over $(0, \frac{C}{2})$ and $(\frac{C}{2}, C)$, we obtain the conditions

$$\gamma - 1 > -1 \quad \text{and} \quad \frac{(d-1)(p-1)-1}{p} > -1$$

to ensure finiteness of the RHS in (4.5). Since $\gamma > 0$ and the second inequality is equivalent to $dp - d > 0$ and we assume $p > 1$, we obtain $w(t) \in W^{1,1}(\mathbb{R}^d)$ for all $t > 0$.

Next, we show $|\nabla w(t)|^{p-2} w(t)^{(q-1)(p-1)} \in W^{1,1}(\mathbb{R}^d)$ for all $t > 0$. From (4.3)-(4.4), we see

$$|\nabla w(t, x)|^{p-2} w(t)^{(q-1)(p-1)} \cong t^{-\frac{d}{\beta}(p-2) - \frac{p(p-2)}{\beta(p-1)} - \frac{d}{\beta}(q-1)(p-1)} \mathbf{1}_{f(t) \geq 0}(x) f(t, x)^{(\gamma-1)(p-2) + \gamma(q-1)(p-1)} |x|^{\frac{p-2}{p-1}},$$

where we recall that by \cong we indicate that we suppress absolute multiplicative constants on the RHS, which are in particular independent of t and x . Note that at least one of the exponents $p-2$ and $(q-1)(p-1)$ is strictly positive, thus we can use our convention $0 \times \infty = 0$ to obtain the factor $\mathbf{1}_{f(t) \geq 0}$ on the RHS above. Indeed, if $p < 2$, then $q(p-1) > 1$ implies $(q-1)(p-1) > 0$. To continue, note that

$$(\gamma-1)(p-2) + \gamma(q-1)(p-1) = 1, \quad (4.6)$$

since this can easily be seen to be equivalent to

$$(p-1)(\gamma-1 + \gamma(q-1)) + 1 - \gamma = 1 \iff (p-1)(\gamma q - 1) = \gamma,$$

which is true by definition of γ . Thus

$$\mathbf{1}_{f(t) \geq 0} f(t)^{(\gamma-1)(p-2) + \gamma(q-1)(p-1)} = f(t)_+ \in (W^{1,1} \cap W^{1,\infty})(\mathbb{R}^d).$$

Since $p > \frac{d+1}{d} > \frac{d+2}{d+1}$, $x \mapsto |x|^{\frac{p-2}{p-1}}$ is in $W_{\text{loc}}^{1,1}(\mathbb{R}^d)$, hence we obtain that $|\nabla w(t)|^{p-2} w(t)^{(q-1)(p-1)}$ is weakly differentiable, and, by the product rule, for a.e. x

$$|\nabla(|\nabla w(t, x)|^{p-2} w(t)^{(q-1)(p-1)})| \lesssim t^{-\frac{d}{\beta}(p-2) - \frac{p(p-2)}{\beta(p-1)} - \frac{d}{\beta}(q-1)(p-1)} \mathbf{1}_{f(t) \geq 0} \left[t^{-\frac{p}{\beta(p-1)}} |x|^{\frac{1}{p-1}} + f(t, x)_+ |x|^{-\frac{1}{p-1}} \right]. \quad (4.7)$$

To obtain that this term is in $L^1(\mathbb{R}^d, \mathbb{R}^d)$ for arbitrary fixed $t > 0$, it suffices to note that $|x|^{-\frac{p}{p-1}} x \in L_{\text{loc}}^1(\mathbb{R}^d, \mathbb{R}^d)$, which is true since $p > \frac{d+1}{d}$. Hence $u(t) \in \mathcal{F}$ for all $t > 0$.

Regarding (2.8), we have $u^q(t, x) = t^{-\frac{qd}{\beta}} f_+^{\gamma q}(t, x)$, with $\gamma q > 1$ by definition of γ , and hence one can simply calculate ∇u^q by the chain rule to deduce (2.8).

Regarding (3.3) we see that, by (4.3)-(4.4)

$$\begin{aligned} & \int_0^T \int_{\mathbb{R}^d} |\nabla w(t, x)|^{p-2} w(t, x)^{(q-1)(p-1)+1} dx dt \\ & \lesssim \int_0^T \int_{B_{R(t)}(0)} t^{-(p-2)(\frac{d}{\beta} + \frac{p}{\beta(p-1)}) - \frac{d}{\beta}[(q-1)(p-1)+1]} f(t, x)^{1+\gamma} |x|^{\frac{p-2}{p-1}} dx dt \\ & \lesssim \int_0^T t^{-(p-2)(\frac{d}{\beta} + \frac{p}{\beta(p-1)}) - \frac{d}{\beta}[(q-1)(p-1)+1]} \int_{B_{R(t)}(0)} |x|^{\frac{p-2}{p-1}} dx dt, \end{aligned} \quad (4.8)$$

where we used

$$(\gamma - 1)(p - 2) + \gamma(q - 1)(p - 1) + \gamma = 1 + \gamma$$

(see (4.6)) and $1 + \gamma > 0$. Using polar coordinates and $p > \frac{d+2}{d+1}$, we find

$$\int_{B_{R(t)}(0)} |x|^{\frac{p-2}{p-1}} dx \lesssim \int_0^{R(t)} r^{\frac{p-2}{p-1} + d - 1} dr \lesssim t^{\frac{p-2}{\beta(p-1)} + \frac{d}{\beta}}.$$

Thus the RHS of (4.8) is further estimated from above by

$$\int_0^T t^{-(p-2)(\frac{d}{\beta} + \frac{p}{\beta(p-1)}) - \frac{d}{\beta}[(q-1)(p-1)] + \frac{p-2}{\beta(p-1)}} dt$$

i.e. to conclude its finiteness we need

$$-(p-2) \left(\frac{d}{\beta} + \frac{p}{\beta(p-1)} \right) - \frac{d}{\beta}[(q-1)(p-1)] + \frac{p-2}{\beta(p-1)} > -1.$$

By a direct calculation, this inequality can be seen to be equivalent to $p - 2 < p$ and is hence true. This concludes the proof of (3.3).

Finally, regarding (3.4), we consider (4.7) to estimate

$$\begin{aligned} & \int_0^T \int_{\mathbb{R}^d} |\nabla(|\nabla w|^{p-2} w^{(q-1)(p-1)})| w dx dt \\ & \lesssim \int_0^T \int_{B_{R(t)}(0)} t^{-\frac{d}{\beta}(p-1) - \frac{p(p-2)}{\beta(p-1)} - \frac{d}{\beta}(q-1)(p-1)} \left[t^{-\frac{p}{\beta(p-1)}} |x|^{\frac{1}{p-1}} + |x|^{-\frac{1}{p-1}} \right] dx dt =: I + II. \end{aligned}$$

We further estimate, using again $p > \frac{d+1}{d}$ to have $|x|^{-\frac{1}{p-1}} \in L_{\text{loc}}^1(\mathbb{R}^d)$,

$$I \lesssim \int_0^T t^{-\frac{d}{\beta}(p-2) - \frac{p(p-1)-1}{\beta(p-1)} - \frac{d}{\beta}(q-1)(p-1)} dt$$

and

$$II \lesssim \int_0^T t^{-\frac{d}{\beta}(p-2) - \frac{p(p-2)+1}{\beta(p-1)} - \frac{d}{\beta}(q-1)(p-1)} dt.$$

Elementary calculations show

$$\min \left(-\frac{d}{\beta}(p-2) - \frac{p(p-1)-1}{\beta(p-1)} - \frac{d}{\beta}(q-1)(p-1), -\frac{d}{\beta}(p-2) - \frac{p(p-2)+1}{\beta(p-1)} - \frac{d}{\beta}(q-1)(p-1) \right) > -1,$$

whereby (3.4) follows and the proof is complete. \square

Since the computations from the previous proof obviously remain valid for w^y instead of w , we obtain:

Corollary 4.5. *Let $d \in \mathbb{N}$, $p > 1$, $q > 0$ such that $q(p-1) > 1$, $p > \frac{1+d}{d}$ and $y \in \mathbb{R}^d$. There is a probabilistically weak solution $X^y = (X^y(t))_{t \geq 0}$ to (3.1) such that $\mathcal{L}_{X^y(t)}(dx) = w^y(t, x)dx$ for all $t \geq 0$ (where again with a slight abuse of notation we write $w^y(0, x)dx = \delta_y(dx)$).*

For later use, we state the following remark.

Remark 4.6. *Definitions 2.1, 2.4 and 3.1 can be modified to initial times $s > 0$ instead of $s = 0$ in the obvious way, and all previously stated results remain valid accordingly. In particular, Corollary 4.5 generalizes in the sense that for all initial pairs (s, y) , there is a probabilistically weak solution $X^{s,y} = (X^{s,y}(t))_{t \geq s}$ to (3.1) such that $\mathcal{L}_{X^{s,y}(t)} = w^y(t-s, x)dx$ for all $t \geq s$.*

5 Associated nonlinear Markov process

Our next goal is to construct a nonlinear Markov process consisting of the path laws of the solutions $X^{s,y}$ from the previous remark, and to show that this Markov process is uniquely determined by the Barenblatt solutions to (1.1). First, we briefly recall the linear case, i.e. $(p, q) = (2, 1)$.

5.1 The classical linear case

In the linear case, i.e. $p = 2$, $q = 1$, when (1.1) is the heat equation, its fundamental solutions given by the classical heat kernel $p(t, x, y)$, $t > 0, x, y \in \mathbb{R}^d$, uniquely determine a Markov process $(P_y)_{y \in \mathbb{R}^d} \subseteq \mathcal{P}(C([0, \infty), \mathbb{R}^d))$ with one-dimensional time marginals $P_y \circ \pi_t^{-1} = p(t, x, y)dx$ (for the definition of a Markov process, see below). We recall that here and below $\pi_t : C(\mathbb{R}_+, \mathbb{R}^d) \rightarrow \mathbb{R}^d$, $\pi_t(w) = w(t)$, denotes the canonical projection at time $t \geq 0$. This Markov process is Brownian motion, P_0 is the *standard Wiener measure*, and any stochastic process $(X^y(t))_{t \geq 0}$ with path law P_y is a Brownian motion started in $y \in \mathbb{R}^d$. Of course, P_y is the unique solution law of the associated SDE

$$dX^y(t) = dW(t), \quad t \geq 0, \quad X^y(0) = y.$$

5.2 Nonlinear Markov processes

The following definition of nonlinear Markov processes is inspired by McKean [49] and was first studied in [52]. It extends the usual notion of Markov property in a suitable way in order to cover path laws of McKean–Vlasov solutions. We refer to [52] for more details.

We use the following notation. For $0 \leq s \leq t$, we denote by π_t^s the projection $\pi_t^s : C([s, \infty), \mathbb{R}^d) \rightarrow \mathbb{R}^d$, $\pi_t^s(w) := w(t)$, and we set $\mathcal{F}_{s,t} := \sigma(\pi_u^s, s \leq u \leq t)$.

Definition 5.1. Let $\mathcal{P}_0 \subseteq \mathcal{P}$. A *nonlinear Markov process* is a family $(P_{s,\zeta})_{(s,\zeta) \in [0,\infty) \times \mathcal{P}_0}$, $P_{s,\zeta} \in \mathcal{P}(C([s, \infty), \mathbb{R}^d))$, such that for all $0 \leq s \leq r \leq t$, $\zeta \in \mathcal{P}_0$

- (i) $\mu_t^{s,\zeta} := P_{s,\zeta} \circ (\pi_t^s)^{-1} \in \mathcal{P}_0$,
- (ii) the *nonlinear Markov property* holds, i.e.

$$P_{s,\zeta}(\pi_t^s \in A | \mathcal{F}_{s,r})(\cdot) = p_{(s,\zeta),(r,\pi_r^s(\cdot))}(\pi_t^r \in A) \quad P_{s,\zeta}\text{-a.s. for all } A \in \mathcal{B}(\mathbb{R}^d), \quad (5.1)$$

where $(p_{(s,\zeta),(r,z)})_{z \in \mathbb{R}^d}$ is a regular conditional probability kernel from \mathbb{R}^d to $\mathcal{B}(C([r, \infty), \mathbb{R}^d))$ of $P_{r,\mu_r^{s,\zeta}}[\cdot | \pi_r^r = z]$, $z \in \mathbb{R}^d$ (i.e. in particular $p_{(s,\zeta),(r,z)} \in \mathcal{P}(C([r, \infty), \mathbb{R}^d))$ and $p_{(s,\zeta),(r,z)}(\pi_r^r = z) = 1$).

The case of a classical (time-inhomogeneous) Markov process $(P_{s,y})_{y \in \mathbb{R}^d}$ is contained in the previous definition. In this case, $(P_{s,\zeta})_{s \geq 0, \zeta \in \mathcal{P}}$, where $P_{s,\zeta} := \int_{\mathbb{R}^d} P_{s,y} d\zeta(y)$, satisfies the previous definition with $\mathcal{P}_0 = \mathcal{P}$ and $p_{(s,y)(r,z)} = P_{r,z}$, for all $0 \leq s \leq r$, $z \in \mathbb{R}^d$, $\zeta \in \mathcal{P}$. Then (5.1) reduces to the usual time-inhomogeneous Markov property, and the *Chapman–Kolmogorov equations* for the one-dimensional time marginals $\mu_t^{s,y} = P_{s,y} \circ (\pi_t^s)^{-1}$ hold, i.e.

$$\mu_t^{s,y} = \int_{\mathbb{R}^d} \mu_t^{r,z} d\mu_r^{s,y}(z), \quad \forall 0 \leq s \leq r \leq t, \quad y \in \mathbb{R}^d. \quad (5.2)$$

In the nonlinear case the map $\zeta \rightarrow P_{s,\zeta}$ is not linear on \mathcal{P}_0 , even if $\mathcal{P}_0 = \mathcal{P}$ (which we do not assume). Thus one loses the Chapman–Kolmogorov equations, but the one-dimensional time marginals of a nonlinear Markov process still satisfy the *flow property*

$$\mu_t^{s,\zeta} \in \mathcal{P}_0, \quad \mu_t^{s,\zeta} = \mu_t^{r,\mu_r^{s,\zeta}}, \quad \forall 0 \leq s \leq r \leq t, \quad \zeta \in \mathcal{P}_0, \quad (5.3)$$

(in the linear case, this follows from (5.2)). A family $\{\mu^{s,\zeta}\}_{s \geq 0, \zeta \in \mathcal{P}_0}$, $\mu^{s,\zeta} = (\mu_t^{s,\zeta})_{t \geq s}$ of weakly continuous solutions to the nonlinear FPE (1.3) in the sense of Definition C.2 with initial condition $\mu_s^{s,\zeta} = \zeta$ and satisfying (5.3) is called a \mathcal{P}_0 -*solution flow* (1.3).

The following result from [52], which allows to construct nonlinear Markov processes consisting of solution laws to McKean–Vlasov SDEs with one-dimensional time marginals given by a prescribed family of solution curves to the associated nonlinear FPE, is the key for our purposes in this chapter.

Theorem 5.2. [52, Theorems 3.4+3.8] *Let $\mathcal{P}_0 \subseteq \mathcal{P}$ and $\{\mu^{s,\zeta}\}_{s \geq 0, \zeta \in \mathcal{P}_0}$ be a \mathcal{P}_0 -solution flow to (1.3) such that the following hypothesis holds.*

(\mathcal{P}_0 -lin_{ex}). $\mu^{s,\zeta}$ is an extreme point of the set of all weakly continuous probability solutions to the linear FPE with coefficients $(t,x) \mapsto a_{ij}(t,x,\mu_t^{s,\zeta})$ and $(t,x) \mapsto b_i(t,x,\mu_t^{s,\zeta})$ with initial condition (s,ζ) in the sense of Definition C.1 for each $(s,\zeta) \in \mathbb{R}_+ \times \mathcal{P}_0$.

Then, for each $(s,\zeta) \in \mathbb{R}_+ \times \mathcal{P}_0$ there is a unique solution law $P_{s,\zeta}$ to the McKean–Vlasov SDE (1.4) such that

$$P_{s,\zeta} \circ (\pi_t^s)^{-1} = \mu_t^{s,\zeta}, \quad \forall 0 \leq s \leq t, \quad \zeta \in \mathcal{P}_0, \quad (5.4)$$

and $(P_{s,\zeta})_{s \geq 0, \zeta \in \mathcal{P}_0}$ is a nonlinear Markov process in the sense of Definition 5.1. In particular, this nonlinear Markov process is uniquely determined by its one-dimensional time marginals $(\mu_t^{s,\zeta})_{0 \leq s \leq t, \zeta \in \mathcal{P}_0}$ and equation (1.4).

Remark 5.3. A relation of extremality and Markov property was discovered in connection with constructing Markov selections (see [41] and also [59, 27]) in the situation where a (linear) martingale problem has more than one solution for Dirac measures as initial conditions (see [60]). But this is a result on path space and in the opposite direction, namely proving that the Markov selection consists of extremal measures in the convex set of all solutions to the (linear) martingale problem with initial condition a Dirac measure. So, our result above is a converse result, but for the time marginals, i.e. on state space, and which also holds in the nonlinear case.

Regarding applications, the equivalence of the above extremality condition and the following restricted linearized uniqueness result has turned out very useful.

Lemma 5.4 ([52, Lemma 3.5]). *Let $(s,\zeta) \in \mathbb{R}_+ \times \mathcal{P}_0$. In the situation of the previous theorem, $\mu^{s,\zeta}$ satisfies the claimed extremality condition if and only if $(\mu_t^{s,\zeta})_{t \geq s}$ is the unique weakly continuous probability solution to the linear FPE with coefficients $(t,x) \mapsto a_{ij}(t,x,\mu_t^{s,\zeta})$ and $(t,x) \mapsto b_i(t,x,\mu_t^{s,\zeta})$ and initial condition (s,ζ) in the sense of Definition C.1 in the class*

$$\{(\nu_t)_{t \geq s} \subseteq \mathcal{P} : \nu_t \leq C \mu_t^{s,\zeta}, \quad t \geq s, \text{ for some } C > 0\}$$

("restricted linearized distributional uniqueness").

In fact, we need the following corollary of the previous theorem. The intuition is that the assertion of the previous theorem remains true when condition $(\mathcal{P}_0\text{-lin}_{\text{ex}})$ holds for less initial data, at the price of a uniqueness assertion for the corresponding solution path laws for less initial data. However, as we will see, in our situation, a posteriori we can restore the uniqueness assertion to all initial data.

Corollary 5.5. [52, Corollary 3.10] *Let $\mathfrak{P}_0 \subseteq \mathcal{P}_0 \subseteq \mathcal{P}$ and suppose $\{\mu^{s,\zeta}\}_{s \geq 0, \zeta \in \mathcal{P}_0}$ is a \mathcal{P}_0 -solution flow to (1.3) such that the following two hypotheses hold.*

($\mathfrak{P}_0\text{-lin}_{\text{ex}}$). $\mu^{s,\zeta}$ is an extreme point in the set of all weakly continuous probability solutions to the linear FPE with coefficients $(t, x) \mapsto a_{ij}(t, x, \mu_t^{s,\zeta})$ and $(t, x) \mapsto b_i(t, x, \mu_t^{s,\zeta})$ with initial condition (s, ζ) in the sense of Definition C.1 for each $(s, \zeta) \in \mathbb{R}_+ \times \mathfrak{P}_0$

and

($\mathfrak{P}_0\text{-smoothing}$). $\mu_t^{s,\zeta} \in \mathfrak{P}_0$ for all $0 \leq s < t$, $\zeta \in \mathcal{P}_0$.

Then, there is a nonlinear Markov process $(P_{s,\zeta})_{s \geq 0, \zeta \in \mathcal{P}_0}$ with (5.4), consisting of solution path laws to (1.4). In this case, the uniqueness-assertion for $P_{s,\zeta}$ of Theorem 5.2 holds for all $(s, \zeta) \in \mathbb{R}_+ \times \mathfrak{P}_0$.

In the next subsection we apply this corollary to the NLFPE (1.7) and the associated McKean–Vlasov SDE (3.1).

5.3 Application to Leibenson equation and its Barenblatt solutions

Let

$$\mathcal{P}_0 := \{w^y(\delta, x)dx : y \in \mathbb{R}^d, \delta \geq 0\},$$

where with some abuse of notation we write $w^y(0, x)dx = \delta_y(dx)$. For each $\zeta \in \mathcal{P}_0$, the pair $(\delta, y) \in [0, \infty) \times \mathbb{R}^d$ such that $\zeta = w^y(\delta, x)dx$ is unique.

Theorem 5.6. *Let $d \geq 2, p > \frac{d}{d-1}, q > 0$ such that $q(p-1) > 1$. If $p < 2$, assume additionally that $q(p-1) > \frac{2-p+d}{d}(> 1)$.*

- (i) *Let $(s, \zeta) \in [0, \infty) \times \mathcal{P}_0$, $\zeta = w^y(\delta, x)dx$. The set of solution laws to the McKean–Vlasov SDE (3.1) with one-dimensional time marginals $w^y(\delta + t - s, x)dx$, $t \geq s$, and initial condition (s, ζ) contains exactly one element $P_{s,\zeta}$. The family $(P_{s,\zeta})_{s \geq 0, \zeta \in \mathcal{P}_0}$ is a nonlinear Markov process in the sense of Definition 5.1. In particular, this nonlinear Markov process is uniquely determined by (3.1) and $w^y(t)$, $y \in \mathbb{R}^d$, $t \geq 0$.*

- (ii) *$(P_{s,\zeta})_{s \geq 0, \zeta \in \mathcal{P}_0}$ is time-homogeneous, i.e. $P_{s,\zeta} = P_{0,\zeta} \circ \hat{\Pi}_s^{-1}$ for all $(s, \zeta) \in [0, \infty) \times \mathcal{P}_0$, where*

$$\hat{\Pi}_s : C([0, \infty), \mathbb{R}^d) \rightarrow C([s, \infty), \mathbb{R}^d), \quad \hat{\Pi}_s : (\omega(t))_{t \geq 0} \mapsto (\omega(t-s))_{t \geq s}. \quad (5.5)$$

Moreover, for $\zeta = w^y(\delta, x)dx$, we have $P_{0,\zeta} = P_{0,y} \circ (\Pi_\delta^0)^{-1}$ (the map Π_δ^0 is defined in (5.7) below).

Remark 5.7. *The p -Laplace-case of this result, i.e. $p > 2, q = 1$, was proven in [3].*

We stress that in the following proof the times s and δ are not related. In particular, for the initial condition $\zeta = w^y(\delta, x)dx$, we not only consider the initial pair (δ, ζ) , but necessarily *any* (s, ζ) , $s \geq 0$. Set

$$\mathfrak{P}_0 := \{w^y(\delta, x)dx : y \in \mathbb{R}^d, \delta > 0\} = \mathcal{P}_0 \setminus \{\delta_y : y \in \mathbb{R}^d\}. \quad (5.6)$$

A crucial ingredient of the proof is Theorem 7.4, which is formulated and proven in Section 7 below.

Proof of Theorem 5.6.

(i) Setting, for $\zeta = w^y(\delta, x)dx$, $\delta \geq 0$, $y \in \mathbb{R}^d$,

$$\mu_t^{s,\zeta} := w^y(\delta + t - s, x)dx, \quad t \geq s,$$

it is straightforward to check that the family of probability measures $\{\mu_t^{s,\zeta}\}_{s \geq 0, t \geq s, \zeta \in \mathcal{P}_0}$ has the flow property (5.3). From Lemmas 4.3 and 2.6 and from what was shown in the proof of Theorem 4.4 it follows that $(\mu_t^{s,\zeta})_{t \geq s}$ is a weakly continuous probability solution to the nonlinear FPE (1.7) with initial condition (s, ζ) in the sense of Definition C.2. By Theorem 7.4, Lemma 5.4 and Remark 7.6 below, condition $(\mathfrak{P}_0\text{-lin}_{\text{ex}})$ holds. Moreover, by definition of \mathcal{P}_0 and \mathfrak{P}_0 , we have $\mu_t^{s,\zeta} \in \mathfrak{P}_0$ for all (s, t, ζ) such that either $s \leq t$ and $\zeta \in \mathfrak{P}_0$ or $s < t$ and $\zeta \in \mathcal{P}_0$. Thus, Corollary 5.5 applies and yields:

There is a nonlinear Markov process $(P_{s,\zeta})_{(s,\zeta) \in [0,\infty) \times \mathcal{P}_0}$, $P_{s,\zeta} \in \mathcal{P}(C([s,\infty), \mathbb{R}^d))$, such that

- (I) $P_{s,\zeta} \circ (\pi_t^s)^{-1} = \mu_t^{s,\zeta}$, $t \geq s$;
- (II) $P_{s,\zeta}$ is a solution path law to the McKean–Vlasov SDE (3.1) on $[s, \infty)$;
- (III) For $s \geq 0$ and $\zeta \in \mathfrak{P}_0$, $P_{s,\zeta}$ is *unique* with properties (I)–(II).

Therefore, since $\mathcal{P}_0 \setminus \mathfrak{P}_0 = \{\delta_y : y \in \mathbb{R}^d\}$, it remains to prove the following claim.

Claim. For $(s, y) \in [0, \infty) \times \mathbb{R}^d$, there is a *unique* path law P_{s,δ_y} with properties (I)–(II).

Proof of Claim. Let P^1, P^2 have properties (I)–(II) for $s \geq 0, \zeta = \delta_y, y \in \mathbb{R}^d$. For $s, r \geq 0$, we define the map $\Pi_r^s : C([s, \infty), \mathbb{R}^d) \rightarrow C([s, \infty), \mathbb{R}^d)$ via

$$\Pi_r^s : \omega(t)_{t \geq s} \mapsto \omega(t+r)_{t \geq s}. \quad (5.7)$$

For any $r > 0$, $i \in \{1, 2\}$, we have

$$(P^i \circ (\Pi_r^s)^{-1}) \circ (\pi_t^s)^{-1} = P^i \circ (\pi_{t+r}^s)^{-1} = \mu_{t+r}^{s,\delta_y} = w^y(t+r-s, x)dx, \quad \forall t \geq s. \quad (5.8)$$

It is straightforward to check that $P^i \circ (\Pi_r^s)^{-1} \in \mathcal{P}(C([s, \infty), \mathbb{R}^d))$, $i \in \{1, 2\}$, is a solution law to (3.1) with initial condition $(s, w^y(r, x)dx)$. As $w^y(r, x)dx \in \mathfrak{P}_0$, (5.8) and (III) yield

$$P^1 \circ (\Pi_r^s)^{-1} = P^2 \circ (\Pi_r^s)^{-1}.$$

Now let $s \leq u_1 < \dots < u_n$, $n \in \mathbb{N}$. First assume $u_1 > s$. Then for $i \in \{1, 2\}$

$$P^i \circ (\pi_{u_1}^s, \dots, \pi_{u_n}^s)^{-1} = (P^i \circ (\Pi_{u_1-s}^s)^{-1}) \circ (\pi_s^s, \dots, \pi_{u_n+s-u_1}^s)^{-1}$$

and by the previous part of the proof the right hand side coincides for $i = 1$ and $i = 2$, since $u_1 - s > 0$. Now assume $s = u_1 < \dots < u_n$. Then, since $P^i \circ (\pi_s^s)^{-1} = \delta_y$, we find

$$P^i \circ (\pi_{u_1}^s, \dots, \pi_{u_n}^s)^{-1} = \delta_y \otimes (P^i \circ (\pi_{u_2}^s, \dots, \pi_{u_n}^s)^{-1})$$

(where $\mu \otimes \nu$ denotes the product measure of the measures μ and ν). Since $u_2 > s$, the argument of the first case again yields that the right hand side coincides for $i = 1$ and $i = 2$. Hence we have proven

$$P^1 \circ (\pi_{u_1}^s, \dots, \pi_{u_n}^s)^{-1} = P^2 \circ (\pi_{u_1}^s, \dots, \pi_{u_n}^s)^{-1}$$

for all $s \leq u_1 < \dots < u_n$, $n \in \mathbb{N}$, i.e. $P^1 = P^2$, which proves the claim and, thereby, the assertion.

(ii) First note

$$P_{s,\zeta} \circ (\pi_t^s)^{-1} = \mu_t^{s,\zeta} = \mu_{t-s}^{0,\zeta} = (P_{0,\zeta} \circ (\hat{\Pi}_s)^{-1}) \circ (\pi_t^s)^{-1}, \quad \forall t \geq s,$$

with $\hat{\Pi}_s$ as in (5.5). Both $P_{s,\zeta}$ and $P_{0,\zeta} \circ (\hat{\Pi}_s)^{-1}$ are solution laws to (3.1) on $[s, \infty)$ with initial condition (s, ζ) , hence, by (i), they coincide.

For the final statement, note that for $\zeta = w^y(\delta, x)dx$ the measures $P_{0,\zeta}$ and $P_{0,y} \circ (\Pi_\delta^0)^{-1}$ have identical one-dimensional time marginals $w^y(\delta+t, x)dx$, $t \geq 0$, and both are solution laws to (3.1) with initial condition $(0, w^y(\delta, x)dx)$. Thus the assertion follows from the uniqueness assertion in (i). \square

Remark 5.8. Theorem 5.6 (ii) implies that the nonlinear Markov process $(P_{s,\zeta})_{(s,\zeta) \in \mathbb{R}_+ \times \mathcal{P}_0}$ from (i) is uniquely determined by

$$(P_y)_{y \in \mathbb{R}^d}, \quad P_y := P_{0,\delta_y}.$$

Therefore, we also refer to $(P_y)_{y \in \mathbb{R}^d}$ as the unique nonlinear Markov process determined by (3.1) and the one-dimensional time marginals $w^y(t)$, $y \in \mathbb{R}^d$, $t \geq 0$. Note that P_y is the path law of the solution X^y from Corollary 4.5.

Finally, we arrive at the following definition.

Definition 5.9. Let $d \geq 2$, $p > \frac{d}{d-1}$, $q > 0$ such that $q(p-1) > 1$. If $p < 2$, assume additionally that $q(p-1) > \frac{2-p+d}{d} (> 1)$. We call the nonlinear Markov process $(P_y)_{y \in \mathbb{R}^d}$ from the previous remark the *Leibenson process*.

In particular, the Barenblatt solutions to (1.1) have a probabilistic representation as the one-dimensional time marginal densities of the Leibenson process, which they uniquely determine as a nonlinear Markov process. The construction of this process and the derivation of the corresponding stochastic equation, i.e. (3.1), is completely analogous to the linear special case $(p, q) = (2, 1)$ of the heat equation and Brownian motion, compare Section 5.1.

The special case $d \geq 2$, $p > 2$, $q = 1$ was studied in [52]; therein the corresponding nonlinear Markov process was called *p-Brownian motion*.

6 The Leibenson process is a functional of Brownian motion

In Section 5 we constructed the Leibenson process, which consists of path laws of probabilistically weak solutions (X, W) to the McKean–Vlasov SDE (3.1) with $u(t) = w^y(t + \delta)$, $t > 0$, $\delta \geq 0$. In this section, we prove that for $\delta > 0$ these solutions are in fact probabilistically strong solutions, that is, X is a measurable adapted functional of the driving Brownian motion W and its initial condition $X(0)$. For simplicity, we only consider the case $y = 0$. The general case $y \in \mathbb{R}^d$ can be proved analogously.

First, we reduce the question of whether there exists a probabilistically strong solution to (1.2) to a problem for an ordinary SDE. For $\delta > 0$, we set for $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$

$$w_\delta(t, x) := w^0(t + \delta, x), \quad \varrho_\delta(t, x) := |\nabla w_\delta(t, x)|^{p-2} w_\delta(t, x)^{(p-1)(q-1)}, \quad R_\delta(t) := R(t + \delta),$$

where again w^0 denotes the Barenblatt solution from (4.1), and R is defined as in (4.2). As computed in the proof of Theorem 4.4, we have

$$C_\varrho(t, x) = C_\varrho(t + \delta) f_+(t + \delta, x) |x|^{\frac{p-2}{p-1}}, \quad (t, x) \in \mathbb{R}_+ \times \mathbb{R}^d, \quad (6.1)$$

where we set

$$C_\varrho(t + \delta) := \left(\frac{\gamma \kappa p}{p-1} \right)^{p-2} (t + \delta)^{-\frac{d(p-2)}{\beta} - \frac{p(p-2)}{\beta(p-1)} - \frac{d(p-1)(q-1)}{\beta}}, \quad (t, x) \in \mathbb{R}_+ \times \mathbb{R}^d. \quad (6.2)$$

Consider the ordinary SDE

$$\begin{cases} dX(t) &= q^{p-1} \nabla \varrho_\delta(t, X(t)) dt + \sqrt{2q^{p-1} \varrho_\delta(t, X(t))} dW(t), \\ \mathcal{L}_{X(t)}(dx) &= w_\delta(t, x) dx. \end{cases} \quad (6.3)$$

Note that (6.3) is *equivalent* to (3.1) and $\mathcal{L}_{X(t)}(dx) = w_\delta(t, x)dx$; it is obtained by fixing $u = w_\delta$ in the coefficients in (3.1).

Now we define when a probabilistically weak solution to (3.1) (this includes (6.3)) is said to be strong. We introduce the following notation. Let $t \in [0, \infty)$. We set $\mathcal{B}_t(C([0, \infty); \mathbb{R}^d)) := \sigma(\pi_s : s \in [0, t])$ and, correspondingly, $\mathcal{B}_t(C([0, \infty); \mathbb{R}^d)_0) := \sigma(\pi_s : s \in [0, t]) \cap C([0, \infty); \mathbb{R}^d)_0$. Furthermore, P^W denotes the Wiener measure on $(C([0, \infty); \mathbb{R}^d)_0, \mathcal{B}(C([0, \infty); \mathbb{R}^d)_0))$.

Definition 6.1. A probabilistically weak solution $(X, W, (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}))$ to (1.2) with initial condition $\nu \in \mathcal{P}(\mathbb{R}^d)$ is called a *probabilistically strong solution*, if there exists

$$F_\nu : \mathbb{R}^d \times C([0, \infty); \mathbb{R}^d)_0 \rightarrow C([0, \infty); \mathbb{R}^d),$$

which is $\overline{\mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(C([0, \infty); \mathbb{R}^d)_0)}^{\nu \otimes P^W} / \mathcal{B}(C([0, \infty); \mathbb{R}^d))$ -measurable such that for ν -a.e. $x \in \mathbb{R}^d$, $F_\nu(x, \cdot)$ is $\overline{\mathcal{B}_t(C([0, \infty); \mathbb{R}^d)_0)}^{P^W} / \mathcal{B}_t(C([0, \infty); \mathbb{R}^d))$ -measurable for every $t \in [0, \infty)$, and

$$X = F_\nu(X(0), W) \text{ } \mathbb{P}\text{-a.s.}$$

Here, $\overline{\mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(C([0, \infty); \mathbb{R}^d)_0)}^{\nu \otimes P^W}$ denotes the completion of $\mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(C([0, \infty); \mathbb{R}^d)_0)$ with respect to $\nu \otimes P^W$, and $\overline{\mathcal{B}_t(C([0, \infty); \mathbb{R}^d)_0)}^{P^W}$ denotes the completion of $\mathcal{B}_t(C([0, \infty); \mathbb{R}^d)_0)$ with respect to P^W in $\mathcal{B}(C([0, \infty); \mathbb{R}^d)_0)$.

Before we state the main result of this section, we recall the definition of pathwise uniqueness for (6.3).

Definition 6.2. *Pathwise uniqueness* holds for (6.3) if for every pair of probabilistically weak solutions $(X, W), (Y, W)$ to (6.3) defined on the same stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ with respect to the same (\mathcal{F}_t) -Brownian motion W and with $X(0) = Y(0)$ \mathbb{P} -a.s., we have $\sup_t |X(t) - Y(t)| = 0$ \mathbb{P} -a.s.

In Theorem 4.4, we proved there exists a probabilistically weak solution (X, W) to (6.3) such that $\mathcal{L}_{X(t)} = w_\delta(t, x)dx, t \in \mathbb{R}_+$. The following main result of this section shows that under quite general conditions these solutions are actually strong.

Theorem 6.3 (The Leibenson process is a functional of Brownian motion). *Let $\delta > 0$. Let $d \geq 2$, $p > 1, q > 0$ such that $p > \frac{d}{d-1}$ and $q > \frac{|p-2|+d}{d(p-1)} \left(> \frac{1}{p-1} \right)$. Then, the probabilistically weak solution to (6.3) constructed in Theorem 4.4 is a probabilistically strong solution. Furthermore, for all $T > 0$, pathwise uniqueness holds among weak solutions (X, W) to (6.3) on $[0, T]$.*

Remark 6.4. Note that Theorem 6.3 includes the case of p -Brownian motion, i.e., $p > 2, q = 1$.

As already mentioned above, the case of arbitrary initial points $y \in \mathbb{R}^d$ can be obtained analogously:

Corollary 6.5. *Let $y \in \mathbb{R}^d$ and $\delta > 0$. Let $d \geq 2$, $p > 1, q > 0$ such that $p > \frac{d}{d-1}$ and $q > \frac{|p-2|+d}{d(p-1)}$. Then, the probabilistically weak solution to (6.3) constructed in Corollary 4.5 is a probabilistically strong solution. Furthermore, for all $T > 0$, pathwise uniqueness holds among weak solutions (X, W) to (6.3) on $[0, T]$.*

6.1 The restricted Yamada–Watanabe theorem

Our result is based on the *restricted Yamada–Watanabe theorem*, which we will recall here in a form tailored to the ordinary SDE (6.3). For the general formulation for stochastic partial differential equations in the variational framework, we refer to [33, Theorem 1.3.1].

Theorem 6.6. *Assume*

(i) There exists a probabilistically weak solution (X, W) to (6.3);

(ii) Pathwise uniqueness holds for (6.3).

Then the weak solution (X, W) to (6.3) from (i) is a probabilistically strong solution.

Proof. The assertion follows from [33, Theorem 1.3.1] by setting (using the notation of [33])

$$P_{w_\delta(0,x)dx} := \{Q \in \mathcal{P}(C([0, \infty); \mathbb{R}^d)) : Q \circ \pi_t^{-1} = w_\delta(t, x)dx \quad \forall t \in \mathbb{R}_+\}.$$

□

By Theorem 4.4, Assumption (i) of Theorem 6.6 is satisfied. Therefore, the rest of the proof of Theorem 6.3 boils down to a pathwise uniqueness result for (6.3). The proof of Theorem 6.3 is carried out in Section 6.4. In the following two subsections we present some technical preparations.

6.2 About the Schwartz distributional derivative of $\nabla \varrho_\delta(t, \cdot)$

Here we discuss regularity and integrability aspects of the drift coefficient of (6.3), which will be essential for the proof of Theorem 6.3.

Lemma 6.7. *Let $d \geq 2$, $p > 1$, $q > 0$ such that $q(p-1) > 1$ and $p > \frac{d}{d-1}$. Let $\delta > 0$ and $t \geq 0$. Then,*

$$\nabla \varrho_\delta(t, \cdot) \in BV(\mathbb{R}^d; \mathbb{R}^d)$$

with Schwartz distributional derivatives given by the finite signed Borel measures (on \mathbb{R}^d)

$$\partial_i \partial_j \varrho_\delta(t, A) = \int_A g_\delta^{i,j}(t, z) dz + C(t + \delta)^{-\frac{d(p-2)}{\beta} - \frac{p}{\beta} - \frac{d(p-1)(q-1)}{\beta}} \int_{A \cap \partial B_{R_\delta(t)}(0)} z_i z_j S(dz) \quad \forall A \in \mathcal{B}(\mathbb{R}^d), \quad (6.4)$$

where $1 \leq i, j \leq d$, $C > 0$ is a constant not depending on t , and $g_\delta^{i,j}(t, \cdot) \in L^1(\mathbb{R}^d)$ denotes the classical second order derivative of $\varrho_\delta(t, \cdot)$ with respect to the coordinates x_i and x_j on $\mathbb{R}^d \setminus (\partial B_{R_\delta(t)}(0) \cup \{0\})$.

Proof. In the proof of Theorem 4.4 we already showed that $\varrho_\delta(t, \cdot) \in L^1(\mathbb{R}^d)$ and that $\varrho_\delta(t, \cdot)$ is weakly differentiable on \mathbb{R}^d with spatial weak derivative (and also classical derivative in $x \in \mathbb{R}^d \setminus (\partial B_{R_\delta(t)}(0) \cup \{0\})$)

$$\nabla \varrho_\delta(t, x) = C_\varrho(t + \delta) \left(\frac{p-2}{p-1} f_+(t + \delta, x) |x|^{-\frac{p}{p-1}} - \frac{\kappa p}{p-1} (t + \delta)^{-\frac{p}{\beta(p-1)}} \mathbb{1}_{B_{R_\delta(t)}(0)}(x) \right) x, \quad (6.5)$$

where C_ϱ is as in (6.2). Note that $\nabla \varrho_\delta(t, \cdot)$ has compact support, so that $\nabla \varrho_\delta(t, \cdot) \in L^1(\mathbb{R}^d)$. Furthermore, it is clear that $\varrho_\delta(t, \cdot)$ is twice continuously differentiable on $\mathbb{R}^d \setminus (\partial B_{R_\delta(t)}(0) \cup \{0\})$. Recall that we denote the ordinary second-order derivative of $\varrho_\delta(t, \cdot)$ in directions x_i and x_j by $g_\delta^{i,j}(t, \cdot)$. A direct computation yields for all $x \in \mathbb{R}^d \setminus (\partial B_{R_\delta(t)}(0) \cup \{0\})$,

$$\begin{aligned} g_\delta^{i,j}(t, x) &= C_\varrho(t + \delta) \mathbb{1}_{B_{R_\delta(t)}(0)}(x) \left[\frac{p-2}{p-1} \left(-\frac{\kappa p}{p-1} (t + \delta)^{-\frac{p}{\beta(p-1)}} |x|^{\frac{2-p}{p-1}} x_i \right) |x|^{-\frac{p}{p-1}} x_j \right. \\ &\quad \left. + \frac{p-2}{p-1} f(t + \delta, x) \left(-\frac{p}{p-1} |x|^{\frac{2-3p}{p-1}} x_i x_j + |x|^{-\frac{p}{p-1}} \delta_{ij} \right) - \frac{\kappa p}{p-1} (t + \delta)^{-\frac{p}{\beta(p-1)}} \delta_{ij} \right] \\ &= C_\varrho(t + \delta) \mathbb{1}_{B_{R_\delta(t)}(0)}(x) \left[-\frac{\kappa p(p-2)}{(p-1)^2} (t + \delta)^{-\frac{p}{\beta(p-1)}} |x|^{-2} x_i x_j \right. \\ &\quad \left. + \frac{p-2}{p-1} f(t + \delta, x) \left(-\frac{p}{p-1} |x|^{\frac{2-3p}{p-1}} x_i x_j + |x|^{-\frac{p}{p-1}} \delta_{ij} \right) - \frac{\kappa p}{p-1} (t + \delta)^{-\frac{p}{\beta(p-1)}} \delta_{ij} \right], \end{aligned}$$

where δ_{ij} denotes the usual Kronecker delta. Note that $g_\delta^{i,j}(t, \cdot) \in L^1(\mathbb{R}^d)$, since $p > \frac{d}{d-1}$. Now, we are prepared to prove (6.4). Let $\varphi \in C_c^\infty(\mathbb{R}^d)$. By Lebesgue's dominated convergence theorem, we have

$$\int_{B_{R_\delta(t)}(0)} \partial_i \varrho_\delta(t, x) \partial_j \varphi(x) dx = \lim_{\varepsilon \downarrow 0} \int_{B_{R_\delta(t)-\varepsilon}(0) \setminus B_\varepsilon(0)} \partial_i \varrho_\delta(t, x) \partial_j \varphi(x) dx.$$

Let $0 < \varepsilon < R_\delta(t)/2$. Then integration by parts yields

$$\begin{aligned} \int_{B_{R_\delta(t)-\varepsilon}(0) \setminus B_\varepsilon(0)} \partial_i \varrho_\delta(t, x) \partial_j \varphi(x) dx &= - \int_{B_{R_\delta(t)-\varepsilon}(0) \setminus B_\varepsilon(0)} g_\delta^{i,j}(t, x) \varphi(x) dx \\ &\quad + \int_{\partial B_{R_\delta(t)-\varepsilon}(0)} \partial_i \varrho_\delta(t, x) \varphi(x) \frac{x_j}{R_\delta(t) - \varepsilon} dS(x) - \int_{\partial B_\varepsilon(0)} \partial_i \varrho_\delta(t, x) \varphi(x) \frac{x_j}{\varepsilon} dS(x). \end{aligned} \quad (6.6)$$

First, the third summand on the right-hand side in (6.6) vanishes as $\varepsilon \rightarrow 0$, since

$$\left| \int_{\partial B_\varepsilon(0)} \partial_i \varrho_\delta(t, x) \varphi(x) \frac{x_j}{\varepsilon} dS(x) \right| \lesssim_\delta \left(\varepsilon^{-\frac{1}{p-1}} + \varepsilon \right) \varepsilon^{d-1} \rightarrow 0, \text{ as } \varepsilon \rightarrow 0,$$

where we used $p > \frac{d}{d-1}$. Secondly, due to (6.5), the second summand on the right-hand side in (6.6) can be written as

$$\begin{aligned} \int_{\partial B_{R_\delta(t)-\varepsilon}(0)} \frac{\partial_i \varrho_\delta(t, x) x_j}{R_\delta(t) - \varepsilon} \varphi(x) dS(x) &= \frac{C_\varrho(t + \delta)(p-2)}{(R_\delta(t) - \varepsilon)(p-1)} \int_{\partial B_{R_\delta(t)-\varepsilon}(0)} f_+(t + \delta, x) |x|^{-\frac{p}{p-1}} x_i x_j \varphi(x) dS(x) \\ &\quad - \frac{C_\varrho(t + \delta)(t + \delta)^{-\frac{p}{p-1}} \kappa p}{(p-1)(R_\delta(t) - \varepsilon)} \int_{\partial B_{R_\delta(t)-\varepsilon}(0)} x_i x_j \varphi(x) dS(x). \end{aligned}$$

Using the transformation $x \mapsto (R_\delta(t) - \varepsilon)x$, it is easy to see that the first summand on the right-hand side of the last equality vanishes as $\varepsilon \rightarrow 0$. A similar argument regarding the last integral on the right-hand side yields

$$\int_{\partial B_{R_\delta(t)-\varepsilon}(0)} x_i x_j \varphi(x) dS(x) \rightarrow \int_{\partial B_{R_\delta(t)}(0)} x_i x_j \varphi(x) dS(x), \text{ as } \varepsilon \rightarrow 0.$$

Hence, letting $\varepsilon \rightarrow 0$ in (6.6), we conclude that the spatial Schwartz distributional derivatives of the components of $\nabla \varrho_\delta$ can be written as

$$\mathcal{D}' \langle \partial_j \partial_i \varrho_\delta(t, \cdot), \varphi \rangle_{\mathcal{D}} = \int_{\mathbb{R}^d} g_\delta^{i,j}(t, x) \varphi(x) dx + \frac{C_\varrho(t + \delta)(t + \delta)^{-\frac{p}{p-1}} \kappa p}{(p-1)(R_\delta(t) - \varepsilon)} \int_{\partial B_{R_\delta(t)}(0)} x_i x_j \varphi(x) dS(x),$$

for $\varphi \in C_c^\infty(\mathbb{R}^d)$ and $1 \leq i, j \leq d$, where $\mathcal{D}' \langle \cdot, \cdot \rangle_{\mathcal{D}}$ denotes the usual action of Schwartz distributions on \mathbb{R}^d , denoted by \mathcal{D}' , on $\mathcal{D} \equiv C_c^\infty(\mathbb{R}^d)$. This finishes the proof. \square

We have the following lemma.

Lemma 6.8. *Let $d \geq 2$ with $\delta > 0$. Let $p > 1, q > 0$ such that $q(p-1) > 1$ and $p > \frac{d}{d-1}$. Then*

$$\int_0^T \int \mathbb{M} |D_x \nabla \varrho_\delta(t, \cdot)|(x) w_\delta(t, x) dx dt < \infty \quad \forall T \in (0, \infty), \quad (6.7)$$

where $D_x \nabla \varrho_\delta(t, \cdot)$ denotes the matrix consisting of all second order Schwartz distributional derivatives of $\varrho_\delta(t, \cdot)$ in form of finite signed Borel measures, according to Lemma 6.7.

Proof. Let $T > 0$. According to and using the notation of Lemma 6.7, we may estimate as follows.

$$\int_0^T \int M |D_x \nabla \varrho_\delta(t, \cdot)|(x) w_\delta(t, x) dx dt \lesssim_\delta \int_0^T \int M \left(\max_{1 \leq i, j \leq d} |g_\delta^{i, j}(t, \cdot)| \right) (x) w_\delta(t, x) dx dt \quad (6.8)$$

$$+ \int_0^T \int M |S(\cdot \cap \partial B_{R_\delta(t)}(0))|(x) w_\delta(t, x) dx dt. \quad (6.9)$$

According to Lemma 6.7, we have

$$\max_{1 \leq i, j \leq d} |g_\delta^{i, j}(t, x)| \lesssim_\delta 1 + |x|^{-\frac{p}{p-1}}.$$

Note that $|\cdot|^{-\frac{p}{p-1}} \in A_1$, since $p > \frac{d}{d-1}$. Hence, we obtain

$$\begin{aligned} \int_0^T \int M \left(\max_{1 \leq i, j \leq d} |g_\delta^{i, j}(t, \cdot)| \right) (x) w_\delta(t, x) dx dt &\lesssim_\delta 1 + \int_0^T \int M(|\cdot|^{-\frac{p}{p-1}})(x) w_\delta(t, x) dx dt \\ &\lesssim 1 + \int_0^T \int |x|^{-\frac{p}{p-1}} w_\delta(t, x) dx dt < \infty. \end{aligned}$$

Furthermore, by Cauchy's surface area formula, we have

$$S(B_r(x) \cap \partial B_{R_\delta(t)}(0)) \leq S(\partial B_r(0)) \lesssim r^{d-1}. \quad (6.10)$$

Also, for each $x \in \mathbb{R}^d$ such that $|x| \neq R_\delta(t)$ and all $0 < r \leq ||x| - R_\delta(t)|$ we have

$$S(B_r(x) \cap \partial B_{R_\delta(t)}(0)) = 0.$$

Hence, for all $x \in \mathbb{R}^d$ such that $|x| \neq R_\delta(t)$, we obtain

$$M |S(\cdot \cap \partial B_{R_\delta(t)}(0))|(x) \lesssim \frac{1}{||x| - R_\delta(t)|}. \quad (6.11)$$

Claim: For all $\rho \geq 0$ there exists $C_\rho \geq 1$ such that for all $r \in [0, 1]$

$$1 - r^\rho \leq C_\rho(1 - r). \quad (6.12)$$

Proof of Claim: By l'Hospital's rule, the function

$$[0, 1) \ni r \mapsto \frac{1 - r^\rho}{1 - r}$$

can be extended to a continuous function on the interval $[0, 1]$, which clearly has a maximum $C_\rho \geq 1$ on $[0, 1]$. \square

Using (6.11), the transformation $x \mapsto R_\delta(t)x$, integration in polar coordinates, and (6.12), we obtain

$$\begin{aligned} \int_0^T \int M |S(\cdot \cap \partial B_{R_\delta(t)}(0))|(x) w_\delta(t, x) dx dt &\lesssim_\delta \int_0^T \int (||x| - R_\delta(t)|)^{-1} f_+(t + \delta, x)^\gamma dx dt \\ &\lesssim_\delta \int_0^1 \frac{r^{d-1}}{1 - r} (1 - r^{\frac{p}{p-1}})^\gamma dr \lesssim \int_0^1 (1 - r)^{\gamma-1} dr < \infty. \end{aligned}$$

Hence, since $\gamma > 0$, (6.7) follows. \square

Remark 6.9. At first sight, the estimate (6.11) seems quite rough. But, in fact, elementary geometrical considerations show that, for $d = 3$ and fixed $t \geq 0$,

$$M |D_x \nabla \varrho_\delta(t, \cdot)|(x) \cong \frac{1}{|x| ||x| - R_\delta(t)|},$$

for all x in a neighborhood of $\partial B_{R_\delta(t)}(0)$.

6.3 Weak derivatives (of powers) of $\sqrt{\varrho_\delta}$

Here we discuss regularity and integrability aspects of the diffusion coefficient of (6.3), which will be essential in the proof of Theorem 6.3.

Lemma 6.10. *Let $d \geq 2$, $p > 1$, $q > 0$ with $q(p-1) > 1$, and*

$$\frac{(p-2)(1+\gamma)}{2(p-1)} > 1-d. \quad (6.13)$$

Then for each $t > 0$, $\varrho_\delta^{\frac{1+\gamma}{2}}(t, \cdot)$ is weakly differentiable.

Remark 6.11. *Consider*

(i)

$$p > \frac{2d}{2d-1} \text{ and } q > \frac{2d(p-1) - p - (p-2)(p-1)}{(p-1)(2d(p-1) - p)}; \quad (6.14)$$

(ii)

$$(2 \geq) \frac{2d}{2d-1} > p > 1 \text{ and } q < \frac{2d(p-1) - p - (p-2)(p-1)}{(p-1)(2d(p-1) - p)}. \quad (6.15)$$

In fact, (6.13) holds if and only if (6.14) or (6.15) is fulfilled.

Proof. By (6.1), clearly $\varrho_\delta^{\frac{1+\gamma}{2}}(t, \cdot) \in L^1(\mathbb{R}^d)$. Note that $\varrho_\delta^{\frac{1+\gamma}{2}}(t, \cdot)$ is (infinitely many times) continuously differentiable on $\mathbb{R}^d \setminus (\partial B_{R_\delta(t)}(0) \cup \{0\})$. Let $\varphi \in C_c^\infty(\mathbb{R}^d)$ and $0 < \varepsilon < R_\delta(t)/2$. By Lebesgue's dominated convergence, we obtain

$$\int \varrho_\delta^{\frac{1+\gamma}{2}}(t, x) \partial_i \varphi(x) dx = \lim_{\varepsilon \rightarrow 0} \int_{B_{R_\delta(t)-\varepsilon} \setminus B_\varepsilon(0)} \varrho_\delta^{\frac{1+\gamma}{2}}(t, x) \partial_i \varphi(x) dx.$$

Integrating by parts, we obtain for $1 \leq i \leq d$

$$\begin{aligned} \int_{B_{R_\delta(t)-\varepsilon} \setminus B_\varepsilon(0)} \varrho_\delta^{\frac{1+\gamma}{2}}(t, x) \partial_i \varphi(x) dx &= - \int_{B_{R_\delta(t)-\varepsilon} \setminus B_\varepsilon(0)} \partial_i \varrho_\delta^{\frac{1+\gamma}{2}}(t, x) \varphi(x) dx \\ &\quad + \int_{\partial B_{R_\delta(t)-\varepsilon}} \varrho_\delta^{\frac{1+\gamma}{2}}(t, x) \varphi(x) \frac{x_i}{R_\delta(t) - \varepsilon} S(dx) \\ &\quad - \int_{\partial B_\varepsilon(0)} \varrho_\delta^{\frac{1+\gamma}{2}}(t, x) \varphi(x) \frac{x_i}{R_\delta(t) - \varepsilon} S(dx), \end{aligned}$$

where $\partial_i \varrho_\delta^{\frac{1+\gamma}{2}}(t, \cdot)$ denotes the dx -a.e. existing pointwise partial derivative of $\varrho_\delta^{\frac{1+\gamma}{2}}(t, \cdot)$ in direction x_i . By Lebesgue's dominated convergence theorem

$$\begin{aligned} &\int_{\partial B_{R_\delta(t)-\varepsilon}} \varrho_\delta^{\frac{1+\gamma}{2}}(t, x) \varphi(x) \frac{x_i}{R_\delta(t) - \varepsilon} S(dx) \\ &= \int_{\partial B_1(0)} \varrho_\delta^{\frac{1+\gamma}{2}}(t, (R_\delta(t) - \varepsilon)x) \varphi((R_\delta(t) - \varepsilon)x) x_i (R_\delta(t) - \varepsilon)^{d-1} S(dx) \xrightarrow{\varepsilon \rightarrow 0} 0. \end{aligned} \quad (6.16)$$

Furthermore, by (6.13),

$$\int_{\partial B_\varepsilon(0)} \varrho_\delta^{\frac{1+\gamma}{2}}(t, x) \varphi(x) \frac{x_i}{R_\delta(t) - \varepsilon} S(dx) \lesssim_\delta \varepsilon^{\frac{(p-2)(1+\gamma)}{2(p-1)} + d-1} \xrightarrow{\varepsilon \rightarrow 0} 0. \quad (6.17)$$

Using chain rule as well as (6.1) and (6.5), an easy calculation yields that the pointwise partial derivatives of $\varrho_\delta^{\frac{1+\gamma}{2}}(t, \cdot)$ on $\mathbb{R}^d \setminus (\partial B_{R_\delta(t)}(0) \cup \{0\})$ satisfy the bound

$$|\partial_i \varrho_\delta^{\frac{1+\gamma}{2}}(t, x)| \lesssim_\delta f_+(t + \delta, x)^{\frac{\gamma-1}{2}} |x|^{\frac{(p-2)(1+\gamma)+2}{2(p-1)}} + f_+(t + \delta, x)^{\frac{\gamma+1}{2}} |x|^{\frac{(p-2)(1+\gamma)}{2(p-1)}-1}, \quad (6.18)$$

where $i \in \{1, \dots, d\}$. Under condition (6.13), it is clear that $\partial_i \varrho_\delta^{\frac{1+\gamma}{2}}(t, \cdot) \in L^1(\mathbb{R}^d)$. Hence, Lebesgue's dominated convergence theorem yields

$$\int_{B_{R_\delta(t)-\varepsilon} \setminus B_\varepsilon(0)} \partial_i \varrho_\delta^{\frac{1+\gamma}{2}}(t, x) \varphi(x) dx \xrightarrow{\varepsilon \rightarrow 0} \int \partial_i \varrho_\delta^{\frac{1+\gamma}{2}}(t, x) \varphi(x) dx. \quad (6.19)$$

This completes the proof. \square

We have the following lemma.

Lemma 6.12. *Let $d \geq 2$, $\delta > 0$, and $p > 1, q > 0$ be such that $p > \frac{d}{d-1}$ and $q > \frac{|p-2|+d}{d(p-1)}$. Then*

$$\int_0^T \int_{\mathbb{R}^d} \left(M |\nabla \varrho_\delta^{\frac{1+\gamma}{2}}(t, \cdot)|(x) \right)^2 |x|^{-\frac{p-2}{q(p-1)-1}} dx dt < \infty \quad \forall T > 0. \quad (6.20)$$

Proof. Let $T > 0$. Note

$$\frac{2d(p-1) - p - (p-2)(p-1)}{(p-1)(2d(p-1) - p)} < \frac{|p-2|+d}{d(p-1)}. \quad (6.21)$$

Hence, by Lemma 6.10 and Remark 6.11, we have $\nabla \varrho_\delta^{\frac{1+\gamma}{2}}(t, \cdot) \in L^1(\mathbb{R}^d)$. Since $q > \frac{|p-2|+d}{d(p-1)}$, we have

$$|\cdot|^{-\frac{p-2}{q(p-1)-1}} \in A_2$$

(for the definition of the Muckenhoupt class A_2 , we refer to Appendix D.2). Hence, in order to show (6.20), it is sufficient to prove

$$\int_0^T \int_{\mathbb{R}^d} \left| \nabla \varrho_\delta^{\frac{1+\gamma}{2}}(t, x) \right|^2 |x|^{-\frac{p-2}{q(p-1)-1}} dx dt < \infty.$$

By the proof of Lemma 6.10 (more precisely (6.18))

$$\left| \nabla \varrho_\delta^{\frac{1+\gamma}{2}}(t, x) \right| \lesssim_\delta f_+(t + \delta, x)^{\frac{\gamma-1}{2}} |x|^{\frac{(p-2)(1+\gamma)+2}{2(p-1)}} + f_+(t + \delta, x)^{\frac{\gamma+1}{2}} |x|^{\frac{(p-2)(1+\gamma)}{2(p-1)}-1}, \quad (6.22)$$

whence

$$\begin{aligned} & \int_0^T \int_{\mathbb{R}^d} \left| \nabla \varrho_\delta^{\frac{1+\gamma}{2}} \right|^2 |x|^{-\frac{p-2}{q(p-1)-1}} dx dt \\ & \lesssim \int_0^T \int_{\mathbb{R}^d} f_+(t + \delta, x)^{\gamma-1} |x|^{\frac{p}{p-1}} + f_+(t + \delta, x)^{1+\gamma} |x|^{-\frac{p}{p-1}} dx dt < \infty, \end{aligned}$$

where we used $p > \frac{d}{d-1}$ and $\gamma > 0$. This completes the proof. \square

6.4 Proof of Theorem 6.3

We show that Theorem 6.3 follows from Theorem 6.6. Therefore, we need to show that both conditions of the latter theorem are fulfilled.

In fact, assumption (ii) of Theorem 6.6 follows from Theorem 4.4, and (i) follows from Theorem 6.13 below.

The proof of the following theorem is inspired by the technique in [54].

Theorem 6.13. *Let $\delta > 0$ and $T > 0$. Let $d \geq 2$, $p > 1$, $q > 0$ such that $p > \frac{d}{d-1}$ and $q > \frac{|p-2|+d}{d(p-1)}$. Let (X, W) and (Y, W) be two probabilistically weak solutions to (6.3) on $[0, T]$ on a common stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$ with respect to the same (\mathcal{F}_t) -Brownian motion $(W(t))_{t \in [0, T]}$ such that $X(0) = Y(0)$ \mathbb{P} -a.s. Then $X = Y$ \mathbb{P} -a.s.*

Proof. Due to Proposition E.2, we may assume without loss of generality

$$\nabla \varrho_\delta(t, 0) = 0, \quad \forall t \in [0, T]. \quad (6.23)$$

From Lemma 6.7 and the proof of Theorem 4.4, it follows that $\sqrt{\varrho_\delta}, |\nabla \varrho_\delta| \in L^1([0, T] \times \mathbb{R}^d)$. For $n \in \mathbb{N}$ we set

$$\tau_n := \inf\{t \in [0, \infty) : |X(t)| \vee |Y(t)| \geq n\}.$$

Note that τ_n is an (\mathcal{F}_t) -stopping time. For $\varepsilon > 0$, we set $h_\varepsilon(x) := \ln(1 + |x|^2/\varepsilon^2)$, $x \in \mathbb{R}^d$. Clearly, for all $x \in \mathbb{R}^d$

$$|\nabla h_\varepsilon(x)| \lesssim \frac{1}{|x| + \varepsilon}, \quad |D^2 h_\varepsilon(x)| \lesssim \frac{1}{|x|^2 + \varepsilon^2}.$$

Let $n \in \mathbb{N}, \varepsilon > 0$. Using Itô's formula, we obtain for each $t \in [0, T]$ (cf. proof of [54, Theorem 1.1])

$$\begin{aligned} & \mathbb{E} \left[\ln \left(1 + \frac{|X(t \wedge \tau_n) - Y(t \wedge \tau_n)|^2}{\varepsilon^2} \right) \right] \\ & \lesssim \mathbb{E} \left[\int_0^{t \wedge \tau_n} \frac{2 \langle \nabla \varrho_\delta(s, X(s)) - \nabla \varrho_\delta(s, Y(s)), X(s) - Y(s) \rangle_{\mathbb{R}^d} + |\sqrt{\varrho_\delta}(s, X(s)) - \sqrt{\varrho_\delta}(s, Y(s))|^2}{|X(s) - Y(s)|^2 + \varepsilon^2} ds \right] \\ & \lesssim \mathbb{E} \left[\int_0^{t \wedge \tau_n} \frac{|\nabla \varrho_\delta(s, X(s)) - \nabla \varrho_\delta(s, Y(s))|}{|X(s) - Y(s)| + \varepsilon} ds \right] + \mathbb{E} \left[\int_0^{t \wedge \tau_n} \frac{|\sqrt{\varrho_\delta}(s, X(s)) - \sqrt{\varrho_\delta}(s, Y(s))|^2}{|X(s) - Y(s)|^2 + \varepsilon^2} ds \right] \\ & =: (\star_1) + (\star_2). \end{aligned} \quad (6.24)$$

Claim: For all $x, z \in \mathbb{R}^d$

$$|\nabla \varrho_\delta(t, x) - \nabla \varrho_\delta(t, z)| \lesssim (M |D_x \nabla \varrho_\delta(t, \cdot)|(x) + M |D_x \nabla \varrho_\delta(t, \cdot)|(z)) |x - z|. \quad (6.25)$$

Proof of Claim: By (6.23) and the continuity of $\nabla \varrho_\delta(t, \cdot)$ on $\mathbb{R}^d \setminus (\partial B_{R(t)}(0) \cup \{0\})$, it is easy to see that for any standard radial Dirac sequence $(\varphi_\varepsilon)_{\varepsilon > 0}$,

$$(\nabla \varrho_\delta(t, \cdot) * \varphi_\varepsilon)(x) \rightarrow \nabla \varrho_\delta(t, x) \quad \forall x \in \mathbb{R}^d \setminus \partial B_{R(t)}(0), \quad (6.26)$$

where the convolution is meant componentwise.

Furthermore, $\partial B_{R_\delta(t)}(0)$ consists of all *approximate jump points* of $\varrho_\delta(t, \cdot)$ in the sense of [2, Definition 3.67]. Hence, by Lemma D.2 and Remark D.3, the assertion follows. Indeed: Fix an arbitrary $t \geq 0$. For all $x_0 \in B_{R_\delta(t)}(0)$, we set (in the notation of [2]), $\nu := \frac{x_0}{|x_0|}$, $a := -\frac{\kappa p}{p-1}(t+\delta)^{-\frac{p}{p-1}} x_0$, $b := 0$. Furthermore, let $B_r^-(x, \nu) := \{y \in B_r(x) : \langle y - x, \nu \rangle < 0\}$ and $B_r^+(x, \nu) := \{y \in B_r(x) : \langle y - x, \nu \rangle > 0\}$. Then a simple transformation of the Lebesgue integral shows (here $|A|$ denotes the Lebesgue measure of a set $A \in \mathcal{B}(\mathbb{R}^d)$)

$$\frac{1}{|B_r^-(x_0, \nu)|} \int_{B_r^-(x_0, \nu)} |\nabla \varrho_\delta(t, y) - a| dy \lesssim \int_{B_1^-(0, \nu)} |\nabla \varrho_\delta(t, x_0 + ry) - a| dy \rightarrow 0, \text{ as } r \downarrow 0,$$

where we used the fact that $\nabla \varrho_\delta(t, y) \rightarrow a$, whenever $y \rightarrow x_0$ with $y \in B_{R_\delta(t)}(0)$, and that $\nabla \varrho_\delta(t, \cdot)$ is bounded on $\mathbb{R}^d \setminus B_\varepsilon(0)$, for all $\varepsilon > 0$. Furthermore, since $\text{supp}(|\nabla \varrho_\delta(t, \cdot)|) = \overline{B_{R_\delta(t)}(0)}$, we have

$$\frac{1}{|B_r^+(x_0, \nu)|} \int_{B_r^+(x_0, \nu)} |\nabla \varrho_\delta(t, y) - b| dy = 0 \quad \forall r > 0.$$

This completes the proof of the claim. \square

Proof of Theorem 6.13 continued. Using (6.25), we may estimate (\star_1) by

$$\begin{aligned} (\star_1) &\lesssim \mathbb{E} \left[\int_0^{t \wedge \tau_n} M |D_x \nabla \varrho_\delta(s, \cdot)|(X(s)) + M |D_x \nabla \varrho_\delta(s, \cdot)|(Y(s)) ds \right] \\ &\leq 2 \int_0^T \int_{B_n(0)} M |D_x \nabla \varrho_\delta(s, \cdot)|(x) w_\delta(s, x) dx ds. \end{aligned} \quad (6.27)$$

By Lemma 6.8, the right-hand side of (6.27) is finite.

Before we estimate (\star_2) , we remark that due to the mean value theorem, the following inequalities are true (under the convention $0 \cdot \infty = 0$) for all $x, y \in \mathbb{R}^d$ if $p > 2$, and for all $x, y \in \mathbb{R}^d \setminus \{0\}$ if $p < 2$:

$$\begin{aligned} |\sqrt{\varrho_\delta}(t, x) - \sqrt{\varrho_\delta}(t, y)|^2 &= \left| \left(\varrho_\delta^{\frac{1+\gamma}{2}}(t, x) \right)^{\frac{1}{1+\gamma}} - \left(\varrho_\delta^{\frac{1+\gamma}{2}}(t, y) \right)^{\frac{1}{1+\gamma}} \right|^2 \\ &\lesssim \left(\int_0^1 \left[\theta \max \left(\varrho_\delta^{\frac{1+\gamma}{2}}(t, x), \varrho_\delta^{\frac{1+\gamma}{2}}(t, y) \right) \right]^{-\frac{\gamma}{1+\gamma}} d\theta \right)^2 \cdot \left| \varrho_\delta^{\frac{1+\gamma}{2}}(t, x) - \varrho_\delta^{\frac{1+\gamma}{2}}(t, y) \right|^2 \\ &\lesssim_\delta \left(\max \left(w_\delta(t, x) |x|^{\frac{(p-2)}{q(p-1)-1}}, w_\delta(t, y) |y|^{\frac{(p-2)}{q(p-1)-1}} \right) \right)^{-1} \cdot \left| \varrho_\delta^{\frac{1+\gamma}{2}}(t, x) - \varrho_\delta^{\frac{1+\gamma}{2}}(t, y) \right|^2. \end{aligned} \quad (6.28)$$

We estimate (\star_2) as follows.

$$\begin{aligned} &\mathbb{E} \left[\int_0^T \frac{|\sqrt{\varrho_\delta}(t, X(t)) - \sqrt{\varrho_\delta}(t, Y(t))|^2}{|X(t) - Y(t)|^2 + \varepsilon^2} dt \right] \\ &\lesssim_\delta \mathbb{E} \left[\int_0^T \frac{\left| \varrho_\delta^{\frac{1+\gamma}{2}}(t, X(t)) - \varrho_\delta^{\frac{1+\gamma}{2}}(t, Y(t)) \right|^2}{(|X(t) - Y(t)|^2 + \varepsilon^2) \max \left(w_\delta(t, X(t)) |X(t)|^{\frac{(p-2)}{q(p-1)-1}}, w_\delta(t, Y(t)) |Y(t)|^{\frac{(p-2)}{q(p-1)-1}} \right)} dt \right]. \end{aligned} \quad (6.29)$$

Recall that, by Lemma 6.10, $\varrho_\delta^{\frac{1+\gamma}{2}}(t, \cdot)$ is weakly differentiable. Clearly, $\varrho_\delta^{\frac{1+\gamma}{2}}(t, \cdot) \in C(\mathbb{R}^d)$ if $p > 2$, and $\varrho_\delta^{\frac{1+\gamma}{2}}(t, \cdot) \in C(\mathbb{R}^d \setminus \{0\})$ if $p < 2$. Here, we emphasize that in the latter case, since X, Y satisfy the integrability condition (3.2), respectively, $\{\min(|X|, |Y|) = 0\}$ is a $dt \otimes \mathbb{P}$ -zero set. Hence, by (6.25) where we formally replace $\nabla \varrho_\delta(t, \cdot)$ by $\varrho_\delta^{\frac{1+\gamma}{2}}(t, \cdot)$ (which, in the case $p > 2$, then holds for all $x, z \in \mathbb{R}^d$, and, if $p < 2$, for all $x, z \in \mathbb{R}^d \setminus \{0\}$ due to the function's respective continuity), we obtain

$$\begin{aligned} &\mathbb{E} \left[\int_0^T \frac{\left| \varrho_\delta^{\frac{1+\gamma}{2}}(t, X(t)) - \varrho_\delta^{\frac{1+\gamma}{2}}(t, Y(t)) \right|^2}{(|X(t) - Y(t)|^2 + \varepsilon^2) \max \left(w_\delta(t, X(t)) |X(t)|^{\frac{(p-2)}{q(p-1)-1}}, w_\delta(t, Y(t)) |Y(t)|^{\frac{(p-2)}{q(p-1)-1}} \right)} dt \right] \\ &\lesssim \mathbb{E} \left[\int_0^T \frac{\left(M |\nabla \varrho_\delta^{\frac{1+\gamma}{2}}(t, \cdot)|(X(t)) \right)^2}{w_\delta(t, X(t)) |X(t)|^{\frac{(p-2)}{q(p-1)-1}}} dt \right] + \mathbb{E} \left[\int_0^T \frac{\left(M |\nabla \varrho_\delta^{\frac{1+\gamma}{2}}(t, \cdot)|(Y(t)) \right)^2}{w_\delta(t, Y(t)) |Y(t)|^{\frac{(p-2)}{q(p-1)-1}}} dt \right] \\ &\lesssim 2 \int_0^T \int_{\mathbb{R}^d} \left(M |\nabla \varrho_\delta^{\frac{1+\gamma}{2}}(t, \cdot)|(x) \right)^2 |x|^{-\frac{p-2}{q(p-1)-1}} dx dt. \end{aligned} \quad (6.30)$$

Note that, according to Lemma 6.12, the right-hand side of (6.30) is finite. Hence, by the monotone convergence theorem, we obtain

$$\mathbb{E} \left[\lim_{\varepsilon \downarrow 0} \ln \left(1 + \frac{|X(t \wedge \tau_n) - Y(t \wedge \tau_n)|^2}{\varepsilon^2} \right) \right] = \lim_{\varepsilon \downarrow 0} \mathbb{E} \left[\ln \left(1 + \frac{|X(t \wedge \tau_n) - Y(t \wedge \tau_n)|^2}{\varepsilon^2} \right) \right] < \infty.$$

Since X and Y are \mathbb{P} -a.e. continuous, we conclude \mathbb{P} -a.s.

$$X(t \wedge \tau_n) = Y(t \wedge \tau_n) \quad \forall t \in [0, T]. \quad (6.31)$$

Again, by the \mathbb{P} -a.s. continuity of X, Y , $\tau_n \rightarrow \infty$, as $n \rightarrow \infty$, \mathbb{P} -a.s. Therefore, letting $n \rightarrow \infty$ in (6.31), we obtain \mathbb{P} -a.s.

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

This completes the proof. \square

7 A restricted uniqueness result for the linearized Leibenson Fokker–Planck equation

The goal of this section is to prove that, in the situation of Theorem 5.6, condition $(\mathfrak{P}_0\text{-lin}_{\text{ex}})$ from Corollary 5.5, more precisely its equivalent condition from Lemma 5.4, is satisfied with \mathfrak{P}_0 as defined in (5.6), see Theorem 7.4. As in Section 6, we set for $\delta > 0$

$$w_\delta(t, x) := w(t + \delta, x), \quad \varrho_\delta(t, x) := |\nabla w_\delta(t, x)|^{p-2} w_\delta(t, x)^{(p-1)(q-1)},$$

where again w denotes the Barenblatt solution from (4.1) with initial datum δ_0 . As computed in the proof of Theorem 4.4, we have

$$\varrho_\delta(t, x) \cong (t + \delta)^{-\frac{d}{\beta}(p-2) - \frac{p(p-2)}{\beta(p-1)} - \frac{d}{\beta}(q-1)(p-1)} f_+(t + \delta, x) |x|^{\frac{p-2}{p-1}}.$$

Here we study the w_δ -linearized version of (1.7), i.e. the linear FPE

$$\partial_t u = q^{p-1} (\Delta(\varrho_\delta u) - \operatorname{div}(\nabla \varrho_\delta u)), \quad (7.1)$$

obtained by fixing w_δ in the nonlinear variable of the coefficients in (1.7). We set $Q_T := (0, T) \times \mathbb{R}^d$ for $T > 0$.

Definition 7.1. A (distributional) solution to (7.1) on $(0, T)$ with initial condition $\nu \in \mathcal{M}_b^+$ is a function $u \in L^1(Q_T)$ such that $t \mapsto u(t, x)dx$ is a weakly continuous curve of (signed) Borel measures with

$$\int_0^T \int_{\mathbb{R}^d} (\varrho_\delta + |\nabla \varrho_\delta|) |u| dx dt < \infty,$$

such that for all $\psi \in C_c^\infty(\mathbb{R}^d)$

$$\int_{\mathbb{R}^d} \psi u(t) dx - \int_{\mathbb{R}^d} \psi d\nu = \int_0^t \int_{\mathbb{R}^d} (\varrho_\delta \Delta \psi + \nabla \varrho_\delta \cdot \nabla \psi) u dx dt, \quad \forall 0 < t < T. \quad (7.2)$$

u is called *probability solution*, if $u \geq 0$ and $\nu, u(t, x)dx \in \mathcal{P}$ for all $t \in (0, T)$.

This definition is consistent with Definition C.1 for the linear Fokker–Planck equation (7.1) (except for the additional assumption $u \in L^1(Q_T)$, and on $(0, T)$ instead of \mathbb{R}_+). Clearly, $u(t, x) = w_\delta(t, x)$ is a probability solution with initial datum $w(\delta, x)dx$ to (7.1).

Remark 7.2. An equivalent condition to (7.2) is

$$\int_{(0, T) \times \mathbb{R}^d} \left(\partial_t \varphi + q^{p-1} \operatorname{div}(\varrho_\delta \nabla \varphi) \right) u dx dt + \int_{\mathbb{R}^d} \varphi(0) d\nu = 0 \quad (7.3)$$

for all $\varphi \in C_c^\infty([0, T) \times \mathbb{R}^d)$.

We collect basic properties of w_δ and ϱ_δ used in the sequel, of which some were already stated in Lemma 4.3.

Lemma 7.3. *Let $\delta > 0$. Then w_δ and ϱ_δ are nonnegative functions with the following properties.*

- (i) $w_\delta \in \bigcap_{R>0} C_c([0, R] \times \mathbb{R}^d) \cap L^\infty([0, \infty) \times \mathbb{R}^d)$.
- (ii) For $p \geq 2$, we have $\varrho_\delta \in \bigcap_{R>0} C_c([0, R] \times \mathbb{R}^d)$ and $\{x \in \mathbb{R}^d : \varrho_\delta(t, x) > 0\} = \{x \in \mathbb{R}^d : 0 \leq |x| < (\frac{C}{\kappa})^{\frac{p-1}{p}}(t + \delta)^{\frac{1}{\beta}}\}$, for all $t \geq 0$, where $\beta = p + d(q(p-1) - 1)$ as above.
- (iii) $\nabla \varrho_\delta \in L^\infty_{\text{loc}}([0, \infty), L^r(\mathbb{R}^d; \mathbb{R}^d))$ for $r \in [1, d(p-1))$ (this interval is nonempty if $p > \frac{d+1}{d}$).

Proof. The properties of w_δ follow from Lemma 4.3. Nonnegativity and regularity of ϱ_δ is obvious from its definition and it is clear that the support of $\varrho_\delta(t)$ equals the support of $w_\delta(t)$, which is stated in Lemma 4.3. Regarding (iii), by taking into account (4.7) and considering d -dimensional spherical coordinates, it suffices to note that $\int_0^1 r^{-\frac{r}{p-1} + d-1} dr < \infty$ holds. Indeed, $-\frac{r}{p-1} + d > 0 \iff r < d(p-1)$. \square

Consider the class

$$\mathcal{A}_{\delta, T} := \{u \in L^1 \cap L^\infty(Q_T) : t \mapsto u(t, x) dx \in C([0, T], \mathcal{P}), \exists C \geq 1 : u(t) \leq C w_\delta(t) \text{ dx-a.s., } \forall t \in [0, T]\},$$

where $C([0, T], \mathcal{P})$ is the set of weakly continuous maps $t \mapsto \mu_t \in \mathcal{P}$. Clearly $w_\delta \in \mathcal{A}_{\delta, T}$. Comparing with Theorem 5.2 (more precisely, with Corollary 5.5 and Lemma 5.4) and the situation of Theorem 5.6, the following result is the remaining piece needed to close the proof of Theorem 5.6. For simplicity, we only formulate the result for $y = 0$ and $s = 0$. The proof for the general case is identical (note that we need this result indeed for all $s \geq 0$ and $y \in \mathbb{R}^d$).

Theorem 7.4. *Let $T \in (0, \infty)$, $\delta > 0$, $d \geq 2$, $q > 0$ and $q(p-1) > 1$. Assume additionally either*

- (i) $p > 2$ and $q(p-1) < 1 + d(p-1)^2$

or

- (ii) $p > \frac{d}{d-1}$ and $q > \frac{|p-2|+d}{d(p-1)}$.

Then w_δ is the unique distributional solution to (7.1) on $(0, T)$ in the sense of Definition 7.1 with initial condition $w_\delta(0, x) dx$ in $\mathcal{A}_{\delta, T}$.

Notably, as shown below, the proof of the assertion under assumption (i) follows by analytic methods, while for case (ii) the probabilistic tools developed in Section 6 are employed. Note that (i) does not imply (ii) and (ii) does not imply (i).

Remark 7.5. *In fact, a stronger uniqueness assertion holds in the situation of the previous theorem. Precisely, for every initial condition $\nu \in \mathcal{P}$ there is at most one distributional solutions to (7.1) in $\mathcal{A}_{\delta, T}$ with initial condition ν . In the proofs below we prove this stronger assertion.*

Remark 7.6. *If $p > 2$, then $p > \frac{d}{d-1}$ for $d \geq 2$ and we have*

$$\frac{|p-2|+d}{d} < 1 + d(p-1)^2.$$

Hence for $p > 2$ the assertion of the previous theorem holds for all $q > 0$ such that $q(p-1) > 1$.

Remark 7.7. *Even though this restricted uniqueness result concerns a linear PDE, we could not obtain its proof by a standard result from the literature. Note that the diffusion coefficient ϱ_δ is degenerate (it is compactly supported in \bar{Q}_T), which renders the assertion rather delicate. For instance, we cannot apply [16, Thm.9.8.2], since this requires $\varrho_\delta^{-\frac{1}{2}} \nabla \varrho_\delta$ to be bounded, which is not true. Our proofs below heavily use the explicit form of w_δ and ϱ_δ .*

7.1 Analytic proof of Theorem 7.4 (i)

In this case, the proof is an extension of the proof of Theorem 6.5 in [3]. For the convenience of the reader, we give all details, also for those parts which are very similar to [3].

Let $T \in (0, \infty)$, $\delta > 0$, $d \geq 2$, $p > 2$, and $q(p-1) > 1$, $q < d(p-1) + \frac{1}{p-1}$. For convenience, below we omit the absolute factor q^{p-1} from the RHS of (7.1). It is clear that this factor does not influence the proof at all.

Remark 7.8. When $u \in (L^1 \cap L^\infty)(\mathbb{R}^d)$, then in Remark 7.2, since

$$\operatorname{div}(\varrho_\delta \nabla \varphi) = \varrho_\delta \Delta \varphi + \nabla \varrho_\delta \cdot \nabla \varphi$$

in $L^1(Q_T)$, by a standard localization argument we can replace $C_c^\infty([0, T] \times \mathbb{R}^d)$ by

$$C_{c,b}^\infty([0, T] \times \mathbb{R}^d) := \{\varphi \in C_b^\infty([0, T] \times \mathbb{R}^d) \mid \exists \lambda > 0 \text{ such that } \varphi(t, x) = 0, \forall (t, x) \in [T - \lambda, T] \times \mathbb{R}^d\}.$$

We give more details on this claim in Appendix A. In the forthcoming proof, we abbreviate partial time derivatives $\frac{df}{dt}$ by f_t . Constants depending only on absolute constants are denoted $C_i, i = 1, 2, \dots$

Proof of Theorem 7.4 (i). Since there is no change in the argument, we prove the stronger assertion mentioned in Remark 7.5. Let $u, \bar{u} \in \mathcal{A}_{\delta, T}$ be distributional solutions to (7.1) with any common initial condition $\nu \in \mathcal{P}$ and set

$$v := u - \bar{u}.$$

Since $u, \bar{u} \in \mathcal{A}_{\delta, T}$, there are nonnegative measurable maps $g, \bar{g} : Q_T \rightarrow \mathbb{R}$ such that $u(t, x)dx = g(t, x)w_\delta(t, x)dx$ and $\bar{u}(t, x)dx = \bar{g}(t, x)w_\delta(t, x)dx$ for all $t \in [0, T]$, and $g(t, x), \bar{g}(t, x) \leq C$ dtdx-a.s. for some $C > 0$. Hence $v(t, x) = h(t, x)w_\delta(t, x)$ dtdx-a.s. for $h := g - \bar{g}$. In particular, $-C \leq h \leq C$.

v is a (signed) distributional solution to (7.1) in the sense of Definition 7.1 with initial condition the zero measure. Let $f \in C_c^\infty(Q_T)$, set and for $\varepsilon \in (0, 1)$ consider the equation

$$\begin{aligned} (\varphi_\varepsilon)_t + \operatorname{div}((\varrho_\delta + \varepsilon)\nabla \varphi_\varepsilon) &= f \text{ in } Q_T \\ \varphi_\varepsilon(T, x) &= 0, \quad x \in \mathbb{R}^d. \end{aligned} \tag{7.4}$$

By standard existence theory for linear parabolic equations (see, e.g., [19, Thm.10.9, p.341]) it follows that equation (7.4) has a unique solution

$$\varphi_\varepsilon \in C([0, T]; L^2) \cap L^2(0, T; H^1),$$

with $(\varphi_\varepsilon)_t \in L^2(0, T; H^{-1})$. Moreover, we have

$$|\varphi_\varepsilon(t)|_2^2 + \int_0^T \int_{\mathbb{R}^d} (\varrho_\delta(t, x) + \varepsilon) |\nabla \varphi_\varepsilon(t, x)|^2 dt dx \leq \int_0^T \int_{\mathbb{R}^d} |f(t, x)|^2 dt dx. \tag{7.5}$$

In particular, (7.4) is equivalent to

$$\begin{aligned} (\varphi_\varepsilon)_t + (\varrho_\delta + \varepsilon)\Delta \varphi_\varepsilon + \nabla \varrho_\delta \cdot \nabla \varphi_\varepsilon &= f \text{ on } Q_T, \\ \varphi_\varepsilon(T, x) &= 0, \quad x \in \mathbb{R}^d. \end{aligned} \tag{7.6}$$

Claim 1. We have $\varphi_\varepsilon \in W_2^{1,2}(Q_T)$, that is,

$$(\varphi_\varepsilon)_t, \partial_i \varphi_\varepsilon, \partial_i \partial_j \varphi_\varepsilon \in L^2(Q_T), \quad i, j = 1, \dots, d.$$

Proof of Claim 1. We set $g_\varepsilon = \nabla \varrho_\delta \cdot \nabla \varphi_\varepsilon$. By (7.5) we have $\varphi_\varepsilon, |\nabla \varphi_\varepsilon| \in L^2(Q_T)$. By Lemma 7.3 (iii), we know that $\nabla \varrho_\delta \in L^\infty(0, T; L^r)$, $r \in [1, d(p-1))$. Fix $r \in (d, d(p-1))$ (this interval is nonempty, since $p > 2$). Then, by Hölder's inequality,

$$g_\varepsilon \in L^{\gamma_1}(Q_T), \quad \text{for } \gamma_1 := \frac{2r}{2+r} \in (1, 2). \tag{7.7}$$

Taking into account that, by (7.6),

$$(\varphi_\varepsilon)_t + (\varrho_\varepsilon + \varepsilon)\Delta\varphi_\varepsilon = f - g_\varepsilon \in L^{\gamma_1}(Q_T),$$

we get by [42, Theorem 9.1, p. 341] that $\varphi_\varepsilon \in W_{\gamma_1}^{1,2}(Q_T)$, that is,

$$(\varphi_\varepsilon)_t, \partial_i \varphi_\varepsilon, \partial_i \partial_j \varphi_\varepsilon \in L^{\gamma_1}(Q_T), \quad i, j = 1, \dots, d.$$

On the other hand, by the Sobolev–Nirenbergh–Gagliardo theorem (see, e.g., [19, p.283]) we have (recall $d \geq 2$)

$$|\nabla \varphi_\varepsilon| \in L^{\alpha_1}(Q_T) \quad \text{for } \alpha_1 := \frac{\gamma_1 d}{d - \gamma_1},$$

and this yields as above

$$g_\varepsilon \in L^{\gamma_2}(Q_T), \quad \text{for } \gamma_2 := \frac{\alpha_1 r}{\alpha_1 + r} \in (1, \infty).$$

Then, again by [42, Theorem 9.1, p. 341], it follows that $\varphi_\varepsilon \in W_{\gamma_2}^{1,2}(Q_T)$ and, therefore, again by Sobolev–Nirenbergh–Gagliardo theorem,

$$|\nabla \varphi_\varepsilon| \in L^{\alpha_2}(Q_T) \quad \text{for } \alpha_2 := \frac{\gamma_2 d}{d - \gamma_2}.$$

Continuing, we obtain sequences γ_i, α_i such that

$$g_\varepsilon \in L^{\gamma_i}(Q_T), \quad \varphi_\varepsilon \in W_{\gamma_i}^{1,2}(Q_T), \tag{7.8}$$

and

$$\begin{aligned} \gamma_1 &= \frac{2r}{2+r}, \quad \gamma_{i+1} = \frac{\alpha_i r}{\alpha_i + r}, \quad i \in \mathbb{N}, \\ \alpha_i &= \frac{\gamma_i d}{d - \gamma_i}, \quad \text{as long as } \gamma_i < d. \end{aligned}$$

This yields the recursive formula

$$\gamma_1 = \frac{2r}{2+r} \in (1, 2), \quad \gamma_{i+1} = \gamma_i \frac{rd}{\gamma_i d + r(d - \gamma_i)}, \quad i \in \mathbb{N}, \quad \text{as long as } \gamma_i < d.$$

This iterative procedure stops after the smallest i such that $\gamma_i \geq d$. Since $r > d$, we have for all i with $\gamma_{i+1} < d$ that

$$\gamma_{i+1} = \gamma_i \frac{rd}{rd + \gamma_i(d - r)} > \gamma_i. \tag{7.9}$$

Suppose

$$\gamma_i < 2 \quad \text{for all } i \in \mathbb{N}. \tag{7.10}$$

Then, by (7.9) there exists

$$\gamma := \lim_{i \rightarrow \infty} \gamma_i \leq 2$$

and, passing to the limit in (7.9), we obtain

$$\gamma = \gamma \frac{rd}{rd + \gamma(d - r)} > \gamma.$$

This contradiction implies that (7.10) is wrong, so there exists $i \in \mathbb{N}$ such that $\gamma_i \geq 2$ and Claim 1 follows by [42, Theorem 9.1, p. 341], because $g_\varepsilon \in L^2(Q_T)$ by (7.8), (7.7) and interpolation, since $\gamma_1 \in (1, 2)$. \square

Next, by the maximum principle

$$\sup_{\varepsilon \in (0,1)} |\varphi_\varepsilon|_{L^\infty(Q_T)} \leq C_0 |f|_{L^\infty(Q_T)}. \quad (7.11)$$

This follows in a standard way from (7.6) by multiplying the equation with $(\varphi_\varepsilon - (T-t)|f|_{L^\infty(Q_T)})_+$ and $(\varphi_\varepsilon + (T-t)|f|_{L^\infty(Q_T)})_-$, respectively, and integrating over Q_T .

Setting $\varrho_\delta^\varepsilon := \varrho_\delta + \varepsilon$ and multiplying (7.6) by $-\varphi_\varepsilon$, integrating over $(t, T) \times \mathbb{R}^d$ and using that $\varphi_\varepsilon(T, \cdot) = 0$, we obtain by Gronwall's lemma (see (7.5))

$$|\varphi_\varepsilon(t)|_{L^2(\mathbb{R}^d)}^2 + \int_t^T \int_{\mathbb{R}^d} \varrho_\delta^\varepsilon |\nabla \varphi_\varepsilon|^2 dx ds \leq C_1 \int_t^T \int_{\mathbb{R}^d} |f|^2 dx ds, \quad \forall t \in (0, T). \quad (7.12)$$

Define for $\lambda \in (0, 1)$

$$\varphi_\varepsilon^\lambda(t, x) := \eta_\lambda(t) \varphi_\varepsilon(t, x), \quad (t, x) \in Q_T,$$

where $\eta_\lambda(t) = \eta\left(\frac{t}{\lambda}\right) \eta\left(\frac{T-t}{\lambda}\right)$ and $\eta \in C^2([0, \infty))$ is such that

$$\eta(r) = 0 \text{ for } r \in [0, 1], \quad \eta(r) = 1 \text{ for } r > 2.$$

Claim 2. We have

$$\int_0^T {}_{H^1} \langle \varrho_\delta^\varepsilon(t) \nabla \varphi_\varepsilon(t), \nabla(\varphi_\varepsilon)_t(t) \rangle_{H^{-1}} dt \leq -\frac{1}{2} \int_{Q_T} (\varrho_\delta)_t |\nabla \varphi_\varepsilon(t, x)|^2 dt dx, \quad (7.13)$$

where ${}_{H^1} \langle \cdot, \cdot \rangle_{H^{-1}}$ is the duality pairing on $H^1(\mathbb{R}^d; \mathbb{R}^d) \times H^{-1}(\mathbb{R}^d; \mathbb{R}^d)$.

Proof of Claim 2. By Claim 1 and its proof we have $\nabla(\varphi_\varepsilon)_t \in L^2((0, T); H^{-1})$ and $\varrho_\delta^\varepsilon \nabla \varphi_\varepsilon \in L^2((0, T); H^1)$, and so the LHS of (7.13) is well defined (indeed, in the proof of Claim 1 we showed $g_\varepsilon = \nabla \varrho_\delta \cdot \nabla \varphi_\varepsilon \in L^2(Q_T)$). Now choose a sequence $\{\varphi_\nu^\varepsilon\} \subset C^1([0, T]; H^1)$ such that $\varphi_\nu^\varepsilon(T, \cdot) = 0$ and, for $\nu \rightarrow 0$,

$$\nabla \varphi_\nu^\varepsilon \rightarrow \nabla \varphi_\varepsilon \text{ strongly in } L^2((0, T); H^1) \quad (7.14)$$

$$\nabla(\varphi_\nu^\varepsilon)_t \rightarrow \nabla(\varphi_\varepsilon)_t \text{ strongly in } L^2((0, T); H^{-1}) \quad (7.15)$$

An example for such a sequence is

$$\varphi_\nu^\varepsilon(t) = (\varphi_\varepsilon * \theta_\nu)(t) - (\varphi_\varepsilon * \theta_\nu)(T),$$

where $\theta_\nu = \theta_\nu(t)$, $\nu > 0$, is a standard mollifier sequence on \mathbb{R} . Here, for technical purposes, we define $\varphi_\varepsilon(r)$ by $\varphi_\varepsilon(0)$ and $\varphi_\varepsilon(T)$ for $r \in (-1, 0)$ and $r \in (T, T+1)$, respectively. Then, since $\nabla \varphi_\nu^\varepsilon(T) = 0$, we have

$$\begin{aligned} \int_0^T {}_{H^1} \langle \varrho_\delta^\varepsilon(t) \nabla \varphi_\nu^\varepsilon(t), \nabla(\varphi_\nu^\varepsilon)_t(t) \rangle_{H^{-1}} dt &= \frac{1}{2} \int_{Q_T} \varrho_\delta^\varepsilon(t, x) |\nabla \varphi_\nu^\varepsilon(t, x)|_t^2 dt dx \\ &= -\frac{1}{2} \int_{\mathbb{R}^d} \varrho_\delta^\varepsilon(0, x) |\nabla \varphi_\nu^\varepsilon(0, x)|^2 dx - \frac{1}{2} \int_{Q_T} \varrho_\delta(t, x) |\nabla \varphi_\nu^\varepsilon(t, x)|^2 dt dx \\ &\leq -\frac{1}{2} \int_{Q_T} (\varrho_\delta)_t(t, x) |\nabla \varphi_\nu^\varepsilon(t, x)|^2 dt dx. \end{aligned}$$

Letting $\nu \rightarrow 0$, we get by (7.14)-(7.15) that (7.13) holds, as claimed. \square

Now, by (7.6) and (7.13), we have

$$\begin{aligned} \int_{Q_T} |(\varphi_\varepsilon)_t(t, x)|^2 dt dx &= \int_0^T {}_{H^1} \langle \varrho_\delta^\varepsilon(t) \nabla \varphi_\varepsilon(t), \nabla(\varphi_\varepsilon)_t(t) \rangle_{H^{-1}} dx + \int_{Q_T} f(t, x) (\varphi_\varepsilon)_t(t, x) dt dx \\ &\leq -\frac{1}{2} \int_{Q_T} (\varrho_\delta)_t(t, x) |\nabla \varphi_\varepsilon(t, x)|^2 dt dx + \int_{Q_T} f(t, x) (\varphi_\varepsilon)_t(t, x) dt dx. \end{aligned}$$

Since $(\varrho_\delta)_t \in L^\infty(Q_T)$, this implies by Young' inequality

$$\int_{Q_T} |(\varphi_\varepsilon)_t|^2 dt dx \leq - \int_{Q_T} (\varrho_\delta)_t |\nabla \varphi_\varepsilon|^2 dt dx + \int_0^T \|f(t)\|_{L^2(\mathbb{R}^d)}^2 dt.$$

From the definition of ϱ_δ we infer

$$-(\varrho_\delta)_t \leq C_2 \delta^{-1} \varrho_\delta, \quad (7.16)$$

and hence we obtain

$$\int_{Q_T} |(\varphi_\varepsilon)_t|^2 dt dx \leq C \int_{Q_T} \varrho_\delta |\nabla \varphi_\varepsilon|^2 dt dx + \int_0^T \|f(t)\|_{L^2(\mathbb{R}^d)}^2 dt, \quad (7.17)$$

where $C > 0$ only depends on p, q, d and δ . Then, by (7.6) we have

$$\begin{aligned} (\varphi_\varepsilon^\lambda)_t + \operatorname{div}((\varrho_\delta + \varepsilon) \nabla \varphi_\varepsilon^\lambda) &= f \eta_\lambda + \eta'_\lambda \varphi_\varepsilon \text{ on } Q_T \\ \varphi_\varepsilon^\lambda(T, x) &= 0, \quad \forall x \in \mathbb{R}^d. \end{aligned} \quad (7.18)$$

From what we derived above, we have $\varphi_\varepsilon^\lambda \in W_2^{2,1}(Q_T)$ and

$$\varphi_\varepsilon^\lambda \in L^2((0, T); H^2) \cap C([0, T]; L^2), \quad (\varphi_\varepsilon^\lambda)_t \in L^2((0, T); L^2), \quad \varphi_\varepsilon^\lambda \in L^\infty(Q_T). \quad (7.19)$$

By Remark 7.8, v satisfies (7.3) for all $\varphi \in C_{c,b}^\infty([0, T] \times \mathbb{R}^d)$. Hence, we infer by density that v satisfies (7.3) also for all functions $\varphi = \varphi_\varepsilon^\lambda$ with properties (7.19). Therefore, we have

$$\int_{Q_T} v ((\varphi_\varepsilon^\lambda)_t + \operatorname{div}(\varrho_\delta \nabla \varphi_\varepsilon^\lambda)) dt dx = 0, \quad \forall \varepsilon, \lambda \in (0, 1) \quad (7.20)$$

(see Appendix B for details).

Next, we get by (7.18) the following equality

$$\frac{1}{2} \|\varphi_\varepsilon^\lambda(t)\|_{L^2(\mathbb{R}^d)}^2 + \varepsilon \int_t^T \|\nabla \varphi_\varepsilon^\lambda(s)\|_{L^2(\mathbb{R}^d)}^2 ds + \int_t^T \int_{\mathbb{R}^d} \varrho_\delta |\nabla \varphi_\varepsilon^\lambda|^2 ds dx = - \int_t^T \int_{\mathbb{R}^d} (f \eta_\lambda + \eta'_\lambda \varphi_\varepsilon) \varphi_\varepsilon^\lambda dt dx.$$

Taking into account (7.11), we get

$$\|\varphi_\varepsilon^\lambda(t)\|_{L^2(\mathbb{R}^d)}^2 + \varepsilon \int_t^T \|\nabla \varphi_\varepsilon^\lambda(s)\|_{L^2(\mathbb{R}^d)}^2 ds + \int_{Q_T} \varrho_\delta |\nabla \varphi_\varepsilon^\lambda|^2 dt dx \leq C_3, \quad \forall \varepsilon, \lambda \in (0, 1). \quad (7.21)$$

Now, we have as in (7.13) that

$$\int_0^T {}_{H^1} \langle \varrho_\delta^\varepsilon(t) \nabla \varphi_\varepsilon^\lambda(t), \nabla (\varphi_\varepsilon^\lambda(t))_t \rangle_{H^{-1}} dt \leq -\frac{1}{2} \int_{Q_T} (\varrho_\delta)_t(t, x) |\nabla \varphi_\varepsilon^\lambda(t, x)|^2 dt dx,$$

and this yields

$$\int_{Q_T} |(\varphi_\varepsilon^\lambda)_t|^2 dt dx \leq - \int_{Q_T} (\varrho_\delta)_t |\nabla \varphi_\varepsilon^\lambda|^2 dt dx + C_4 \int_{Q_T} (|f \eta_\lambda|^2 + |\eta'_\lambda \varphi_\varepsilon|^2) dt dx, \quad \forall \varepsilon \in (0, 1).$$

Then, by (7.12), (7.16) and (7.21) we get

$$\int_{Q_T} |(\varphi_\varepsilon^\lambda)_t|^2 dt dx \leq \int_{Q_T} \varrho_\delta |\nabla \varphi_\varepsilon^\lambda|^2 dt dx + C_4 \int_{Q_T} (|f|^2 + \lambda^{-1} |\varphi_\varepsilon|^2) dt dx \leq \left(1 + \frac{1}{\lambda}\right) C_5, \quad \forall \varepsilon, \lambda \in (0, 1). \quad (7.22)$$

(Here, we have denoted by C_i , $i = 0, 1, 2, \dots$, several positive constants independent of ε, λ .) Now, we fix a sequence $\lambda_n \rightarrow 0$. Hence, by (7.16), (7.17), (7.21) and (7.22) and a diagonal argument, we find a subsequence $\varepsilon_k \rightarrow 0$ such that, for all $n \in \mathbb{N}$ as $k \rightarrow \infty$,

$$\begin{cases} \varphi_{\varepsilon_k} \rightarrow \varphi, \varphi_{\varepsilon_k}^{\lambda_n} \rightarrow \varphi^{\lambda_n} & \text{weakly in } L^2(Q_T) \\ (\varphi_{\varepsilon_k})_t \rightarrow (\varphi)_t, (\varphi_{\varepsilon_k}^{\lambda_n})_t \rightarrow (\varphi^{\lambda_n})_t & \text{weakly in } L^2(Q_T) \\ \operatorname{div}((\varrho_\delta + \varepsilon_k) \nabla \varphi_{\varepsilon_k}^{\lambda_n}) \rightarrow \zeta_{\lambda_n} & \text{weakly in } L^2(Q_T). \end{cases} \quad (7.23)$$

To finish the proof, we need the following three claims.

Claim 3. $t \mapsto u(t)$ and $t \mapsto \bar{u}(t)$ are weakly continuous from $[0, T]$ to $L^2(\mathbb{R}^d)$.

Proof of Claim 3. Since $\sup_{t \in [0, T]} \{|u(t)|_{L^1(\mathbb{R}^d)}, |u(t)|_{L^\infty(\mathbb{R}^d)}\} < \infty$, also $\sup_{t \in [0, T]} \{|u(t)|_{L^2(\mathbb{R}^d)}\} < \infty$. Hence, for every sequence $\{t_n\}_{n \in \mathbb{N}} \subseteq [0, T]$ with limit $t \in [0, T]$ there is a subsequence $\{t_{n_k}\}_{k \in \mathbb{N}}$ such that $u(t_{n_k})$ has a weak limit in $L^2(\mathbb{R}^d)$. Due to the weak continuity (in the sense of measures) of $t \mapsto u(t, x)dx$, this limit is $u(t)$, which implies the $L^2(\mathbb{R}^d)$ -weak continuity of $[0, T] \ni t \mapsto u(t)$. The same argument applies to \bar{u} . \square

Claim 4. We have

$$\lim_{n \rightarrow \infty} \lim_{k \rightarrow \infty} \int_{Q_T} v \eta'_{\lambda_n} \varphi_{\varepsilon_k} dt dx = 0.$$

Proof of Claim 4. We first note that by (7.23) for every $n \in \mathbb{N}$

$$\lim_{k \rightarrow \infty} \int_{Q_T} v \eta'_{\lambda_n} \varphi_{\varepsilon_k} dt dx = \int_{Q_T} v \eta'_{\lambda_n} \varphi dt dx,$$

and that $\varphi \in C([0, T]; L^2)$. Furthermore, by (7.12), $\varphi(T, \cdot) = 0$. Then, for every $\lambda \in (0, 1)$,

$$\begin{aligned} \int_{Q_T} v \eta'_{\lambda} \varphi dt dx &= \frac{1}{\lambda} \int_{\lambda}^{2\lambda} \eta' \left(\frac{t}{\lambda} \right) \int_{\mathbb{R}^d} v(t, x) \varphi(t, x) dt dx + \frac{1}{\lambda} \int_{T-2\lambda}^{T-\lambda} \eta' \left(\frac{T-t}{\lambda} \right) \int_{\mathbb{R}^d} v(t, x) \varphi(t, x) dx \\ &= \int_1^2 \eta'(\tau) d\tau \int_{\mathbb{R}^d} v(\lambda\tau, x) \varphi(\lambda\tau, x) dx + \int_1^2 \eta'(\tau) d\tau \int_{\mathbb{R}^d} v(T - \lambda\tau, x) \varphi(T - \lambda\tau, x) dx, \end{aligned}$$

where, as $\lambda \rightarrow 0$, both terms converge to zero by Claim 3 and Lebesgue's dominated convergence theorem. \square

Claim 5. There is $\alpha \in (0, 1)$ such that for every $\lambda \in (0, 1)$ there exists $C_\lambda \in (0, \infty)$ such that

$$\int_{Q_T} \varepsilon |\Delta \varphi_\varepsilon^\lambda|^{1+\alpha} w_\delta dt dx \leq C_\lambda, \quad \forall \varepsilon \in (0, 1). \quad (7.24)$$

Hence, $\{\varepsilon^{\frac{1}{1+\alpha}} \Delta \varphi_\varepsilon^\lambda \mid \varepsilon \in (0, 1)\}$ is equi-integrable, hence weakly relatively compact in $L^1(Q_T; w_\delta dt dx)$. Therefore, selecting another subsequence if necessary, for every $n \in \mathbb{N}$ as $k \rightarrow \infty$, we find

$$\varepsilon_k \Delta \varphi_{\varepsilon_k}^{\lambda_n} \rightarrow 0 \text{ weakly both in } L^1(Q_T; w_\delta dt dx) \text{ and } L^1(Q_T; h w_\delta dt dx), \quad (7.25)$$

where h is as in the beginning of the proof.

Proof of Claim 5. By the de la Vallée Poussin theorem and a diagonal argument, (7.25) follows from (7.24), so we only have to prove (7.24). To this purpose, fix

$$s \in \left(\frac{2d(p-1)}{p}, d(p-1) \right) \quad (7.26)$$

such that

$$s > \frac{1}{\gamma} = \frac{q(p-1) - 1}{p-1} \quad (7.27)$$

and

$$s > \frac{2d(p-1)}{d(p-1) - (p-2)} \quad (7.28)$$

This is possible since $q < d(p-1) + \frac{1}{p-1}$ implies $\frac{1}{\gamma} < d(p-1)$ and since $d \geq 2, p > 2$, respectively. We also note that the interval in (7.26) is not empty and that its left boundary point is strictly bigger than 2, since $d \geq 2, p > 2$. Then,

$$\frac{p-2}{p-1} \frac{s}{s-2} < d,$$

and so we may choose $\alpha \in (0, \frac{1}{3})$ such that

$$0 < \frac{p-2}{p-1} \frac{(3\alpha+1)s}{s(1-\alpha) - 2(\alpha+1)} < d \quad (7.29)$$

and also such that

$$\alpha < \frac{d(p-1) - p}{d(p-1) + p} \quad \left(< \frac{d(p-1)}{(d+2)(p-2) + d} \right) \quad (7.30)$$

as well as

$$\alpha < \gamma, \quad (7.31)$$

where $\gamma > 0$ is defined in Definition 4.1, and α sufficiently small such that

$$2s\gamma - (3\alpha+1)s > 2(\alpha+1) - s(1-\alpha). \quad (7.32)$$

The final inequality is true for $\alpha > 0$ sufficiently small due to $s > \frac{1}{\gamma}$. Then, we multiply (7.18) by $\text{sign } \Delta\varphi_\varepsilon^\lambda |\Delta\varphi_\varepsilon^\lambda|^\alpha w_\delta$ and integrate over Q_T to obtain after rearranging

$$\begin{aligned} \int_{Q_T} (\varrho_\delta + \varepsilon) |\Delta\varphi_\varepsilon^\lambda|^{1+\alpha} w_\delta \, dt dx &\leq \int_{Q_T} |\nabla \varrho_\delta| |\nabla \varphi_\varepsilon^\lambda| |\Delta\varphi_\varepsilon^\lambda|^\alpha w_\delta \, dt dx \\ &+ \int_{Q_T} |(\varphi_\varepsilon^\lambda)_t| |\Delta\varphi_\varepsilon^\lambda|^\alpha w_\delta \, dt dx + \int_{Q_T} |f\eta_\lambda + \eta'_\lambda \varphi_\varepsilon| |\Delta\varphi_\varepsilon^\lambda|^\alpha w_\delta \, dt dx. \end{aligned} \quad (7.33)$$

We now show, using (7.19), that the right hand side of (7.33) is finite, hence so is its left hand side. Indeed, by Young inequality, for any $r \in (0, 1)$ and a large enough constant $C_r > 1$ (both independent of ε and λ), the second term on the right hand side of (7.33) can be estimated by

$$|(\varphi_\varepsilon^\lambda)_t|_{L^2(Q_T)}^2 + r \int_{Q_T} |\Delta\varphi_\varepsilon^\lambda|^{1+\alpha} \varrho_\delta w_\delta \, dt dx + C_r \int_{Q_T} w_\delta^{\frac{2}{1-\alpha}} \varrho_\delta^{-\frac{2\alpha}{1-\alpha}} \, dt dx. \quad (7.34)$$

Likewise, we can estimate the last term in (7.33) by

$$|f\eta_\lambda + \eta'_\lambda \varphi_\varepsilon|_{L^2(Q_T)}^2 + r \int_{Q_T} |\Delta\varphi_\varepsilon^\lambda|^{1+\alpha} \varrho_\delta w_\delta \, dt dx + C_r \int_{Q_T} w_\delta^{\frac{2}{1-\alpha}} \varrho_\delta^{-\frac{2\alpha}{1-\alpha}} \, dt dx. \quad (7.35)$$

We note that, for $\alpha \in (0, \frac{1}{3})$ satisfying (7.30) and (7.31), the definition of ω_δ and ϱ_δ implies that the last term in (7.34) and (7.35) are finite. Furthermore, by (7.22) the respective first terms in (7.34), (7.35) are uniformly bounded in $\varepsilon \in (0, 1)$. Finally, for any $r \in (0, 1)$ and $C_r > 1$ sufficiently large, both independent of ε and λ , the first term on the right hand side of (7.33) can be estimated by

$$r \int_{Q_T} |\Delta\varphi_\varepsilon^\lambda|^{\alpha+1} \varrho_\delta w_\delta \, dt dx + C_r \int_{Q_T} \frac{|\nabla \varrho_\delta|^{\alpha+1}}{\varrho_\delta^\alpha} |\nabla \varphi_\varepsilon^\lambda|^{\alpha+1} w_\delta \, dx,$$

where the second integral is up to a constant bounded by

$$\int_{Q_T} |\nabla \varrho_\delta|^{\frac{2(\alpha+1)}{1-\alpha}} \varrho_\delta^{-\frac{3\alpha+1}{1-\alpha}} w_\delta^{\frac{2}{1-\alpha}} \, dt dx + \int_{Q_T} |\nabla \varphi_\varepsilon^\lambda|^2 \varrho_\delta \, dt dx$$

of which the second integral is uniformly bounded in $\lambda, \varepsilon \in (0, 1)$ by (7.21) and, by (7.30) and Young inequality, the first integral is up to a constant bounded by

$$\int_{Q_T} |\nabla \varrho_\delta|^s dt dx + \int_{Q_T} w_\delta^{\frac{2s}{s(1-\alpha)-2(\alpha+1)}} \varrho_\delta^{-\frac{(3\alpha+1)s}{s(1-\alpha)-2(\alpha+1)}} dt dx,$$

where by (7.26), (7.29) and (7.28) as well as (7.32) this quantity is finite. Choosing r small enough, we hence get from (7.33) and the nonnegativity of ϱ_δ the estimate (7.24), and Claim 5 follows. \square

Finally, we conclude the proof as follows. We have by Claim 4 and (7.18)

$$\begin{aligned} \int_{Q_T} f v dt dx &= \lim_{n \rightarrow \infty} \lim_{k \rightarrow \infty} \int_{Q_T} (f \eta_{\lambda_n} + \eta'_{\lambda_n} \varphi_{\varepsilon_k}) v dt dx \\ &= \lim_{n \rightarrow \infty} \lim_{k \rightarrow \infty} \int_{Q_T} (\varphi_{\varepsilon_k}^{\lambda_n})_t + \operatorname{div}((\varrho_\delta + \varepsilon_k) \nabla \varphi_{\varepsilon_k}^{\lambda_n}) v dt dx, \end{aligned}$$

which, taking into account that $v = h w_\delta$, by Claim 5 is equal to

$$\lim_{n \rightarrow \infty} \lim_{k \rightarrow \infty} \int_{Q_T} (\varphi_{\varepsilon_k}^{\lambda_n})_t + \operatorname{div}(\varrho_\delta \nabla \varphi_{\varepsilon_k}^{\lambda_n}) v dt dx,$$

which in turn equals zero by (7.20). Hence, $\int_{Q_T} v f dt dx = 0$ and so, because $f \in C_c^\infty(Q_T)$ was arbitrary, the assertion follows. \square

7.2 Probabilistic proof of Theorem 7.4 (ii)

We start with the following claim.

Claim: Let $T > 0$. Pathwise uniqueness holds for (6.3) on $[0, T]$ among probabilistically weak solutions (X, W) with $\mathcal{L}_X \in P^{\mathcal{A}_{\delta, T}}$, where we set

$$P^{\mathcal{A}_{\delta, T}} := \{Q \in \mathcal{P}(C([0, T]; \mathbb{R}^d)) : (Q \circ \pi_t^{-1})_{t \in [0, T]} \in \mathcal{A}_{\delta, T}\}.$$

Proof of Claim. Let $(X, W), (Y, W)$ be two probabilistically weak solutions to (6.3) (on $[0, T]$) on the same stochastic basis, with the same Brownian motion W , $X(0) = Y(0)$ a.s., and $\mathcal{L}_X, \mathcal{L}_Y \in P^{\mathcal{A}_{\delta, T}}$ with $\mathcal{L}_{X(t)}(dx) = u_X(t, x) dx$ and $\mathcal{L}_{Y(t)}(dx) = u_Y(t, x) dx$, for all $t \in [0, T]$. Imitating the proof of Theorem 6.13, we may similarly estimate for all $n \in \mathbb{N}$ and $\varepsilon > 0$

$$\begin{aligned} &\mathbb{E} \left[\ln \left(1 + \frac{|X(t \wedge \tau_n) - Y(t \wedge \tau_n)|^2}{\varepsilon^2} \right) \right] \\ &\lesssim_\delta \int_0^T \int_{B_n(0)} \mathbb{M} |D_x \nabla \varrho_\delta(s, \cdot)|(x) u_X(s, x) dx ds + \int_0^T \int_{B_n(0)} \mathbb{M} |D_x \nabla \varrho_\delta(s, \cdot)|(x) u_Y(s, x) dx ds \\ &\quad + \int_0^T \int_{\mathbb{R}^d} \left(\mathbb{M} |\nabla \varrho_\delta^{\frac{1+\gamma}{2}}(s, \cdot)|(x) \right)^2 w_\delta(s, x)^{-1} |x|^{-\frac{p-2}{q(p-1)-1}} u_X(s, x) dx ds \\ &\quad + \int_0^T \int_{\mathbb{R}^d} \left(\mathbb{M} |\nabla \varrho_\delta^{\frac{1+\gamma}{2}}(s, \cdot)|(x) \right)^2 w_\delta(s, x)^{-1} |x|^{-\frac{p-2}{q(p-1)-1}} u_Y(s, x) dx ds \\ &\leq 2 \int_0^T \int_{B_n(0)} \mathbb{M} |D_x \nabla \varrho_\delta(s, \cdot)|(x) w_\delta(s, x) dx ds + 2 \int_0^T \int_{\mathbb{R}^d} \left(\mathbb{M} |\nabla \varrho_\delta^{\frac{1+\gamma}{2}}(s, \cdot)|(x) \right)^2 |x|^{-\frac{p-2}{q(p-1)-1}} dx ds. \end{aligned}$$

By the same arguments as in Theorem 6.13, we conclude the finiteness of the right-hand side of the above chain of inequalities, and, finally, that $X \equiv Y$ a.s., by letting $\varepsilon \downarrow 0$. This finishes the proof of the claim. \square

Proof of Theorem 7.4 under assumption (ii). Let $u, \bar{u} \in \mathcal{A}_{\delta,T}$ be distributional solutions to (7.1) in the sense of Definition 7.1, both with initial condition $w_{\delta}(0, x)dx$. By Theorem 3.3, there exist probabilistically weak solutions $(X, W^1), (Y, W^2)$ to (6.3) such that $\mathcal{L}_{X(t)}(dx) = u(t, x)dx$ and $\mathcal{L}_{Y(t)}(dx) = \bar{u}(t, x)dx$, for all $t \in [0, T]$, respectively. In particular, $\mathcal{L}_X, \mathcal{L}_Y \in P^{\mathcal{A}_{\delta,T}}$. Since the previous claim implies weak uniqueness among probabilistically weak solutions whose laws are elements in $P^{\mathcal{A}_{\delta,T}}$ (cf. [33, Proof of Theorem 1.3.1]), we obtain $\mathcal{L}_X = \mathcal{L}_Y$. It follows that $u(t, \cdot) = \bar{u}(t, \cdot)$ dx-a.s., for all $t \in [0, T]$. This completes the proof. \square

A Proof of Remark 7.8

Clearly, $C_c^{\infty}([0, T] \times \mathbb{R}^d)$ is dense in $C_{c,b}^{\infty}([0, T] \times \mathbb{R}^d)$ (defined in Remark 7.8) with respect to uniform convergence of all partial derivatives (including zero derivatives) on $[0, T] \times \mathbb{R}^d$. Let $\varphi \in C_{c,b}^{\infty}([0, T] \times \mathbb{R}^d)$ and $(\varphi_k)_{k \in \mathbb{N}} \subseteq C_c^{\infty}([0, T] \times \mathbb{R}^d)$, such that $\varphi_k \rightarrow \varphi$ as $k \rightarrow \infty$ in the above sense. Then, in particular,

- (i) $\varphi_k(0) \rightarrow \varphi(0)$ uniformly on \mathbb{R}^d .
- (ii) $(\varphi_k)_t \rightarrow \varphi_t$, $\nabla \varphi_k \rightarrow \nabla \varphi$, $\Delta \varphi_k \rightarrow \Delta \varphi$ uniformly on $[0, T] \times \mathbb{R}^d$.

Since $u \in L^1(Q_T) \cap L^{\infty}(Q_T)$, $\varrho_{\delta} \in L^{\infty}(Q_T)$, $\nabla \varrho_{\delta} \in L^1(Q_T)$, (i)–(ii) yield

$$0 = \lim_k \left(\int_{Q_T} u((\varphi_k)_t + \operatorname{div}(\varrho_{\delta} \nabla \varphi_k)) dt dx + \int_{\mathbb{R}^d} \varphi_k(0) d\nu \right) = \int_{Q_T} u(\varphi_t + \operatorname{div}(\varrho_{\delta} \nabla \varphi)) dt dx + \int_{\mathbb{R}^d} \varphi(0) d\nu,$$

which proves Remark 7.8.

B Details for (7.20)

We know $\varphi_{\varepsilon}^{\lambda} \in W_2^{2,1}(Q_T)$. It is standard that $C_{c,b}^{\infty}([0, T] \times \mathbb{R}^d)$ (as introduced in Remark 7.8) is dense in $W_2^{2,1}(Q_T) \cap \{g \in W_2^{2,1}(Q_T) : g(T) = 0\}$ with respect to the usual norm

$$(|g|_2^{2,1})^2 := |g|_{L^2(Q_T)}^2 + |g_t|_{L^2(Q_T)}^2 + |\nabla g|_{L^2(Q_T)}^2 + |\Delta g|_{L^2(Q_T)}^2.$$

Now, let $(\varphi_k)_{k \in \mathbb{N}} \subseteq C_{c,b}^{\infty}([0, T] \times \mathbb{R}^d)$ such that $\varphi_k \xrightarrow{k} \varphi_{\varepsilon}^{\lambda}$ with respect to this norm. Since ν is the zero measure and since $v \in L^1 \cap L^{\infty}(Q_T)$, $\varrho_{\delta} \in L^{\infty}(Q_T)$, $\nabla \varrho_{\delta} \in L^2(Q_T)$, (cf. Lemma 7.3) and $\operatorname{div}(\varrho_{\delta} \nabla \varphi_{\varepsilon}^{\lambda}) = \varrho_{\delta} \Delta \varphi_{\varepsilon}^{\lambda} + \nabla \varrho_{\delta} \cdot \nabla \varphi_{\varepsilon}^{\lambda}$, we deduce

$$0 = \lim_k \int_{Q_T} v((\varphi_k)_t + \operatorname{div}(\varrho_{\delta} \nabla \varphi_k)) dt dx = \int_{Q_T} v((\varphi_{\varepsilon}^{\lambda})_t + \operatorname{div}(\varrho_{\delta} \nabla \varphi_{\varepsilon}^{\lambda})) dt dx,$$

which proves (7.20).

C Linear and nonlinear Fokker–Planck equations

Here we recall the definition of linear and nonlinear Fokker–Planck equations and their standard notion of distributional solution. The linear FPE associated with Borel measurable coefficients $a_{ij}, b_i : (0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R}$, $1 \leq i, j \leq d$, is the second-order parabolic differential equation for measures

$$\partial_t \mu_t = \partial_{ij}(a_{ij}(t, x) \mu_t) - \partial_i(b_i(t, x) \mu_t), \quad (t, x) \in (0, \infty) \times \mathbb{R}^d. \quad (\text{C.1})$$

Usually, an initial condition $\mu_0 = \nu \in \mathcal{M}_b^+$ is imposed.

Definition C.1. A (distributional) solution to (C.1) with initial condition $\nu \in \mathcal{M}_b^+$ is a weakly continuous curve $(\mu_t)_{t \geq 0}$ of signed locally finite Borel measures on \mathbb{R}^d such that

$$\int_0^T \int_{\mathbb{R}^d} |a_{ij}(t, x)| + |b_i(t, x)| d\mu_t(x) dt < \infty, \quad \forall T > 0,$$

and

$$\int_{\mathbb{R}^d} \psi d\mu_t(x) - \int_{\mathbb{R}^d} \psi d\nu = \int_0^t \int_{\mathbb{R}^d} a_{ij}(s, x) \partial_{ij} \psi(x) + b_i(s, x) \partial_i \psi(x) d\mu_s(x) ds, \quad \forall t \geq 0,$$

for all $\psi \in C_c^2(\mathbb{R}^d)$. A solution is called *probability solution*, if ν and each μ_t , $t \geq 0$, are probability measures. Instead of the initial time 0, one may consider an initial time $s > 0$. It is obvious how to generalize the definition in this regard. In this case, the initial condition is the pair (s, ν) and the solution is defined on $[s, \infty)$.

For the nonlinear FPE (1.3), the notion of distributional solution is similar.

Definition C.2. A (distributional) solution to (1.3) with initial condition ν is a weakly continuous curve $(\mu_t)_{t \geq 0}$ of signed locally finite Borel measures on \mathbb{R}^d such that $(t, x) \mapsto a_{ij}(t, x, \mu_t)$ and $(t, x) \mapsto b_i(t, x, \mu_t)$ are product Borel measurable on $(0, \infty) \times \mathbb{R}^d$,

$$\int_0^T \int_{\mathbb{R}^d} |a_{ij}(t, x, \mu_t)| + |b_i(t, x, \mu_t)| d\mu_t(x) dt < \infty, \quad \forall T > 0,$$

and

$$\int_{\mathbb{R}^d} \psi d\mu_t(x) - \int_{\mathbb{R}^d} \psi d\nu = \int_0^t \int_{\mathbb{R}^d} a_{ij}(s, x, \mu_s) \partial_{ij} \psi(x) + b_i(s, x, \mu_s) \partial_i \psi(x) d\mu_s(x) ds, \quad \forall t \geq 0,$$

for all $\psi \in C_c^2(\mathbb{R}^d)$. The notion of probability solution and the extension to initial times $s > 0$ is as in the linear case.

D On the Hardy–Littlewood maximal operators

Let $d, n \in \mathbb{N}$, and denote by λ^d the d -dimensional Lebesgue measure.

D.1 Definitions and basic properties

We recall the definition and properties of (local) Hardy–Littlewood maximal operators on the set of (extended) signed Borel measures on \mathbb{R}^d , where the latter are denoted as $M_{\text{loc}}(\mathbb{R}^d; \mathbb{R}^n)$ in the following. Furthermore, if $\mu \in M_{\text{loc}}(\mathbb{R}^d; \mathbb{R}^n)$, then $|\mu|$ denotes its variation measure.

The following definition and lemmatas are essentially taken from [21].

Definition D.1 ([21, Definition A.1]). Let $\mu \in M_{\text{loc}}(\mathbb{R}^d; \mathbb{R}^n)$ and $R \in (0, \infty]$. We define the (local) maximal function as

$$M_R |\mu|(x) := \sup_{0 < r < R} \frac{1}{\lambda^d(B_r(0))} \int_{B_r(x)} |\mu|(dx), \quad x \in \mathbb{R}^d,$$

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

Lemma D.2 ([20, Lemma 3.1]). Let $f \in BV(\mathbb{R}^d; \mathbb{R}^n)$. 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 $x, y \in N^c$

$$|f(x) - f(y)| \leq C_d (M |Df|(x) + M |Df|(y)) |x - y|, \quad (\text{D.1})$$

where $Df = (\partial_{x_j} f^i)_{1 \leq i \leq n, 1 \leq j \leq d}$ denotes the matrix of all Schwartz distributional derivatives of f 's component-functions in form of finite signed Borel measures on \mathbb{R}^d .

Remark D.3. Consider the situation of Lemma D.2. Let $\varphi_\varepsilon = \varepsilon^{-d}\varphi(\cdot/\varepsilon)$, $\varepsilon > 0$, be a standard Dirac sequence with respect to some nonnegative $\varphi \in C_c^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ with $\int \varphi(x)dx = 1$. Let $x, y \in \mathbb{R}^d$. If for all $z \in \{x, y\}$ we have

$$f(z) = \lim_{\varepsilon \rightarrow 0} (f * \varphi_\varepsilon)(z), \quad (\text{D.2})$$

then (D.1) holds for this pair x, y . This follows from the proof of [20, Lemma 3.1]. In particular, if $f \in C(\mathbb{R}^d; \mathbb{R}^n)$, then (D.1) holds for all $x, y \in \mathbb{R}^d$.

Furthermore, it follows from [2, Proposition 3.92 (b)], that if $x \in \mathbb{R}^d$ is an approximate jump point of f (in the sense of [2, Definition 3.67]), then $M|Df|(x) = +\infty$. Then, with the convention $0 \cdot \infty = 0$, (D.1) is trivially satisfied with respect to such x , and every $y \in \mathbb{R}^d$.

Lemma D.4 ([21, Lemma A.2]). (i) Let $\mu \in \mathcal{M}_{\text{loc}}(\mathbb{R}^d; \mathbb{R}^n)$ and $R \in (0, \infty]$. Then for λ^d -a.e. $x \in \mathbb{R}^d$, $M_R|\mu|(x) < \infty$.

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

$$\int_{B_\rho(0)} (M_R|f|(x))^p dx \leq C_{d,p} \int_{B_{\rho+R}(0)} |f(x)|^p dx.$$

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 $f \in L_{\text{loc}}^1(\mathbb{R}^d; \mathbb{R}^n)$

$$\lambda^d(\{x \in B_\rho(0) : M_R|f|(x) > \alpha\}) \leq \frac{C_d}{\alpha} \int_{B_{\rho+R}(0)} |f(x)| dx.$$

D.2 Muckenhoupt weights

Definition D.5 ([57, p. 194]). Let $w \in L_{\text{loc}}^1(\mathbb{R}^d)$. Let $p, p' \in (1, \infty)$ such that $p^{-1} + (p')^{-1} = 1$. If there exists $C > 0$ such that for all balls $B \subset \mathbb{R}^d$

$$\frac{1}{\lambda^d(B)} \int_B w(x) dx \cdot \left[\frac{1}{\lambda^d(B)} \int_B w(x)^{-\frac{p'}{p}} dx \right]^{\frac{p}{p'}} \leq C < \infty, \quad (\text{D.3})$$

then $w \in A_p$.

Theorem D.6 ([57, p. 201, Theorem 1]). Suppose $1 < p < \infty$ and $w \in A_p$. Then there exists a constant $C > 0$ such that

$$\int_{\mathbb{R}^d} (M|f|(x))^p w(x) dx \leq C \int_{\mathbb{R}^d} |f(x)|^p w(x) dx. \quad (\text{D.4})$$

D.3 Maximal operator of the surface measure restricted to the boundary of a ball

We consider the case $d = 3$. We set $w_3 := \lambda^3(B_1(0))$.

Lemma D.7. Let $R > 0$ and $x \in \mathbb{R}^3 \setminus (\partial B_R(0) \cup \{0\})$. Then

$$M|S(\cdot \cap B_R(0))|(x) = \begin{cases} \frac{2\pi R}{w_3 \sqrt{27}} (|x||x| - R)^{-1}, & \text{if } \sqrt{3}||x| - R| \leq |x| + R; \\ \frac{4\pi R^2}{w_3} (|x| + R)^{-1}, & \text{if } |x| + R < \sqrt{3}||x| - R|. \end{cases} \quad (\text{D.5})$$

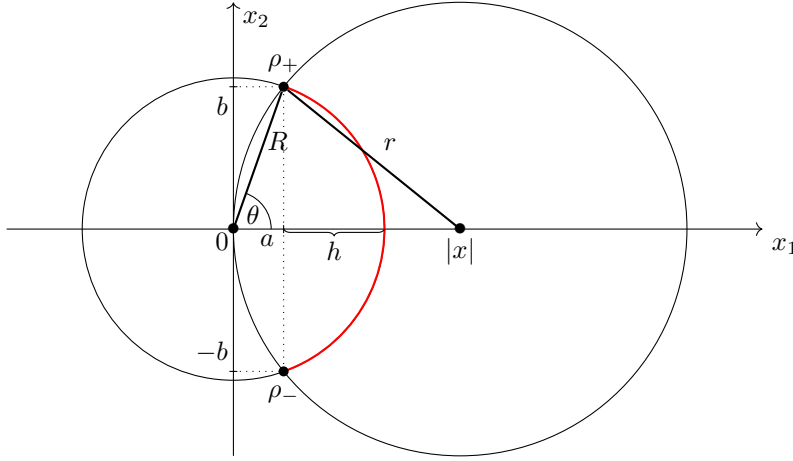


Figure 1: Projection of two intersecting balls onto the x_1/x_2 plane. The projection of the spherical cap $\partial B_R(0) \cap B_r(x)$ is highlighted in red.

Proof. Let us assume for a moment that $B_r(x) \cap \partial B_R(0) \neq \emptyset$ and $B_r(x) \cap \partial B_R(0) \neq \partial B_R(0)$, which is exactly the case when $||x| - R| < r \leq |x| + R$. Then $S(B_r(x) \cap \partial B_R(0))$ can be interpreted geometrically as the area of the *spherical cap* $B_r(x) \cap \partial B_R(0)$. For further calculations we wish to express the height h of the spherical cap as a function of r with parameters $|x|$ and R , more precisely as

$$h \equiv h_{R,|x|} : (||x| - R|, |x| + R] \rightarrow [0, 2R]. \quad (\text{D.6})$$

We note that the surface measure S is invariant under rotation. Hence, we may assume without loss of generality $x = (|x|, 0, 0)$. Moreover, there are exactly two different vectors ρ_+, ρ_- in the x_1/x_2 -plane¹ of the form $\rho_+ = (a, b, 0)$, $\rho_- = (a, -b, 0)$ where $a, b \in (0, R)$ are determined by the following two equations

$$\begin{cases} a^2 + b^2 &= R^2, \\ (a - |x|)^2 + b^2 &= r^2. \end{cases}$$

By subtracting the second equation from the first, one easily deduces that $a = \frac{R^2 + |x|^2 - r^2}{2|x|}$. Here, we refer to [67] for the detailed elementary calculation and a similar geometric visualization as in Figure 1, where the latter is, however, tailored to our setting. Hence, the height of the spherical cap is $h(r) = R - \frac{R^2 + |x|^2 - r^2}{2|x|}$, $r \in (||x| - R|, |x| + R]$. (In fact, the previous arguments translate to higher dimensions, and the formula for the height of a general hyperspherical cap is the same.) Let us extend h (continuously) to a function on $[0, \infty)$ by setting

$$h(r) := 2R, \text{ for } r > |x| + R, \text{ and } h(r) := 0, \text{ for } 0 < r \leq ||x| - R|.$$

Using polar coordinates, it is standard to note

$$S(B_r(x) \cap \partial B_R(0)) = 2\pi R h_{R,|x|}(r) \quad \forall r \geq 0. \quad (\text{D.7})$$

We aim to identify the maximum of the function

$$f(r) := 2\pi R r^{-3} h_{R,|x|}(r) = 2\pi R r^{-3} \left(R - \frac{|x|^2 + R^2 - r^2}{2|x|} \right), r \in [||x| - R|, |x| + R].$$

¹The hyperplane spanned by $(1, 0, 0), (0, 1, 0)$ in \mathbb{R}^3 .

Note that f is twice differentiable in $(|x| - R, |x| + R)$ with derivatives

$$\begin{aligned} f'(r) &= \frac{\pi R(3(|x| - R)^2 - r^2)}{|x|r^4}, \\ f''(r) &= \frac{2\pi R(r^2 - 6(|x| - R)^2)}{|x|r^5}. \end{aligned}$$

So, if $\sqrt{3}||x| - R| \in (|x| - R, |x| + R)$, f has a global maximum point only at $\sqrt{3}||x| - R|$. In this case, we have

$$f(\sqrt{3}||x| - R|) = \frac{2\pi R}{\sqrt{27}}(|x||x| - R|)^{-1}.$$

Since

$$\max(f(|x| - R), f(|x| + R)) = \frac{4\pi R^2}{(|x| + R)^3},$$

(D.5) is evident. This finishes the proof. \square

E Details for the proof of Theorem 6.13

For $1 \leq i, j \leq d$, let

$$b_i, \sigma_{ij} : [0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R}$$

be Borel measurable functions. Consider the SDE

$$\begin{cases} dX(t) &= b(t, X(t))dt + \sigma(t, X(t))dW(t), \quad t \geq 0, \\ \mathcal{L}_{X(t)} &\ll dx, \text{ for } dt\text{-a.e. } t \geq 0, \end{cases} \quad (\text{E.1})$$

where we denote the absolute continuity of a Borel probability measure μ on \mathbb{R}^d with respect to d -dimensional Lebesgue measure by $\mu \ll dx$.

Definition E.1. A *probabilistically weak solution* to (E.1) is an adapted stochastic process $X = (X(t))_{t \geq 0}$ on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ with an (\mathcal{F}_t) -standard Brownian motion W , such that

$$\mathbb{E} \left[\int_0^T |b(t, X(t))| + |\sigma(t, X(t))\sigma^T(t, X(t))| dt \right] < \infty \quad \forall T > 0, \quad (\text{E.2})$$

and \mathbb{P} -a.s.

$$X(t) = X(0) + \int_0^t b(s, X(s))ds + \int_0^t \sigma(s, X(s))dW(s) \quad \forall t \geq 0.$$

Proposition E.2. Let (X, W) be a probabilistically weak solution to (E.1). Then, (X, W) is also a probabilistically weak solution to (E.1) where b, σ are replaced by Borel measurable functions $\bar{b}, \bar{\sigma}$ with components

$$\bar{b}_i, \bar{\sigma}_{ij} : [0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R},$$

$1 \leq i, j \leq d$, satisfying

$$\begin{aligned} b &= \bar{b} \quad dt \otimes dx\text{-a.s.}, \\ \sigma &= \bar{\sigma} \quad dt \otimes dx\text{-a.s.} \end{aligned}$$

In particular, probabilistically weak solutions to (E.1) do not depend on the $dt \otimes dx$ -version of the coefficients b, σ .

Proof. For $t \geq 0$, we set

$$Y(t) := X(0) + \int_0^t \bar{b}(s, X(s)) ds + \int_0^t \bar{\sigma}(s, X(s)) dW(s).$$

We have

$$\mathbb{E}[|X(t) - Y(t)|] \leq \mathbb{E} \left[\left| \int_0^t b(s, X(s)) - \bar{b}(s, X(s)) ds \right| \right] + \left(\mathbb{E} \left[\left| \int_0^t \sigma(s, X(s)) - \bar{\sigma}(s, X(s)) dW(s) \right|^2 \right] \right)^{\frac{1}{2}} \quad (\text{E.3})$$

$$\leq \int_0^t \mathbb{E} [|b(s, X(s)) - \bar{b}(s, X(s))|] ds + \left(\int_0^t \mathbb{E} [|\sigma(s, X(s)) - \bar{\sigma}(s, X(s))|^2] ds \right)^{\frac{1}{2}} = 0, \quad (\text{E.4})$$

where we used Jensen's inequality in the first inequality, and Itô's isometry and Fubini's theorem in the second inequality. For the last equality, we used the assumption $\mathcal{L}_{X(s)} \ll dx$ for ds -a.e. $s \geq 0$. By the (a.s.-)continuity of X and Y , we conclude $X \equiv Y$ a.s. This finishes the proof. \square

References

- [1] L. Ambrosio. *Transport equation and Cauchy problem for non-smooth vector fields*, volume 1927 of *Lecture Notes in Math.*, pages 1–41. Springer, Berlin, 2008.
- [2] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000.
- [3] V. Barbu, M. Rehmeier, and M. Röckner. p -Brownian motion and the p -Laplacian. *to appear in The Annals of Probability, available under arXiv preprint 2409.18744*, 2024.
- [4] V. Barbu and M. Röckner. Probabilistic representation for solutions to nonlinear Fokker–Planck equations. *SIAM J. Math. Anal.*, 50(4):4246–4260, 2018.
- [5] V. Barbu and M. Röckner. From nonlinear Fokker–Planck equations to solutions of distribution dependent SDE. *Ann. Probab.*, 48(4):1902–1920, 2020.
- [6] V. Barbu and M. Röckner. Solutions for nonlinear Fokker–Planck equations with measures as initial data and McKean–Vlasov equations. *J. Funct. Anal.*, 280(7):Paper No. 108926, 35, 2021.
- [7] V. Barbu and M. Röckner. The evolution to equilibrium of solutions to nonlinear Fokker–Planck equation. *Indiana Univ. Math. J.*, 72(1):89–131, 2023.
- [8] V. Barbu and M. Röckner. Uniqueness for nonlinear Fokker–Planck equations and for McKean–Vlasov SDEs: the degenerate case. *J. Funct. Anal.*, 285(4):Paper No. 109980, 37, 2023.
- [9] V. Barbu and M. Röckner. Nonlinear Fokker–Planck equations with fractional Laplacian and McKean–Vlasov SDEs with Lévy noise. *Probab. Theory Related Fields*, 189(3-4):849–878, 2024.
- [10] V. Barbu and M. Röckner. *Nonlinear Fokker–Planck flows and their probabilistic counterparts*, volume 2353 of *Lecture Notes in Mathematics*. Springer, Cham, 2024.
- [11] V. Barbu, M. Röckner, and D. Zhang. Uniqueness of distributional solutions to the 2D vorticity Navier–Stokes equation and its associated nonlinear Markov process. *J. Eur. Math. Soc.*, 2025.

- [12] G. I. Barenblatt. On some unsteady motions of a liquid and gas in a porous medium. *Akad. Nauk SSSR. Prikl. Mat. Meh.*, 16:67–78, 1952.
- [13] G. I. Barenblatt. *Scaling*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 2003.
- [14] J. Bliedtner and W. Hansen. *Potential theory*. Universitext. Springer-Verlag, Berlin, 1986. An analytic and probabilistic approach to balayage.
- [15] R. M. Blumenthal and R. K. Gettoor. *Markov processes and potential theory*, volume Vol. 29 of *Pure and Applied Mathematics*. Academic Press, New York-London, 1968.
- [16] V. I. Bogachev, N. V. Krylov, M. Röckner, and S. V. Shaposhnikov. *Fokker–Planck–Kolmogorov equations*, volume 207 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2015.
- [17] V. I. Bogachev, M. Röckner, and S. V. Shaposhnikov. On the Ambrosio–Figalli–Trevisan superposition principle for probability solutions to Fokker–Planck–Kolmogorov equations. *J. Dynam. Differential Equations*, 33(2):715–739, 2021.
- [18] M. Bossy and J.-F. Jabir. On the wellposedness of some McKean models with moderated or singular diffusion coefficient. In *Frontiers in stochastic analysis—BSDEs, SPDEs and their applications*, volume 289 of *Springer Proc. Math. Stat.*, pages 43–87. Springer, Cham, 2019.
- [19] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011.
- [20] N. Champagnat and P.-E. Jabin. Strong solutions to stochastic differential equations with rough coefficients. *Ann. Probab.*, 46(3):1498–1541, 2018.
- [21] G. Crippa and C. De Lellis. Estimates and regularity results for the DiPerna–Lions flow. *J. Reine Angew. Math.*, 616:15–46, 2008.
- [22] J. L. Doob. *Classical potential theory and its probabilistic counterpart*. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1984 edition.
- [23] E. B. Dynkin. *Markov processes. Vols. I, II*, volume Band 121, 122 of *Die Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin-Göttingen-Heidelberg; Academic Press, Inc., Publishers, New York, 1965. Translated with the authorization and assistance of the author by J. Fabius, V. Greenberg, A. Maitra, G. Majone.
- [24] S. N. Ethier and T. G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.
- [25] B. Fehrman and B. Gess. Non-equilibrium large deviations and parabolic-hyperbolic PDE with irregular drift. *Invent. Math.*, 234(2):573–636, 2023.
- [26] A. Figalli. Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients. *J. Funct. Anal.*, 254(1):109–153, 2008.
- [27] F. Flandoli and M. Romito. Markov selections for the 3D stochastic Navier–Stokes equations. *Probability Theory and Related Fields*, 140(3):407–458, 2008.
- [28] M. Freidlin. *Markov processes and differential equations: asymptotic problems*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 1996.

- [29] M. Fukushima, Y. Oshima, and M. Takeda. *Dirichlet forms and symmetric Markov processes*, volume 19 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, extended edition, 2011.
- [30] A. Grigor'yan and P. Sürig. Finite propagation speed for Leibenson's equation on Riemannian manifolds. *Comm. Anal. Geom.*, 32(9):2467–2504, 2024.
- [31] A. Grigor'yan and P. Sürig. Sharp propagation rate for solutions of Leibenson's equation on Riemannian manifolds. *Ann. Sc. Norm. Super. Pisa Cl. Sci.*, 2024.
- [32] S. Grube. Strong solutions to McKean–Vlasov SDEs with coefficients of Nemytskii-type. *Electron. Commun. Probab.*, 28:Paper No. 11, 13, 2023.
- [33] S. Grube. *Strong solutions to McKean–Vlasov stochastic differential equations with coefficients of Nemytskii-type*. PhD thesis. Bielefeld University, 2023.
- [34] S. Grube. Strong solutions to degenerate SDEs and uniqueness for degenerate Fokker–Planck equations. *arXiv preprint 2409.17135*, 2024.
- [35] S. Grube. Strong solutions to McKean–Vlasov SDEs associated to a class of degenerate Fokker–Planck equations with coefficients of Nemytskii-type. *arXiv preprint 2404.18918*, 2024.
- [36] S. Grube. Strong solutions to McKean–Vlasov SDEs with coefficients of Nemytskii type: the time-dependent case. *J. Evol. Equ.*, 24(2):37, 2024.
- [37] K. Ishige. On the existence of solutions of the Cauchy problem for a doubly nonlinear parabolic equation. *SIAM J. Math. Anal.*, 27(5):1235–1260, 1996.
- [38] A. V. Ivanov. Regularity for doubly nonlinear parabolic equations. *Journal of Mathematical Sciences*, 83(1):22–37, 1997.
- [39] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.
- [40] V. N. Kolokoltsov. *Markov processes, semigroups and generators*, volume 38 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 2011.
- [41] N. V. Krylov. On the selection of a markov process from a system of processes and the construction of quasi-diffusion processes. *Mathematics of the USSR-Izvestiya*, 7(3):691, jun 1973.
- [42] O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Ural'ceva. *Linear and quasilinear equations of parabolic type*, volume Vol. 23 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1968. Translated from the Russian by S. Smith.
- [43] H. Lee. On the pathwise uniqueness for a class of degenerate Itô-stochastic differential equations. *Osaka J. Math.*, 60(3):533–543, 2023.
- [44] L. S. Leibenson. General problem of the movement of a compressible fluid in a porous medium. *Bull. Acad. Sci. URSS. Sér. Géograph. Géophys. [Izvestia Akad. Nauk SSSR]*, 9:7–10, 1945.
- [45] L. S. Leibenson. Turbulent movement of gas in a porous medium. *Bull. Acad. Sci. URSS. Sér. Géograph. Géophys. [Izvestia Akad. Nauk SSSR]*, 9:3–6, 1945.
- [46] T. M. Liggett. *Continuous time Markov processes*, volume 113 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. An introduction.
- [47] D. Luo. Pathwise uniqueness of multi-dimensional stochastic differential equations with Hölder diffusion coefficients. *Front. Math. China*, 6(1):129–136, 2011.

- [48] D. Luo. Uniqueness of degenerate Fokker–Planck equations with weakly differentiable drift whose gradient is given by a singular integral. *Electron. Commun. Probab.*, 19:no. 43, 14, 2014.
- [49] H. P. McKean. A class of Markov processes associated with nonlinear parabolic equations. *Proc. Nat. Acad. Sci. U.S.A.*, 56:1907–1911, 1966.
- [50] D. Nualart. *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, Berlin, second edition, 2006.
- [51] M. Rehmeier. Flow selections for (nonlinear) Fokker–Planck–Kolmogorov equations. *J. Differential Equations*, 328:105–132, 2022.
- [52] M. Rehmeier and M. Röckner. On nonlinear Markov processes in the sense of McKean. *J. Theoret. Probab.*, 38(3):Paper No. 60, 36, 2025.
- [53] M. Rehmeier and M. Romito. 2D vorticity Euler equations: Superposition solutions and nonlinear Markov processes. *arXiv preprint 2407.16609*, 2024.
- [54] M. Röckner and X. Zhang. Weak uniqueness of Fokker–Planck equations with degenerate and bounded coefficients. *C. R. Math. Acad. Sci. Paris*, 348(7-8):435–438, 2010.
- [55] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales. Vol. 2*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. Itô calculus, Reprint of the second (1994) edition.
- [56] M. Sharpe. *General theory of Markov processes*, volume 133 of *Pure and Applied Mathematics*. Academic Press, Inc., Boston, MA, 1988.
- [57] E. M. Stein. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, volume 43 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.
- [58] D. W. Stroock. *An introduction to Markov processes*, volume 230 of *Graduate Texts in Mathematics*. Springer, Heidelberg, second edition, 2014.
- [59] D. W. Stroock and S. R. S. Varadhan. *Multidimensional diffusion processes*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1979 edition.
- [60] D. W. Stroock and M. Yor. On extremal solutions of martingale problems. *Ann. Sci. École Norm. Sup. (4)*, 13(1):95–164, 1980.
- [61] S. Sturm. Existence of weak solutions of doubly nonlinear parabolic equations. *J. Math. Anal. Appl.*, 455(1):842–863, 2017.
- [62] K. Suresh Kumar. A class of degenerate stochastic differential equations with non-Lipschitz coefficients. *Proc. Indian Acad. Sci. Math. Sci.*, 123(3):443–454, 2013.
- [63] P. Sürig. Finite extinction time for subsolutions of the weighted Leibenson equation on Riemannian manifolds. *arXiv preprint 2412.06496*, 2024.
- [64] D. Trevisan. Well-posedness of multidimensional diffusion processes with weakly differentiable coefficients. *Electron. J. Probab.*, 21:Paper No. 22, 41, 2016.
- [65] J. L. Vázquez. *The porous medium equation*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, Oxford, 2007. Mathematical theory.
- [66] Z. Wang and X. Zhang. Existence and uniqueness of degenerate SDEs with Hölder diffusion and measurable drift. *J. Math. Anal. Appl.*, 484(1):123679, 11, 2020.
- [67] E. W. Weisstein. Circle-circle intersection. *From MathWorld—A Wolfram Web Resource*, 2025. <https://mathworld.wolfram.com/Circle-CircleIntersection.html>; Accessed August 18, 2025.