p-Brownian motion and the p-Laplacian

Viorel Barbu[*]   Marco Rehmeier[†]   Michael Röckner[‡§]

October 1, 2024

Abstract

In this paper we construct a stochastic process, more precisely, a (nonlinear) Markov process, which is related to the parabolic p-Laplace equation in the same way as Brownian motion is to the classical heat equation given by the (2-) Laplacian.

Keywords: p-Laplacian; Brownian motion; Barenblatt solution; nonlinear Fokker—Planck equation; Markov process; McKean—Vlasov stochastic differential equation.
2020 MSC: Primary: 60J65; Secondary: 35K55, 35Q84, 60J25.

1   Introduction

As it has been known for more than 100 years (where, e.g, the articles by A. Einstein [17], M. von Smoluchowski [50] and N. Wiener [52] are among the pioneering papers), there is a close relationship between Brownian motion and the Laplace operator, more precisely the classical heat equation, i.e.,

$$\frac{\partial}{\partial t} u(t,x) = \Delta u(t,x), \ (t,x) \in (0,\infty) \times \mathbb{R}^d, \tag{1.1}$$

where $\Delta = \operatorname{div} \nabla$. But it has been an open question whether the same is true for the p-Laplace operator, i.e. whether there exists a "p-Brownian motion" related to the parabolic p-Laplace equation, i.e., for $p > 2$,

$$\frac{\partial}{\partial t} u(t,x) = \operatorname{div}(|\nabla u(t,x)|^{p-2}\nabla u(t,x)), \ (t,x) \in (0,\infty) \times \mathbb{R}^d, \tag{1.2}$$

in an analogous way as in the case $p = 2$, i.e. (1.1). The p-Laplacian

$$\Delta_p u = \operatorname{div}(|\nabla u|^{p-2}\nabla u)$$

has been for many years and still is an extensively studied operator in a large number of contexts, be it from the point of view of partial differential equations (PDEs) (see, e.g., [13], [14], [24], [30], [32], [33], [49] and the references therein), within nonlinear functional analysis, e.g. as the duality map between $W^{1,p}$ and its dual $W^{-1,p'}$ (see, e.g., [11], [34]), in nonlinear potential theory (see, e.g., [1], [23], [37] and the references therein) or its applications in physics (see, e.g., [29], [38]).

It should be mentioned that linear potential theory has played a crucial role in developing and exploiting the relation between the (2-) Laplacian and Brownian motion and, more generally, between large classes of linear partial (and pseudo) differential operators and their associated Markov processes for more than 60 years (see, e.g., [6], [7], [15], [16], [18], [20], [21], [26], [31], [43], [44], [45] and

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

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

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

[§]Academy of Mathematics and System Sciences, CAS, Beijing

[46]). Hence, one could expect that non-linear potential theory, which also has a long and exciting history, would pave the way for discovering the probabilistic counterpart of the p-Laplacian. One approach in this direction was developed in the fundamental papers [40], [39], where a deep relation of a stochastic game, the "tug-of-war" game with noise, was discovered and exploited to find a beautiful probabilistic description of the p-harmonic function solving the Dirichlet problem for the p-Laplacian on a bounded domain in $\mathbb{R}^d$. In this paper we propose a different approach, namely to construct the desired probabilistic counterpart to the p-Laplacian as a Markov process which is related to the p-Laplace equation in the same way as Brownian motion is to the classical heat equation (1.1), and which then may be called a "p-Brownian motion".

The starting point in our approach is to recall that the relation between the classical heat equation and Brownian motion is a special case of the relation between a Fokker–Planck equation (FPE) (see, e.g., [8]) and its associated Markov process. In fact, this relation is a one-to-one correspondence, which turns out to extend to large classes of FPEs and their canonically associated Markov processes, and even to nonlinear FPEs and their canonically associated nonlinear Markov processes (see [36], [5], [42]). We will explain this below in more detail and only mention here that such nonlinear FPEs can be shown to have an associated family of stochastic processes given as the solutions to a corresponding McKean–Vlasov stochastic differential equation (SDE), which turns out to constitute a nonlinear Markov process (though the latter is extremely hard to prove in concrete cases).

Key Step 1 in our construction is to identify the parabolic p-Laplace equation as a nonlinear FPE, and Key Step 2 to solve the corresponding McKean–Vlasov SDE. The third and hardest Key Step is then to prove that the solutions of the latter constitute a nonlinear Markov process which is thus associated to the parabolic p-Laplace equation in the same way as Brownian motion is to the classical heat equation.

Let us now explain in more detail the relation of nonlinear FPEs and nonlinear Markov processes, where the latter have already been defined in McKean's classical paper [36] (see also [42]). Let $\mathcal{P}(\mathbb{R}^d)$ denote the space of Borel probability measures on $\mathbb{R}^d$, and for $1 \leq i, j \leq d$ consider measurable maps

$$b_i, a_{ij} : [0, \infty) \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$$

such that the matrix $(a_{ij})_{i,j}$ is pointwise symmetric and nonnegative definite. Then, a nonlinear FPE is an equation of type

$$\frac{\partial}{\partial t} \mu_t = \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} (a_{ij}(t, x, \mu_t)\mu_t) - \frac{\partial}{\partial x_i} (b_i(t, x, \mu_t)\mu_t), \ (t, x) \in (0, \infty) \times \mathbb{R}^d, \tag{1.3}$$

where the solution $[0, \infty) \ni t \mapsto \mu_t$ is a curve in $\mathcal{P}(\mathbb{R}^d)$ with some specified initial condition μ_0. Here, we use Einstein's summation convention, (1.3) is meant as a weak equation in the sense of Schwartz distributions (see Definition D.2 in Appendix D) and, henceforth, we call solutions to (1.3) *distributional solutions*. The (in space) dual operator to the operator on the right hand side of (1.3) is called the corresponding Kolmogorov operator L, i.e. its action on test functions $\varphi \in C_0^\infty((0, \infty) \times \mathbb{R}^d)$ is given as

$$L\varphi(t, x) = a_{ij}(t, x, \mu_t)\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} \varphi(t, x) + b_i(t, x, \mu_t)\frac{\partial}{\partial x_i} \varphi(t, x), \ (t, x) \in (0, \infty) \times \mathbb{R}^d. \tag{1.4}$$

In turn, this operator determines the corresponding McKean–Vlasov SDE

$$dX(t) = b(t, X(t), \mu_t)dt + \sqrt{2}\,\sigma(t, X(t), \mu_t)dW(t), \ t > 0, \tag{1.5}$$

$$\mathcal{L}_{X(t)} = \mu_t, \ t \geq 0, \tag{1.6}$$

where $\sigma = (\sigma_{ij})_{ij}$ with $\sigma\sigma^\top = (a_{ij})_{ij}$, $b = (b_1, ..., b_d)$, $W(t)$, $t \geq 0$, is a (standard) d-dimensional Brownian motion on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and the maps $X(t) : \Omega \to \mathbb{R}^d$, $t \geq 0$, form the continuous in t solution process to (1.5) such that its one-dimensional time marginals

$$\mathcal{L}_{X(t)} := X(t)_*\mathbb{P}, \ t \geq 0, \tag{1.7}$$

2

(i.e. the push forward or image measures of $\mathbb{P}$ under $X(t)$) satisfy (1.6). We refer to [12], [35], [47], [51] and the references therein for details on McKean–Vlasov SDEs. The correspondence between the nonlinear FPE (1.3) and the McKean–Vlasov SDE (1.5), (1.6) is now given as follows. Suppose one has a solution to (1.5), (1.6), then by a simple application of Itô's formula $\mu_t := \mathcal{L}_{X(t)}$, $t \geq 0$, solves (1.3). Conversely, suppose one has a distributional solution to (1.3), then the nonlinear version in [3], [4] of the superposition principle [48, Theorem 2.5] (which in turn is a generalization of [19], see also [28] and [10]) together with [22, Proposition 2.2.3] (which is a slight generalization of [25, Proposition 4.11]) gives rise to a solution to (1.5), (1.6). We would like to emphasize that in [4] (see also [5]) apart from some integrability conditions the coefficients are only required to be measurable (in contrast to the cases studied in [12]).

Obviously, the special case of classical Brownian motion and the classical heat equation is the case where $a_{ij}(t, x, \mu) = \delta_{ij}$ (= Kronecker delta), $b_i(t, x, \mu) = 0$, i.e. (1.3) turns into

$$\frac{\partial}{\partial t} \mu_t = \Delta \mu_t, \ (t, x) \in (0, \infty) \times \mathbb{R}^d, \tag{1.8}$$

and (1.5), (1.6) into

$$dX(t) = dW(t), \ t > 0, \tag{1.9}$$
$$\mathcal{L}_{X(t)} = \mathcal{L}_{W(t)} = \mu_t, \ t \geq 0, \tag{1.10}$$

and, of course, μ_t is absolutely continuous with respect to Lebesgue measure dx on $\mathbb{R}^d$ with density $u(t, x)$, so (1.8) is really (1.1).

To see that also (1.2) is of type (1.3), we recall that the coefficients in (1.3) only need to be measurable in the variable μ, so in case the solutions μ_t, $t \geq 0$, have densities $u(t, \cdot)$, $t \geq 0$, we might, e.g., have dependencies as follows

$$a_{ij}(t, x, \mu_t) = \tilde{a}_{ij}(t, x, \Gamma_1(u)(t, x)),$$
$$b_i(t, x, \mu_t) = \tilde{b}_i(t, x, \Gamma_2(u)(t, x)), \tag{1.11}$$

where $\tilde{b}_i, \tilde{a}_{ij} : [0, \infty) \times \mathbb{R}^d \times \mathbb{R}^k \to \mathbb{R}$ and each Γ_i is a functional on the space of distributional solutions whose values are again measurable functions of t and x. Then, rewriting (1.2) as

$$\frac{\partial}{\partial t} u(t, x) = \Delta(|\nabla u(t, x)|^{p-2} u(t, x)) - \mathrm{div}(\nabla(|\nabla u(t, x)|^{p-2}) u(t, x)), \ (t, x) \in (0, \infty) \times \mathbb{R}^d, \tag{1.12}$$

we see that (1.2), respectively (1.12), is of type (1.3) with a_{ij}, b_i as in (1.11) with

$$\tilde{a}_{ij}(t, x, \Gamma_1(u)(t, x)) = \delta_{ij} |\nabla u(t, x)|^{p-2}$$
$$\tilde{b}(t, x, \Gamma_2(u)(t, x)) = \nabla(|\nabla u(t, x)|^{p-2}). \tag{1.13}$$

We point out that while the already challenging so-called *Nemytskii-case*, where $\Gamma_i(u)(t, x) = u(t, x)$, has received more and more attention in the last years (see, for instance, [2], [3], [4], [5] and the references therein), the coefficients in (1.13) even depend on u via its first- and second-order derivatives. To the best of our knowledge, the relation of such nonlinear FPEs to McKean–Vlasov SDEs has not been studied before. The Kolmogorov operator associated with the coefficients (1.13) is then given by

$$L\varphi(t, x) = |\nabla u(t, x)|^{p-2} \Delta \varphi(t, x) + \nabla(|\nabla u(t, x)|^{p-2}) \cdot \nabla \varphi(t, x), \ (t, x) \in (0, \infty) \times \mathbb{R}^d, \tag{1.14}$$

and the corresponding McKean–Vlasov SDE by

$$dX(t) = \nabla(|\nabla u(t, X(t))|^{p-2}) dt + \sqrt{2} |\nabla u(t, X(t))|^{\frac{p-2}{2}} dW(t), \ t > 0, \tag{1.15}$$
$$\mathcal{L}_{X(t)} = u(t, x) dx, \ t > 0. \tag{1.16}$$

This completes Key Steps 1 and 2.

3

For Key Step 3, i.e. to obtain the nonlinear Markov process from (1.15), (1.16), we need to take into account initial conditions which we choose to be Dirac measures δ_y, $y \in \mathbb{R}^d$. So, we impose in (1.12) that $u(0,x)dx = \delta_y$ and in (1.16) that $\mathcal{L}_{X(0)} = \delta_y$. In both cases $p = 2$ and $p > 2$, for such initial conditions the fundamental solution to (1.1) and (1.12) (that is, to (1.2)) are explicitly known, namely in case $p = 2$, for $y \in \mathbb{R}^d$ it is given by the classical Gaussian heat kernel

$$u^y(t,x) := \frac{1}{(4\pi t)^{\frac{d}{2}}} \exp\left(-\frac{1}{4t}|x-y|^2\right), \ (t,x) \in (0,\infty) \times \mathbb{R}^d, \tag{1.17}$$

and in case $p > 2$ by the famous Barenblatt solution (see [24])

$$w^y(t,x) = t^{-k}\left(C_1 - qt^{-\frac{kp}{d(p-1)}}|x-y|^{\frac{p}{p-1}}\right)_+^{\frac{p-1}{p-2}}, \ (t,x) \in (0,\infty) \times \mathbb{R}^d, \tag{1.18}$$

where $k := \left(p - 2 + \frac{p}{d}\right)^{-1}$, $q := \frac{p-2}{p}\left(\frac{k}{d}\right)^{\frac{1}{p-1}}$, $C_1 \in (0,\infty)$ is the unique constant such that $|w^y(t)|_{L^1(\mathbb{R}^d)}$ = 1 for all $t > 0$, and $f_+ := \max(f,0)$. Then, we consider the *path laws* of the corresponding solutions $X^y(t)$, $t \geq 0$, of (1.9), (1.10) and (1.15), (1.16), respectively, namely

$$P_y := (X^y)_* \mathbb{P}, \ y \in \mathbb{R}^d, \tag{1.19}$$

i.e. the push forward or image measure of $\mathbb{P}$ under the map $X^y : \Omega \to C([0,\infty), \mathbb{R}^d)$ (= all continuous paths in $\mathbb{R}^d$). Then, if $p = 2$, P_y, $y \in \mathbb{R}^d$, form a Markov process in the sense of e.g. [7], [16], [44], also called *Brownian motion*, which is uniquely determined by $u^y(t,x)$, $y \in \mathbb{R}^d$, $t > 0$. And in this paper we prove that P_y, $y \in \mathbb{R}^d$, form a nonlinear Markov process in the sense of McKean [36], which is uniquely determined by $w^y(t,x)$, $y \in \mathbb{R}^d$, $t > 0$ and (1.15)–(1.16). In contrast to the results of Sections 2–4, which hold for all $d \geq 1$, $p > 2$, for this part, i.e. Key Step 3, we need $d \geq 2$, $p > 2\left(1 + \frac{1}{d}\right)$. By analogy, we call this nonlinear Markov process *p-Brownian motion*.

The fact that the probability measures P_y, $y \in \mathbb{R}^d$, in (1.19) form a nonlinear Markov process is of fundamental importance, since it pins down their interrelation and is the analogue on path space to the flow property of the fundamental solutions to the parabolic p-Laplace equation on state space (see (5.3) below). The proof of this fact is quite involved, because the diffusion coefficient in front of the noise in (1.15) with w^y replacing u is very degenerate (see (4.3) below). One of the proof's main ingredients is Theorem 6.5 below. It uses very heavily the explicit form of the Barenblatt solution w^y in (1.18) and of the coefficients $|\nabla w^y(t,x)|^{p-2}$ and $\nabla(|\nabla w^y(t,x)|^{p-2})$ in (1.13) with $u = w^y$, see (4.3) and (4.4) below. This completes Key Step 3.

All these three Key Steps are implemented in detail in the main text of the paper, including stating or recalling all required definitions and notions.

We would also like to make a remark concerning the above quoted references. The literature about the p-Laplacian, Markov processes, nonlinear potential theory, Fokker–Planck equations and McKean–Vlasov equations is so huge that it is impossible to give credit to it in an appropriate way in this paper. Therefore, above we tried to make a reasonably sized selection which is very probably not at all optimal. But this was necessary, in regard to the length of this paper. Nevertheless, we would like to comment on one more reference, namely [27], to avoid confusion. We point out that the notion of nonlinear Markov processes therein is different from McKean's, and hence from ours in this paper, since it is defined there as a family of linear Markov processes. We refer to [27] for details.

We firmly hope that the relation between the p-Brownian motion and the p-Laplacian, or more generally nonlinear Markov processes and nonlinear FPEs, will turn out to be as fruitful as in the linear theory, which profited a lot from transferring an analytic problem for the PDE into a probabilistic problem for the stochastic process and vice versa.

Furthermore, we would like to mention that for the heat equation or more general linear FPEs, considering Dirac initial conditions is in a sense generic, since by linearity in the initial condition a solution with another initial probability measure is a simple linear superposition of the solutions with Dirac initial conditions. This is completely different for the p-Laplacian equation, which by

4

its nonlinearity produces a richer structure of solutions, if one considers general initial probability measures. Nevertheless, our general approach above can be also applied to more general initial conditions, and the corresponding McKean–Vlasov SDEs can be solved as above. This will be the subject of a forthcoming paper.

Finally, it is obvious that Key Steps 1 and 2 above apply to much more general nonlinear parabolic PDEs than the parabolic p-Laplace equation (1.2). The hard part is Key Step 3, more precisely the extremality of the solution to the nonlinear PDE in the class of distributional solutions to the linearized PDE, analogous to Theorem 6.5 (see also Remark 6.6 (ii)) below. Also this will be the subject of our future work.

Let us now summarize the structure of the paper. While we have already presented parts of Key Step 1, namely the identification of the parabolic p-Laplace equation as the nonlinear FPE (1.12) above, this step is completed in Section 2, where we define our notion of distributional solutions to (1.12) (see Definition 2.3) and compare it with the usual notion of (weak) solution to the parabolic p-Laplace equation (see Definition 2.1). Key Step 2 is implemented in Sections 3 and 4. In Section 3, we solve the corresponding McKean–Vlasov equation (1.15), (1.16) (see Definition 3.1 and Theorem 3.3), and Section 4 is devoted to the study of the solutions to the nonlinear FPE (1.12) and to the construction of solutions to (1.15), (1.16) imposing Dirac initial conditions (see Proposition 4.1 and Theorem 4.2). Key Step 3 is presented in Sections 5 and 6. In Section 5 we recall McKean's definition of a nonlinear Markov process (see Definition 5.1) and then prove that the path laws P_y, $y \in \mathbb{R}^d$, (see (1.19)) of the solutions to (1.15), (1.16) form such a process (see Theorem 5.4) and we give the definition of the p-Brownian motion (see Definition 5.7). The crucial ingredient already mentioned above, namely, the restricted distributional uniqueness result for the linearized parabolic p-Laplace equation (see Theorem 6.5) is subsequently proved in Section 6. The Appendix contains some specified details of proofs in the main text (see Appendices A–C) and some basic facts on FPEs (see Appendix D).

1.1 Notation

Standard notation. On $\mathbb{R}^d$ we write $|\cdot|$, $x \cdot y$ and dx (or dt in one dimension) for the Euclidean norm, its inner product and Lebesgue measure. δ_{ij} denotes the Kronecker delta, and supp f the support of a function f on a topological space.

Measures. $\mathcal{M}_b^+$ is the space of finite nonnegative Borel measures on $\mathbb{R}^d$. For a topological space X, we write $\mathcal{P}(X)$ for the probability measures on $\mathcal{B}(X)$, the Borel σ-algebra of X. For $X = \mathbb{R}^d$, we write $\mathcal{P}(\mathbb{R}^d) = \mathcal{P}$. A curve $t \mapsto \mu_t \in \mathcal{M}_b^+$ on an interval $I \subseteq (-\infty, +\infty)$ is *vaguely*, respectively *weakly continuous*, if $t \mapsto \int_{\mathbb{R}^d} \psi \, d\mu_t$ is continuous for all continuous compactly supported, respectively continuous bounded $\psi : \mathbb{R}^d \to \mathbb{R}$. The symbol $\mathbb{E}$ denotes expectation with respect to a given probability measure on a prescribed measurable space.

Function spaces. For $U \subseteq \mathbb{R}^d$, a measure μ on $\mathcal{B}(U)$ and a normed space $(X, |\cdot|_X)$, the spaces $L^p(U; X; \mu)$, $p \in [1, \infty]$, denote the usual spaces of p-integrable measurable functions $f : U \to X$ with norm $|f|_{L^p(U;X;\mu)}^p = \int_U |f|_X^p \, d\mu$ for $p < \infty$ and $|f|_{L^\infty(U;X;\mu)} = \inf\{C > 0 : |f|_X \le C \, \mu - a.s.\}$. When $X = \mathbb{R}^1$ or $\mu = dx$, we omit X or dx from the notation. The corresponding local spaces are $L^p_{\mathrm{loc}}(U; X; \mu)$.

For $m \in \mathbb{N}$ and $p \in [1, \infty]$, $W^{m,p}_{(\mathrm{loc})}(\mathbb{R}^d; \mathbb{R}^k)$ are the usual Sobolev spaces of functions $g : \mathbb{R}^d \to \mathbb{R}^k$ with (locally) p-integrable weak partial derivatives up to order m. If $k = 1$, we write $W^{m,p}_{(\mathrm{loc})}(\mathbb{R}^d)$. For $p = 2$, $H^m(\mathbb{R}^d, \mathbb{R}^k) := W^{m,2}(\mathbb{R}^d, \mathbb{R}^k)$ are Hilbert spaces with their usual norm $|\cdot|_{H^m}$ and dual space $H^{-m}(\mathbb{R}^d, \mathbb{R}^k)$. For an open interval $I \subseteq (-\infty, +\infty)$, $W^{2,1}_2(I \times \mathbb{R}^d)$ denotes the space of functions $g : I \times \mathbb{R}^d \to \mathbb{R}$ such that g and its weak partial derivatives $\frac{d}{dt}g$, $\frac{\partial}{\partial x_i}g$ and $\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}g$, $1 \le i, j \le d$, belong to $L^2(I \times \mathbb{R}^d)$.

For $U \subseteq \mathbb{R}^d$, a Banach space X and $m \in \mathbb{N} \cup \{0\}$, $C^m(U, X)$ $(C^m_b(U, X), C^m_0(U, X))$ are the spaces of continuous (bounded, compactly supported, respectively) functions $g : U \to X$ with continuous partial derivatives up to order m, written $C(U, X)$ $(C_b(U, X), C_0(U, X))$ if $m = 0$. If $X = \mathbb{R}^1$, we write $C^m(U), C^m_b(U)$ and $C^m_0(U)$, and we set $C^\infty(U) = \bigcap_{m>0} C^m(U)$ (likewise for $C^\infty_b(U)$ and $C^\infty_0(U)$).

On $C([0, \infty), \mathbb{R}^d)$, π_t denotes the projection $\pi_t(f) := f(t)$. In time-dependent situations, we denote by π_t^s, $t \geq s$, the same map on $C([s, \infty), \mathbb{R}^d)$ and write $\pi_t^0 = \pi_t$. For an interval $I \subseteq (-\infty, +\infty)$, $C_0^{2,1}(I \times \mathbb{R}^d)$ is the space of continuous functions $g : I \times \mathbb{R}^d \to \mathbb{R}$ with compact support in $I \times \mathbb{R}^d$ such that the pointwise derivatives $\frac{d}{dt}g$, $\frac{\partial}{\partial x_i}g$ and $\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}g$, $1 \leq i, j \leq d$, are continuous on $I \times \mathbb{R}^d$.

2 The p-Laplace equation and its formulation as a Fokker–Planck equation

Regarding our Key Step 1 mentioned in the introduction, we begin by recalling the usual definition of solution to (1.2) and to its Fokker–Planck formulation (1.12). Let $d \geq 1$ and $p > 2$.

Definition 2.1. $u \in L_{\text{loc}}^1((0, \infty); W_{\text{loc}}^{1,1}(\mathbb{R}^d))$ is a *solution* to (1.2), if $\nabla u \in L_{\text{loc}}^{p-1}((0, \infty); L_{\text{loc}}^{p-1}(\mathbb{R}^d; \mathbb{R}^d))$ and for all $\varphi \in C_0^1((0, \infty) \times \mathbb{R}^d)$

$$\int_0^\infty \int_{\mathbb{R}^d} \left(-u \frac{d}{dt}\varphi + |\nabla u|^{p-2}\nabla u \cdot \nabla \varphi \right) dx dt = 0. \tag{2.1}$$

u *has initial condition* $\nu \in \mathcal{M}_b^+$, if for every $\psi \in C_0(\mathbb{R}^d)$ there is a set $J_\psi \subseteq (0, \infty)$ of full dt-measure such that

$$\int_{\mathbb{R}^d} \psi \, d\nu = \lim_{t \to 0, t \in J_\varphi} \int_{\mathbb{R}^d} u(t)\psi \, dx. \tag{2.2}$$

A solution u with initial condition ν is called *probability solution*, if ν and dt-a.e. $u(t, x)dx$ belong to $\mathcal{P}$.

Condition (2.2) for the initial condition is similar to [8, (6.1.3)]. Note that in both parts of the following remark one needs the local integrability of $t \mapsto \nabla u(t) \in L_{(\text{loc})}^{p-1}(\mathbb{R}^d; \mathbb{R}^d)$ on $[0, \infty)$, whereas in Definition 2.1 it is only required on $(0, \infty)$.

Remark 2.2.

(i) If u is a solution with initial condition ν such that $t \mapsto u(t, x)dx$ is vaguely continuous on $(0, \infty)$, then $J_\psi = (0, \infty)$ for all $\psi \in C_0(\mathbb{R}^d)$ and $t \mapsto u(t, x)dx$ extends vaguely continuous in $t = 0$ with value ν. In this case, if also $\nabla u \in L_{\text{loc}}^{p-1}([0, \infty); L_{\text{loc}}^{p-1}(\mathbb{R}^d; \mathbb{R}^d))$, (2.1)+(2.2) are equivalent to

$$\int_{\mathbb{R}^d} \psi \, u(t) \, dx = \int_{\mathbb{R}^d} \psi \, d\nu - \int_0^t \int_{\mathbb{R}^d} |\nabla u|^{p-2}\nabla u \cdot \nabla \psi \, dx dt, \quad \forall t \geq 0, \tag{2.3}$$

for every $\psi \in C_0^1(\mathbb{R}^d)$. This can be proven as in the proof of [8, Prop.6.1.2].

(ii) Every nonnegative solution u with initial condition $\nu \in \mathcal{M}_b^+$, ess $\sup_{t>0}|u(t)|_{L^1(\mathbb{R}^d)} < \infty$ and $\nabla u \in L_{\text{loc}}^{p-1}([0, \infty); L_{\text{loc}}^{p-1}(\mathbb{R}^d; \mathbb{R}^d))$ has a unique vaguely continuous dt-version $\tilde{u}$ on $[0, \infty)$. This follows from [41, Lemma 2.3]. Moreover, if $\nabla u \in L_{\text{loc}}^{p-1}([0, \infty), L^{p-1}(\mathbb{R}^d))$, then $|u(t)|_{L^1(\mathbb{R}^d)} = \nu(\mathbb{R}^d)$ for all $t \geq 0$. The latter follows by considering (2.3) for an increasing sequence $(\psi_n)_{n \in \mathbb{N}} \subseteq C_0^1(\mathbb{R}^d)$ such that $0 \leq \psi_n \leq 1$, $\psi_n(x) = 1$ for $|x| < n$, $\sup_n |\nabla \psi_n(x)| \leq C$ for all $x \in \mathbb{R}^d$ for C not depending on n or x and by letting $n \to \infty$. In this case, $t \mapsto \tilde{u}(t, x)dx$ is weakly continuous.

Now, we turn to the nonlinear Fokker–Planck formulation (1.12) of (1.2) and its notion of distributional solution.

Definition 2.3. $u \in L_{\text{loc}}^1((0, \infty); W_{\text{loc}}^{1,1}(\mathbb{R}^d))$ is a *distributional solution* to (1.12), if

$$|\nabla u|^{p-2} \in L_{\text{loc}}^1((0, \infty); W_{\text{loc}}^{1,1}(\mathbb{R}^d)), \quad |\nabla u|^{p-2}u \in L_{\text{loc}}^1((0, \infty) \times \mathbb{R}^d),$$
$$\nabla(|\nabla u|^{p-2})u \in L_{\text{loc}}^1((0, \infty) \times \mathbb{R}^d; \mathbb{R}^d), \tag{2.4}$$

6

and for all $\varphi \in C_0^2((0,\infty) \times \mathbb{R}^d)$

$$\int_0^\infty \int_{\mathbb{R}^d} u \left(\frac{d}{dt}\varphi + |\nabla u|^{p-2}\Delta\varphi + \nabla(|\nabla u|^{p-2}) \cdot \nabla\varphi \right) dx dt = 0. \tag{2.5}$$

u has initial condition $\nu \in \mathcal{M}_b^+$, *if for every* $\psi \in C_0(\mathbb{R}^d)$ *there is a set* J_ψ *as in Definition 2.1 such that (2.2) holds. A distributional solution* u *with initial condition* ν *is a* probability solution, *if* ν *and dt-a.e.* $u(t,x)dx$ *belong to* $\mathcal{P}$.

Remark 2.4. Remark 2.2 (i) holds accordingly: if u is a distributional solution to (1.12) in the sense of Definition 2.3 with initial condition ν such that $t \mapsto u(t,x)dx$ is vaguely continuous, then $J_\varphi = \emptyset$ and, in this case, if the second and third condition in (2.4) hold with $L_{\mathrm{loc}}^1([0,\infty) \times \mathbb{R}^d)$ and $L_{\mathrm{loc}}^1([0,\infty) \times \mathbb{R}^d; \mathbb{R}^d)$, replacing $L_{\mathrm{loc}}^1((0,\infty) \times \mathbb{R}^d)$ and $L_{\mathrm{loc}}^1((0,\infty) \times \mathbb{R}^d; \mathbb{R}^d)$, respectively, then (2.5) and u having initial condition ν are equivalent to

$$\int_{\mathbb{R}^d} \psi\, u(t)\, dx = \int_{\mathbb{R}^d} \psi\, d\nu + \int_0^t \int_{\mathbb{R}^d} \left(|\nabla u|^{p-2}\Delta\psi + \nabla(|\nabla u|^{p-2}) \cdot \nabla\psi \right) u\, dx dt, \ \forall t \geq 0 \tag{2.6}$$

for every $\psi \in C_0^2(\mathbb{R}^d)$.

Lemma 2.5. *Assume* $u \in L_{\mathrm{loc}}^1((0,\infty); W_{\mathrm{loc}}^{1,1}(\mathbb{R}^d))$ *satisfies* $\nabla u \in L_{\mathrm{loc}}^{p-1}((0,\infty); L_{\mathrm{loc}}^{p-1}(\mathbb{R}^d; \mathbb{R}^d))$ *and (2.4). Then* u *satisfies Definition* 2.1 *with initial condition* ν *if and only if* u *satisfies Definition* 2.3 *with initial condition* ν.

If for $\mu_0 := \nu$, $\mu_t(dx) := u(t,x)dx$, $t > 0$, *in addition, we have that* $[0,\infty) \ni t \mapsto \mu_t$ *is vaguely continuous and the above local in time integrability conditions hold on* $[0,\infty)$ *instead of* $(0,\infty)$, *then* u *satisfies (2.3) for all* $\psi \in C_0^1(\mathbb{R}^d)$ *if and only if it satisfies (2.6) for all* $\psi \in C_0^2(\mathbb{R}^d)$.

Proof. The second assertion follows from the first, since by Remarks 2.2 and 2.4, under the additional assumption, (2.3) and (2.6) are equivalent to Definitions 2.1 and 2.3, respectively. The first assertion follows, since under the assumption, the integrals on the left hand side of (2.1) and (2.5) are equal for each $\varphi \in C_0^2((0,\infty) \times \mathbb{R}^d)$, as can be seen by integration by parts. $\qquad\square$

Remark 2.6. We point out that the fundamental solutions w^y from (1.18) solve both (1.2) and (1.12) in the sense of Definitions 2.1 and 2.3, respectively, and all additional assumptions from Remarks 2.2 and 2.4 are satisfied. Details are given in Section 4.

3 Corresponding McKean–Vlasov SDE and its solution

We now turn to Key Step 2 from the introduction, i.e. we associate with (1.2) (more precisely, with (1.12)) the McKean–Vlasov SDE (1.15)–(1.16). Equation (1.12) is a nonlinear Fokker–Planck equation with coefficients as in (1.13). These coefficients are defined for those measures μ in $\mathcal{M}^+$ such that $\mu = u(x)dx$ with u, $|\nabla u|^{p-2} \in W_{\mathrm{loc}}^{1,1}(\mathbb{R}^d)$.

Definition 3.1.

(i) A *probabilistically weak solution* (short: *solution*) to (1.15)–(1.16) is a triple consisting of a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, an $(\mathcal{F}_t)$-adapted $\mathbb{R}$-valued stochastic process $X = (X(t))_{t\geq 0}$ and an $(\mathcal{F}_t)$-Brownian motion W such that

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

$$\mathcal{L}_{X(t)} = u(t,x)dx, \ \forall t > 0,$$

and $\mathbb{P}$-a.s.

$$X(t) = X_0 + \int_0^t \nabla(|\nabla u(s,X(s))|^{p-2})ds + \sqrt{2}\int_0^t |\nabla u(s,X(s))|^{\frac{p-2}{2}} dW(s), \quad \forall t \geq 0.$$

We say X *has initial condition* $\nu \in \mathcal{P}$, if $\mathcal{L}_{X(0)} = \nu$. Often, we shortly refer to X as the solution.

7

(ii) We call the path law $X_*\mathbb{P} \in \mathcal{P}(C([0, \infty), \mathbb{R}^d)$ of a solution to (1.15)–(1.16) a *solution path law* to (1.15)–(1.16). We say P *is the unique solution path law* (with initial condition $\nu \in \mathcal{P} = \mathcal{P}(\mathbb{R}^d)$), if $X_*\mathbb{P} = P$ for all weak solutions X with initial condition ν.

Remark 3.2.

(i) It is obvious how to generalize the previous definition to initial times $s > 0$. In this case, the initial condition is the pair $(s, \nu) \in [0, \infty) \times \mathcal{P}$, and the solution is defined on $[s, \infty)$.

(ii) As part of the definition it is assumed that for dt-a.e. $t > 0$ the measure $\mathcal{L}_{X(t)} \in \mathcal{P}$ has a Radon–Nikodym density $u(t, x)$ with respect to dx with sufficient Sobolev regularity in order to make sense of the appearing integrands, an aspect we explicitly deal with for the fundamental solutions from (1.18) in Section 4.

By a fundamental result by Trevisan [48, Theorem 2.5], known as *superposition principle*, in conjunction with [22, Proposition 2.2.3] and [25, Proposition 4.11], for any weakly continuous (in general measure-valued) probability solution $(\mu_t)_{t\geq 0}, \mu_t \in \mathcal{P}$, to a linear Fokker–Planck equation with measurable coefficients $a_{ij}, b_i : (0, \infty) \times \mathbb{R}^d \to \mathbb{R}, 1 \leq i, j \leq d$, in the sense of Definition D.1 there exists a (probabilistically weak) solution $X = (X(t))_{t\geq 0}$ to the SDE

$$dX(t) = b(t, X(t))dt + \sqrt{2}\, \sigma(t, X(t))dW(t),\ t \geq 0,$$

where $b = (b_i)_{i\leq d}$, $\sigma = (\sigma_{ij})_{i,j\leq d}$ with $(\sigma\sigma^\top)_{ij} = a_{ij}$, and W a d-dimensional Brownian motion, with one-dimensional time marginals $\mathcal{L}_{X(t)} = \mu_t, t \geq 0$.

In order to associate with solutions to (1.2) (more precisely, (1.12)) solutions to (1.15)–(1.16), we follow the approach developed in [3], [4], i.e. we use this superposition principle in the nonlinear case by first fixing a probability solution u to (1.12) in the sense of Definition 2.3 in the coefficients (1.13) to obtain linear (that is, (t, x)-depending) coefficients. Then, we apply Trevisan's result to the same solution u for the resulting linear equation. Clearly, the process $X = (X(t))_{t\geq 0}$ obtained this way solves (1.15)–(1.16). Thus, we obtain the following new result.

Theorem 3.3. *Let u be a weakly continuous probability solution to* (1.2) *in the sense of Definition 2.1 such that*

$$|\nabla u|^{p-2} \in L^1_{\mathrm{loc}}((0, \infty); W^{1,1}_{\mathrm{loc}}(\mathbb{R}^d)).$$

If

$$\int_0^T \int_{\mathbb{R}^d} \left(|\nabla u|^{p-2} + |\nabla(|\nabla u|^{p-2})|\right) u\, dxdt < \infty, \quad \forall T > 0, \tag{3.1}$$

then there exists a solution $X = (X(t))_{t\geq 0}$ to the McKean–Vlasov SDE (1.15)–(1.16) *in the sense of Definition* 3.1 *such that $u(t, x)dx = \mathcal{L}_{X(t)}$ for all $t \geq 0$.*

Proof. By (3.1) and Lemma 2.5, u solves equation (1.12) in the sense of Remark (2.4). Thus, $u(t, x)dx$ solves the linear Fokker–Planck equation with coefficients

$$a_{ij}(t, x) = \delta_{ij}|\nabla u(t, x)|^{p-2},\ b_i(t, x) = \frac{\partial}{\partial x_i}(|\nabla u(t, x)|^{p-2}),\ 1 \leq i, j \leq d, \tag{3.2}$$

in the sense of Definition D.1. By [48, Theorem 2.5], there is a solution $X = (X(t))_{t\geq 0}$ to the corresponding (non-distribution dependent) SDE with coefficients b and $\sqrt{2}\,\sigma$, where $b = (b_1, ..., b_d)$ is as in (3.2) and $(\sigma\sigma^\top)_{ij} = a_{ij}$, with a_{ij} as in (3.2), such that

$$\mathcal{L}_{X(t)} = u(t, x)dx,\ t \geq 0.$$

Clearly, $X = (X(t))_{t\geq 0}$ solves the McKean–Vlasov SDE (1.15)–(1.16). $\qquad\square$

By Theorem 3.3, solutions to the p-Laplace equation (1.2) are represented as the one-dimensional time marginal density curves of stochastic processes, which solve (1.15)–(1.16).

4 Dirac initial conditions and fundamental solutions

Now we implement Key Step 2 specifically for the fundamental solutions w^y from (1.18). As said before, w^y, $y \in \mathbb{R}^d$, solves both (1.2) and (1.12) in the sense of Definition 2.1 and 2.3, respectively. This follows from part (i) of the following proposition and from the proof of Theorem 4.2 together with Lemma 2.5, respectively. We collect some properties of w^y in the following proposition.

Proposition 4.1. *Let $y \in \mathbb{R}^d$.*

(i) *w^y is a weakly continuous probability solution to (1.2) with initial condition $\nu = \delta_y$ in the sense of Definition 2.1.*

(ii) *$w^y(t)$ is radially symmetric and compactly supported for every $t > 0$. More precisely $\operatorname{supp} w^y(t)$ is contained in the ball centered at y with radius $\left(\frac{C_1 t^{\frac{kp}{d(p-1)}}}{q} \right)^{\frac{p-1}{p}}$.*

(iii) *w^y is the unique nonnegative vaguely continuous solution to (1.2) in the sense of Definition 2.1 with initial condition δ_y.*

We note that w^y has the expected parabolic regularity, i.e. while the initial condition is a degenerate measure, at each time $t > 0$, $w^y(t)$ is a continuous, weakly differentiable, compactly supported function on $\mathbb{R}^d$.

Proof. (i) and (ii) can directly be inferred from the definition of w^y. For (iii), we refer to [24]. $\quad\square$

Next, we show that Theorem 3.3 applies to w^y.

Theorem 4.2. *Let $d \geq 1$, $p > 2$, and $y \in \mathbb{R}^d$. There exists a solution $X^y = (X^y(t))_{t \geq 0}$ to the McKean–Vlasov SDE (1.15)–(1.16) with $X^y(0) = y$ such that $\mathcal{L}_{X^y(t)} = w^y(t,x)dx$ for all $t > 0$.*

Corollary 4.3. *Let $s, \delta \geq 0$, $y \in \mathbb{R}^d$, $\nu = w^y(\delta, x)dx$ (with the convention $w^y(0,x)dx = \delta_y$). There exists a solution $X^{s,(\delta,y)} = (X^{s,(\delta,y)}(t))_{t \geq s}$ to (1.15), (1.16) on $[s, \infty)$ such that $\mathcal{L}_{X^{s,(\delta,y)}(t)} = w^y(\delta + t - s, x)dx$ for all $t \geq s$.*

Proof of Theorem 4.2. Since $w^y(t,x) = w^0(t, x-y)$, we only consider the case $y = 0$. We show that Theorem 3.3 applies to w^0. By Proposition 4.1, w^0 is a weakly continuous probability solution to (1.2). Thus it is sufficient to prove $|\nabla w^0|^{p-2} \in L^1_{\text{loc}}((0, \infty); W^{1,1}_{\text{loc}}(\mathbb{R}^d))$ and

$$|\nabla w^0|^{p-2} w^0, \ \nabla(|\nabla w^0|^{p-2}) w^0 \in L^1_{\text{loc}}([0, \infty); L^1(\mathbb{R}^d)). \tag{4.1}$$

Below let $C > 0$ denote a constant possibly changing from line to line, only depending on d, p, C_1 and T. A direct calculation yields

$$\nabla w^0(t,x) = -C t^{-k(1 + \frac{p}{d(p-1)})} \left(C_1 - q t^{\frac{-kp}{d(p-1)}} |x|^{\frac{p}{p-1}} \right)_+^{\frac{1}{p-2}} |x|^{\frac{2-p}{p-1}} x, \tag{4.2}$$

and, therefore,

$$|\nabla w^0(t,x)|^{p-2} = C t^{-k(p-2)(1 + \frac{p}{d(p-1)})} \left(C_1 - q t^{\frac{-kp}{d(p-1)}} |x|^{\frac{p}{p-1}} \right)_+ |x|^{\frac{p-2}{p-1}}. \tag{4.3}$$

Due to Proposition 4.1 (ii) and since $\left(C_1 - q t^{\frac{-k}{d}} |x|^{\frac{p}{p-1}} \right)_+$ is uniformly bounded on $(0, \infty) \times \mathbb{R}^d$, we find for each $T > 0$

$$\int_0^T \int_{\mathbb{R}^d} |\nabla w^0|^{p-2}(t,x) w^0(t,x) \, dxdt \leq C \int_0^T t^{-k(p-2)(1 + \frac{p}{d(p-1)})} \int_{\{|x| \leq \beta t^{\frac{k}{d}}\}} |x|^{\frac{p-2}{p-1}} w^0(t,x)dx \, dt$$

$$\leq C \int_0^T \left(t^{-k(p-2)(1 + \frac{p}{d(p-1)}) + \frac{k}{d}\frac{p-2}{p-1}} \int_{\mathbb{R}^d} w^0(t,x)dx \right) dt = \int_0^T t^{-k(p-2)(1 + \frac{p}{d(p-1)}) + \frac{k}{d}\frac{p-2}{p-1}} dt,$$

where we set $\beta := (C_1 q^{-1})^{\frac{p-1}{p}}$ and used that $w^0(t,x)dx$ is a probability measure for each $t > 0$ for the last equality. Since $-k(p-2)(1 + \frac{p}{d(p-1)}) + \frac{k}{d}\frac{p-2}{p-1} > -1$ for $p > 2$, $|\nabla w^0|^{p-2}w^0 \in L^1_{\text{loc}}([0,\infty); L^1(\mathbb{R}^d))$ follows. Likewise, since even $-k(p-2)(1 + \frac{p}{d(p-1)}) > -1$, we have $|\nabla w^0|^{p-2} \in L^1_{\text{loc}}([0,\infty); L^1(\mathbb{R}^d))$. Furthermore, (4.3) implies $|\nabla w^0(t)|^{p-2} \in W^{1,1}_{\text{loc}}(\mathbb{R}^d)$ for all $t > 0$, with

$$\nabla |\nabla w^0(t,x)|^{p-2}$$
$$= Ct^{-k(p-2)(1+\frac{p}{d(p-1)})} \left[\frac{p-2}{p-1} \left(C_1 - qt^{\frac{-kp}{d(p-1)}} |x|^{\frac{p}{p-1}} \right)_+ |x|^{-\frac{p}{p-1}} x - \frac{qp}{p-1} t^{\frac{-kp}{d(p-1)}} x \mathbb{1}_{|x| \leq \beta t^{\frac{k}{d}}}(x) \right]. \quad (4.4)$$

The estimate

$$|\nabla |\nabla w^0(t,x)|^{p-2}| \leq C \mathbb{1}_{|x| \leq \beta t^{\frac{k}{d}}}(x) t^{-k(p-2)(1+\frac{p}{d(p-1)})} \left[|x|^{-\frac{1}{p-1}} + t^{\frac{-kp}{d(p-1)}} |x| \right] \quad (4.5)$$

implies $|\nabla w^0|^{p-2} \in L^1_{\text{loc}}([0,\infty); W^{1,1}_{\text{loc}}(\mathbb{R}^d))$. Finally, (4.5) also yields

$$\nabla(|\nabla w^0|^{p-2})w^0 \in L^1_{\text{loc}}([0,\infty); L^1(\mathbb{R}^d)).$$

Thus, Theorem 3.3 applies to w^0 and the assertion follows. $\qquad\square$

For the convenience of the reader, in Appendix A we give more detailed calculations for some claims from the previous proof.

5 Nonlinear Markov processes and p-Brownian motion

5.1 The classical case: Heat equation and Brownian motion

For $p = 2$, (1.2) is the heat equation (1.1), with fundamental solution u^y, $y \in \mathbb{R}^d$, as in (1.17). The kernels $u^y(t,x)dx$, $t \geq 0, y \in \mathbb{R}^d$, where we set $u^y(0,x)dx := \delta_y$, determine a unique Markov process $(P_y)_{y \in \mathbb{R}^d} \subseteq \mathcal{P}(C([0,\infty), \mathbb{R}^d))$ with one-dimensional time marginals $(\pi_t)_* P_y = u^y(t,x)dx$ (for the definition of a Markov process, see Subsection 5.2 below). This Markov process is Brownian motion, P_0 is the *classical Wiener measure*, and P_y is obtained from P_0 as

$$P_y = (T_y)_* P_0,$$

where $T_y : C([0,\infty), \mathbb{R}^d) \to C([0,\infty), \mathbb{R}^d)$, $T_y(\omega) := \omega + y$, is the translation of a path by y. Equivalently, any stochastic process $X^y = (X^y(t))_{t \geq 0}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with path law $(X^y)_* \mathbb{P} = P_y$ is called Brownian motion with start in y ($y = 0$: *standard Brownian motion*). The associated SDE is

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

which is probabilistically weakly well-posed, i.e. the solution path law of any solution equals P_y.

5.2 Linear and nonlinear Markov processes

The following definition of nonlinear Markov processes is inspired by McKean [36] and elaborated on in [42]. As mentioned in the introduction, it is the natural extension of the usual Markov property (satisfied by the family of solution path laws to a weakly well-posed (non-distribution dependent) SDE) to nonlinear FPEs and their associated McKean–Vlasov SDEs. For more details on the motivation and properties of nonlinear Markov processes, we refer to [36] and [42].

Since equations (1.2) and (1.12) are time-dependent due to the dependence of the coefficients on $\nabla u(t,x)$, we present a definition of nonlinear Markov processes in the time-inhomogeneous case. In this regard, we also refer to Remark 3.2 (i).

Definition 5.1. Let $\mathcal{P}_0 \subseteq \mathcal{P} \ (= \mathcal{P}(\mathbb{R}^d))$. 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

(i) $(\pi_t^s)_* P_{s,\zeta} =: \mu_t^{s,\zeta} \in \mathcal{P}_0$ for all $0 \le s \le t$ and $\zeta \in \mathcal{P}_0$.

(ii) The *nonlinear Markov property* holds: for all $0 \le s \le r \le t$, $\zeta \in \mathcal{P}_0$

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

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 $p_{(s,\zeta),(r,z)}(\pi_r^r=z)=1$), and $\mathcal{F}_{s,r}$ is the σ-algebra generated by π_u^s, $s \le u \le r$.

The case of a classical (we also say *linear*) time-inhomogeneous Markov process is contained in the previous definition. In this case,

$$\mathcal{P}_0 = \mathcal{P}, \ P_{s,\zeta} = \int_{\mathbb{R}^d} P_{s,y}\, d\zeta(y), \ P_{s,y} := P_{s,\delta_y},$$

and $p_{(s,y)(r,z)} = P_{r,z}$, for all $0 \le s \le r$, $z \in \mathbb{R}^d$, $\zeta \in \mathcal{P}$. In this case, (5.1) is the standard time-inhomogeneous Markov property. Since in this linear case the map $\zeta \mapsto P_{s,\zeta}$ is linear on its domain $\mathcal{P}$ for every $s \ge 0$, the measures $(P_{s,y})_{s\ge0,y\in\mathbb{R}^d}$ determine the entire Markov process and, moreover, the *Chapman–Kolmogorov equations* for the one-dimensional time marginals $\mu_t^{s,\zeta}$ (see (i) of the previous definition) hold, i.e.

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

In contrast to this, in the general nonlinear case the map $\zeta \to P_{s,\zeta}$ is not linear on $\mathcal{P}_0$, even if $\mathcal{P}_0 = \mathcal{P}$ (which we do not assume). As a consequence, one loses the Chapman–Kolmogorov equations, but the one-dimensional time marginals still satisfy the *flow property*

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

(in the linear case, this follows from (5.2)).

The following theorem from [42] is the key for our main result of this section, Theorem 5.4 below.

Theorem 5.2. [42, Thms. 3.4+3.8] *Let $\mathcal{P}_0 \subseteq \mathcal{P}$ be a class of initial conditions and $(\mu_t^{s,\zeta})_{0\le s\le t,\zeta\in\mathcal{P}_0} \subseteq \mathcal{P}_0$ such that*

(i) *For all $(s,\zeta) \in [0,\infty)\times\mathcal{P}_0$, $(\mu_t^{s,\zeta})_{t\ge s}$ is a weakly continuous probability solution to the nonlinear FPE (1.3) with initial condition (s,ζ) in the sense of Definition D.2, satisfying the flow property (5.3).*

(ii) *For all $(s,\zeta) \in [0,\infty)\times\mathcal{P}_0$, $(\mu_t^{s,\zeta})_{t\ge s}$ is the unique weakly continuous probability solution to the linear FPE (D.1) 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 D.1 in the class*

$$\{(\nu_t)_{t\ge s} \subseteq \mathcal{P} : \nu_t \le C\mu_t^{s,\zeta}, \quad t \ge s, \ \text{for some } C > 0\} \qquad (5.4)$$

("restricted linearized distributional uniqueness").

Then, for each $(s,\zeta) \in [0,\infty)\times\mathcal{P}_0$ there is a unique solution path law $P_{s,\zeta}$ to the McKean–Vlasov SDE (1.5)–(1.6) such that

$$(\pi_t^s)_* P_{s,\zeta} = \mu_t^{s,\zeta}, \quad \forall 0 \le s \le t, \zeta \in \mathcal{P}_0, \qquad (5.5)$$

and the family $(P_{s,\zeta})_{s\ge0,\zeta\in\mathcal{P}_0}$ 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 $(\mu_t^{s,\zeta})_{0\le s\le t,\zeta\in\mathcal{P}_0}$ and equations (1.5), (1.6).

11

More precisely, in the present work we will not apply the previous theorem, but the following corollary. The idea behind it is that the statement of the previous theorem remains true if the restricted linearized distributional uniqueness condition is known for fewer initial data, at the price of a uniqueness assertion for the corresponding solution path laws, which holds for fewer initial data as well.

Corollary 5.3. [42, Corollary 3.10] *Let $\mathfrak{P}_0 \subseteq \mathcal{P}_0 \subseteq \mathcal{P}$ and suppose $(\mu_t^{s,\zeta})_{0 \leq s \leq t, \zeta \in \mathcal{P}_0}$ satisfies (i) from the previous theorem, but (ii) only for $(s,\zeta) \in [0,\infty) \times \mathfrak{P}_0$. Further, assume $\mu_t^{s,\zeta} \in \mathfrak{P}_0$ for all triples (s,t,ζ) such that either $s \leq t$, $\zeta \in \mathfrak{P}_0$ or $s < t$, $\zeta \in \mathcal{P}_0$. Then, there is a nonlinear Markov process $(P_{s,\zeta})_{s \geq 0, \zeta \in \mathcal{P}_0}$ with (5.5), consisting of solution path laws to (1.5), (1.6). The uniqueness-assertion for $P_{s,\zeta}$ of Theorem 5.2 remains true for initial data (s,ζ) with $\zeta \in \mathfrak{P}_0$.*

For the proof of Theorem 5.4 below, we apply Corollary 5.3 to the nonlinear FPE (1.12) and the corresponding McKean–Vlasov SDE (1.15), (1.16). Before we do this, let us describe the main idea of the proof of the general Theorem 5.2 from above.

Idea of proof of Theorem 5.2. For the uniqueness assertion, let $(s,\zeta) \in [0,\infty) \times \mathcal{P}_0$ and P_1, P_2 be two solution path laws to (1.5), (1.6) with

$$(\pi_t^s)_* P_i = \mu_t^{s,\zeta}, \quad \forall t \geq s, i \in \{1,2\}. \tag{5.6}$$

It suffices to prove

$$\mathbb{E}_{P_1}[H] = \mathbb{E}_{P_2}[H] \tag{5.7}$$

for all $H = \Pi_{i=1}^n h_i(\pi_{t_i}^s)$, $n \in \mathbb{N}$, $s \leq t_1 < \cdots < t_n$, $h_i : \mathbb{R}^d \to \mathbb{R}$ measurable, $c_i \leq h_i \leq C_i$ for some $0 < c_i < C_i$ (indeed, the set of such H is closed under pointwise multiplication and generates the Borel σ-algebra of $C([s,\infty),\mathbb{R}^d)$ with respect to the topology of locally uniform convergence). We call n the *length* of H. By (5.6), (5.7) holds for all H of length 1, and we proceed via induction over n. Assuming (5.7) is valid for all H of length n, consider an arbitrary $H = \Pi_{i=1}^{n+1} h_i(\pi_{t_i}^s)$ as above. By [48, Lemma 2.6], the curves

$$\eta_t^i := (\pi_{t_n}^s)_*(\varrho P_i), \quad t \geq t_n,$$

where ϱ, defined via

$$\varrho := \frac{\Pi_{i=1}^n h_i(\pi_{t_i}^s)}{\mathbb{E}_{P^1}[\Pi_{i=1}^n h_i(\pi_{t_i}^s)]},$$

is a probability density with respect to P_i, $i \in \{1,2\}$, such that there is $0 < c < C$ with $c \leq \varrho \leq C$, are weakly continuous probability solutions to the linear FPE with coefficients $a_{ij}(t,x,\mu_t^{s,\zeta})$ and $b_i(t,x,\mu_t^{s,\zeta})$ from the initial condition (t_n,η) in the sense of Definition D.1, where

$$\eta := (\pi_{t_n}^s)_*(\varrho P_1) = (\pi_{t_n}^s)_*(\varrho P_2) \in \mathcal{P}.$$

Indeed, the latter equality follows from the induction assumption. From the definition of ϱ it follows that $(\eta_t^i)_{t \geq t_n}$, $i \in \{1,2\}$, belong to the set (5.4), with s and ζ replaced by t_n and $\mu_{t_n}^{s,\zeta}$, respectively. Moreover, the definition of ϱ entails that η is equivalent to $\mu_{t_n}^{s,\zeta}$ with a density bounded between two strictly positive constants. Due to the flow property, we have $\mu_t^{t_n, \mu_{t_n}^{s,\zeta}} = \mu_t^{s,\zeta}$, $t \geq t_n$, and so assumption (ii) and [42, Lemma 3.7(ii)] imply $\eta_t^1 = \eta_t^2$, $t \geq t_n$, i.e. in particular $\eta_{t_{n+1}}^1 = \eta_{t_{n+1}}^2$. Thus, we infer

$$\frac{\mathbb{E}_{P^i}\left[\Pi_{i=1}^{n+1} h_i(\pi_{t_i}^s)\right]}{\mathbb{E}_{P^1}\left[\Pi_{i=1}^n h_i(\pi_{t_i}^s)\right]} = \int_{C([s,\infty),\mathbb{R}^d)} \varrho(w) h_{n+1}(\pi_{t_{n+1}}^s(w)) \, dP^i(w) = \int_{\mathbb{R}^d} h_{n+1}(x) \, d\eta_{t_{n+1}}^i(x),$$

which together with the induction assumption concludes the proof of the uniqueness assertion.

Regarding the nonlinear Markov property, for $0 \leq s \leq r \leq t$ and $\zeta \in \mathcal{P}_0$, it suffices to prove for all $A \in \mathcal{B}(\mathbb{R}^d)$, $n \in \mathbb{N}$, $s \leq t_1 < \cdots < t_n \leq r$, and $h : (\mathbb{R}^d)^n \to \mathbb{R}$ measurable with $c \leq h \leq C$ for some $0 < c < C$

$$\mathbb{E}_{s,\zeta}\left[h(\pi_{t_1}^s, \ldots, \pi_{t_n}^s) \mathbb{1}_{\pi_t^s \in A}\right] = \int_{C([s,\infty),\mathbb{R}^d)} p_{(s,\zeta),(r,\pi_r^s(\omega))}(\pi_t^r \in A) h(\pi_{t_1}^s(\omega), \ldots, \pi_{t_n}^s(\omega)) \, dP_{s,\zeta}(\omega) \tag{5.8}$$

12

(where $\mathbb{E}_{s,\zeta}$ denotes expectation with respect to $P_{s,\zeta}$ and $p_{(s,\zeta),(r,z)}$, $z \in \mathbb{R}^d$, is as in Definition 5.1, and to apply a monotone class argument. To prove (5.8), define the probability measures Q_g and Q^θ on $C([r,\infty),\mathbb{R}^d)$ as follows. First, define

$$Q_g := \int_{\mathbb{R}^d} p_{(s,\zeta),(r,z)} g(z) \, d\mu_r^{s,\zeta}(z),$$

i.e.

$$Q_g(C) = \int_{\mathbb{R}^d} p_{(s,\zeta),(r,z)}(C) g(z) \, d\mu_r^{s,\zeta}(z), \quad \forall C \in \mathcal{B}(C([r,\infty),\mathbb{R}^d)),$$

where $g = c_0 \tilde{g}$, with $\tilde{g} : \mathbb{R}^d \to \mathbb{R}$ the $\mu_r^{s,\zeta}$-a.s. unique map such that

$$\mathbb{E}_{s,\zeta}\big[h(\pi_{t_1}^s, \ldots, \pi_{t_n}^s)|\pi_r^s\big] = \tilde{g}(\pi_r^s) \quad \mathbb{P}_{s,\zeta} - a.s.$$

obtained from the factorization lemma and c_0 a normalizing constant such that $\int_{\mathbb{R}^d} g \, d\mu_r^{s,\zeta} = 1$. Second, set

$$Q^\theta := (\Lambda_r^s)_*(\theta P_{s,\zeta}),$$

where $\Lambda_r^s := C([s,\infty),\mathbb{R}^d) \to C([r,\infty),\mathbb{R}^d)$, $\Lambda_r^s(\omega) := \omega_{|[r,\infty)}$ is the projection of a path $\omega = (\omega(t))_{t \geq s}$ to $(\omega(t))_{t \geq r}$, and

$$\theta : C([s,\infty),\mathbb{R}^d) \to \mathbb{R}, \quad \theta := c_0 h(\pi_{t_1}^s, \ldots, \pi_{t_n}^s)$$

with c_0 as above. By [48, Lemma 2.8] and [48, Proposition 2.6], $t \mapsto (\pi_t^r)_* Q_g$ and $t \mapsto (\pi_t^r)_* Q^\theta$, respectively, are weakly continuous probability solutions to the linear FPE with coefficients $(t,x) \mapsto a_{ij}(t,x,\mu_t^{s,\zeta})$ and $(t,x) \mapsto b_i(t,x,\mu_t^{s,\zeta})$ in the sense of Definition D.1, and their common initial condition is $(r, g\mu_r^{s,\zeta})$. Since g and θ are bounded between two strictly positive constants, both of the above $\mathcal{P}$-valued curves belong to the set (5.4), with (s,ζ) replaced by $(r,\mu_r^{s,\zeta})$. Using the uniqueness assumption from (ii) for the initial condition $(r,\mu_r^{s,\zeta})$ instead of (s,ζ) as well as [42, Lemma 3.7 (ii)], we obtain

$$(\pi_t^r)_* Q_g = (\pi_t^r)_* Q^\theta, \quad \forall t \geq r.$$

We note that the flow property of $(\mu_t^{s,\zeta})_{s \leq t, \zeta \in \mathcal{P}_0}$ was implicitly used several times in the above arguments. Now we conclude (5.8) via the definitions of Q^θ and Q_g by noting

$$\mathbb{E}_{s,\zeta}\big[h(\pi_{t_1}^s, \ldots, \pi_{t_n}^s) \mathbb{1}_{\pi_t^s \in A}\big] = c_0^{-1} Q^\theta \circ (\pi_t^r)^{-1}(A) = c_0^{-1} Q_g \circ (\pi_t^r)^{-1}(A)$$

$$= c_0^{-1} \int_{C([s,\infty),\mathbb{R}^d)} p_{(s,\zeta),(r,\pi_r^s(\omega))}(\pi_t^r \in A) g(\pi_r^s(\omega)) \, dP_{s,\zeta}(\omega)$$

$$= \int_{C([s,\infty),\mathbb{R}^d)} p_{(s,\zeta),(r,\pi_r^s(\omega))}(\pi_t^r \in A) h(\pi_{t_1}^s(\omega), \ldots, \pi_{t_n}^s(\omega)) \, dP_{s,\zeta}(\omega).$$

5.3 p-Brownian motion

Now, in analogy to the case $p = 2$ presented in Section 5.1, we implement Key Step 3 as follows: We consider the solutions $X^y = (X^y(t))_{t \geq 0}$, $y \in \mathbb{R}^d$, to (1.15), (1.16) constructed in Theorem 4.2, more precisely their path laws, and prove that these laws are uniquely determined by $w^y(t,x)$, $y \in \mathbb{R}^d$, $t > 0$, and (1.15), (1.16), as well as that they satisfy the nonlinear Markov property. To do so, we apply Corollary 5.3 to (1.12), (1.15), (1.16) and w^y, $y \in \mathbb{R}^d$, as follows. Set

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

with the convention $w^y(0,x)dx = \delta_y$. We note that for each $\zeta \in \mathcal{P}_0$, the pair $(\delta,y) \in [0,\infty) \times \mathbb{R}^d$ for the representation $\zeta = w^y(\delta,x)dx$ is unique.

Theorem 5.4. Let $d \geq 2$, $p > 2\left(1 + \frac{1}{d}\right)$.

13

(i) *Let $(s, \zeta) \in [0, \infty) \times \mathcal{P}_0$, $\zeta = w^y(\delta, x)dx$. The set of solution path laws to the McKean–Vlasov SDE (1.15)–(1.16) 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 (1.15)–(1.16) and $\omega^y(t)$, $y \in \mathbb{R}^d$, $t \geq 0$.*

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

$$\hat{\Pi}_s : C([0, \infty), \mathbb{R}^d) \to C([s, \infty), \mathbb{R}^d), \quad \hat{\Pi}_s : (\omega(t))_{t \geq 0} \mapsto (\omega(t - s))_{t \geq s}. \tag{5.9}$$

Moreover, for $\zeta = w^y(\delta, x)dx$, we have $P_{0,\zeta} = (\Pi_\delta^0)_* P_{0,y}$ (the map Π_δ^0 is defined in (5.10) below).

Remark 5.5. For $\zeta = w^y(\delta, x)dx$, $P_{s,\zeta}$ from the previous theorem is the path law of the solution $X^{s,(\delta,y)}$ from Corollary 4.3, i.e. when the latter is defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, then

$$P_{s,\zeta} = (X^{s,(\delta,y)})_* \mathbb{P}.$$

To avoid confusion, we remark 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 do not only consider the initial pair (δ, ζ), but necessarily *any* (s, ζ), $s \geq 0$. Set

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

In the following proof, a crucial ingredient is Theorem 6.5, which is formulated and proven in Section 6.

Proof of Theorem 5.4.

(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). By Lemma 2.5, Proposition 4.1 and (4.1), $(\mu_t^{s,\zeta})_{t \geq s}$ is a weakly continuous probability solution to the nonlinear FPE (1.12) for each $(s, \zeta) \in [0, \infty) \times \mathcal{P}_0$ in the sense of Definition D.2. Moreover, by Theorem 6.5 below, condition (ii) of Theorem 5.2 holds for all $(\mu_t^{s,\zeta})_{t \geq s}$ with $(s, \zeta) \in [0, \infty) \times \mathfrak{P}_0$. Since also the final condition of Corollary 5.3 is satisfied, we may apply Corollary 5.3 to obtain:

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) $(\pi_t^s)_* P_{s,\zeta} = \mu_t^{s,\zeta}, t \geq s$;

(II) $P_{s,\zeta}$ is a solution path law to the McKean–Vlasov SDE (1.15)–(1.16) on $[s, \infty)$;

(III) For $s \geq 0$ and $\zeta \in \mathfrak{P}_0$, $P_{s,\zeta}$ is the *unique* element from $\mathcal{P}(C([s, \infty), \mathbb{R}^d))$ 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 $s \geq 0, \zeta = \delta_y$, $y \in \mathbb{R}^d$. We define for $\omega = \omega(t)_{t \geq s}$ the map $\Pi_r^s : C([s, \infty), \mathbb{R}^d) \to C([s, \infty), \mathbb{R}^d)$ via

$$\Pi_r^s : \omega(t)_{t \geq s} \mapsto \omega(t + r)_{t \geq s}. \tag{5.10}$$

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

$$(\pi_t^s)_* ((\Pi_r^s)_* P^i) = (\pi_{t+r}^s)_* P^i = \mu_{t+r}^{s,\delta_y} = w^y(t + r - s, x)dx, \quad \forall t \geq s, \tag{5.11}$$

14

where the second equality is due to (I). It is straightforward to check that $(\Pi_r^s)_* P^i \in \mathcal{P}(C([s,\infty), \mathbb{R}^d))$, $i \in \{1,2\}$, is a solution path law to the McKean–Vlasov SDE (1.15)–(1.16) with initial condition $(s, w^y(r,x)dx)$. As $w^y(r,x)dx \in \mathfrak{P}_0$, (5.11) and (III) yield

$$(\Pi_r^s)_* P^1 = (\Pi_r^s)_* P^2.$$

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

$$(\pi_{u_1}^s, ..., \pi_{u_n}^s)_* P^i = (\pi_s^s, ..., \pi_{u_n+s-u_1}^s)_*((\Pi_{u_1-s}^s)_* P^i),$$

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 $s = u_1 < \cdots < u_n$. Then, since $(\pi_s^s)_* P^i = \delta_y$, we find

$$(\pi_{u_1}^s, ..., \pi_{u_n}^s)_* P^i = \delta_y \otimes ((\pi_{u_2}^s, ..., \pi_{u_n}^s)_* P^i)$$

(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

$$(\pi_{u_1}^s, ..., \pi_{u_n}^s)_* P^1 = (\pi_{u_1}^s, ..., \pi_{u_n}^s)_* P^2$$

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

(ii) First note

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

with $\hat{\Pi}_s$ as in (5.9). Both $P_{s,\zeta}$ and $(\hat{\Pi}_s)_* P_{0,\zeta}$ are solution path laws to the McKean–Vlasov equation (1.15)–(1.16) on $[s,\infty)$ with initial condition (s,ζ), hence, by (i),

$$P_{s,\zeta} = (\hat{\Pi}_s)_* P_{0,\zeta}.$$

For the final statement, note that for $\zeta = w^y(\delta, x)dx$ the measures $P_{0,\zeta}$ and $(\Pi_\delta^0)_* P_{0,y}$ have identical one-dimensional time marginals $w^y(\delta + t, x)dx$, $t \geq 0$, and both are solution path laws to (1.15)–(1.16) with initial condition $(0, w^y(\delta, x)dx)$. Thus the assertion follows from the uniqueness assertion in (i). $\square$

Remark 5.6. Theorem 5.4 (ii) implies that the nonlinear Markov process $(P_{s,\zeta})_{(s,\zeta) \in \mathbb{R}_+ \times \mathcal{P}_0}$ is uniquely determined by

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

Therefore, we also refer to $(P_y)_{y \in \mathbb{R}^d}$ as the *unique nonlinear Markov process determined by* (1.15)–(1.16) *and the one-dimensional time marginals* $w^y(t), y \in \mathbb{R}^d, t \geq 0$.

In analogy to the linear case (discussed in Section 5.1), we now define p-Brownian motion as the unique nonlinear Markov process, as mentioned in the previous remark.

Definition 5.7. Let $d \geq 2$, $p > 2\left(1 + \frac{1}{d}\right)$. We call the family of path laws $(P_y)_{y \in \mathbb{R}^d}$ from Remark 5.6 *p-Brownian motion*.

Remark 5.8.

(i) Moreover, in analogy to the linear case $p = 2$, we call P_0 the *p-Wiener measure*.

(ii) As in the case $p = 2$, we also call any stochastic process $X^y = (X^y(t))_{t \geq 0}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ a *p-Brownian motion with start in y*, if $(X^y)_* \mathbb{P} = P_y$, with P_y as in Remark 5.6. The one-dimensional time marginals of X^y are $\mathcal{L}_{X^y(t)} = w^y(t,x)dx$, $t > 0$, $\mathcal{L}_{X^y(0)} = \delta_y$.

Remark 5.9. We point out that for $p > 2$, unlike in the case $p = 2$, the measures P_y are not given as the image measure of P_0 under the translation map $T_y : C([0,\infty), \mathbb{R}^d) \to C([0,\infty), \mathbb{R}^d), T_y(\omega) := \omega + y$ (compare Section 5.1). This follows, for instance, from the fact that $(T_y)_* P_0$ is *not* the solution path law of the McKean–Vlasov SDE (1.15)–(1.16) wtih initial condition δ_y.

6 A crucial uniqueness result for the linearized p-Laplace equation

As mentioned before, a crucial result used for the proof of Theorem 5.4 is a restricted distributional uniqueness result for the linearized p-Laplace equation. Here, we formulate and prove this result (see Theorem 6.5 below) and, thereby, complete Key Step 3.

6.1 A restricted linearized uniqueness result

Here, we set for $\delta > 0$,

$$w_\delta(t, x) := w^0(t + \delta, x), \quad \varrho_\delta(t, x) := |\nabla w_\delta(t)|^{p-2}(x), \ (t, x) \in [0, \infty) \times \mathbb{R}^d.$$

The results of this subsection hold for w^y instead of w^0 for every initial value $y \in \mathbb{R}^d$, but for simplicity we only mention the case $y = 0$. On $Q_T := (0, T) \times \mathbb{R}^d$, where $T > 0$ or $T = \infty$, we consider the linearized version of equation (1.12)

$$\frac{d}{dt} u = \Delta(\varrho_\delta u) - \operatorname{div}(\nabla \varrho_\delta u), \tag{6.1}$$

which is a linear Fokker–Planck equation obtained by fixing in the nonlinear equation (1.12) a priori the coefficients ϱ_δ and $\nabla \varrho_\delta$ in place of $|\nabla u|^{p-2}$ and $\nabla(|\nabla u|^{p-2})$, see also Appendix D.

Definition 6.1. By a *distributional solution to* (6.1) *with initial condition* $\nu \in \mathcal{M}_b^+$ we mean 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

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

such that for all $\psi \in C_0^\infty(\mathbb{R}^d)$

$$\int_{\mathbb{R}^d} \psi u(t) \, dx - \int_{\mathbb{R}^d} \psi \, d\nu = \int_0^t \int_{\mathbb{R}^d} (\varrho_\delta \Delta \psi + \nabla \varrho_\delta \cdot \nabla \psi) \, u \, dxdt, \quad \forall 0 < t < T. \tag{6.2}$$

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

Remark 6.2. An equivalent condition to (6.2) is

$$\int_{(0,T)\times\mathbb{R}^d} \left(\frac{d\varphi}{dt} + \operatorname{div}(\varrho_\delta \nabla \varphi) \right) u \, dxdt + \int_{\mathbb{R}^d} \varphi(0) \, d\nu = 0 \tag{6.3}$$

for all $\varphi \in C_0^\infty([0, T) \times \mathbb{R}^d)$.

Distributional solutions to (6.1) are not necessarily related to w_δ, but clearly $u(t, x) = w_\delta(t, x)$ is a distributional solution to (6.1) in the sense of the previous definition with initial condition $w_\delta(0, x)dx = w^0(\delta, x)dx$.

We collect some basic properties of w_δ and ϱ_δ used in the sequel.

Lemma 6.3. *Let* $\delta > 0$. *Then* w_δ *and* ϱ_δ *are nonnegative functions with the following properties.*

(i) $w_\delta \in \bigcap_{R>0} C_0([0, R] \times \mathbb{R}^d) \cap L^\infty([0, \infty) \times \mathbb{R}^d)$.

(ii) $\varrho_\delta \in \bigcap_{R>0} C_0([0, R] \times \mathbb{R}^d)$ *and* $\{x \in \mathbb{R}^d : \varrho_\delta(t, x) > 0\} = \{x \in \mathbb{R}^d : 0 < |x| \leq \beta(t + \delta)^{\frac{k}{d}}\}$ *for all* $t \geq 0$, *where* $\beta = (\frac{C_1}{q})^{\frac{p-1}{p}}$.

(iii) $\nabla \varrho_\delta \in L_{\text{loc}}^\infty([0, \infty), L^q(\mathbb{R}^d; \mathbb{R}^d))$ *for* $q \in [1, d(p-1))$.

Proof. The nonnegativity of w_δ and ϱ_δ as well as (i) and (ii) follow from their definitions. Regarding (iii), by taking into account (4.5) and considering d-dimensional spherical coordinates, it suffices to note that $\int_0^1 r^{-\frac{q}{p-1}+d-1} dr < \infty$ holds. Indeed, $-\frac{q}{p-1} + d > 0 \iff q < d(p-1)$. $\qquad\square$

Comparing with Theorem 5.2 and the sets (5.4), we are interested in uniqueness of distributional solutions to (6.1) in the sense of Definition 6.1 in the class of probability solutions

$$\mathcal{A}_{\delta,T} := \left\{ u \in L^1 \cap L^\infty(Q_T) : t \mapsto u(t,x)dx \in C([0,T],\mathcal{P}), \exists C \geq 1 : u(t,x)dx \leq Cw_\delta(t,x)dx, \forall t \in [0,T] \right\},$$

where the inequality $u(t,x)dx \leq Cw_\delta(t,x)dx$ is understood set-wise, i.e.

$$\int_A u(t,x)dx \leq C \int_A w_\delta(t,x)dx, \quad \forall A \in \mathcal{B}(\mathbb{R}^d),$$

and $C([0,T],\mathcal{P})$ is the set of weakly continuous maps $t \mapsto \mu_t \in \mathcal{P}$. If $T = \infty$, the interval $[0,T]$ in the definition of $\mathcal{A}_{\delta,\infty}$ is understood as $[0,\infty)$.

Remark 6.4. Clearly, $w_\delta \in \mathcal{A}_{\delta,T}$. Note that $u \in \mathcal{A}_{\delta,T}$ implies $u(t,x)dx = (\eta(t,x)w_\delta(t,x))dx$ with $0 \leq \eta \leq C$ for some $C \geq 1$ $dxdt$-a.s. and, for each $t \in [0,T]$, $\operatorname{supp}\eta(t) \subseteq \operatorname{supp} w_\delta(t)$.

The following result was crucially used in the proof of Theorem 5.4.

Theorem 6.5. *Let $d \geq 2$, $p > 2(1+\frac{1}{d})$, $\delta > 0$, $0 < T < \infty$ and $v_1 \in \mathcal{A}_{\delta,T}$ be a distributional solution to equation (6.1) in the sense of Definition 6.1 with initial condition $w_\delta(0,x)dx = w^0(\delta,x)dx$. Then $v_1 = w_\delta$ $dxdt$-a.s. on Q_T. In particular, w_δ is the only distributional solution to (6.1) on $(0,\infty) \times \mathbb{R}^d$ in $\mathcal{A}_{\delta,\infty}$ with initial condition $w_\delta(0,x)dx$.*

As explained in Section 5, for the proof of Theorem 5.4, we need the result of Theorem 6.5 for all initial times $s \geq 0$. More precisely, for $s \geq 0$, one considers $w_\delta(t-s)$ and $\varrho_\delta(t-s)$ instead of $w_\delta(t)$ and $\varrho_\delta(t)$ in equation (6.1) and considers (6.1) on $Q_{s,T} := (s,T) \times \mathbb{R}^d$ instead of Q_T. It is obvious how to extend Definition 6.1 and the definition of the sets $\mathcal{A}_{\delta,T}$ in this regards, and also that Theorem 6.5 holds accordingly.

Remark 6.6.

(i) We would like to mention here that Theorem 6.5 appears to be really new, though it is a uniqueness result for a linear Fokker–Planck equation which, however, is degenerate. For uniqueness results for the latter we refer to [8, Section 9.8] and the references therein. It should be noted that [8, Theorem 9.8.2] (see also Theorem 1 and its complete proof in the original work [9]) is close to what we need for our case. But it requires that $\varrho_\delta^{-\frac{1}{2}} \nabla \varrho_\delta$ is bounded, which, however, is not true in our case (see (4.4)), so it does not apply. Our proof of Theorem 6.5 below is strongly based on the explicit form of w_δ and ϱ_δ. In this respect we, in particular, refer to the proof of Claim 4 below.

(ii) As an easy consideration shows, the assertion of Theorem 6.5 is equivalent to the statement that w_δ is an extreme point in the convex set of all probability solutions to (6.1) with initial condition $w^0(\delta,x)$, see also [42, Lemma 3.5].

6.2 Proof of Theorem 6.5

Let $T \in (0,\infty)$, $\delta > 0$, $d \geq 2$, and $p > 2\left(1+\frac{1}{d}\right)$.

Remark 6.7. We note that in Remark 6.2, since

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

by a standard localization argument we can replace $C_0^\infty([0,T) \times \mathbb{R}^d)$ by

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

For the convenience of the reader we include a proof in Appendix B.

Proof of Theorem 6.5. Let $v_1 \in \mathcal{A}_{\delta,T}$, i.e. there is $\eta \in L^1(Q_T) \cap L^\infty(Q_T)$ such that $\eta \geq 0$, $|\eta(t)|_{L^1(\mathbb{R}^d)} = 1$ for all $t \in [0,T]$, and $v_1 = \eta w_\delta$. Define

$$v := w_\delta - v_1 = (1 - \eta)w_\delta.$$

Note that v is a (signed) distributional solution to (6.1) in the sense of Definition 6.1 with initial condition $v(0) \equiv 0$. Let $f \in C_0^\infty(Q_T)$ and for $\varepsilon \in (0,1)$ consider the equation

$$\frac{\partial \varphi_\varepsilon}{\partial t} + (\varrho_\delta + \varepsilon)\Delta\varphi_\varepsilon + \nabla\varrho_\delta \cdot \nabla\varphi_\varepsilon = f \quad \text{on } Q_T,$$
$$\varphi_\varepsilon(T, x) = 0, \ x \in \mathbb{R}^d. \tag{6.4}$$

By Lemma 6.3 (iii) and because $p > 2\left(1 + \frac{1}{d}\right)$, we have

$$|\nabla\varrho_\delta| \in L^\infty((0,T); L^{d+2}), \tag{6.5}$$

so we may apply [30, Theorem 9.1, p. 341] to obtain a unique solution φ_ε to (6.4) with

$$\frac{\partial\varphi_\varepsilon}{\partial t}, \frac{\partial}{\partial x_i}\varphi_\varepsilon, \frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j}\varphi_\varepsilon \in L^2(Q_T), \ \forall i,j = 1, ..., d. \tag{6.6}$$

Hence, there exists a dt-version $\widetilde{\varphi}_\varepsilon$, again denoted by φ_ε, such that $\varphi_\varepsilon : [0,T] \to L^2$ is absolutely continuous, $\frac{d\varphi_\varepsilon}{dt} \in L^2((0,T); L^2)$ and, in particular, it follows that

$$\varphi_\varepsilon \in C([0,T], L^2). \tag{6.7}$$

Moreover, by the maximum principle

$$\sup_{\varepsilon \in (0,1)} |\varphi_\varepsilon|_{L^\infty(Q_T)} \leq C_0 |f|_{L^\infty(Q_T)}. \tag{6.8}$$

This follows in a standard way from (6.4) 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 (6.4) 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

$$\frac{1}{2}|\varphi_\varepsilon(t)|^2_{L^2(\mathbb{R}^d)} + \int_t^T \int_{\mathbb{R}^d} \varrho_\delta^\varepsilon |\nabla\varphi_\varepsilon|^2 dxds \leq C_1 \int_t^T \int_{\mathbb{R}^d} |f|^2 dxds, \ \forall t \in (0,T). \tag{6.9}$$

Now, we define for $\lambda \in (0,1)$

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

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

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

Below, we abbreviate $\frac{d}{dt}f$ by f_t.

Claim 1. We have

$$\int_0^T {}_{H^1}\langle \varrho_\delta^\varepsilon(t)\nabla\varphi_\varepsilon(t), \nabla(\varphi_\varepsilon)_t(t)\rangle_{H^{-1}} dt \leq -\frac{1}{2}\int_{Q_T} (\varrho_\delta)_t |\nabla\varphi_\varepsilon(t, x)|^2 dt dx, \tag{6.10}$$

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

18

Proof of Claim 1. We set $p^* := \frac{2d}{d-2}$, $q^* := d$ if $d \geq 3$, and $p^* := q^* := 4$ if $d = 2$, and note that by (6.6), we have $\nabla(\varphi_\varepsilon)_t \in L^2((0,T); H^{-1})$ and, for a.e. $t \in (0,T)$,

$$|\varrho_\delta^\varepsilon(t)\nabla\varphi_\varepsilon(t)|_{H^1} \leq |\nabla\varrho_\delta(t) \cdot \nabla\varphi_\varepsilon(t)|_{L^2(\mathbb{R}^d)} + |\varrho_\delta^\varepsilon(t)\Delta\varphi_\varepsilon(t)|_{L^2(\mathbb{R}^d)}$$

$$\leq |\varrho_\delta^\varepsilon(t)|_{L^\infty(\mathbb{R}^d)}|\Delta\varphi_\varepsilon(t)|_{L^2(\mathbb{R}^d)} + \left(\int_{\mathbb{R}^d} |\nabla\varrho_\delta(t,x)|^2|\nabla\varphi_\varepsilon(t,x)|^2 dx\right)^{\frac{1}{2}}$$

$$\leq |\varrho_\delta^\varepsilon(t)|_{L^\infty(\mathbb{R}^d)}|\Delta\varphi_\varepsilon(t)|_{L^2(\mathbb{R}^d)} + C_2|\nabla\varphi_\varepsilon(t)|_{L^{p^*}(\mathbb{R}^d)}|\nabla\varrho_\delta(t)|_{L^{q^*}(\mathbb{R}^d)} \leq C_3|\varphi_\varepsilon(t)|_{H^2(\mathbb{R}^d)},$$

because, by Lemma 6.3 (iii), $|\nabla\varrho_\delta|_{L^{q^*}} \in L^\infty(0,T)$. Since (6.6) yields $|\varphi_\varepsilon|_{H^2} \in L^2(0,T)$, we infer that $\varrho_\delta^\varepsilon\nabla\varphi_\varepsilon \in L^2((0,T); H^1)$ and so the left hand side of (6.10) is well defined. We choose a sequence $\{\varphi_\nu^\varepsilon\} \subset C^1([0,T]; H^1)$ such that $\varphi_\nu^\varepsilon(T, \cdot) = 0$ and, for $\nu \to 0$,

$$\begin{array}{rll} \nabla\varphi_\nu^\varepsilon &\to& \nabla\varphi_\varepsilon \qquad \text{strongly in } L^2((0,T); H^1) \\ \nabla(\varphi_\nu^\varepsilon)_t &\to& \nabla(\varphi_\varepsilon)_t \quad \text{strongly in } L^2((0,T); H^{-1}) \end{array} \tag{6.11}$$

Such a sequence might be, for instance,

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

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

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

$$= -\frac{1}{2}\int_{\mathbb{R}^d} \varrho_\delta^\varepsilon(0,x)|\nabla\varphi_\nu^\varepsilon(0,x)|^2 dx - \frac{1}{2}\int_{Q_T} \varrho_\delta(t,x)|\nabla\varphi_\nu^\varepsilon(t,x)|^2 dtdx$$

$$\leq -\frac{1}{2}\int_{Q_T} (\varrho_\delta)_t(t,x)|\nabla\varphi_\nu^\varepsilon(t,x)|^2 dtdx.$$

Letting $\nu \to 0$, we get by (6.11) that (6.10) holds, as claimed. $\qquad\square$

Now, by (6.4) and (6.10), we have

$$\int_{Q_T} |(\varphi_\varepsilon)_t(t,x)|^2 dtdx = \int_0^T {}_{H^1}\langle\varrho_\delta^\varepsilon(t)\nabla\varphi_\varepsilon(t), \nabla(\varphi_\varepsilon)_t(t)\rangle_{H^{-1}} dx + \int_{Q_T} f(t,x)(\varphi_\varepsilon)_t(t,x) dtdx$$

$$\leq -\frac{1}{2}\int_{Q_T} (\varrho_\delta)_t(t,x)|\nabla\varphi_\varepsilon(t,x)|^2 dtdx + \int_{Q_T} f(t,x)(\varphi_\varepsilon)_t(t,x) dtdx.$$

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

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

By the definition of ϱ_δ we have

$$-(\varrho_\delta)_t \leq C_2\delta^{-1}\varrho_\delta, \tag{6.12}$$

and hence we obtain

$$\int_{Q_T} |(\varphi_\varepsilon)_t|^2 dtdx \leq \int_{Q_T} \varrho_\delta|\nabla\varphi_\varepsilon|^2 dtdx + \int_0^T |f(t)|_{L^2(\mathbb{R}^d)}^2 dt. \tag{6.13}$$

19

Then, by (6.4) we have

$$\frac{\partial}{\partial t} \varphi_\varepsilon^\lambda + \operatorname{div}((\varrho_\delta + \varepsilon)\nabla\varphi_\varepsilon^\lambda) = f\eta_\lambda + \eta_\lambda'\varphi_\varepsilon \text{ on } Q_T$$
$$\varphi_\varepsilon^\lambda(T, x) = 0, \ \forall x \in \mathbb{R}^d. \tag{6.14}$$

By (6.4)–(6.8) we have $\varphi_\varepsilon^\lambda \in W_2^{2,1}(Q_T)$ and

$$\varphi_\varepsilon^\lambda \in L^2((0,T); H^2(\mathbb{R}^d)) \cap C([0,T]; L^2), \ \frac{d}{dt}\varphi_\varepsilon^\lambda \in L^2((0,T); L^2), \ \varphi_\varepsilon^\lambda \in L^\infty(Q_T). \tag{6.15}$$

By Remark 6.7, v satisfies (6.3) for all $\varphi \in C_{0,b}^\infty([0,T) \times \mathbb{R}^d)$, and $\nu = v(0,x)dx$ is the zero measure. Hence, we infer by density that v satisfies (6.3) also for all functions $\varphi = \varphi_\varepsilon^\lambda$ with properties (6.15). Therefore, we have

$$\int_{Q_T} v\left(\frac{\partial\varphi_\varepsilon^\lambda}{\partial t} + \operatorname{div}(\varrho_\delta\nabla\varphi_\varepsilon^\lambda)\right) dt dx = 0, \ \forall\varepsilon,\lambda \in (0,1) \tag{6.16}$$

(see Appendix C for details).

Next, we get by (6.14) 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 (6.8), we get

$$|\varphi_\varepsilon^\lambda(t)|_{L^2(\mathbb{R}^d)}^2 + \varepsilon\int_t^T |\nabla\varphi_\varepsilon^\lambda(s)|_{L^2(\mathbb{R}^d)}^2 ds + \int_{Q_T}\varrho_\delta|\nabla\varphi_\varepsilon^\lambda|^2 dt dx \le C_3, \ \forall\varepsilon,\lambda \in (0,1). \tag{6.17}$$

Now, we have as in (6.10) that

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

and this yields

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

Then, by (6.9), (6.12) and (6.17) we get

$$\int_{Q_T}|(\varphi_\varepsilon^\lambda)_t|^2 dt dx \le \int_{Q_T}\varrho_\delta|\nabla\varphi_\varepsilon^\lambda|^2 dt dx + C_4\int_{Q_T}(|f|^2 + \lambda^{-1}|\varphi_\varepsilon|^2)dt dx \le \left(1 + \frac{1}{\lambda}\right)C_5, \tag{6.18}$$
$$\forall\varepsilon,\lambda \in (0,1).$$

(Here, we have denoted by C_i, $i = 0, 1, 2, ...$, several positive constants independent of ε, λ.) Now, we fix a sequence $\lambda_n \to 0$. Hence, by (6.12), (6.13), (6.17) and (6.18) and a diagonal argument, we find a subsequence $\varepsilon_k \to 0$ such that, for all $n \in \mathbb{N}$ as $k \to \infty$,

$$\begin{cases} \varphi_{\varepsilon_k} \to \varphi, \ \varphi_{\varepsilon_k}^{\lambda_n} \to \varphi^{\lambda_n} & \text{weakly in } L^2(Q_T) \\ (\varphi_{\varepsilon_k})_t \to (\varphi)_t, \ (\varphi_{\varepsilon_k}^{\lambda_n})_t \to (\varphi^{\lambda_n})_t & \text{weakly in } L^2(Q_T) \\ \operatorname{div}((\varrho_\delta + \varepsilon_k)\nabla\varphi_{\varepsilon_k}^{\lambda_n}) \to \zeta_{\lambda_n} & \text{weakly in } L^2(Q_T). \end{cases} \tag{6.19}$$

To finish the proof of Theorem 6.5, we need the following three claims. □

Claim 2. The maps $t \mapsto w_\delta(t)$ and $t \mapsto v_1(t)$ are weakly continuous from $[0,T]$ to $L^2(\mathbb{R}^d)$.

Proof of Claim 2. Since $\sup_{t\in[0,T]} \{|w_\delta(t)|_{L^1(\mathbb{R}^d)}, |w_\delta(t)|_{L^\infty(\mathbb{R}^d)}\} < \infty$, also $\sup_{t\in[0,T]} \{|w_\delta(t)|_{L^2(\mathbb{R}^d)}\} < \infty$.
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 $w_\delta(t_{n_k})$ has a weak limit in $L^2(\mathbb{R}^d)$. Due to the weak continuity (in the sense of measures) of $t \mapsto w_\delta(t,x)dx$, this limit is $w_\delta(t)$, which implies the $L^2(\mathbb{R}^d)$-weak continuity of $[0,T] \ni t \mapsto w_\delta(t)$. The same argument applies to v_1. □

Claim 3. *We have*

$$\lim_{n\to\infty} \lim_{k\to\infty} \int_{Q_T} v\eta'_{\lambda_n} \varphi_{\varepsilon_k} \, dtdx = 0. \tag{6.20}$$

Proof of Claim 3. We first note that by (6.19) for every $n \in \mathbb{N}$

$$\lim_{k\to\infty} \int_{Q_T} v\eta'_{\lambda_n} \varphi_{\varepsilon_k} \, dtdx = \int_{Q_T} v\eta'_{\lambda_n} \varphi \, dtdx,$$

and that $\varphi \in C([0,T]; L^2)$. Furthermore, by (6.9), $\varphi(T,\cdot) = 0$. Then, for every $\lambda \in (0,1)$,

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

where, as $\lambda \to 0$, both terms converge to zero by Claim 2 and Lebesgue's dominated convergence theorem. □

Claim 4. There is $\alpha \in (0,1)$ such that for every $\lambda \in (0,1)$ there exists $C_\lambda \in (0,\infty)$ such that

$$\int_{Q_T} \varepsilon |\Delta\varphi_\varepsilon^\lambda|^{1+\alpha} w_\delta \, dtdx \leq C_\lambda, \ \forall \varepsilon \in (0,1). \tag{6.21}$$

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

$$\varepsilon_k \Delta\varphi_{\varepsilon_k}^{\lambda_n} \to 0 \text{ weakly both in } L^1(Q_T; w_\delta \, dtdx) \text{ and } L^1(Q_T; \eta w_\delta \, dtdx). \tag{6.22}$$

Proof of Claim 4. By the de la Vallée Poussin theorem and a diagonal argument, (6.22) follows from (6.21), so we only have to prove (6.21). To this purpose, fix

$$s \in \left(\frac{2d(p-1)}{p}, d(p-1)\right). \tag{6.23}$$

We note that this interval 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 $\alpha \in \left(0, \frac{1}{3}\right)$ such that

$$0 < \frac{p-2}{p-1} \frac{(3\alpha+1)s}{s(1-\alpha) - 2(\alpha+1)} < d \tag{6.24}$$

and also that

$$\alpha < \frac{d(p-1)}{(d+2)(p-2)+d}. \tag{6.25}$$

Then, we multiply (6.14) by $sign\ \Delta\varphi^\lambda_\varepsilon |\Delta\varphi^\lambda_\varepsilon|^\alpha w_\delta$ and integrate over Q_T to obtain after rearranging

$$\int_{Q_T} (\varrho_\delta + \varepsilon)|\Delta\varphi^\lambda_\varepsilon|^{1+\alpha} w_\delta \, dtdx \leq \int_{Q_T} |\nabla\varrho_\delta|\, |\nabla\varphi^\lambda_\varepsilon|\, |\Delta\varphi^\lambda_\varepsilon|^\alpha \, w_\delta \, dtdx$$

$$+ \int_{Q_T} |(\varphi^\lambda_\varepsilon)_t|\, |\Delta\varphi^\lambda_\varepsilon|^\alpha \, w_\delta \, dtdx + \int_{Q_T} |f\eta_\lambda + \eta'_\lambda \varphi_\varepsilon|\, |\Delta\varphi^\lambda_\varepsilon|^\alpha \, w_\delta \, dtdx. \tag{6.26}$$

We point out that by (6.15) the right hand side of (6.26) is finite, hence so is its left hand side. The second term on the right hand side of (6.26) can be estimated by

$$|(\varphi^\lambda_\varepsilon)_t|^2_{L^2(Q_T)} + r\int_{Q_T} |\Delta\varphi^\lambda_\varepsilon|^{1+\alpha} \varrho_\delta \, w_\delta \, dtdx + C_r \int_{Q_T} w_\delta^{\frac{2}{1-\alpha}} \varrho_\delta^{-\frac{2\alpha}{1-\alpha}} \, dtdx. \tag{6.27}$$

Likewise, we can estimate the last term in (6.26) by

$$|f\eta_\lambda + \eta'_\lambda \varphi_\varepsilon|^2_{L^2(Q_T)} + r\int_{Q_T} |\Delta\varphi^\lambda_\varepsilon|^{1+\alpha} \varrho_\delta \, w_\delta \, dtdx + C_r \int_{Q_T} w_\delta^{\frac{2}{1-\alpha}} \varrho_\delta^{-\frac{2\alpha}{1-\alpha}} \, dtdx, \tag{6.28}$$

where $r \in (0,1)$ and C_r is a large enough constant independent of ε, λ. We note that, for $\alpha \in \left(0, \frac{1}{3}\right)$ satisfying (6.25), the definition of ω_δ and ϱ_δ implies that the last terms in (6.27) and (6.28) are finite. Furthermore, by (6.18) the respective first terms in (6.27), (6.28) are uniformly bounded in $\varepsilon \in (0,1)$. Finally, the first term on the right hand side of (6.26) can be estimated by

$$r\int_{Q_T} |\Delta\varphi^\lambda_\varepsilon|^{\alpha+1} \varrho_\delta \, w_\delta \, dtdx + C_r \int_{Q_T} \frac{|\nabla\varrho_\delta|^{\alpha+1}}{\varrho_\delta^\alpha} |\nabla\varphi^\lambda_\varepsilon|^{\alpha+1} \, w_\delta \, dx,$$

for $r \in (0,1)$ and C_r as above, where the second integral is up to a constant bounded by

$$\int_{Q_T} |\nabla\varrho_\delta|^{\frac{2(\alpha+1)}{1-\alpha}} \varrho_\delta^{-\frac{3\alpha+1}{1-\alpha}} w_\delta^{\frac{2}{1-\alpha}} \, dtdx + \int_{Q_T} |\nabla\varphi^\lambda_\varepsilon|^2 \varrho_\delta \, dtdx$$

of which the second integral is uniformly bounded in $\lambda, \varepsilon \in (0,1)$ by (6.17) and the first integral is up to a constant bounded by

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

where by (6.23) and (6.24) this quantity is finite. Choosing r small enough, we hence get from (6.26) and the nonnegative of ϱ_δ the estimate (6.21), and Claim 4 follows. $\qquad\square$

Now, we can finish the proof of Theorem 6.5. We have by Claim 3 and (6.14)

$$\int_{Q_T} fv \, dtdx \;=\; \lim_{n\to\infty} \lim_{k\to\infty} \int_{Q_T} (f\eta_{\lambda_n} + \eta'_{\lambda_n}\varphi_{\varepsilon_k})v \, dtdx$$

$$=\; \lim_{n\to\infty} \lim_{k\to\infty} \int_{Q_T} \left(\frac{d}{dt}\, \varphi^{\lambda_n}_{\varepsilon_k} + \mathrm{div}((\varrho_\delta + \varepsilon_k)\nabla\varphi^{\lambda_n}_{\varepsilon_k})\right) v \, dtdx,$$

which, taking into account that $v = (1-\eta)w_\delta$, by Claim 4 is equal to

$$\lim_{n\to\infty} \lim_{k\to\infty} \int_{Q_T} \left(\frac{\partial}{\partial t}\, \varphi^{\lambda_n}_{\varepsilon_k} + \mathrm{div}\left(\varrho_\delta \nabla\varphi^{\lambda_n}_{\varepsilon_k}\right)\right) v \, dtdx,$$

which in turn equals zero by (6.16). Hence, $\int_{Q_T} vf \, dtdx = 0$ and so, because $f \in C^\infty_0(Q_T)$ was arbitrary, the assertion follows. $\qquad\square$

A   Details to proof of Theorem 4.2

Here we give more details on some assertions made in the proof of Theorem 4.2. Below, we denote by $\lesssim$ an inequality where we supress a multiplicative constant on the right hand side which only depends on d, p and T. First, by definition of $k = (p - 2 + \frac{p}{d})^{-1}$, we have

$$-k(p-2)(1 + \frac{p}{d(p-1)}) > -1 \iff p - 2 + \frac{p(p-2)}{d(p-1)} < p - 2 + \frac{p}{d} \iff p - 2 < p - 1.$$

So, since $\frac{k}{d}\frac{p-2}{p-1} > 0$, also $-k(p-2)(1 + \frac{p}{d(p-1)}) + \frac{k}{d}\frac{p-2}{p-1} > -1$. Next, $\nabla|\nabla w_0(t)|^{p-2} \in L^1_{\text{loc}}([0, \infty); L^1_{\text{loc}}(\mathbb{R}^d))$ follows from (4.5). Indeed, $|x|^{-\frac{1}{p-1}} \in L^1_{\text{loc}}(\mathbb{R}^d)$ and both exponents of t appearing on the right hand side of (4.5) are greater than -1. For the first exponent, this was already shown above; for the second, we estimate $\mathbf{1}_{|x| \leq \beta t^{\frac{k}{d}}}|x|$ by $\beta t^{\frac{k}{d}}$ and have

$$- k(p-2)(1 + \frac{p}{d(p-1)}) - \frac{kp}{d(p-1)} + \frac{k}{d} > -1 \iff p - 2 + \frac{p(p-2)}{d(p-1)} + \frac{p}{d(p-1)} - \frac{1}{d} < p - 2 + \frac{p}{d}$$

$$\frac{p(p-2) + p}{p-1} < p + 1 \iff p < p + 1.$$

Finally, we give details regarding $\nabla(|\nabla w_0|^{p-2})w_0 \in L^1_{\text{loc}}([0, \infty); L^1(\mathbb{R}^d))$. Let $T > 0$. We have

$$\int_0^T \int_{\mathbb{R}^d} |\nabla(|\nabla w_0|^{p-2})|w_0 \, dxdt \tag{A.1}$$

$$\lesssim \int_0^T t^{-k(p-2)(1 + \frac{p}{d(p-1)})}\left(\int_{\{|x| \leq \beta t^{\frac{k}{d}}\}} |x|^{-\frac{1}{p-1}} w_0 dx + t^{-\frac{kp}{d(p-1)}}\int_{\{|x| \leq \beta t^{\frac{k}{d}}\}} |x| \, w_0 \, dx\right) dt$$

$$\lesssim \int_0^T t^{-k(p-2)(1 + \frac{p}{d(p-1)})}\left(t^{-k}\int_{\{|x| \leq \beta t^{\frac{k}{d}}\}} |x|^{-\frac{1}{p-1}} \, dx + t^{-\frac{kp}{d(p-1)} + \frac{k}{d}}\int_{\mathbb{R}^d} w_0 \, dx\right) dt,$$

where for the second inequality we used $w_0(t, x) \lesssim t^{-k}$. Note

$$\int_{\{|x| \leq \beta t^{\frac{k}{d}}\}} |x|^{-\frac{1}{p-1}} \, dx \leq C_t \int_0^{\beta t^{\frac{k}{d}}} r^{-\frac{1}{p-1} + d - 1} dr = C(T)t^{\frac{k(d(p-1)-1)}{d(p-1)}},$$

where C_t is a t-dependent constant related to the surface area of the d-dimensional ball with radius $\beta t^{\frac{k}{d}}$. Since d is fixed, $C(T) := \sup_{t \in (0,T)} C_t$ is finite. Since we find

$$- k(p-2)(1 + \frac{p}{d(p-1)}) + \frac{k(d(p-1)-p)}{d(p-1)} - k > -1 \iff p - 2 + \frac{p(p-2)}{d(p-1)} + \frac{1}{d(p-1)} < p - 2 + \frac{p}{d}$$

$$\frac{p(p-2) + 1}{p-1} < p \iff 1 - p < 0,$$

the first integral term on the right hand side of (A.1) is hence finite. Regarding the second integral, we observe, similarly as above,

$$-k(p-2)(1 + \frac{p}{d(p-1)}) - \frac{kp}{d(p-1)} + \frac{k}{d} > -1 \iff 1 - p < 0,$$

hence the finiteness of the second integral follows from $\int_{\mathbb{R}^d} w_0(t, x)dx = 1$ for all $t > 0$. Consequently, the right hand side of (A.1) is finite for every $T > 0$ and we are done.

B Proof of Remark 6.7

Clearly, $C_0^\infty([0,T] \times \mathbb{R}^d)$ is dense in $C_{0,b}^\infty([0,T] \times \mathbb{R}^d)$ (defined in Remark 6.7) with respect to uniform convergence of all partial derivatives (including zero derivatives) on $[0,T] \times \mathbb{R}^d$. Let $\varphi \in C_{0,b}^\infty([0,T] \times \mathbb{R}^d)$ and $(\varphi_k)_{k \in \mathbb{N}} \subseteq C_0^\infty([0,T] \times \mathbb{R}^d)$, such that $\varphi_k \to \varphi$ as $k \to \infty$ in the above sense. Then, in particular,

(i) $\varphi_k(0) \to \varphi(0)$ uniformly on $\mathbb{R}^d$.

(ii) $(\varphi_k)_t \to \varphi_t$, $\nabla \varphi_k \to \nabla \varphi$, $\Delta \varphi_k \to \Delta \varphi$ uniformly on $[0,T] \times \mathbb{R}^d$, where we abbreviate $\frac{d}{dt} \varphi$ by φ_t.

Since $v \in L^1(Q_T) \cap L^\infty(Q_T)$ and $v_0 \in L^1(\mathbb{R}^d)$, $\varrho_\delta \in L^\infty$, $\nabla \varrho_\delta \in L^2(Q_T)$, (i)–(ii) yield

$$0 = \lim_k \left(\int_{Q_T} v((\varphi_k)_t + \operatorname{div}(\varrho_\delta \nabla \varphi_k)) dt dx + \int_{\mathbb{R}^d} \varphi_k(0) v_0 dx \right) = \int_{Q_T} v(\varphi_t + \operatorname{div}(\varrho_\delta \nabla \varphi)) dt dx + \int_{\mathbb{R}^d} \varphi(0) v_0 dx,$$

which proves Remark 6.7.

C Details for (6.16)

We know that $\varphi_\varepsilon^\lambda \in W_2^{2,1}(Q_T)$. It is standard that $C_{0,b}^\infty([0,T] \times \mathbb{R}^d)$ (as introduced in Remark 6.7) 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 $(\varphi_k)_{k \in \mathbb{N}} \subseteq C_{0,b}^\infty([0,T] \times \mathbb{R}^d)$ such that $\varphi_k \xrightarrow{k} \varphi_\varepsilon^\lambda$ with respect to this norm. Since $v_0 \equiv 0$ and since $v \in L^2(Q_T)$, $\varrho_\delta \in L^\infty(Q_T)$, $\nabla \varrho_\delta \in L^2(Q_T)$, $v \in L^\infty(Q_T)$, 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 (6.16).

D Linear, linearized and nonlinear Fokker–Planck equations

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

$$\frac{d}{dt} \mu_t = \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} (a_{ij}(t,x)\mu_t) - \frac{\partial}{\partial x_i} (b_i(t,x)\mu_t), \ (t,x) \in (0,\infty) \times \mathbb{R}^d. \tag{D.1}$$

The equation is linear, since in contrast to (1.3) the coefficients do not depend on the solution μ_t. Usually, an initial condition $\mu_0 = \nu$ is imposed for solutions to (D.1), where $\nu \in \mathcal{M}_b^+$.

Definition D.1. A *distributional solution to* (D.1) *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

$$\int_0^T \int_{\mathbb{R}^d} |a_{ij}(t,x)| + |b_i(t,x)| d\mu_t(x) dt < \infty, \ \forall T > 0, \tag{D.2}$$

and

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

for all $\psi \in C_0^\infty(\mathbb{R}^d)$. A distributional 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 D.2. A *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 measurable on $(0, \infty) \times \mathbb{R}^d$,

$$\int_0^T \int_{\mathbb{R}^d} |a_{ij}(t, x, \mu_t)| + |b_i(t, x, \mu_t)| d\mu_t(x) dt < \infty, \ \forall T > 0, \tag{D.3}$$

and

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

for all $\psi \in C_0^\infty(\mathbb{R}^d)$. The notion of probability solution and the extension to initial times $s > 0$ is as in the linear case.

Linearized equations. For the nonlinear FPEs (1.3), a standard way to linearize them is as follows. Fix a solution $(\mu_t)_{t \geq 0}$ in the measure-coefficient of the nonlinear coefficients, i.e. consider the *linear* coefficients

$$\tilde{a}_{ij}(t, x) := a_{ij}(t, x, \mu_t), \quad \tilde{b}_i(t, x) := b_i(t, x, \mu_t),$$

and consider the linear FPE with coefficients $\tilde{a}_{ij}, \tilde{b}_i$. Clearly, $(\mu_t)_{t \geq 0}$ is a distributional solution to this linear FPE in the sense of Definition D.1. In general, there can be further solutions, which are not related to the fixed solution $(\mu_t)_{t \geq 0}$.

This way, we obtain the linearized p-Laplace equation (6.1) by fixing in the coefficients (1.13) the fundamental solution w^0 from (1.18) and Definition 6.1 is exactly Definition D.1 for the linear FPE (6.1).

References

[1] Adams, D.R., Hedberg, L.I., *Function Spaces and Potential Theory*, Springer-Verlag, 1996.

[2] Barbu, V., *Semigroup Approach to Nonlinear Diffusion Equations*, World Scientific, Singapore, 2021.

[3] Barbu, V., Röckner, M., Probabilistic representation for solutions to nonlinear Fokker–Planck equations, *SIAM J. Math. Anal.*, 50 (4) (2018), 4246-4260.

[4] Barbu, V., Röckner, M., From nonlinear Fokker–Planck equations to solutions of distribution dependent SDE, *Ann. Probab.*, 50 (4) (2020), 1902-1920.

[5] Barbu, V., Röckner, M., *Nonlinear Fokker–Planck Flows and their Probabilistic Counterparts*, Springer, 2024.

[6] Bliedtner, J., Hansen, W., *Potential Theory: An analytic and probabilistic approach to balayage*, Universitext., Springer-Verlag, Berlin, 1986.

[7] Blumenthal, R.M., Getoor, R.K., Markov Processes and Potential Theory, *Pure and Applied Mathematics*, vol. 29, Academic Press, New York-London, 1968.

[8] Bogachev, V.I., Krylov, N.V., Röckner, M., Shaposhnikov, S.V., *Fokker–Planck–Kolmogorov Equations*, Mathematical Surveys and Monographs, 207. American Mathematical Society, 2015.

[9] Bogachev, V.I., Röckner, M., Shaposhnikov, S.V., On uniqueness of solutions to the Cauchy problem for degenerate Fokker–Planck–Kolmogorov equations, *J. Evol. Eqn.*, 13 (2013), 577-593.

[10] Bogachev, V.I., Röckner, M., Shaposhnikov, S.V., On the Ambrosio–Figalli–Trevisan superposition principle for probability solutions to Fokker–Planck–Kolmogorov equations, *Journal of Dynamics and Differential Equations*, 33 (2) (2021), 715-739.

[11] Brezis, H., *Analyse fonctionnelle: Théorie et applications*, Masson, Paris, 1983.

[12] Carmona, R., Delarue, F., *Probabilistic theory of Mean Field Games with Applications*, I–II, Springer, Berlin, 2017.

[13] Di Benedetto, E., *Degenerate Parabolic Equations*, Springer, 1993.

[14] Di Benedetto, E., Herrero, M.A., On the Cauchy Problem and Initial Traces for a Degenerate Parabolic Equation, *Transactions of the American Mathematical Society*, Vol. 314, No. 1 (1989), 187-224.

[15] Doob, J.L., *Classical potential theory and its probabilistic counterpart*, Reprint of the 1984 edition Classics Math., Springer-Verlag, Berlin, 2001. xxvi+846 pp.

[16] Dynkin, E.B., *Markov Processes*, Vols. I, II, *Die Grundlehren der Mathematischen Wissenschaften*, vol. 122, vol. 121, Academic Press, Inc., Publishers, New York, 1965, Springer-Verlag, Berlin-Göttingen-Heidelberg, translated with the authorization and assistance of the author by J. Fabius, V. Greenberg, A. Maitra, G. Majone.

[17] Einstein, A., *Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen.* In: *Annalen der Physik*, Band 322, Nr. 8, 1905, S. 549–560.

[18] Ethier, S.N., Kurtz, T.G., *Markov Processes: Characterization and convergence*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1986.

[19] Figalli, A., Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients, *J. Funct. Anal.*, 254 (1) (2008), 109-153.

[20] Freidlin, M., *Markov Processes and Differential Equations: Asymptotic Problems*, Lectures in Mathematics ETH Zürich., Birkhäuser Verlag, Basel, 1996.

[21] Fukushima, M., Oshima, Y., Takeda, M., *Dirichlet Forms and Symmetric Markov Processes*, extended edition, De Gruyter Studies in Mathematics, vol. 19, Walter de Gruyter & Co., Berlin, 2011.

[22] Grube, S., *Strong solutions to McKean-–Vlasov stochastic differential equations with coefficients of Nemytskii-type*, PhD thesis. Bielefeld University, 2023, https://pub.uni-bielefeld.de/record/2979897.

[23] Heinonen, J., Kilpeläinen, T., Martio, O., *Nonlinear Potential Theory of Degenerate Elliptic Equations*, Clarendon Press (Oxford Univ. Press), 1993.

[24] Kamin, S., Vázquez, J.L., Fundamental solutions and asymptotic behaviour for the p-Laplacian equation, *Revista Matemática Iberoamericana*, 4 (2) (1988), 339-354.

[25] Karatzas, I., Shreve, S.E., *Brownian motion and stochastic calculus*, vol. 113 of Graduate Texts in Mathematics. Springer-Verlag, New York, second edition, 1991.

[26] Kolokoltsov, V.N., *Markov Processes, Semigroups and Generators*, De Gruyter Studies in Mathematics, vol. 38, Walter de Gruyter & Co., Berlin, 2011.

[27] Kolokoltsov, V.N., *Nonlinear Markov Processes and Kinetic Equations*, Cambridge University Press, 2010. Online publication date July 2014. https://doi.org/10.1017/CBO9780511760303.

[28] Kurtz, T.G., Equivalence of Stochastic Equations and Martingale Problems, In: Crisan, D. (eds.), *Stochastic Analysis*, Springer, Berlin, Heidelberg, 2010. https://doi.org/10.1007/978-3-642-15358-7_6

[29] Ladyzenskaya, O.A., New equations for the description of the motions of viscous incompressible fluids, and global solvability for their boundary value problems, *Proc. Steklov Inst. Math.*, 102 (1967) 95–118; English transl. in: *Tr. Mat. Inst. Steklova*, 102 (1967), 85–104.

[30] Ladyzhenskaya, O.A., Sollonikov, V.A., Uralceva, N.N., *Linear and Quasilinear Equations of Parabolic Type*, AMS Translations, 1988.

[31] Liggett, T.M., *Continuous Time Markov Processes: An introduction*, Graduate Studies in Mathematics, vol. 113, American Mathematical Society, Providence, RI, 2010.

[32] Lindqvist, P., *Notes on the p-Laplace Equation*, Rep. Univ. Jyväskylä Dep. Math. Stat., 102, University of Jyväskylä, 2006, ii+90 pp.

[33] Lindqvist, P., *Notes on the Stationary p-Laplace Equation*, Springer, 2019.

[34] Lions, J.L., *Quelques Méthodes de Résolution des Problèmes aux Limites Non Linéaires*, Dunod–Gauthier Villars, Paris, 1969.

[35] McKean, Jr., H.P., Propagation of chaos for a class of non-linear parabolic equations. In: *Stochastic Differential Equations*, Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967, pp. 41–57. Air Force Office Sci. Res., Arlington, Va.

[36] McKean, Jr., H.P., A class of Markov processes associated with nonlinear parabolic equations, *Proceedings of the National Academy of Sciences of the United States of America*, 56 (6) (1966), 1907-1911.

[37] Mingione, G., Recent progress in nonlinear potential theory. (English summary), European Congress of Mathematics, 501–524. European Mathematical Society (EMS), Zürich, 2018.

[38] Pelissier, M.C., Sur quelques problèmes non linéaires en glaciologie, Thèse, *Publ. Mat.*, d'Orsay, 110 (1975).

[39] Peres, Y., Schramm, O., Sheffield, S., Wilson, D.B., Tug-of-war and the infinity Laplacian, *J. Amer. Math. Soc.*, 22 (1) (2009), 167–210.

[40] Peres, Y., Sheffield, S., Tug-of-war with noise: a game theoretic view of the p-Laplacian. *Duke Math. J.*, 145 (2008), 91–120.

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

[42] Rehmeier, M., Röckner, M., On nonlinear Markov processes in the sense of McKean, *arXiv preprint* 2212.12424, 2022.

[43] Rogers, L.C.G., Williams, D., *Diffusions, Markov Processes, and Martingales*, Cambridge Mathematical Library, vol. 2, *Itô calculus*, Cambridge University Press, Cambridge, 2000, reprint of the second (1994) edition.

[44] Sharpe, M., *General Theory of Markov Processes*, Pure and Applied Mathematics, vol. 133, Academic Press, Inc., Boston, MA, 1988.

[45] Stroock, D.W., *An Introduction to Markov Processes*, second edition, Graduate Texts in Mathematics, vol. 230, Springer, Heidelberg, 2014.

[46] Stroock D.W., Srinivasa Varadhan, S.R., *Multidimensional diffusion processes*, Springer, Berlin. New York, 1997.

[47] Sznitman, A.-S., *Topics in propagation of chaos*, in: *Ecole d'Et'e de Probabilité de Saint-Flour XIX-1989*, in: *Lecture Notes in Math.*, vol. 1464, Springer, Berlin, 1991, pp. 165–251.

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

[49] Vázquez, J.L., *The Porous Medium Equation: Mathematical Theory*, Oxford Mathematical Monographs, Clarendon, 2007.

[50] von Smoluchowski, M., *Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen.* In: *Annalen der Physik*, Band 326, Nr. 14, 1906, S. 756–780, doi:10.1002/andp.19063261405 (freier Volltext).

[51] Vlasov, A.A., The vibrational properties of an electron gas, *Phys. Usp.*, 10 (6) (1968), 721–733.

[52] Wiener, N., Differential-Space, *J. Math. Phys.* 2 (1923), 131–174.