STRONG SOLUTIONS TO STOCHASTIC DIFFERENTIAL EQUATIONS WITH ROUGH COEFFICIENTS

BY NICOLAS CHAMPAGNAT[*,†] AND PIERRE-EMMANUEL JABIN[‡,1]

Université de Lorraine*, INRIA Nancy-Grand Est† and University of Maryland‡

We study strong existence and pathwise uniqueness for stochastic differential equations in $\mathbb{R}^d$ with rough coefficients, and without assuming uniform ellipticity for the diffusion matrix. Our approach relies on direct quantitative estimates on solutions to the SDE, assuming Sobolev bounds on the drift and diffusion coefficients, and L^p bounds for the solution of the corresponding Fokker–Planck PDE, which can be proved separately. This allows a great flexibility regarding the method employed to obtain these last bounds. Hence we are able to obtain general criteria in various cases, including the uniformly elliptic case in any dimension, the one-dimensional case and the Langevin (kinetic) case.

1. Introduction. We investigate the well-posedness of the stochastic differential equation (SDE) in $\mathbb{R}^d$, $d \geq 1$,

$$(1.1) \qquad dX_t = F(t, X_t)\, dt + \sigma(t, X_t)\, dW_t, \qquad X_0 = \xi,$$

where $F : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d \times r}$ are Borel measurable function, $(W_t, t \geq 0)$ is a r-dimensional standard Brownian motion on some given complete filtered probability space $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$, and ξ is a $\mathcal{F}_0$-measurable random variable.

When σ and F are bounded, the law $u(t, dx)$ of X_t belongs to the set M_1 of functions from $\mathbb{R}_+$ with value in the set P_1 of probability measures on $\mathbb{R}^d$ such that, for all Borel subset Γ of $\mathbb{R}^d$, $t \mapsto u(t, \Gamma)$ is measurable. It is standard to deduce from Itô's formula that $u(t, dx)$ is a (weak, measure) solution to the Fokker–Planck PDE on $\mathbb{R}_+ \times \mathbb{R}^d$

$$(1.2) \quad \partial_t u + \nabla_x \cdot (Fu) = \nabla_x^2 : (au) = \sum_{1 \leq i, j \leq d} \frac{\partial^2 (a_{ij} u)}{\partial x_i\, \partial x_j}, \qquad u(t = 0, dx) = u^0,$$

where $a = \frac{1}{2} \sigma \sigma^*$ and u^0 is the law of the initial r.v. ξ.

We first recall some classical terminology: weak existence holds for (1.1) if one can construct a filtered probability space $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$, an adapted Brownian motion W and an adapted process X on this space solution to (1.1). Uniqueness in law holds if every solution X to (1.1), possibly on different probability space, has the same law. Strong existence means that one can find a solution to (1.1) on any given filtered probability space equipped with any given adapted Brownian motion. Finally, pathwise uniqueness means that, on any given filtered probability space equipped with any given Brownian motion, any two solutions to (1.1) with the same given $\mathcal{F}_0$-measurable initial condition ξ coincide. Our goal is to study *strong existence and pathwise uniqueness* for rough σ and F, through quantitative estimates on the difference between solutions and a priori bounds on the solutions to (1.2).

This question has been the object of many works aiming to improve the original result of Itô [15]. Veretennikov [32] proved strong existence and pathwise uniqueness for bounded measurable drifts F and certain nondegenerate diffusion matrices like $a = \text{Id}$. The case of unbounded drifts was then studied by Krylov and Röckner [21] under the assumption that $F \in L_{t,\text{loc}}^q(L_x^p)$ with $d/p + 2/q < 1$, and the case where $\sigma(t, x)$ is uniformly continuous with respect to x and $\nabla\sigma$ also belongs to $L_{t,\text{loc}}^q(L_x^p)$ by Zhang [36, 38]. All these works assume that the matrix a is uniformly elliptic, that is, that $a(x) - c\text{Id}$ is positive definite for all x for some constant $c > 0$. The time-independent one-dimensional case was also deeply studied by Engelbert and Schmidt [8–10] (see also [26, 34, 35]).

The main tools used in all the previous works are Krylov's inequality [17] and its extensions (see, e.g., [7, 21, 22, 38]), Zvonkin's transformation [40] to remove the drift, and a priori estimates on solutions of the backward Kolmogorov equation or Fokker–Planck PDE (1.2) [20, 21, 33, 36]. Of great importance is also the result of Yamada and Watanabe [34, 35], which proves that strong existence holds as soon as pathwise uniqueness and weak existence hold for all initial condition. Since general conditions for weak existence are well known (see [7, 13, 18, 22, 25, 29, 31]), one only has to prove pathwise uniqueness to obtain strong existence. In dimension one, a key tool to prove pathwise uniqueness is the local time.

Most of these works also use estimates on functionals of the difference between two solutions of (possibly regularizations of) (1.1). Recently, a particular form of functional, inspired by the method of Crippa and De Lellis [5] to obtain an alternative proof of the results of Di Perna and Lions [6] on well-posedness for ODEs, has been used in [28, 37, 39]. The functional of [5] was originally used and adapted to obtain several extensions [4, 16] of the result of DiPerna and Lions for deterministic systems.

Other approaches also exist, like the one of Le Bris and Lions in [23–25], based on well-posedness and stability properties for the backward Kolmogorov equation and the Kolmogorov equation obtained from a doubling of variable technique. This approach gives different criteria for strong existence and pathwise uniqueness, involving boundedness conditions on $\text{div}\,\sigma$ and $(D\sigma)^2$. Several works also studied

the existence of a stochastic flow for the SDE [11, 12, 14, 27] (e.g., [11, 12, 27] prove regularity properties of the flow in the settings of [21, 32, 36], resp.), or a (weaker) almost-everywhere stochastic flow [24, 37].

In this work, we make use of simple quantitative estimates on functionals of solutions to (1.1) following the estimates for the deterministic ODEs in Crippa and DeLellis [5], which already inspired the works in [28, 37, 39] for the uniformly elliptic cases. A main difference with our current work is in the way we use the Crippa–DeLellis ideas. We show how those estimates can be used to prove directly strong existence and pathwise uniqueness (without making use of Yamada–Watanabe classical results). This lets the method be used in more general settings, such as cases where pathwise uniqueness can be proved only for particular initial conditions (e.g., with bounded density) or with hypoelliptic diffusions, where weak existence has been less studied (note that the two probabilistic notions of strong and weak existence coincide when $\sigma = 0$).

The originality of the method we develop here is that it allows to decouple the various questions involved in well-posedness. We do not prove any bounds on the solution u of the PDE (1.2). Instead assuming that such bounds have already been obtained through other means, we show how to use the Crippa–DeLellis estimates to directly prove strong existence and pathwise uniqueness for (1.1) without using any other probabilistic ideas or methods.

This leads to very explicit and constructive proofs, based on quantitative stability estimates. The first advantage is to offer a simple and unifying framework to formulate our assumptions. We are able to identify explicit norms, whose definition depends on u; s.t. one has well-posedness if the drift F and diffusion σ are bounded in those norms. This does not require any ellipticity assumptions on σ and is compatible with degenerate diffusion or even with the deterministic theory for $\sigma = 0$.

As a good illustration of this unified framework, we want to emphasize also the results that we obtain in dimension 1. By using suitable and new, modified estimates, we are able to obtain the currently optimal $1d$ result (with $\sigma \in H^{1/2}$ for instance) through essentially the same procedure.

The second main advantage is the flexibility that one then enjoys as it is possible to choose the best method to deal with (1.2) according to any additional structure. If σ satisfies an ellipticity condition, we can then recover now classical results. But depending on the precise structure of equation (1.2), one can have much better results. A very good example is second-order equations with degenerate diffusion, as shown in Corollary 1.4.

The final advantage of the method is its simplicity as it relies on some direct and self-contained quantitative estimates on the solutions.

To give a better idea, let us present a typical result that we obtain. For existence, we consider sequence of approximations to (1.1)

$$(1.3) \qquad dX_t^n = F_n(t, X_t^n)\, dt + \sigma_n(t, X_t^n)\, dW_t, \qquad X_0^n = \xi,$$

with the same Brownian motion W_t for any n. And we introduce the corresponding approximation for (1.2)

(1.4)
$$\partial_t u_n + \nabla_x \cdot (F_n u_n)$$
$$= \sum_{1 \le i, j \le d} \frac{\partial^2}{\partial x_i \, \partial x_j} \big(a_{ij}^n(t, x) u_n(t, x) \big), \qquad u_n(t = 0, dx) = u^0,$$

with $a^n = \frac{1}{2} \sigma_n \sigma_n^*$ and $u_n \in M_1$.

The next result is not the most general we obtain, but it does not require any additional definition and illustrates the type of assumptions we need. We use the classical notation for L^p and Sobolev spaces with different exponents for space and time. For example, $L^q_{t,\text{loc}}(W^{1,p}_x)$ for $1 \le p, q \le \infty$ is the set of measurable functions f of the variables $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$, such that, for almost all $t \ge 0$, $f(t, \cdot) \in W^{1,p}(\mathbb{R}^d)$ and $t \mapsto \|f(t, \cdot)\|_{W^{1,p}(\mathbb{R}^d)} \in L^q([0, T])$ for all $T > 0$. We also call weak topology on the set M_1 of measurable functions of time with values in the set P_1 of probability measures on $\mathbb{R}^d$, the topology of weak-* convergence in time for the tight topology of probability measures on $\mathbb{R}^d$. In other words, $u_n \to u$ for the weak topology of M_1 iff $\langle u_n, f \rangle \to \langle u, f \rangle$ for all bounded continuous function f on $\mathbb{R}_+ \times \mathbb{R}^d$ with support included in $[0, T] \times \mathbb{R}^d$ for some $T > 0$.

THEOREM 1.1. *Assume $d \ge 2$. One has*:

(i) *Existence: Assume that there exists a sequence of smooth $F_n, \sigma_n \in L^\infty$ converging in the sense of distributions to F and σ, respectively, such that the solution $u_n \in M_1$ to (1.4) satisfies for $1 \le p, q \le \infty$, with $1/p' + 1/p = 1$, $1/q + 1/q' = 1$*

$$\sigma_n - \sigma \to 0 \qquad \text{in } L^q_{t,\text{loc}}(L^p_x) \quad \text{and} \quad F_n - F \to 0 \qquad \text{in } L^q_{t,\text{loc}}(L^p_x),$$

$$\sup_n \big(\|\sigma_n\|_{L^{2q}_{t,\text{loc}}(W^{1,2p}_x)} + \|F_n\|_{L^q_{t,\text{loc}}(W^{1,p}_x)} + \|F_n\|_{L^\infty} + \|\sigma_n\|_{L^\infty} \big) < \infty,$$

$$\sup_n \|u_n\|_{L^{q'}_{t,\text{loc}}(L^{p'}_x)} < \infty, \qquad u_n \to u \text{ in the weak topology of } M_1.$$

Then there exists a strong solution X_t to (1.1) and $(X^n_t - \xi, t \in [0, T])_n$ converges in $L^p(\Omega, L^\infty([0, T]))$ for all $p > 1$ and $T > 0$ to $(X_t - \xi, t \in [0, T])$, with X^n_t the solutions to (1.3). In addition, $u(t, dx)$ is the law of X_t for almost all $t \ge 0$.

(ii) *Uniqueness: Let X and Y be two solutions to (1.1) with one-dimensional time marginals $u_X(t, x)\, dx$ and $u_Y(t, x)\, dx$ both in $L^{q'}_{t,\text{loc}}(L^{p'}_x)$. Assume that $F, \sigma \in L^\infty$, $X_0 = Y_0$ a.s. and that*

$$\|F\|_{L^q_{t,\text{loc}}(W^{1,p}_x)} + \|\sigma\|_{L^{2q}_{t,\text{loc}}(W^{1,2p}_x)} < \infty$$

with $1/p + 1/p' = 1$ and $1/q + 1/q' = 1$. Then one has pathwise uniqueness: $\sup_{t \ge 0} |X_t - Y_t| = 0$ a.s.

We obtain better results in the one-dimensional case.

THEOREM 1.2.   *Assume $d = 1$.*

(i) *The existence result of Theorem* 1.1(i) *holds under the same assumptions on F_n, σ_n, u_n, except that the assumption $\sup_n \|\sigma_n\|_{L^{2q}_{t,loc}(W^{1,2p}_x)} < \infty$ can be replaced by*

$$\sup_n \|\sigma_n\|_{L^{2q}_{t,\mathrm{loc}}(W^{1/2,2p}_x)} < \infty$$

and in the case $p = 1$, the assumption $\sup_n \|F_n\|_{L^q_{t,loc}(W^{1,p}_x)} < \infty$ must be replaced by

$$\sup_n \|F_n\|_{L^q_{t,\mathrm{loc}}(W^{1,1+\varepsilon}_x)} < \infty$$

for some $\varepsilon > 0$.

(ii) *The uniqueness result of Theorem* 1.1(ii) *holds true under the same assumptions on F, σ, u_X, u_Y, except that $\|\sigma\|_{L^{2q}_{t,loc}(W^{1,2p}_x)} < \infty$ can be replaced by*

$$\|\sigma\|_{L^{2q}_{t,\mathrm{loc}}(W^{1/2,2p}_x)} < \infty$$

and in the case $p = 1$, the assumption $\|F\|_{L^q_{t,loc}(W^{1,p}_x)} < \infty$ must be replaced by

$$\|F\|_{L^q_{t,\mathrm{loc}}(W^{1,1+\varepsilon}_x)} < \infty$$

for some $\varepsilon > 0$.

Note that no assumption of uniform ellipticity is needed in Theorems 1.1 and 1.2, provided one can prove a priori estimates on the various solutions u_n, u_X, u_Y to (1.4) and (1.2). Note also that pathwise uniqueness is proved only for particular solutions to (1.1), so we cannot use directly the result of Yamada and Watanabe to deduce strong existence. Hence our method proves separately strong existence and pathwise uniqueness.

Of course, as they are laws, u_n, u_X and u_Y all have bounded mass so Theorems 1.1 and 1.2 really depend on whether it is possible to obtain higher integrability for a solution of (1.2). From Theorem 1.1, we may for instance simply deduce the following.

COROLLARY 1.3.   *Assume that $d \geq 2$, $u^0 \in L^1 \cap L^\infty$, $F, \sigma \in L^\infty$, $F \in L^1_{t,\mathrm{loc}}(W^{1,1}_x)$ and $\nabla\sigma \in L^q_{t,\mathrm{loc}}(L^p_x)$, where $2/q + d/p = 1$ with $p > d$. Assume as well that σ is uniformly elliptic. Then one has existence of a strong solution to (1.1) with marginal distributions $u(t, dx)$ in $L^\infty_{t,\mathrm{loc}}(L^\infty_x)$. In addition, pathwise uniqueness holds among all solutions with marginal distributions in $L^\infty_{t,\mathrm{loc}}(L^\infty_x)$.*

Hence we recover similar (actually, slightly better) assumptions on σ as in [36], but different assumptions on F [neither stronger nor weaker since the sets $L^1_{t,\text{loc}}(W^{1,1}_x)$ and $L^p_{t,\text{loc}}(L^q_x)$ are not included in one another]. This difference comes from the fact that our general method does not assume *a priori* that σ is uniformly elliptic. Hence the minimal conditions on F we can expect are those under which existence and uniqueness hold in the case $\sigma = 0$, that is, those of DiPerna and Lions [6]: $F \in L^1_{t,\text{loc}}(W^{1,1}_x)$ and some bounds on div F. (Actually, $F \in L^1_{t,\text{loc}}(BV_x)$ was shown to be sufficient by Ambrosio [2], but we do not consider this case.)

However, in many physical cases, uniform ellipticity is not necessary. For instance, in the phase space problem,

$$dX_t = V_t \, dt, \qquad dV_t = F(t, X_t) \, dt + \sigma(t, X_t) \, dW_t, \qquad (X_0, V_0) = \xi,$$

one obtains an even better result.

COROLLARY 1.4. *Assume that $\sigma \in L^\infty \cap L^2_{t,\text{loc}}(H^1_x)$ and $F \in L^1_{t,\text{loc}}(W^{1,1}_x)$. Assume also that the law $u_0 \in L^\infty$. Then one has both existence of a strong solution to* (2.25) *and pathwise uniqueness among all solutions with marginal distributions in $L^\infty_{t,\text{loc}}(L^\infty_x)$.*

The goal of Section 2 is to give the statement of all our results. We start in Section 2.1 by defining the norms and Banach spaces needed to state our most general results in Section 2.2. Theorems 1.1 and 1.2 are then obtained as corollaries of these general results. In Section 2.3, several corollaries of Theorems 1.1 and 1.2 are stated in various situations, including the uniformly elliptic case (Corollary 1.3), the nondegenerate one-dimensional case and the kinetic (Langevin) case (Corollary 1.4). The rest of the paper is devoted to the proofs of all the results stated in Section 2, and the organization of the rest of the paper is given in the end of Section 2.

2. Statement of the results. As usual, one needs regularity assumptions on F and σ to ensure strong existence and pathwise uniqueness for (1.1). In our case, these are Sobolev norms with respect to some $u \in M_1$, defined in Section 2.1. Our general results are then stated in Section 2.2, and several consequences of these results are discussed in Section 2.3.

2.1. *Norms and Banach spaces.* The conditions we shall impose on F and σ can be roughly described as follows. We need σ to be L^2 in time and H^1 in space (in dimension $d \geq 2$) or $H^{1/2}$ in space (in dimension $d = 1$) w.r.t. the measure u solution to (1.2), and F to be L^1 in time and $W^{1,1}$ in space w.r.t. the measure u. Weighted Sobolev spaces have been extensively used and studied, but the key difference here is that no regularity is known on the weight u. It could very well be a

sum of Dirac masses. This is why one must be careful and why maximal functions are required.

The goal of the next subsections is to give the precise definitions and basic properties of our spaces.

2.1.1. *The space $H_T^1(u)$.* Fix first $v \in P_1$. We start with the following definition.

DEFINITION 2.1. The space $H^1(v)$ is defined as the subspace of functions $f \in BV_{\text{loc}}(\mathbb{R}^d)$, the space of functions on $\mathbb{R}^d$ with locally bounded variations, such that

$$(2.1) \qquad \|f\|^2_{H^1(v)} := \int_{\mathbb{R}^d} \left((M|f|(x))^2 + (M|\nabla_x f|(x))^2 \right) v(dx) < \infty,$$

where M is the usual maximal operator.

First of all, observe that the definition makes perfect sense. If $f \in BV_{\text{loc}}(\mathbb{R}^d)$, then $|\nabla f|$ is a locally finite measure. This allows to define $M|\nabla f|$ per

$$M|\nabla f|(x) = \sup_{r>0} \frac{1}{|B(0,r)|} \int_{B(0,r)} |\nabla f|(x+dz) \qquad \forall x \in \mathbb{R}^d.$$

In that case, it is well known (see [30]) that $M|\nabla f|$ is a Borel function with value in $\mathbb{R}_+ \cup \{+\infty\}$. It locally belongs in fact to the weak L^1 space, that is, for any $R > 0$, there exists C_R s.t.

$$\left| \{ x \in B(0, R), M|\nabla f|(t, x) > L \} \right| \leq \frac{C_R}{L}.$$

Therefore, the integral of $(M|\nabla f|)^2$ against the Borelian measure v is well defined with value in $\mathbb{R}_+ \cup \{+\infty\}$, thus justifying the definition.

The main point of the definition is that we have a well-behaved space independently of any regularity on v.

THEOREM 2.2. *Assume that v belongs to P_1. Then $H^1(v)$ is a Banach space with norm* (2.1). *Moreover, the norm is lower semicontinuous with respect to convergence in the sense of distribution: If $f_n \to f$ in the sense of distribution, then*

$$(2.2) \qquad \|f\|_{H^1(v)} \leq \liminf_n \|f_n\|_{H^1(v)}.$$

And if for a given $f \in BV_{\text{loc}}(\mathbb{R}^d)$, v_n converges to v in the tight topology of probability measures then

$$\|f\|_{H^1(v)} \leq \liminf_n \|f\|_{H^1(v_n)}.$$

This result is proved in Section 3.

There are several technical reasons why we use $M|\nabla f|$ in the definition of the norm. Note however that the intuitive definition with just ∇f would most certainly be too weak as v could for instance vanish just at the points where ∇f is very large. In particular, without the maximal function in the definition of the norm (2.1), it would be very easy to find counterexamples to (2.2).

Now, given any $u \in M_1$, we give a second definition.

DEFINITION 2.3.   For all $T > 0$, the space $H_T^1(u)$ is defined as the subspace of the set of measurable functions on $[0, T] \times \mathbb{R}^d$ such that, for almost all $t \in [0, T]$, $f(t, \cdot) \in H^1(u(t, \cdot))$ and

$$(2.3) \qquad \|f\|_{H_T^1(u)}^2 = \int_0^T \|f(t, \cdot)\|_{H^1(u(t,\cdot))}^2 \, dt < \infty.$$

In particular, if $u(t, \cdot)$ is the distribution of X_t solution to (1.1), then, for all $T > 0$ and $\sigma \in H_T^1(u)$,

$$(2.4) \qquad \|\sigma\|_{H_T^1(u)}^2 = \mathbb{E}\left(\int_0^T M|\sigma|^2(t, X_t) \, dt\right) + \mathbb{E}\left(\int_0^T \left(M|\nabla\sigma|(t, X_t)\right)^2 dt\right).$$

We then have the following immediate consequence of Theorem 2.2.

COROLLARY 2.4.   *Fix $T > 0$. Assume u belongs to M_1. Then $H_T^1(u)$ is a Banach space with norm* (2.3). *Moreover, the norm is lower semicontinuous with respect to convergence in the sense of distribution*: *If $f_n \to f$ in the sense of distribution, then*

$$(2.5) \qquad \|f\|_{H_T^1(u)} \leq \liminf_n \|f_n\|_{H_T^1(u)}.$$

And if for a given f measurable on $\mathbb{R}_+ \times \mathbb{R}^d$ with $f(t, \cdot) \in BV_{\text{loc}}(\mathbb{R}^d)$ for almost all $t \geq 0$, u_n converges to u for the weak topology in M_1, then

$$\|f\|_{H_T^1(u)} \leq \liminf_n \|f\|_{H_T^1(u_n)}.$$

2.1.2. *The space $H_T^{1/2}(u)$.*   In the one-dimensional case, we can prove strong existence and pathwise uniqueness using $H^{1/2}$ type of assumptions on σ. The definitions and properties of the spaces $H_T^{1/2}(u)$ follow exactly the same steps as before. We first fix $v \in P_1$.

DEFINITION 2.5.   For any function $f \in L_{\text{loc}}^1(\mathbb{R}^d)$, one defines

$$\partial_x^{1/2} f = \mathcal{F}^{-1} |\xi|^{1/2} \mathcal{F} f,$$

with $\mathcal{F}$ the Fourier transform in $\mathbb{R}^d$. The space $H^{1/2}(v)$ is defined as the subspace of functions $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ s.t. $\partial_x^{1/2} f$ is a locally finite Radon measure and

$$\|f\|^2_{H^{1/2}(v)} = \int_{\mathbb{R}^d} ((M|f|(x))^2 + (M|\partial_x^{1/2} f|(x))^2) v(dx) < \infty.$$

As for $H^1(v)$, the maximal function can be extended to measures by

$$M|\partial_x^{1/2} f|(x) = \sup_{r>0} \frac{1}{|B(0,r)|} \int_{B(0,r)} |\partial_x^{1/2} f|(x + dz) \qquad \forall x \in \mathbb{R}^d.$$

One has again that $M|\partial_x^{1/2} f|$ is a Borel function with value in $\mathbb{R}_+ \cup \{+\infty\}$ belonging to the local weak L^1 space. The integral against the Borelian measure v is hence well defined in $\mathbb{R}_+ \cup \{+\infty\}$, independently of the regularity of v.

The next result is proved in Section 3.

THEOREM 2.6. *Assume that v belongs to P_1. Then $H^{1/2}(v)$ is a Banach space with norm* (2.8). *Moreover, the norm is lower semicontinuous with respect to convergence in the sense of distribution*: *If $f_n \to f$ in the sense of distribution, then*

$$(2.6) \qquad \|f\|_{H^{1/2}(v)} \le \liminf_n \|f_n\|_{H^{1/2}(v)}.$$

And if, for a given $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ s.t. $\partial_x^{1/2} f$ is a locally finite Radon measure, v_n converges to v in the tight topology of probability measures on $\mathbb{R}^d$, then

$$\|f\|_{H^{1/2}(v)} \le \liminf_n \|f\|_{H^{1/2}(v_n)}.$$

Given any $u \in M_1$, we give a second definition.

DEFINITION 2.7. For all $T > 0$, the space $H_T^{1/2}(u)$ is defined as the subspace of the set of measurable functions on $[0, T] \times \mathbb{R}^d$ such that, for almost all $t \in [0, T]$, $f(t, \cdot) \in H^{1/2}(u(t, \cdot))$ and

$$(2.7) \qquad \|f\|^2_{H_T^{1/2}(u)} = \int_0^T \|f(t, \cdot)\|^2_{H^{1/2}(u(t,\cdot))} dt < \infty.$$

In particular, if $u(t, \cdot)$ is the distribution of X_t solution to (1.1), then, for all $T > 0$ and $\sigma \in H_T^{1/2}(u)$,

$$(2.8) \quad \|\sigma\|^2_{H_T^{1/2}(u)} = \mathbb{E}\left(\int_0^T M|\sigma|^2(t, X_t) dt\right) + \mathbb{E}\left(\int_0^T (M|\partial_x^{1/2} \sigma|(t, X_t))^2 dt\right).$$

Again, one has the following immediate consequence of Theorem 2.6.

COROLLARY 2.8. *Fix $T > 0$. Assume u belongs to M_1. Then $H_T^{1/2}(u)$ is a Banach space with norm* (2.7). *Moreover, the norm is lower semicontinuous with respect to convergence in the sense of distribution: If $f_n \to f$ in the sense of distribution, then*

$$(2.9) \qquad \|f\|_{H_T^{1/2}(u)} \leq \liminf_n \|f_n\|_{H_T^{1/2}(u)}.$$

And if for a given $f \in L^1(\mathbb{R}_+ \times \mathbb{R}^d)$ s.t. $\partial_x^{1/2} f(t, \cdot)$ is a locally finite Radon measure for almost all $t \in [0, T]$, u_n converges to u for the weak topology in M_1, then

$$\|f\|_{H_T^{1/2}(u)} \leq \liminf_n \|f\|_{H_T^{1/2}(u_n)}.$$

Let us emphasize here that, as it may already be clear from the definition and as it will be seen in the proof, the space $H_T^{1/2}(u)$ is much more intricate than the previous space $H_T^1(u)$. Using this space is key to our improved 1-dimensional result. However, it does require the development of new techniques to make the estimates compatible with this weaker norm.

2.1.3. *The space $W_T^{\phi,\text{weak}}(u)$.* We also need some similar $W^{1,1}$ assumptions on F. Following the definition of $H^1(u)$, a first attempt would be

$$(2.10) \qquad \|F\|_{W_T^{1,1}(u)} = \int_0^T \int_{\mathbb{R}^d} \big(M|F|(t, x) + M|\nabla F|(t, x)\big) u(t, dx)\, dt.$$

Unfortunately, while this definition would work, it is too strong in some cases. This is due to the fact that the maximal operator M is bounded on L^p, $p > 1$, but not on L^1. In particular, if $u \in L^\infty$ then the norm defined in (2.4) would automatically be finite if σ is in the usual H^1 space but the norm defined in (2.10) would *not* be finite if $F \in W^{1,1}$ in general.

Therefore, in order to obtain better assumptions we have to work with a more complicated space. We proceed as before and fix $v \in P_1$. We also introduce a super-linear function ϕ, that is, a function ϕ on $[1, \infty)$ such that $\phi(\xi)/\xi$ is nondecreasing and converges to ∞ as $\xi \to \infty$.

DEFINITION 2.9. For any locally finite Radon measure μ, decomposing μ into a part absolutely continuous with respect to the Lebesgue measure μ_a and the singular part μ_s, one defines

$$M_L \mu = \sqrt{\log L} + \int_{\mathbb{R}^d} \frac{|\mu_a|(z) \mathbb{1}_{|\mu_a(z)| \geq \sqrt{\log L}}\, dz + |\mu_s|(dz)}{(L^{-1} + |x - z|)|x - z|^{d-1}}.$$

For any function $f \in BV_{\text{loc}}(\mathbb{R}^d)$, the decomposition of ∇f, into a part absolutely continuous with respect to the Lebesgue measure $\nabla_a f$ and the singular part $\nabla_s f$, makes $M_L \nabla f$ well defined.

The space $W^{\phi,\text{weak}}(v)$ is hence defined as the subspace of functions $f \in BV_{\text{loc}}(\mathbb{R}^d)$ s.t.

$$\|f\|_{W^{\phi,\text{weak}}(v)} = \int_{\mathbb{R}^d} M|f|(x)v(dx) + \sup_{L \geq 1} \frac{\phi(L)}{L \log L} \int_{\mathbb{R}^d} M_L \nabla f \, v(dx) < \infty.$$

In this definition, the maximal function is regularized so that $M_L \nabla f$ is locally integrable for any fixed L. The supremum is then taken outside.

Obviously, the space heavily depends on the choice of ϕ. Note that $M_L \nabla f \geq \sqrt{\log L}$ so that

$$\|f\|_{W^{\phi,\text{weak}}(v)} \geq \sup_{L \geq 1} \frac{\phi(L)}{L \sqrt{\log L}}.$$

In particular, $\|f\|_{W^{\phi,\text{weak}}(v)} = +\infty$ for all f if $\phi(L) \gg L\sqrt{\log L}$ asymptotically as $L \to +\infty$. On the other hand, we want to choose ϕ superlinear as we need to control the integrability of $|\nabla f|$. This leads to the assumptions:

$$(2.11) \qquad \frac{\phi(L)}{L} \to +\infty, \qquad \frac{\phi(L)}{L\sqrt{\log L}} \to 0 \qquad \text{as } L \to +\infty.$$

Even with this assumption, $W^{\phi,\text{weak}}(v)$ is not a Banach space and in particular $\|\cdot\|_{W^{\phi,\text{weak}}(v)}$ is not a norm. Of course, $\|0\|_{W^{\phi,\text{weak}}(v)} \neq 0$ but this could easily be remedied by considering $\|\cdot\|_{W^{\phi,\text{weak}}(v)} - \alpha_\phi$ instead, for the right constant α_ϕ.

The main problem is that $\|\lambda f\|_{W^{\phi,\text{weak}}(v)} \neq |\lambda| \|f\|_{W^{\phi,\text{weak}}(v)}$ and this cannot easily be corrected. It is in fact the same kind of issue that one has with the definition of so-called Orlicz spaces such as $L \log L$. The solution is similar and would consist in constructing the right norm by duality.

We did not feel that it was appropriate in this article however. Such a construction in the present case would be considerably more complex than for classical Orlicz space. It would also distract from our main goal while bringing very little to our results. It is worth recalling the main reason why we introduce the space $W^{\phi,\text{weak}}$. It is a compromise between two requirements:

- The estimates that we perform later in the text would not work for instance with the simple requirement that

$$\int (|f| + |\nabla f|)v(dx) < \infty,$$

  so the maximal operator is needed.
- We want to recover the classical assumption if v is bounded from below and above. That means that if $1/C \leq v \leq C$, then any $f \in W^{1,1}$ must be included in $W^{\phi,\text{weak}}(v)$ for some well chosen ϕ (depending on f). This is in particular why we do not use the direct extension $W^{1,1}(v)$ of the space $H^1(v)$, given by (2.10).

The above definition of $W^{\phi,\text{weak}}(v)$ fulfills those two goals and, therefore, we study further this space.

THEOREM 2.10.   *Assume that v belongs to P^1, that ϕ is superlinear and continuous and that* (2.11) *holds. Then $W^{\phi,\text{weak}}(v)$ is well defined and $\|\cdot\|_{W^{\phi,\text{weak}}(v)}$ is lower semicontinuous with respect to convergence in the sense of distribution: If $f_n \to f$ in the sense of distribution then*

$$(2.12) \qquad \|f\|_{W^{\phi,\text{weak}}(v)} \leq \liminf_n \|f_n\|_{W^{\phi,\text{weak}}(v)}.$$

And if for a given $f \in BV_{\text{loc}}(\mathbb{R}^d)$, v_n converges to v in the tight topology of probability measures then

$$\|f\|_{W^{\phi,\text{weak}}(v)} \leq \liminf_n \|f\|_{W^{\phi,\text{weak}}(v_n)}.$$

Moreover if $v \geq 1/C$ over a smooth open set Ω and $f \in W^{\phi,\text{weak}}(v)$ then $f \in W^{1,1}(K)$ for any compact set $K \subset \Omega$. Reciprocally, if $v \leq C$ over Ω and $f \in W^{1,1}(\Omega)$ with compact support in Ω, then there exists a superlinear ϕ satisfying (2.11) *s.t. $f \in W^{\phi,\text{weak}}(v)$.*

Now, given $u \in M_1$ and a superlinear function ϕ, we define the following.

DEFINITION 2.11.   For all $T > 0$, the space $W_T^{\phi,\text{weak}}(u)$ is defined as the set of measurable f on $[0, T] \times \mathbb{R}^d$ such that $f(t, \cdot) \in W_T^{\phi,\text{weak}}(u(t, \cdot))$ for almost all $t \in [0, T]$ and

$$\|f\|_{W_T^{\phi,\text{weak}}(u)} = \int_0^T \|f(t, \cdot)\|_{W^{\phi,\text{weak}}(u(t,\cdot))} \, dt < \infty.$$

In particular, if $u(t, \cdot)$ is the distribution of X_t solution to (1.1), then, for all $T > 0$ and $F \in W_T^{\phi,\text{weak}}(u)$,

$$(2.13) \quad \|F\|^2_{W_T^{\phi,\text{weak}}(u)} = \sup_{L \geq 1} \frac{\phi(L)}{L \log L} \mathbb{E}\left(\int_0^T \left(M|F|(t, X_t) + M_L|\nabla F|(t, X_t) \right) dt \right).$$

Then we have the following.

COROLLARY 2.12.   *Fix $T > 0$, assume u belongs to M_1 and that ϕ is superlinear, continuous and satisfies* (2.11). *Then $W_T^{\phi,\text{weak}}(u)$ is well defined and $\|\cdot\|_{W_T^{\phi,\text{weak}}(u)}$ is lower semicontinuous with respect to convergence in the sense of distribution: If $f_n \to f$ in the sense of distribution, then*

$$(2.14) \qquad \|f\|_{W_T^{\phi,\text{weak}}(u)} \leq \liminf_n \|f_n\|_{W_T^{\phi,\text{weak}}(u)}.$$

And if for a given f measurable on $\mathbb{R}_+ \times \mathbb{R}^d$ with $f(t, \cdot) \in BV_{\text{loc}}(\mathbb{R}^d)$ for almost all $t \in [0, T]$, u_n converges to u for the weak topology in M_1, then

$$\|f\|_{W_T^{\phi,\text{weak}}(u)} \leq \liminf_n \|f\|_{W_T^{\phi,\text{weak}}(u_n)}.$$

Moreover, if $u \geq 1/C$ over $[0, T] \times \Omega$ where $\Omega \subset \mathbb{R}^d$ is a smooth open set and $f \in W_T^{\phi, \text{weak}}(u)$ then $f \in L_t^1([0, T], W^{1,1}(K))$ for any compact set $K \subset \Omega$. Reciprocally, if $u \leq C$ over $[0, T] \times \Omega$ and $f \in L_t^1([0, T], W^{1,1}(\Omega))$ with compact support in $[0, T] \times \Omega$, then there exists a superlinear ϕ satisfying (2.11) s.t. $f \in W_T^{\phi, \text{weak}}(u)$.

The first two points of Corollary 2.12 are direct consequences of Theorem 2.10, and the last statements about the cases where u is bounded from above or below can be proved exactly as the similar statement of Theorem 2.10 is proved in Section 3.

Theorem 2.10 and Corollary 2.12 support the introduction of the seminorm and the space $W_T^{\phi, \text{weak}}$. We point out in particular the conclusion of both results (in the time independent and time dependent case) that if the law u is bounded, then any $W^{1,1}$ function belongs to some $W_T^{\phi, \text{weak}}$.

This will allow us to obtain the critical $W^{1,1}$ assumption for the drift. However, it leads to difficulties in the proof. Typically, Crippa–DeLellis estimates naturally work if the drift satisfies an estimate like

$$(2.15) \qquad \bigl|F(x) - F(y)\bigr| \leq C\bigl(h(x) + h(y)\bigr)|x - y|,$$

with h in $L^1(u)$. Such an estimate is essentially equivalent to an L^1 control on the maximal function as in (2.10) and, therefore, not compatible with F only in $W^{1,1}$.

In the purely deterministic case where $\sigma = 0$, well-posedness can be obtained for $F \in W^{1,1}$ by interpolation but this seems to be more complicated if some stochasticity is involved.

Because we work in a weaker space, a direct pointwise bound like (2.15) is not available to us (see Lemma 3.2 in Section 3) and this forces us to work the estimates in a different manner.

Let us finally note that the pointwise bound (2.15) is very close to the assumption found in [28] for uniqueness, which reads

$$(2.16) \qquad \bigl(F(x) - F(y)\bigr) \cdot (x - y) \leq C\bigl(h(x) + h(y)\bigr)|x - y|^2,$$

again for h in $L^1(u)$. Just as for (2.15), (2.16) is in general not satisfied for $F \in W^{1,1}$ even if u is bounded. But obviously, and contrary to our space $W_T^{\phi, \text{weak}}$, (2.16) is only a one-sided bound.

Just as for ODEs, only a one-sided bound is needed for pathwise uniqueness. However, to obtain strong existence as well without any ellipticity assumption, the other one-sided bound is required as well. This is quite similar to the usual well-posedness conditions on ODEs, which is rather natural: If no ellipticity assumptions are made, then the proofs have to be compatible with the deterministic case.

2.2. *General results on strong solutions to* (1.1). In the multidimensional case, our most general result is the following one, proved in Section 4.

THEOREM 2.13. *Assume that $d \geq 2$. One has*:

(i) *Existence*: *Fix $T > 0$ and assume that there exists a sequence of smooth $F_n, \sigma_n \in L^\infty$ converging in the sense of distribution to F and σ respectively, such that the solution $u_n \in M_1$ to* (1.4) *satisfies for some superlinear ϕ*

$$(2.17) \qquad \int_0^T \int_{\mathbb{R}^d} (|\sigma_n - \sigma| + |F_n - F|) \, du_n \, dt \to 0,$$

$$(2.18) \qquad \sup_n (\|F\|_{W_T^{\phi,\text{weak}}(u_n)} + \|\sigma\|_{H_T^1(u_n)} + \|F_n\|_{L^\infty} + \|\sigma_n\|_{L^\infty}) < \infty,$$

$$(2.19) \qquad u_n \to u \qquad \text{for the weak topology of } M_1.$$

Then there exists a strong solution X_t to (1.1) *s.t. $(X_t^n - \xi, t \in [0, T])_n$ converges in $L^p(\Omega, L^\infty([0, T]))$ for all $p > 1$ to $(X_t - \xi, t \in [0, T])$, with X_t^n the solutions to* (1.3). *In addition, $u(t, dx)$ is the law of X_t for almost all $t \in [0, T]$.*

(ii) *Uniqueness*: *Let X and Y be two solutions to* (1.1) *with one-dimensional time marginals $u_X(t, \cdot)$ and $u_Y(t, \cdot)$ on $[0, T]$. Assume that $F, \sigma \in L^\infty$, $X_0 = Y_0$ a.s. and that*

$$(2.20) \qquad \|F\|_{W_T^{\phi,\text{weak}}(u_X)} + \|F\|_{W_T^{\phi,\text{weak}}(u_Y)} + \|\sigma\|_{H_T^1(u_X)} + \|\sigma\|_{H_T^1(u_Y)} < \infty$$

for some superlinear function ϕ. Then one has pathwise uniqueness on $[0, T]$, that is, $\sup_{t \in [0, T]} |X_t - Y_t| = 0$ a.s.

Note that we do not require any ellipticity on σ for this result. In that sense, we cannot hope to have any smoothing effect from the Wiener process and the assumption on F must be enough to provide well-posedness in the purely deterministic setting ($\sigma = 0$). In this case, taking any $u_0 \in L^\infty$, our result gives that there exists a unique solution of $\dot{X}_t = F(t, X_t)$ with $X_0 = \xi$ and with law $u \in L^\infty$ provided that there exists a sequence of regularized F_n s.t. $u_n \to u$ for the weak-$*$ topology with $u \in L^\infty$ and a superlinear ϕ s.t.

$$\sup_{L \geq 1} \frac{\phi(L)}{L \log L} \|F + M_L \nabla F\|_{L^1([0,T] \times \mathbb{R}^d)} < \infty.$$

The first point is, for example, implied by the assumption div $F \in L^\infty$ and the second one can be proved to hold if $F \in L_{t,\text{loc}}^1(W_x^{1,1})$ as in the proof of Corollary 1.1 in the Appendix. Hence, we recover the classical results of DiPerna and Lions [6] but not the optimal BV assumption from Ambrosio [2].

REMARK 2.14. In the last result, the law of the diffusion u needs not be absolutely continuous with respect to Lebesgue's measure, so the value of the coefficients F and σ on sets of Lebesgue's measure zero may have some importance.

Our assumptions in fact ensures that all points where u is concentrated are automatically Lebesgue points for F and σ. For this reason, it is straightforward to choose the right representative for the almost everywhere defined functions F and σ (see Remark 3.3 in Section 3.1).

In dimension 1, the result is even better: we recover the $H^{1/2}$ type of assumption from [8–10, 26, 34, 35], but we lose a little bit on F [we have to use (2.10) instead of (2.13)].

THEOREM 2.15.   *Assume that $d = 1$. One has*:

(i) *Existence*: *Fix $T > 0$ and assume that there exists a sequence of smooth $F_n, \sigma_n \in L^\infty$ converging in the sense of distribution to F and σ, respectively, such that the solution u_n to* (1.4) *satisfies*

$$\int_0^T \int_{\mathbb{R}} \left(|\sigma_n - \sigma| + |F_n - F| \right) du_n \, dt \to 0,$$

$$\sup_n \left(\|F\|_{W_T^{1,1}(u_n)} + \|\sigma\|_{H_T^{1/2}(u_n)} + \|F_n\|_{L^\infty} + \|\sigma_n\|_{L^\infty} \right) < \infty,$$

$$u_n \to u \qquad \text{for the weak topology of } M_1.$$

Then there exists a strong solution X_t to (1.1) *s.t.* $(X_t^n - \xi, t \in [0, T])_n$ *converges in $L^p(\Omega, L^\infty([0, T]))$ for all $p > 1$ to $(X_t - \xi, t \in [0, T])$, with X_t^n the solutions to* (1.3). *In addition, $u(t, dx)$ is the law of X_t for almost all $t \in [0, T]$.*

(ii) *Uniqueness*: *Let X and Y be two solutions to* (1.1) *with one-dimensional time marginals $u_X(t, \cdot)$ and $u_Y(t, \cdot)$ on $[0, T]$. Assume that $F, \sigma \in L^\infty$, $X_0 = Y_0$ a.s. and that*

$$\|F\|_{W_T^{1,1}(u_X)} + \|F\|_{W_T^{1,1}(u_Y)} + \|\sigma\|_{H_T^{1/2}(u_X)} + \|\sigma\|_{H_T^{1/2}(u_Y)} < \infty.$$

Then pathwise uniqueness holds on $[0, T]$, that is, $\sup_{t \in [0, T]} |X_t - Y_t| = 0$ a.s.

Of course, while precise, the norms given by (2.4)–(2.13) or (2.8)–(2.10) are not so simple to use. However, it is quite easy to deduce more intuitive results with the more usual $W^{1,p}$ norms. We recall that M is continuous onto every L^p space for $1 < p \leq \infty$, and hence the norms $\| \cdot \|_{H_T^1(u)}$ and $\| \cdot \|_{W_T^{1,1}(u)}$ are controlled by appropriate Sobolev norms if some L^q estimate is available on the law u.

One complication occurs when $u_X \in L^\infty$ and one wants to obtain the close to optimal $W^{1,1}$ assumption on F (instead of $W^{1,p}$ for some $p > 1$) as the maximal function is not bounded onto L^1. This is the reason why we defined (2.13), which can be used following [16] (we recall the main steps in the Appendix).

Therefore, Theorems 1.1 and 1.2 are simple corollaries of Theorems 2.13 and 2.15, respectively, except for the previous complication for Theorem 1.1.

In order to apply Theorems 1.1 and 1.2, we need to consider cases where it is possible to obtain better integrability than L^1 bounds for a solution to (1.2). This

occurs in various situations, some of which will be studied in the next subsection. One difficulty to apply Theorems 1.1(ii) and 1.2(ii) is to obtain pathwise uniqueness without restriction on the set of solutions considered. This will of course be ensured if uniqueness in law is known for (1.1). More precisely, if the conclusion of Theorem 1.1(i) or Theorem 1.2(i) holds, then either $\|F\|_{W_T^{\phi,\text{weak}}(u)} + \|\sigma\|_{H_T^1(u)} < \infty$ (if $d \geq 2$) or $\|F\|_{W_T^{1,1}(u)} + \|\sigma\|_{H_T^{1/2}(u)} < \infty$ (if $d = 1$) by Corollaries 2.4, 2.8 and 2.12. Since there is uniqueness in law for (1.1), then $u_X = u_Y = u$ for all solutions X and Y to (1.1) as in Theorem 1.1(ii) or Theorem 1.2(ii), and hence pathwise uniqueness holds. This argument will be used repeatedly in the next subsection. Note however that condition (2.20) may impose restrictions on the initial distribution. This issue will be studied in Proposition 2.24.

2.3. *Consequences.* Let us first consider the case where σ is uniformly elliptic: for all t, x,

$$(2.21) \qquad \frac{1}{2}\sigma\sigma^*(t, x) = a(t, x) \geq cI$$

for some $c > 0$. For example, if $F = 0$ and σ does not depend on time, then there exists a corresponding stationary measure $\bar{u} > 0$ in $L^{d/(d-1)}$ as per Aleksandrov [1]. In that case, when $u_0 \leq C\bar{u}$, then the unique solution u of (1.2) in $L^2_{t,\text{loc}}(H^1_x)$ satisfies $u(t, dx) \leq C\bar{u}(x)\,dx$ for all $t \geq 0$ by the maximum principle.

COROLLARY 2.16. *Assume that $F = 0$ and $\sigma(x)$ satisfies (2.21) and belongs to $L^\infty \cap W_x^{1,2d}$ (or $L^\infty \cap H^{1/2}$ if $d = 1$). Assume also that $u_0 \leq C\bar{u}$ for some constant $C > 0$. Then one has both existence of a strong solution to (1.1) and pathwise uniqueness.*

Note that pathwise uniqueness holds without additional assumption since $\sigma \in W^{1,2d}$ implies that σ is continuous, and uniqueness in law holds in this case since σ is bounded and uniformly elliptic [31], Theorem 7.2.1.

Those results were later extended by Krylov in the parabolic, time dependent case [17, 19]. We may, for example, use the following version found in [38].

THEOREM 2.17. *Assume that F and σ are bounded and σ satisfies (2.21). Then, for all solution X of (1.1) with any initial distribution, for all $T > 0$ and $p, q > 1$ such that*

$$\frac{d}{p} + \frac{2}{q} < 2,$$

there exists a constant C such that for all $f \in L^q_t(L^p_x)$

$$\mathbb{E}\left[\int_0^T f(t, X_t)\,dt\right] \leq C\|f\|_{L^q_t(L^p_x)}.$$

This result means that

$$u \in L_t^{q'}(L_x^{p'}),$$

where $1/p + 1/p' = 1$ and $1/q + 1/q' = 1$, and we obtain the following corollary.

COROLLARY 2.18.   (i) *Assume that $d \geq 2$, $F, \sigma \in L^\infty$, σ satisfies* (2.21), $F \in L_{t,\text{loc}}^{q/2}(W_x^{1,p/2})$ *and $\sigma \in L_{t,\text{loc}}^q(W_x^{1,p})$ with $2/q + d/p < 1$. Then one has both existence of a strong solution to* (1.1) *and pathwise uniqueness for any initial condition ξ.*

(ii) *Assume that $d = 1$, $F, \sigma \in L^\infty$, σ satisfies* (2.21), $\sigma \in L_{t,\text{loc}}^q(W_x^{1/2,p})$ *with $2/q + 1/p < 1$ and $F \in L_{t,\text{loc}}^{q/2}(W_x^{1,p/2})$ if $p > 2$, $F \in L_{t,\text{loc}}^{q/2}(W^{1,1+\varepsilon})$ for some $\varepsilon > 0$ if $p \leq 2$. Then one has both existence of a strong solution to* (1.1) *and pathwise uniqueness for any initial condition ξ.*

Note that in this case, pathwise uniqueness holds without additional assumption since Krylov's inequality implies that $u \in L_t^{q'}(L_x^{p'})$ for *all* solutions to (1.1).

In our setting, since we need additional regularity on σ, it is easy to obtain better a priori estimates for u than those given by Krylov's inequality. For instance, we have the following.

PROPOSITION 2.19.   *For any $d \geq 1$, assume $u^0 \in L^1 \cap L^\infty$, $F, \sigma \in L^\infty$, σ satisfies* (2.21) *and $\nabla\sigma \in L_{t,\text{loc}}^q(L_x^p)$ satisfying $2/q + d/p = 1$ with $p > d$. Then any u solution to* (1.2), *limit for the weak topology in M_1 of smooth solutions, belongs to $L_t^\infty(L_x^r)$ for any $1 \leq r \leq \infty$.*

This proposition is based on classical energy estimates, and hence we just give a very short proof of it in Section 6. Combined with Theorem 1.1 this gives slightly better conditions for σ and much better conditions for F, assuming additional conditions on the initial distribution. We obtain Corollary 1.3, restated here

COROLLARY 2.20.   *Assume that $d \geq 2$, $u^0 \in L^1 \cap L^\infty$, $F, \sigma \in L^\infty$, $F \in L_{t,\text{loc}}^1(W_x^{1,1})$ and $\nabla\sigma \in L_{t,\text{loc}}^q(L_x^p)$, where $2/q + d/p = 1$ with $p > d$. Assume as well that σ satisfies* (2.21). *Then one has existence of a strong solution to* (1.1) *with marginal distributions $u(t, dx)$ in $L_{t,\text{loc}}^\infty(L_x^\infty)$. In addition, pathwise uniqueness holds among all solutions with marginal distributions in $L_{t,\text{loc}}^\infty(L_x^\infty)$.*

As above, the pathwise uniqueness property could be improved if we could prove uniqueness in law. If $d = 2$, uniqueness in law holds when σ and F are bounded and σ is uniformly elliptic [18]. When $d \geq 3$, by Sobolev embedding, the assumption $\nabla\sigma \in L_{t,\text{loc}}^q(L_x^p)$ implies that $x \mapsto \sigma(t, x)$ is continuous for almost all $t \geq 0$. This condition is not exactly sufficient to use the result of Stroock and Varadhan [31], Theorem 7.2.1, which assumes that $\sup_{t \in [0,T]} |\sigma(t, x) - \sigma(t, y)| \to 0$ when $y \to x$. This is true, for example, if $\nabla\sigma \in L_{t,\text{loc}}^\infty(L_x^p)$ for $p > d$. Hence we obtain the following.

COROLLARY 2.21. *Assume that $d \geq 2$, $u^0 \in L^1 \cap L^\infty$, $F, \sigma \in L^\infty$, $F \in L^1_{t,\mathrm{loc}}(W^{1,1}_x)$ and $\nabla \sigma \in L^q_{t,loc}(L^p_x)$ where $2/q + d/p = 1$ with $p > d$. Assume as well that σ satisfies (2.21), and if $d \geq 3$ that for all x,*

$$\sup_{t \in [0,T]} |\sigma(t,x) - \sigma(t,y)| \to 0 \qquad when \ y \to x.$$

Then one has both existence of a strong solution to (1.1) *and pathwise uniqueness.*

This result can be compared with previous works dealing with the uniformly elliptic case. The best result in this case seem to be the one of [38], where strong existence and pathwise uniqueness are proved under the assumptions $\nabla \sigma \in L^q_{t,\mathrm{loc}}(L^p_x)$, $\sigma(t,x)$ uniformly continuous with respect to x and $F \in L^q_{t,\mathrm{loc}}(L^p_x)$ with $d/p + 2/q < 1$, so we obtain a slightly better condition on σ (we can handle the limit case $d/p + 2/q = 1$ and no uniform continuity is needed for strong existence), and a condition on F which is neither stronger nor weaker, since $L^1_{t,\mathrm{loc}}(W^{1,1}_x)$ neither contains nor is contained in $L^q_{t,\mathrm{loc}}(L^p_x)$ with $d/p + 2/q < 1$.

In dimension 1 in the stationary case, even if (2.21) is not satisfied but instead only

$$(2.22) \qquad \frac{1}{2}\sigma^2(x) = a(x) > 0,$$

then one has the a priori bound

$$u(t,x) \leq \frac{C}{a(x)} e^{\int_0^x \frac{F(y)}{a(y)} dy} \qquad \forall x \in \mathbb{R},$$

for solutions to (1.2) again provided that u^0 satisfies the same bound. Therefore, we obtain the following.

COROLLARY 2.22. *Assume $d = 1$, $\sigma, F \in L^\infty$, σ satisfies (2.22), $F/a \in L^1$,*

$$u_0(x) \leq \frac{C}{a(x)} e^{\int_0^x \frac{F(y)}{a(y)} dy} \qquad \forall x \in \mathbb{R}$$

and

$$(2.23) \qquad \int_{\mathbb{R}} \frac{(M|\partial^{1/2}\sigma|(x))^2}{a(x)} dx < \infty \quad and \quad \int_{\mathbb{R}} \frac{M|\nabla F|}{a(x)} dx < \infty.$$

Then one has both existence of a strong solution to (1.1) *and pathwise uniqueness.*

Note that the assumptions (2.23) imply that $a^{-1} \in L^1_{\mathrm{loc}}$, which is a necessary and sufficient condition for uniqueness in law when F is bounded [8–10].

We will prove in Lemma 3.5 of Section 3 that for all x, y,

$$(2.24) \qquad |\sigma(x) - \sigma(y)| \leq (M|\partial_x^{1/2}\sigma|(x) + M|\partial_x^{1/2}\sigma|(y))|x - y|^{1/2}.$$

This inequality allows us to compare our result with similar results of the literature [8–10, 26, 34, 35, 40]. The best conditions in the time homogeneous case seem to be those of [8–10], Theorem 4.41, where pathwise uniqueness is proved to hold if $F/a \in L^1_{\mathrm{loc}}$, $|\sigma(x) - \sigma(y)|^2 \leq f(x)h(|y - x|)$ for all x, y with $f/a \in L^1_{\mathrm{loc}}$ and $\int_{0+} h^{-1}(u)\,du = +\infty$. Our result gives worse conditions on F, and our condition on σ is slightly worse, since we need to take $h(u) = u$ in (2.24). However, we improve the conditions on σ of all the other references.

We point out that, in higher dimension as well, ellipticity is not always required for bounds on the law. We give the classical example of SDEs in the phase space $\mathbb{R}^{2d}$

$$(2.25) \quad dX_t = V_t\,dt, \qquad dV_t = F(t, X_t)\,dt + \sigma(t, X_t)\,dW_t, \qquad (X_0, V_0) = \xi.$$

The joint law $u(t, x, v)$ of the process $(X_t, V_t)_{t \geq 0}$ solves the kinetic equation

$$\partial_t u(t, x, v) + v \cdot \nabla_x u(t, x, v) + F(t, x) \cdot \nabla_v u(t, x, v)$$

$$(2.26) \qquad = \sum_{1 \leq i, j \leq d} a_{ij}(t, x) \frac{\partial^2 u(t, x, v)}{\partial v_i\,\partial v_j}.$$

Equation (2.26) is in fact better behaved than (1.2) for rough coefficients as its symplectic structure for instance guarantees that it satisfies a maximum principle for all measure-valued solutions that are limit of smooth solutions. In particular, for any initial data $u^0 \in L^\infty(\mathbb{R}^{2d})$, there exists a measure-valued solution $u \in L^\infty(\mathbb{R}_+ \times \mathbb{R}^{2d})$ for σ and F as in the next result. This is true even though the diffusion in (2.25) is degenerate (there is no diffusion in the x direction, and σ can also be degenerate).

Hence in this situation, one may deduce as claimed Corollary 1.4 or the following.

COROLLARY 2.23. *Assume that $\sigma \in L^\infty \cap L^2_{t,\mathrm{loc}}(H^1_x)$ and $F \in L^1_{t,\mathrm{loc}}(W^{1,1}_x)$. Assume also that $u_0 \in L^\infty$. Then one has both existence of a strong solution to* (2.25) *and pathwise uniqueness among all solutions with marginal distributions in $L^\infty_{t,\mathrm{loc}}(L^\infty_x)$.*

Note that other hypoelliptic situations or even subelliptic situations may lead to a better integrability of the solution u to (1.2) than L^1. Several examples are given in [25], Section 4.5, each of which imply a corollary of our result in various situations were σ is degenerate.

To conclude, let us observe that most of the previous results give strong existence for nondeterministic initial distributions. However, one can use the next result to obtain strong existence and pathwise uniqueness for almost all deterministic initial conditions.

PROPOSITION 2.24.   *Under the assumptions of either Corollary* 2.16, *Corollary* 2.21 *or Corollary* 2.22, *for any complete filtered probability space* $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ *equipped with a r-dimensional standard Brownian motion* W, *there is strong existence and pathwise uniqueness for* (1.1) *on* $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}, W)$ *for almost all deterministic initial condition* $\xi = x \in \mathbb{R}^d$.

The proofs of the previous results are organized as follows. We start in Section 3 with some simple technical proofs, including those of Theorems 2.2, 2.6 and 2.10, Section 4 is then devoted to the proof of Theorem 2.13, Section 5 to the proof of Theorem 2.15, Section 6 to the proof of Proposition 2.19, and Section 7 to the proof of Proposition 2.24. The proof of Theorem 1.1 is given in the Appendix.

3. Useful technical results.   The results and proofs presented in this section are mostly easy extensions of well-known techniques, which we need in the following sections, and hence include here for the sake of completeness.

3.1. *Pointwise difference estimates.*   We often need to estimate the difference of the coefficients σ and F at two different points x and y during the proofs. We collect here all the results which allow us to do so and that we later use. In all those estimates, time is only a parameter and we accordingly omit the time variable in most formulas.

We start by recalling the classical inequality (it is, for instance, a direct consequence of [30], Theorem VII.1, and of basic properties of the Poisson Kernel).

LEMMA 3.1.   *Fix* $t \geq 0$ *and assume that* $\sigma(t, \cdot) \in BV(\mathbb{R}^d)$. *Then for any* $x, y \in \mathbb{R}^d$

$$(3.1) \qquad \left| \sigma(t, x) - \sigma(t, y) \right| \leq C_d \big(M |\nabla_x \sigma|(t, x) + M |\nabla_x \sigma|(t, y) \big) |x - y|.$$

This next lemma provides an extension of (3.1) with the operator M_L used in the definition (2.13).

LEMMA 3.2.   *Fix* $t \geq 0$ *and assume that* $F(t, \cdot) \in BV(\mathbb{R}^d)$. *For any* $x \in \mathbb{R}^d$, *if* $h(t, x) < \infty$ *with* $h(t, x) = |F(t, x)| + M_L \nabla F(t, x)$, *then* x *is a Lebesgue point of* F. *Then for any* $x, y \in \mathbb{R}^d$

$$(3.2) \qquad \left| F(t, x) - F(t, y) \right| \leq C_d \big(h(t, x) + h(t, y) \big) \left(|x - y| + \frac{1}{L} \right),$$

for some constant C_d *that depends only on* d.

REMARK 3.3.   Note that inequalities (3.1) and (3.2) hold true for all x and y and not only for almost every x and y. This is true if one chooses a natural representative for the almost everywhere defined σ and F. This is done classically as

follows: recall that a simple definition of a Lebesgue point is that, for any convolution kernel, $K_\varepsilon \star F(x)$ has a unique limit as $\varepsilon \to 0$. F is not continuous at x in that case in general, only approximately continuous in the above sense. Note that this definition is independent of the chosen representative for F. Classically, one then chooses as the right representative of the almost everywhere defined function F at x, the limit of $K_\varepsilon \star F(x)$ as $\varepsilon \to 0$. This only changes F on a negligible set since if $F \in BV$ then every point x is a Lebesgue point except on the jump set of F, which is at most of dimension $d - 1$ (it is σ-finite for the Hausdorff measure $\mathcal{H}^{d-1}$, see [3]). As will appear in the proof, this representative satisfies (3.2) for all Lebesgue points x, y. Since the inequality is obvious when $h(t, x) = \infty$ or $h(t, y) = \infty$, it is also true when x or y is not a Lebesgue point of F.

In the sequel, we shall always assume that the functions F and σ are equal to their natural representative as defined above.

We start with the proof of Lemma 3.2.

PROOF OF LEMMA 3.2.   First, observe that by the definition of h, the result is obvious if $|x - y| \geq 1$. Assume now that $|x - y| \leq 1$. We recall the lemma from [16].

LEMMA 3.4.   *Assume $F \in C^1(\mathbb{R}^d)$. There exists a constant C depending only on d s.t. for any x, $y \in \mathbb{R}^d$,*

$$(3.3) \qquad |F(x) - F(y)| \leq C \int_{B(x,y)} \left(\frac{1}{|x - z|^{d-1}} + \frac{1}{|y - z|^{d-1}} \right) |\nabla F|(dz),$$

where $B(x, y)$ denotes the ball of center $(x + y)/2$ and diameter $|x - y|$.

The first point is to extend inequality (3.3) to any $F \in BV_{\text{loc}}$. Consider a sequence of smooth approximations $K_\varepsilon \star F$ with $K_\varepsilon \geq 0$ a classical convolution kernel with $K(-x) = K(x)$ and support in $B(0, 1)$. At every Lebesgue point x of F, one has that $K_\varepsilon \star F \to F$ and, therefore, if x and y are distinct Lebesgue points apply inequality (3.3) to $K_\varepsilon \star F$ and take the limit $\varepsilon \to 0$ to find

$$|F(x) - F(y)| = \lim_{\varepsilon \to 0} |K_\varepsilon \star F(x) - K_\varepsilon \star F(y)|$$

$$\leq C \limsup_{\varepsilon \to 0} \int_{B(x,y)} \left(\frac{1}{|x - z|^{d-1}} + \frac{1}{|y - z|^{d-1}} \right) |\nabla K_\varepsilon \star F|(dz).$$

We hence have to dominate the right-hand side. First, notice that since $K_\varepsilon \geq 0$,

$$|\nabla K_\varepsilon \star F|(z) = \left| \int K_\varepsilon(z - w) \nabla F(dw) \right| \leq \int K_\varepsilon(z - w) |\nabla F(dw)|$$

$$\leq K_\varepsilon \star |\nabla F|(z).$$

Therefore, provided ε is small with respect to $|x - y|$, which we may always assume as we are considering $\varepsilon \to 0$, and since K_ε has support in the ball of radius ε, we have that

$$\int_{B(x,y)} \left(\frac{1}{|x - z|^{d-1}} + \frac{1}{|y - z|^{d-1}} \right) |\nabla K_\varepsilon \star F|(dz)$$

$$\leq \int_{\tilde{B}(x,y)} K_\varepsilon \star \phi_{x,y}(z) |\nabla F|(dz),$$

where $\tilde{B}(x, y)$ denotes the ball of center $(x + y)/2$ and diameter $2|x - y|$ and with $\phi_{x,y}(z) = \frac{1}{|x-z|^{d-1}} + \frac{1}{|y-z|^{d-1}}$. Now observe that, since w^{-d+1} is integrable, one has for all $z \in \mathbb{R}^d$

$$\int K_\varepsilon(z - w) w^{-d+1} \, dw \leq \frac{C}{(|z| + \varepsilon)^{d-1}} \leq \frac{C}{|z|^{d-1}}.$$

Therefore,

$$K_\varepsilon \star \phi_{x,y}(z) \leq C \phi_{x,y}(z).$$

On the other hand, since $|\nabla F|$ is a positive measure, we have that

$$\int_{\tilde{B}(x,y)} K_\varepsilon \star \phi_{x,y}(z) |\nabla F|(dz) \leq C \int_{\tilde{B}(x,y)} \phi_{x,y}(z) |\nabla F|(dz).$$

Hence we have proved that for any x, y that are Lebesgue points of F,

$$(3.4) \qquad |F(x) - F(y)| \leq C \int_{\tilde{B}(x,y)} \left(\frac{1}{|x - z|^{d-1}} + \frac{1}{|y - z|^{d-1}} \right) |\nabla F|(dz).$$

On the other hand, for $F \in BV_{\mathrm{loc}}$ the set of non-Lebesgue points can be defined (see [2] and the references therein) as the set of x s.t.:

$$\liminf r^{-d+1} \int_{B(x,r)} |\nabla F(dz)| > 0.$$

At such a point x, one has

$$\int_{B(x,r)} \frac{|\nabla F|(dz)}{|x - z|^{d-1}} \geq \sum_{n \geq |\log_2 r|} 2^{n(-d+1)} \int_{2^{-n-1} \leq |x-z| < 2^{-n}} |\nabla F|(dz) = +\infty,$$

and inequality (3.4) is trivial. As $h(t, x) = +\infty$, if

$$\int_{B(x,r)} \frac{|\nabla F|(dz)}{|x - z|^{d-1}} = +\infty,$$

for some $r > 0$ then this also implies that x is necessarily a Lebesgue point of F if $h(t, x) < \infty$. Now $|\nabla F| \leq |\nabla F|_s + \sqrt{\log L} \lambda + |\nabla F|_a \mathbb{1}_{|\nabla F|_a \geq \sqrt{\log L}}$ where λ is Lebesgue's measure on $\mathbb{R}^d$, where $|\nabla F|_a$ and $|\nabla F|_s$ are the absolutely continuous

and singular parts of the measure $|\nabla F|$, and where we identified $|\nabla F|_a$ with its density w.r.t. λ in the indicator function. Thus, if $1/L \leq |x - y| \leq 1$,

$$\frac{1}{|x - y|} \int_{\tilde{B}(x,y)} \frac{|\nabla F|(dz)}{|x - z|^{d-1}}$$

$$\leq C\left(\sqrt{\log L} + \int_{B(x,2)} \frac{|\nabla F|_a(z)\mathbb{1}_{|\nabla F| \geq \sqrt{\log L}}\, dz + |\nabla F|_s(dz)}{(1/L + |x - z|)|x - z|^{d-1}}\right),$$

where $B(x, 2)$ is the ball of radius 2 centered at x and where we used that if $z \in \tilde{B}(x, y)$ then $|x - z| + 1/L \leq 3|x - y|$. Similarly, if $|x - y| \leq 1/L$,

$$\int_{\tilde{B}(x,y)} \frac{|\nabla F|(dz)}{|x - z|^{d-1}}\, dz$$

$$\leq \frac{C}{L}\left(\sqrt{\log L} + \int_{B(x,2)} \frac{|\nabla F|_a(z)\mathbb{1}_{|\nabla F| \geq \sqrt{\log L}}\, dz + |\nabla F|_s(dz)}{(1/L + |x - z|)|x - z|^{d-1}}\right),$$

where we used that if $z \in B(x, y)$, then $|x - z| + 1/L \leq 2/L$. By the definition of M_L, this concludes the proof. $\quad\square$

PROOF OF LEMMA 3.1. This is a classical result for which we can give an easy proof by applying the preliminary work that we have just done. In particular, we recall that we have proved that estimate (3.4) holds for any $F \in BV$ at any points x and y. Applying this inequality to σ, we obtain that for any x, y,

$$|\sigma(x) - \sigma(y)| \leq C \int_{\tilde{B}(x,y)} \left(\frac{1}{|x - z|^{d-1}} + \frac{1}{|y - z|^{d-1}}\right)|\nabla \sigma|(dz).$$

Now for a given x decompose

$$\int_{\tilde{B}(x,y)} \frac{|\nabla \sigma|(dz)}{|x - z|^{d-1}} = \sum_{k=0}^{\infty} \int_{2^{-k} \leq \frac{|x-z|}{|x-y|} \leq 2^{-k+1}} \frac{1}{|x - z|^{d-1}}|\nabla \sigma|(dz)$$

$$\leq \sum_{k=0}^{\infty} 2^{k(d-1)}|x - y|^{1-d} \int_{\frac{|x-z|}{|x-y|} \leq 2^{-k+1}} |\nabla \sigma|(dz)$$

$$\leq \sum_{k=0}^{\infty} 2^{-k+d}|x - y|M|\nabla \sigma|(x) \leq 2^{d+1}|x - y|M|\nabla \sigma|(x),$$

by the definition of the maximal function. This concludes the proof. $\quad\square$

Let us turn now to our last bound which uses $\partial_x^{1/2}\sigma$.

LEMMA 3.5. *Fix $t \geq 0$ and assume that $\sigma(t, \cdot) \in L^1_{\text{loc}}$ and $\partial_x^{1/2}\sigma(t, \cdot)$ is a locally finite Radon measure. Then for any $x, y \in \mathbb{R}^d$,*

$$|\sigma(t, x) - \sigma(t, y)| \leq (M|\partial_x^{1/2}\sigma|(t, x) + M|\partial_x^{1/2}\sigma|(t, y))|x - y|^{1/2}.$$

PROOF. By the definition of $\partial_x^{1/2}\sigma$,

$$\sigma(x) = K \star \partial_x^{1/2}\sigma,$$

for the convolution kernel K with $\mathcal{F}K = |\xi|^{-1/2}$, which implies that

(3.5) $$|K(x)| \leq C|x|^{d-1/2}, \qquad |\nabla K(x)| \leq C|x|^{d+1/2}.$$

Now simply compute

$$\begin{aligned}|\sigma(x) - \sigma(y)| \leq & \int_{|z-x| \geq 2|x-y|} |K(x-z) - K(y-z)||\partial_x^{1/2}\sigma|(dz) \\ & + \int_{|z-x| \leq 2|x-y|} (|K(x-z)| + |K(y-z)|)|\partial_x^{1/2}\sigma|(dz).\end{aligned}$$

Denote $|x - y| = r$. One has by (3.5)

$$\begin{aligned}\int_{|z-x| \leq 2r} |K(x-z)||\partial_x^{1/2}\sigma|(dz) & \leq C \sum_{n \geq -1} \int_{|z-x| \leq 2^{-n}r} \frac{2^{n(d-1/2)}}{r^{d-1/2}}|\partial_x^{1/2}\sigma|(dz) \\ & \leq C \sum_{n \geq -1} 2^{-n/2}r^{1/2}M|\partial_x^{1/2}\sigma|(x) \\ & = Cr^{1/2}M|\partial_x^{1/2}\sigma|(x).\end{aligned}$$

Since $|z - x| \leq 2r$ implies that $|z - y| \leq 3r$, one has the same inequality:

$$\int_{|z-x| \leq 2r} |K(y-z)||\partial_x^{1/2}\sigma|(dz) \leq Cr^{1/2}M|\partial_x^{1/2}\sigma|(y).$$

As for the last term, first note that if $|x - z| \geq 2|x - y|$ then $|y - z| \geq |x - z|/2$. Hence by (3.5) if $|x - z| \geq 2|x - y|$

$$|K(x-z) - K(y-z)| \leq C\frac{|x-y|}{|x-z|^{d+1/2}}.$$

Therefore,

$$\begin{aligned}\int_{|z-x| \geq 2|x-y|} & |K(x-z) - K(y-z)||\partial_x^{1/2}\sigma|(dz) \\ & \leq \sum_{n \geq 1} \int_{|z-x| \geq 2^n r} C\frac{r}{(2^n r)^{d+1/2}}|\partial_x^{1/2}\sigma|(dz) \\ & \leq Cr^{1/2} \sum_{n \geq 1} 2^{-n/2}M|\partial_x^{1/2}\sigma|(x) \\ & \leq Cr^{1/2}M|\partial_x^{1/2}\sigma|(x).\end{aligned}$$

Summing up the three estimates concludes the proof. $\quad\square$

3.2. *Proof of Theorems* 2.2, 2.6 *and* 2.10.

PROOF OF THEOREM 2.2. First of all, $\| \cdot \|_{H^1(v)}$ is indeed a norm on $H^1(v)$. By definition, it is nonnegative and finite on $H^1(v)$. Next, if $\lambda > 0$ then $M(|\lambda f|) = \lambda M|f|$, and thus $\|\lambda f\|_{H^1(v)} = |\lambda| \|f\|_{H^1(v)}$. The triangle inequality is also trivially satisfied as $M(f + g) \leq Mf + Mg$.

Finally, if $\|f\|_{H^1(v)} = 0$ then $M|f| = 0$ on the support of v which contains (at least) one point x_0 since v is a probability measure. But now $M|f|(x_0) = 0$ implies that $f = 0$ by the definition of the maximal function.

We now prove (2.2). Consider a sequence f_n in $H^1(v)$ s.t. f_n converges to some f in the sense of distributions and assume (possibly restricting to a subsequence)

$$\sup_n \|f_n\|_{H^1(v)} < \infty.$$

(Otherwise, there is nothing to prove.)

We notice that f_n is hence uniformly bounded in BV_{loc}. Indeed for any $R > 0$, and any $x \in B(0, R)$

$$|\nabla f_n|\big(B(0, R)\big) \leq (2R)^d M|\nabla f_n|(x),$$

so that by Cauchy–Schwarz,

$$(3.6) \qquad |\nabla f_n|\big(B(0, R)\big) \leq \frac{2^d R^d}{(\int_{B(0,R)} v(dx))^{1/2}} \|f_n\|_{H^1(v)}.$$

As $f_n \to f$ in $\mathcal{D}'$ then f belongs to BV_{loc} as well. Therefore, $M|\nabla f|$ is well defined.

On the other hand, ∇f_n converges to ∇f in $\mathcal{D}'$. Note that, for all $\varphi \in C_c^\infty(\mathbb{R}^d)$ with $\varphi \geq 0$, the map $\mu \mapsto \int \varphi |\mu|$ is convex and continuous on the set of locally finite Radon measures on $\mathbb{R}^d$ for the strong topology of total variation. Hence it is lower semicontinuous for the weak-* topology, and so

$$\int \varphi |\nabla f|(dx) \leq \liminf \int \varphi |\nabla f_n|(dx).$$

Now fix any $c > 1$ and any $r > 0$ and note that the previous inequality implies that

$$\frac{1}{|B(0, r)|} \int_{B(0,r)} |\nabla f|(x + dz) \leq \frac{1}{|B(0, r)|} \liminf \int_{B(0,cr)} |\nabla f_n|(x + dz)$$

$$\leq c^d \liminf M|\nabla f_n|(x).$$

Taking now the supremum in r, we deduce that for any $c > 1$

$$M|\nabla f|(x) \leq c^d \liminf M|\nabla f_n|(x).$$

Apply now Fatou's lemma and let c go to 1 to deduce

$$\int \big(M|\nabla f|(x)\big)^2 u(dx) \leq \liminf \int \big(M|\nabla f_n|(x)\big)^2 u(dx).$$

The same steps can be performed with $M|f_n|$ and $M|f|$ thus proving that $f \in H^1(v)$ and that (2.2) holds.

Let us now prove that $H^1(v)$ is complete which concludes the proof that $H^1(v)$ is a Banach space. Accordingly, consider any Cauchy sequence f_n in $H^1(v)$.

The sequence f_n is then also Cauchy in BV_{loc}. Indeed using (3.6) for $f_n - f_m$, we obtain that for any $R > 0$

$$|\nabla(f_n - f_m)|(B(0, R)) \leq \frac{2^d R^d}{(\int_{B(0,R)} v(dx))^{1/2}} \|f_n - f_m\|_{H^1(v)}.$$

Therefore, there exists $f \in BV_{\mathrm{loc}}$ s.t. f_n converges toward f in BV_{loc}. In particular, f_n converges to f in $\mathcal{D}'$ and we may use (2.2) a first time to deduce that $f \in H^1(v)$.

It remains to show that $\|f_n - f\|_{H^1(v)} \to 0$. For that, fix n and consider the sequence $f_n - f_m$ in m. This sequence converges in the sense of distribution to $f_n - f$. We conclude using again (2.2) that

$$\|f_n - f\|_{H^1(v)} \leq \liminf_{m \to \infty} \|f_n - f_m\|_{H^1(v)}.$$

Let us now turn to the last part of Theorem 2.2. We first recall that if μ is a finite, nonnegative Radon measure, then $M\mu$ is lower semicontinuous. This follows from similar arguments to the ones above: Consider any $x_n \to x$, then for $c > 1$

$$\frac{1}{|B(0, r)|} \int_{B(0,r)} \mu(x + dz) \leq \frac{1}{|B(0, r)|} \liminf \int_{B(0,cr)} \mu(x_n + dz)$$

$$\leq c^d \liminf M\mu(x_n).$$

The lower semicontinuity of $M\mu$ then follows taking the supremum in r and then the infimum in c.

Denote now $g = (M|\nabla f|)^2 + (M|f|)^2$, g is a nonnegative, Borel function with values in $\mathbb{R}_+ \cup \{+\infty\}$. By the previous remark, it is also lower semicontinuous. Note that for any positive measure μ

$$\int g \, d\mu = \int_0^\infty \int \mathbb{1}_{g(x) > \xi} \mu(dx) \, d\xi.$$

Now assume $v_n \to v$ in the tight topology of P^1. Note that for any open set O

$$\int_O dv \leq \liminf \int_O dv_n.$$

Take $O = \{g(x) > \xi\}$ which is open by the lower semicontinuity of g. Therefore, Fatou's lemma entails

$$\int g \, dv \leq \liminf \int g \, dv_n,$$

which completes the proof of Theorem 2.2. $\quad\square$

PROOF OF THEOREM 2.6.   The proof is nearly identical to that of Theorem 2.2, and for this reason, we omit it here. The only difference is that the space BV is replaced by the space of L^1_{loc} functions f s.t. $\partial_x^{1/2} f$ is a locally finite measure. $\square$

PROOF OF THEOREM 2.10.   The first part of the proof concerning the lower semicontinuity follows exactly the same steps as the proof of Theorem 2.2. One uses the same intermediate control through the BV norm as, for all $R, L \geq 1$,

$$M_L |\nabla f|(x_0)$$
$$\geq \sqrt{\log L}$$
$$+ \frac{1}{R^{d-1}(R + L^{-1})} \int_{B(0,R)} \left(|\nabla_a f|(z) \mathbb{1}_{|\nabla_a f|(z) \geq \sqrt{\log L}} \, dz + |\nabla_s f|(dz) \right)$$
$$\geq \frac{1}{C\sqrt{L}(1 + R^d)} \int_{B(0,R)} |\nabla f|(dz).$$

One also has the same type of lower semicontinuity properties as for instance if $f_n \to f$ in the sense of distribution for f_n a sequence uniformly bounded in BV_{loc} then for any $L' < L$

$$M_{L'} \nabla f(x) \leq \liminf_n M_L \nabla f_n(x).$$

Taking the supremum over L leads to (2.12) as ϕ is continuous.

We skip the rest of the details for this first part and instead focus on the connection with $W^{1,1}$ which is the main novel feature of $W^{\phi,\text{weak}}$.

By contradiction, assume that $f \in W^{\phi,\text{weak}}(v)$ and $v \geq 1/C$ over Ω but that $f \notin W^{1,1}(B(x_0, r))$ for some ball s.t. $B(x_0, 2r) \subset \Omega$. Since $f \in BV_{\text{loc}}$, it implies that the singular part $|\nabla_s f|$ does not vanish on $B(x_0, r)$. On the other hand,

$$\int_{\mathbb{R}^d} M_L \nabla f v(dx) \geq \frac{1}{C} \iint_{B(x_0,2r)^2} \frac{|\nabla_s f|(dz)}{(L^{-1} + |z - x|)|z - x|^{d-1}} \, dx.$$

Define the kernel

$$K_L = C_L \frac{\mathbb{1}_{|x| \leq 2r}}{(L^{-1} + |x|)|x|^{d-1}},$$

with C_L s.t. $\|K_L\|_{L^1} = 1$. Observe that K_L is a standard approximation of the identity by convolution so in particular

$$\liminf_{L \to \infty} \int_{B(x_0,2r)} K_L \star (|\nabla_s f|) \, dx \geq \int_{B(x_0,r)} |\nabla_s f|(B(x_0, r)) > 0.$$

As $C_L \sim \log L$, this has for consequence that there exists $C > 0$ s.t. for L large enough

$$\int_{B(x_0,2r)^2} \frac{|\nabla_s f|(dz)}{(L^{-1} + |z - x|)|z - x|^{d-1}} \, dx \geq \frac{\log L}{C}.$$

Therefore,

$$\|f\|_{W^{\phi,\mathrm{weak}}(v)} \geq \frac{1}{C}\sup_L \frac{\phi(L)}{L} = +\infty,$$

giving the desired contradiction.

Reciprocally, assume that $v \leq C$ on Ω and that $f \in W^{1,1}(K)$ compactly supported in $K \subset \Omega$. First, by Sobolev embedding, f and hence Mf belong to L^p for some $p > 1$ and $Mf \in L^\infty(\Omega^c)$. Therefore,

$$\int M|f|(x)v(dx) < \infty.$$

Then for $x \notin \Omega$

$$M_L \nabla f(x) \leq \sqrt{\log L} + \frac{1}{d(x,K)^d}\int_K |\nabla f(z)|\, dz.$$

As a consequence for any ϕ satisfying (2.11), there exists some finite constant C_ϕ s.t.

$$\|f\|_{W^{\phi,\mathrm{weak}}} \leq C_\phi + C\sup_L \frac{\phi(L)}{L\log L}\int_\Omega M_L\nabla f(x)\, dx.$$

Now decompose ∇f in level sets by defining for all $n \in \mathbb{Z}$

$$\omega_n = \{z \in K, 2^n \leq |\nabla f(z)| < 2^{n+1}\}.$$

Then

$$\int_\Omega M_L\nabla f(x)\, dx \leq |\Omega|\sqrt{\log L} + \sum_{n \geq \log_2 L - 1}\iint_{\Omega \times K}\frac{2^{n+1}\mathbb{1}_{z \in \omega_n}\, dz\, dx}{(L^{-1} + |z - x|)|z - x|^{d-1}}$$

$$\leq |\Omega|\sqrt{\log L} + \sum_{n \geq \log_2 L - 1} 2^{n+1}|\omega_n|\log L.$$

Since $\nabla f \in L^1$, one has $\sum 2^n|\omega_n| < \infty$, and thus

$$s_N = \sum_{n \geq N} 2^n|\omega_n| \to 0 \qquad \text{as } N \to \infty.$$

We can now define an appropriate ϕ: Choose any smooth function s.t. $\phi(x)/x$ is nondecreasing and

$$\phi(2^{N+1}) = 2^{N+1}\min(N^{1/4}, s_N^{-1}).$$

Then ϕ satisfies (2.11) while

$$\sup_L \frac{\phi(L)}{L\log L}\int_\Omega M_L\nabla f(x)\, dx \leq 2\sup_N \frac{\phi(2^{N+1})}{2^N}s_N \leq 4,$$

therefore concluding that $f \in W^{\phi,\mathrm{weak}}(v)$. $\quad\square$

4. Proof of Theorem 2.13. We use two types of estimates; one is based on an explicit quantitative estimate which generalizes the one in [5] for ordinary differential equations and one which generalizes the local time which is used in dimension 1 in the classical approach [8–10, 26, 34, 35]. We use the first quantitative estimate to prove existence and the second one to prove uniqueness (though with suitable modifications any one could be used for both existence and uniqueness).

The first method is more precise but more complicated than the second, which makes use of a similar argument as in [28, 39].

4.1. *Existence.* We consider the sequence of solutions to the regularized problem (1.3), and assume it satisfies the assumptions of Theorem 2.13. The proof is based on estimates on the expectation of the family of quantities

$$(4.1) \qquad Q_{nm}^{(\varepsilon)}(t) = \log\left(1 + \frac{|X_t^n - X_t^m|^2}{\varepsilon^2}\right), \qquad \varepsilon \in (0, 1], n, m \geq 1,$$

given in the next lemma.

LEMMA 4.1. *There exists a constant C such that, for all $0 < \varepsilon \leq 1$ and $n, m \geq 1$,*

$$(4.2) \qquad \sup_{t \in [0,T]} \mathbb{E}\big(Q_{nm}^{(\varepsilon)}(t)\big) \leq C\big(1 + |\log \varepsilon| \tilde{\eta}(\varepsilon)\big) + C \frac{\eta(n, m)}{\varepsilon^2},$$

where $\eta(n, m) \to 0$ when $n, m \to +\infty$ and $\tilde{\eta}(\varepsilon) := (\varepsilon \phi(\varepsilon^{-1}))^{-1} \to 0$ when $\varepsilon \to 0$.

PROOF. Note that

$$|\nabla(\log(1 + |x|^2/\varepsilon))| = \left|\frac{2x}{\varepsilon^2 + |x|^2}\right| \leq \frac{C}{\varepsilon + |x|}$$

and

$$|\nabla^2(\log(1 + |x|^2/\varepsilon))| = \left|\nabla\left(\frac{2x}{\varepsilon^2 + |x|^2}\right)\right| \leq \frac{C}{\varepsilon^2 + |x|^2}.$$

By Itô's formula, for any C_b^2 function f,

$$\begin{aligned}
\mathbb{E}(f(X_t^n - X_t^m)) = {}& f(0) + \frac{1}{2}\int_0^t \mathbb{E}[\nabla^2 f(X_s^n - X_s^m)(\sigma_n \sigma_n^*(X_s^n) \\
& + \sigma_m \sigma_m^*(X_s^m) - \sigma_n(X_s^n)\sigma_m^*(X_s^m) - \sigma_m(X_s^m)\sigma_n^*(X_s^n))]\, ds \\
& + \int_0^t \mathbb{E}(\nabla f(X_s^n - X_s^m) \cdot (F_n(s, X_s^n) - F_m(s, X_s^m)))\, ds.
\end{aligned}$$

Since $\sup_n(\|\sigma_n\|_\infty + \|F_n\|_\infty) < +\infty$, we deduce

$$
\begin{aligned}
\mathbb{E}&(f(X_t^n - X_t^m)) \\
&\leq f(0) + \frac{1}{2}\int_0^t \mathbb{E}\Big[|\nabla^2 f(X_s^n - X_s^m)|\Big(|\sigma(X_s^n) - \sigma(X_s^m)|^2 \\
&\quad + \sup_k \|\sigma_k\|_{L^\infty}\big(|\sigma_n(X_s^n) - \sigma(X_s^n)| + |\sigma_m(X_s^m) - \sigma(X_s^m)|\big)\Big)\Big]\,ds \\
&\quad + \int_0^t \mathbb{E}\big(|\nabla f(X_s^n - X_s^m)|\,|F_n(s, X_s^n) - F_m(s, X_s^m)|\big)\,ds.
\end{aligned}
$$

(4.3)

Hence

$$
\begin{aligned}
\mathbb{E}(Q_{nm}^{(\varepsilon)}(t)) &\leq C\int_0^t \mathbb{E}\Big(\frac{|\sigma(s, X_s^n) - \sigma(s, X_s^m)|^2}{\varepsilon^2 + |X_s^n - X_s^m|^2}\Big)\,ds + C\frac{\eta(n, m)}{\varepsilon^2} \\
&\quad + C\int_0^t \mathbb{E}\Big(\frac{|F(s, X_s^n) - F(s, X_s^m)|}{\varepsilon + |X_t^n - X_t^m|}\Big)\,ds,
\end{aligned}
$$

(4.4)

with C a constant independent of n and ε and $\eta(n, m) \to 0$ as $n, m \to \infty$ by Assumption (2.17).

Since $\|\sigma\|_{H_T^1(u_n)} + \|\sigma\|_{H_T^1(u_m)} < \infty$, denoting $h = M|\nabla\sigma|$,

$$
\int_0^T \int h^2(t, x)\big(u_n(t, dx) + u_m(t, dx)\big)\,dt \leq \|\sigma\|_{H_T^1(u_n)} + \|\sigma\|_{H_T^1(u_m)} \leq C,
$$

with C independent of n, m and ε. Now it follows from (3.1) that

$$
\int_0^t \mathbb{E}\Big(\frac{|\sigma(s, X_s^n) - \sigma(s, X_s^m)|^2}{\varepsilon^2 + |X_s^n - X_s^m|^2}\Big)\,ds \leq C\int_0^t \mathbb{E}\big(h^2(s, X_s^n) + h^2(s, X_s^m)\big)\,ds,
$$

and so

$$
\int_0^t \mathbb{E}\Big(\frac{|\sigma(s, X_s^n) - \sigma(s, X_s^m)|^2}{\varepsilon^2 + |X_s^n - X_s^m|^2}\Big)\,ds \leq C.
$$

We now turn to the term involving F and introduce the corresponding $h = |F| + M_{1/\varepsilon}\nabla F$.

By Lemma 3.2,

$$
\int_0^t \mathbb{E}\Big(\frac{|F(s, X_n^s) - F(s, X_m^s)|}{\varepsilon + |X_n^s - X_m^s|}\Big)\,ds \leq C\int_0^t \int h(s, x)\big(u_n(s, x) + u_m(s, x)\big)\,dx\,ds,
$$

and by (2.13),

$$
\begin{aligned}
\int_0^t &\int h(s, x)\big(u_n(s, x) + u_m(s, x)\big)\,dx\,ds \\
&\leq \frac{1/\varepsilon\,\log(1/\varepsilon)}{\phi(1/\varepsilon)}\big(\|F\|_{W_T^{\phi,\mathrm{weak}}(u_n)} + \|F\|_{W_T^{\phi,\mathrm{weak}}(u_m)}\big) \leq C\frac{|\log\varepsilon|}{\varepsilon\phi(\varepsilon^{-1})}.
\end{aligned}
$$

Note that we used the inequality $|F| \leq M|F|$ a.e., which follows from Lebesgue's points theorem since $BV_{\mathrm{loc}}(\mathbb{R}^d) \subset L^1_{\mathrm{loc}}(\mathbb{R}^d)$. The function $\tilde{\eta}(\varepsilon) = (\varepsilon\phi(\varepsilon^{-1}))^{-1} \to 0$ as $\varepsilon \to 0$ since ϕ is superlinear.

Combining the previous inequalities, we obtain (4.2). $\quad\square$

Fix $p > 1$. The next step consists in deducing from Lemma 4.1 that $(X_t^n - \xi)$ is a Cauchy sequence in $L^p(\Omega, L^\infty([0, T]))$. Since F_n and σ_n are uniformly bounded, it is standard to deduce from the Burkholder–Davis–Gundy inequality that $X_t^n - \xi$ are uniformly bounded in $L^p(\Omega, L^\infty([0, T]))$ for all $p > 1$, so we only need to prove the next lemma.

LEMMA 4.2.   *For all $p > 1$,*

$$(4.5) \qquad \mathbb{E}\Big(\sup_{t\in[0,T]} |X_t^n - X_t^m|^p \Big) \to 0 \qquad \text{as } n, m \to +\infty.$$

PROOF.   For fixed t, for any ε and L s.t. $0 < \varepsilon < L$,

$$\mathbb{E}(|X_t^n - X_t^m|^p) \leq \mathbb{E}(|X_t^n - X_t^m|^p; |X_t^n - X_t^m| \geq L) + \varepsilon^{p/2}$$
$$+ L^p \mathbb{P}(|X_t^n - X_t^m| \geq \sqrt{\varepsilon}).$$

Note that

$$(4.6) \quad \mathbb{E}(|X_t^n - X_t^m|^p; |X_t^n - X_t^m| \geq L) \leq \frac{1}{L}\big(\mathbb{E}(|X_t^n - \xi|^{p+1}) + \mathbb{E}(|X_t^m - \xi|^{p+1})\big).$$

Now the inequalities

$$\sup_{n\geq 1, t\in[0,T]} \mathbb{E}(|X_t^n - \xi|^{p+1}) < +\infty$$

and

$$\mathbb{P}(|X_t^n - X_t^m| \geq \sqrt{\varepsilon}) \leq \frac{\mathbb{E}Q_{nm}^{(\varepsilon)}(t)}{|\log\varepsilon|},$$

imply that

$$\mathbb{E}(|X_t^n - X_t^m|^p) \leq C\Big[\frac{1}{L} + \varepsilon^{p/2} + \frac{L^p}{|\log\varepsilon|}\Big(1 + |\log\varepsilon|\tilde{\eta}(\varepsilon) + \frac{\eta(n,m)}{\varepsilon^2}\Big)\Big].$$

Taking, for example, $\varepsilon^2 = \eta(n,m)$ and $L = (\frac{1}{|\log\varepsilon|} + \tilde{\eta}(\varepsilon))^{-1/2p}$, one concludes that

$$\sup_{t\in[0,T]} \mathbb{E}(|X_t^n - X_t^m|^p) \to 0 \qquad \text{as } n, m \to +\infty.$$

In order to pass the supremum inside the expectation, it suffices to observe that the computation of (4.3–4.4) in the proof of Lemma 4.1 can be applied to

$|A_{t\wedge\tau}^n - A_{t\wedge\tau}^m|^2 \vee |M_{t\wedge\tau}^n - M_{t\wedge\tau}^m|^2$, where τ is any stopping time and $X_t^n = \xi + A_t^n + M_t^n$ is Doob's decomposition of the semimartingale X_t^n, that is,

$$A_t^n = \int_0^t F(s, X_s^n)\, ds \quad \text{and} \quad M_t^n = \int_0^t \sigma(s, X_s^n)\, dW_s.$$

Note that to be fully rigorous, one first needs to regularize the supremum $\vee$.

Instead of (4.4), we obtain

$$\mathbb{E}\log\left(1 + \frac{|A_{t\wedge\tau}^n - A_{t\wedge\tau}^m|^2 \vee |M_{t\wedge\tau}^n - M_{t\wedge\tau}^m|^2}{\varepsilon^2}\right)$$

$$\leq C\int_0^t \mathbb{E}\left(\frac{|\sigma(s, X_s^n) - \sigma(s, X_s^m)|^2}{\varepsilon^2 + |A_t^n - A_t^m|^2 \vee |M_t^n - M_t^m|^2}\right) ds + C\frac{\eta(n, m)}{\varepsilon^2}$$

$$+ C\int_0^t \mathbb{E}\left(\frac{|F(s, X_s^n) - F(s, X_s^m)|}{\varepsilon + |A_t^n - A_t^m| \vee |M_t^n - M_t^m|}\right) ds,$$

or

$$\mathbb{E}\log\left(1 + \frac{|A_{t\wedge\tau}^n - A_{t\wedge\tau}^m|^2 \vee |M_{t\wedge\tau}^n - M_{t\wedge\tau}^m|^2}{\varepsilon^2}\right)$$

$$\leq C\int_0^t \mathbb{E}\left(\frac{|\sigma(s, X_s^n) - \sigma(s, X_s^m)|^2}{\varepsilon^2 + \frac{1}{4}|X_s^n - X_s^m|^2}\right) ds + C\frac{\eta(n, m)}{\varepsilon^2}$$

$$+ C\int_0^t \mathbb{E}\left(\frac{|F(s, X_s^n) - F(s, X_s^m)|}{\varepsilon + \frac{1}{2}|X_t^n - X_t^m|}\right) ds.$$

Therefore, the same computation as in Lemma 4.1 gives

$$\sup_{\substack{t\in[0,T],\\ \tau \text{ stopping time}}} \mathbb{E}\left(|A_{t\wedge\tau}^n - A_{t\wedge\tau}^m|^p \vee |M_{t\wedge\tau}^n - M_{t\wedge\tau}^m|^p\right) \to 0 \qquad \text{as } n, m \to +\infty.$$

Since $p > 1$, Doob's inequality entails

$$\mathbb{E}\left(\sup_{t\in[0,T]} |M_t^n - M_t^m|^p\right) \to 0 \qquad \text{as } n, m \to +\infty.$$

Fix $\eta > 0$, and fix n_0 such that

$$\sup_{\substack{t\in[0,T],\\ \tau \text{ stopping time}}} \mathbb{E}\left(|A_{t\wedge\tau}^n - A_{t\wedge\tau}^m|^p\right) \leq \eta$$

for all $n, m \geq n_0$. For all $M > 0$, let $\tau = \inf\{t \geq 0 : |A_t^n - A_t^m| \geq M\}$. Then

$$\mathbb{P}\left(\sup_{t\in[0,T]} |A_t^n - A_t^m| \geq M\right) = \mathbb{P}(\tau \leq T) \leq \frac{\eta}{M^p}.$$

Now, for all $1 < q < p$,

$$\mathbb{E}\Big(\sup_{t \in [0,T]} |A_t^n - A_t^m|^q \Big) = q \int_0^{+\infty} x^{q-1} \mathbb{P}\Big(\sup_{t \in [0,T]} |A_t^n - A_t^m| \ge x \Big) dx$$

$$\le q \int_0^{+\infty} x^{q-1} \Big(\frac{\eta}{x^p} \wedge 1 \Big) dx = \frac{p \eta^{q/p}}{p - q}.$$

Therefore,

$$\mathbb{E}\Big(\sup_{t \in [0,T]} |A_t^n - A_t^m|^q \Big) \to 0 \qquad \text{as } n, m \to +\infty,$$

which concludes the proof of (4.5). $\square$

From the fact that $(X^n - \xi)$ is a Cauchy sequence in $L^p(\Omega, L^\infty([0,T]))$, it is standard to deduce the almost sure convergence for the L^∞ norm of a subsequence of $(X_t^n, t \in [0,T])_n$ to a process $(X_t, t \in [0,T])$ such that $(X_t - \xi, t \in [0,T]) \in L^p(\Omega, L^\infty([0,T]))$ for all $p > 1$. Since the convergence holds for the L^∞ norm, the process X is a.s. continuous and adapted to the filtration $(\mathcal{F}_t)_{t \ge 0}$.

Since u_n converges to u in the weak topology of M_1, we have for all bounded continuous function f on $[0,T] \times \mathbb{R}^d$

$$\mathbb{E} \int_0^T f(t, X_t) \, dt = \int_{\mathbb{R}^d} \int_0^T f(t, x) u(dt, dx),$$

so $u(t, dx)$ is the law of X_t for almost all t.

Defining for all $t \in [0,T]$,

$$Y_t := \int_0^t F(s, X_s) \, ds + \int_0^t \sigma(s, X_s) \, dW_s,$$

it only remains to check that $Y_t = X_t - \xi$ for all $t \in [0,T]$ a.s. As

$$X_t^n - \xi = \int_0^t F_n(s, X_s^n) \, ds + \int_0^t \sigma_n(s, X_s^n) \, dW_s,$$

one has $Y_t = X_t$ provided that

$$\int_0^t \mathbb{E}\big(|F_n(s, X_s^n) - F(s, X_s)| + |\sigma_n(s, X_s^n) - \sigma(x, X_s)|^2 \big) \, ds \to 0.$$

From assumption (2.17) and the L^∞ bounds on F, σ and σ_n, this is implied by: For any fixed $\varepsilon > 0$,

$$\int_0^T \big[\mathbb{P}(|F(s, X_s^n) - F(s, X_s)| > \varepsilon) + \mathbb{P}(|\sigma(s, X_s^n) - \sigma(x, X_s)| > \varepsilon) \big] \, ds \to 0.$$

We prove it for σ, the argument for F being fully similar.

By Corollary 2.4,

$$(4.7) \qquad \int_0^T \int_{\mathbb{R}^d} \left(M |\nabla \sigma(t,x)| \right)^2 \left(u(t,dx) + u_n(t,dx) \right) dt$$

$$\leq \|\sigma\|_{H_T^1(u_n)} + \liminf \|\sigma\|_{H_T^1(u_n)} \leq C.$$

Now by (3.1),

$$\mathbb{P}\big(|\sigma(s, X_s^n) - \sigma(s, X_s)| > \varepsilon \big)$$

$$\leq \mathbb{P}\big((M|\nabla\sigma|(s, X_s^n) + M|\nabla\sigma|(s, X_s)) > \varepsilon / |X_s^n - X_s| \big)$$

$$\leq \mathbb{P}\big(|X_s^n - X_s| > \varepsilon^2 \big) + \mathbb{P}\left(M|\nabla\sigma|(s, X_s^n) \geq \frac{1}{2\varepsilon} \right)$$

$$+ \mathbb{P}\left(M|\nabla\sigma|(s, X_s) \geq \frac{1}{2\varepsilon} \right),$$

and one easily concludes from (4.7) and the fact that $|X_s^n - X_s| \to 0$ almost surely.

4.2. *Uniqueness.*   Consider two solutions X and Y satisfying the assumptions of point (ii) in Theorem 2.13. Define a family of functions $(L_\varepsilon)_\varepsilon$ in $C^\infty(\mathbb{R}^d)$ satisfying

$$L_\varepsilon(x) = 1 \qquad \text{if } |x| \geq \varepsilon, \qquad L_\varepsilon(x) = 0 \qquad \text{if } |x| \leq \varepsilon/2,$$

$$\varepsilon \|\nabla L_\varepsilon\|_{L^\infty} + \varepsilon^2 \|\nabla^2 L_\varepsilon\|_{L^\infty} \leq C,$$

with C independent of ε, and $L_\varepsilon(x) \geq L_{\varepsilon'}(x)$ for all $\varepsilon \leq \varepsilon'$ and $x \in \mathbb{R}^d$. Use Itô's formula,

$$\mathbb{E}\big(L_\varepsilon(X_t - Y_t) \big) = L(0) + \int_0^t \mathbb{E}\big(\nabla L_\varepsilon(X_s - Y_s) \cdot \big(F(s, X_s) - F(s, Y_s) \big) \big) ds$$

$$+ \int_0^t \mathbb{E}\big(\nabla^2 L_\varepsilon(X_s - Y_s)(X_s)$$

$$: \big(\sigma\sigma^* + \sigma\sigma^*(Y_s) - \sigma(X_s)\sigma^*(Y_s) - \sigma(Y_s)\sigma^*(X_s) \big) \big) ds.$$

Hence

$$\mathbb{E}\big(L_\varepsilon(X_t - Y_t) \big) \leq C \int_0^t \mathbb{E}\left(\mathbb{1}_{\varepsilon/2 \leq |X_s - Y_s| \leq \varepsilon} \left(\frac{|\sigma(s, X_s) - \sigma(s, Y_s)|^2}{\varepsilon^2} \right.\right.$$

$$\left.\left. + \frac{|F(s, X_s) - F(s, Y_s)|}{\varepsilon} \right) \right) ds.$$

Now denote $h = M|\nabla\sigma|$ so that

$$\int_0^T \int |h(t,x)|^2 \big(u_X(t,dx) + u_Y(t,dx) \big) dt \leq C < \infty.$$

Define as well $\widetilde{h}_\varepsilon = |F| + M_{1/\varepsilon} \nabla F$ s.t.

$$\int_0^T \int \widetilde{h}_\varepsilon (u_X + u_Y) \, dx \, ds \leq \frac{C|\log \varepsilon|}{\varepsilon \phi(\varepsilon^{-1})}.$$

The corresponding computation involving $\widetilde{h}_\varepsilon$ is now tricky, precisely because of the dependence on ε in $\widetilde{h}_\varepsilon$. To simplify it, we will use a slightly different definition.

First, note that, as observed in Section 2.1.3, one can always assume that ϕ satisfies (2.11), and so $\phi(\xi)/\xi$ is a nondecreasing function which grows note faster than $\log \xi$. In particular, there exists a constant $C > 0$ s.t.

$$\frac{1}{C} \varepsilon \phi(\varepsilon^{-1}) \leq \frac{\phi(\xi)}{\xi} \leq C \varepsilon \phi(\varepsilon^{-1}) \qquad \forall \xi \in \left[\varepsilon^{-1/2}, \varepsilon^{-1} \right].$$

Consider the partition of $(0, 1) = \bigcup_i I_i$ where the $I_i = [a_i, b_i)$ are disjoint with $b_i = \sqrt{a_i}$ (except for $I_0 := [1/2, 1)$). In particular, $|I_i| := b_i - a_i$ satisfies

$$|I_i| \sim \sqrt{a_i} \qquad \text{when } i \to +\infty.$$

Now for any $\varepsilon \in I_i$, choose $\bar{h}_\varepsilon = \widetilde{h}_{a_i}$. One has

$$\int_0^T \int \bar{h}_\varepsilon(t, x)\big(u_X(t, x) + u_Y(t, x)\big) \, dx \, dt \leq C \frac{|\log \varepsilon|}{\varepsilon \phi(\varepsilon)} \leq C' \frac{|\log b_i|}{b_i \phi(b_i^{-1})}.$$

Now by (3.1) and Lemma 3.2,

$$\mathbb{E}\big(L_\varepsilon(X_t - Y_t)\big) \leq C \int_0^t \mathbb{E}\big[\big(h^2(s, X_s) + h^2(s, Y_s)\big) \mathbb{1}_{\varepsilon/2 \leq |X_t - Y_t| \leq \varepsilon}\big] \, ds$$

$$+ C \int_0^t \mathbb{E}\big[\big(\bar{h}_\varepsilon(s, X_s) + \bar{h}_\varepsilon(s, Y_s)\big) \mathbb{1}_{\varepsilon/2 \leq |X_t - Y_t| \leq \varepsilon}\big] \, ds.$$

Denote

$$\alpha_k = \int_0^t \mathbb{E}\big[\big(h^2(s, X_s) + h^2(s, Y_s)\big) \mathbb{1}_{2^{-k-1} \leq |X_t - Y_t| \leq 2^{-k}}\big] \, ds.$$

Note that

$$\sum_k \alpha_k \leq \int_0^t \mathbb{E}\big(\big(h^2(s, X_s) + h^2(s, Y_s)\big)\big) \, ds$$

$$= \int_0^t \int h^2(s, x)\big(u_X(dx, s) + u_Y(dx, s)\big) \, ds \leq C.$$

Therefore, $\alpha_k \to 0$ as $k \to +\infty$.

Denote similarly

$$\beta_k = \int_0^t \mathbb{E}\big(\big(\bar{h}_{2^{-k}}(s, X_s) + \bar{h}_{2^{-k}}(s, Y_s)\big) \mathbb{1}_{2^{-k-1} \leq |X_t - Y_t| \leq 2^{-k}}\big) \, ds.$$

Denote $J_i = \{k, [2^{-k-1}, 2^{-k}) \subset I_i\}$. Note that $|J_i| \geq \frac{1}{C}|\log b_i|$ (in fact, $|J_i| = \frac{|\log b_i|}{2\log 2}$), and since $\bar{h}_\varepsilon$ is fixed on $\varepsilon \in I_i$,

$$\frac{1}{|J_i|} \sum_{k \in J_i} \beta_k \leq \frac{1}{|J_i|} \int_0^t \int \bar{h}_{b_i}(s, x)(u_X(dx, s) + u_Y(dx, s)) \, ds$$

$$\leq \frac{C}{b_i \phi(b_i^{-1})} \to 0 \qquad \text{as } i \to \infty.$$

Therefore, $\beta_{n_k} \to 0$ as $k \to +\infty$ for some subsequence $n_k \to +\infty$. Consequently, since the sequence of functions L_ε is nonincreasing,

$$\sup_{t \in [0,T]} \mathbb{E}(L_\varepsilon(X_t - Y_t)) \to 0 \qquad \text{as } \varepsilon \to 0.$$

On the other hand,

$$\mathbb{E}(L_\varepsilon(X_t - Y_t)) \geq \mathbb{P}(|X_t - Y_t| > \varepsilon),$$

and by taking the limit $\varepsilon \to 0$, we deduce that for any $t \in [0, T]$

$$\mathbb{P}(|X_t - Y_t| > 0) = 0.$$

Since X_t and Y_t have a.s. continuous paths, we finally deduce that

$$\mathbb{P}\Big(\sup_{t \in [0,T]} |X_t - Y_t| = 0 \Big) = 1.$$

5. Proof of Theorem 2.15. This proof follows exactly the same steps as the general multidimensional case given in Section 4. The only differences are the functionals used and accordingly we skip the other parts of the proof which are identical.

Technically, the reason why the one-dimensional case is so special is that $|x|$ is linear except at $x = 0$ (see Section 5.2).

5.1. *Existence.* For $d = 1$, we replace the functional $Q_{nm}^{(\varepsilon)}$ by

$$\tilde{Q}_{nm}^{(\varepsilon)}(t) = e^{-U_t^{n,m}}|X_t^n - X_t^m|\log\Big(1 + \frac{|X_t^n - X_t^m|^2}{\varepsilon^2}\Big),$$

for $U_t^{n,m}$ a nonnegative stochastic process with bounded variation satisfying $dU_t^{n,m} = \lambda_t^{n,m} \, dt$ with $\lambda_t^{n,m}$ an adapted process (measurable function of a continuous, adapted process) to be chosen later.

Note that $f(x) = |x|\log(1 + |x|^2/\varepsilon^2)$ satisfies

$$|f'(x)| \leq 4\log\Big(1 + \frac{|x|^2}{\varepsilon^2}\Big) \quad \text{and} \quad |f''(x)| \leq \frac{C}{\varepsilon + |x|}.$$

Therefore, by Itô's formula,

$$\mathbb{E}(\tilde{Q}_{nm}^{(\varepsilon)}(t)) \leq C + C \int_0^t \mathbb{E}\left(\frac{|\sigma(X_s^n) - \sigma(X_s^m)|^2}{\varepsilon + |X_s^n - X_s^m|}\right) ds + \frac{\eta(n, m)}{\varepsilon}$$

$$+ \int_0^t \mathbb{E}\left(|X_s^n - X_s^m| \log(1 + |X_s^n - X_s^m|^2/\varepsilon^2)\right.$$

$$\times \left.\left(4\frac{|F(s, X_s^n) - F(s, X_s^m)|}{|X_s^n - X_s^m|} - \lambda_t^{n,m}\right)\right) ds.$$

The first term is treated identically as for the multidimensional case. The only difference here is that the careful choice of $\tilde{Q}_{nm}^{(\varepsilon)}$ improved the exponent of $|X_s^n - X_s^m|$ to 1 instead of 2 in the denominator. Therefore, this term can be controlled with the $H_T^{1/2}(u_{n,m})$ norm of σ by using Lemma 3.5 instead of estimate (3.1).

The drawback is that the term with F must be dealt with differently. We introduce $\tilde{h} = M|\nabla F|$ s.t.

$$\int_0^T \int_{\mathbb{R}^d} \tilde{h}(t, x)\left(u_m(t, dx) + u_n(t, dx)\right) dt \leq C.$$

We then choose

$$\lambda_t^{n,m} = 4\left(\tilde{h}(t, X_s^m) + \tilde{h}(t, X_t^m)\right).$$

Therefore, we deduce that

$$\sup_{t \leq T} \mathbb{E}(\tilde{Q}_{nm}^{(\varepsilon)}(t)) \leq C + \frac{\eta(n, m)}{\varepsilon}.$$

Using a similar method as in Theorem 2.13, we write for constants L and K to be chosen later

$$\mathbb{E}(|X_t^n - X_t^m|^p) \leq \mathbb{E}(|X_t^n - X_t^m|^p; |X_t^n - X_t^m| \geq L) + \frac{1}{|\log \varepsilon|^{p/2}}$$

$$+ \mathbb{P}(U_t^{n,m} \geq \log K)$$

$$+ L^p \mathbb{P}\left(|X_t^n - X_t^m| \geq \frac{1}{\sqrt{|\log \varepsilon|}}; U_t^{n,m} \leq \log K\right).$$

Note that

$$\mathbb{E}(U_t^{n,m}) = \mathbb{E}\left(\int_0^t \lambda_s^{n,m} ds\right) \leq 4 \int_0^t \tilde{h}(s, x)\left(u_n(s, dx) + u_m(s, dx)\right) ds \leq C.$$

Consequently

$$\mathbb{P}(U_t^{n,m} \geq \log(K)) \leq \frac{C}{\log K}.$$

In addition, for ε small enough,

$$\mathbb{P}\left(|X_t^n - X_t^m| \geq \frac{1}{\sqrt{|\log \varepsilon|}}; U_t^{n,m} \leq \log K\right) \leq \frac{K\mathbb{E}\tilde{Q}_{nm}^{(\varepsilon)}(t)}{2\sqrt{|\log \varepsilon|}}.$$

Therefore, using (4.6) as in the proof of Lemma 4.2,

$$\mathbb{E}(|X_t^n - X_t^m|^p) \leq C\left(\frac{1}{L} + \frac{1}{|\log \varepsilon|^{p/2}} + \frac{1}{\log K} + \frac{L^p K(1 + \frac{\eta(n,m)}{\varepsilon})}{\sqrt{|\log \varepsilon|}}\right).$$

Taking, for example, $\varepsilon = \eta(n,m)$, $K = |\log \varepsilon|^{1/8}$ and $L = |\log \varepsilon|^{1/8p}$, we deduce that

$$\sup_{t \in [0,T]} \mathbb{E}(|X_t^n - X_t^m|^p) \to 0 \qquad \text{as } n, m \to +\infty.$$

The rest of the proof is similar.

5.2. *Uniqueness.* For simplicity, we assume here that $F = 0$. Otherwise it is necessary to introduce U_t as in the previous subsection but it is handled in exactly the same way.

We similarly change the definition of L_ε in

$$\tilde{L}_\varepsilon(x) = |x| \qquad \text{if } |x| \geq \varepsilon, \qquad \tilde{L}_\varepsilon(x) = 0 \qquad \text{if } |x| \leq \varepsilon/2,$$

$$\|\nabla \tilde{L}_\varepsilon\|_{L^\infty} + \varepsilon\|\nabla^2 \tilde{L}_\varepsilon\| \leq C,$$

with C independent of ε.

Applying Itô's formula,

$$\mathbb{E}(\tilde{L}_\varepsilon(X_t - Y_t)) \leq C \int_0^t \mathbb{E}\left(\mathbb{1}_{\varepsilon/2 \leq |X_t - Y_t| \leq \varepsilon} \frac{|\sigma(X_s) - \sigma(Y_s)|^2}{\varepsilon}\right) ds.$$

By using as before the assumptions, Lemma 3.5 and the corresponding definition of $H_T^{1/2}(u_X)$ and $H_T^{1/2}(u_Y)$, one deduces that

$$\mathbb{E}(\tilde{L}_\varepsilon(X_t - Y_t)) \to 0 \qquad \text{as } \varepsilon \to 0.$$

This is slightly less strong than before $[L_\varepsilon(\varepsilon) \gg \tilde{L}_\varepsilon(\varepsilon)$ when $\varepsilon \to 0]$ but still enough. In particular, one has if $\alpha \geq \varepsilon$

$$\mathbb{P}(|X_t - Y_t| \geq \alpha) \leq \frac{1}{\alpha}\mathbb{E}(\tilde{L}_\varepsilon(X_t - Y_t)).$$

Therefore, by taking $\varepsilon \to 0$, one still obtains that for any $t \in [0, T]$,

$$\mathbb{P}(|X_t - Y_t| > 0) = 0,$$

which allows us to conclude as before.

6. Proof of Proposition 2.19. We simply use the energy estimates. The computations below are formal but could easily be made rigorous by taking a regularization of σ, F, and hence a and then pass to the limit

$$\frac{d}{dt} \int u^\alpha(t, x) \, dx = -\alpha(\alpha - 1) \int u^{\alpha-1}(t, x) \nabla u(t, x) \cdot F(t, x) \, dx$$

$$- \alpha(\alpha - 1) \int u^{\alpha-2}(t, x) \nabla u(t, x)^T a(t, x) \nabla u(t, x) \, dx$$

$$- \alpha(\alpha - 1) \int u^{\alpha-1}(t, x) \sum_{1 \le i, j \le d} \frac{\partial u(t, x)}{\partial x_i} \frac{\partial a_{ij}(t, x)}{\partial x_j} \, dx.$$

Note that by (2.21)

$$\int u^{\alpha-2}(t, x) \nabla u(t, x)^T a(t, x) \nabla u(t, x) \, dx \ge C \|\nabla u^{\alpha/2}\|_{L^2}^2.$$

On the other hand,

$$\int u^{\alpha-1}(t, x) \nabla u(t, x) \cdot F(t, x) \, dx \le \|\nabla u^{\alpha/2}\|_{L^2} \|u^{\alpha/2}\|_{L^2} \|F\|_{L^\infty}$$

$$\le \frac{C}{4} \|\nabla u^{\alpha/2}\|_{L^2}^2 + C' \int u^\alpha(t, x) \, dx.$$

And

$$\int u^{\alpha-1}(t, x) \sum_{1 \le i, j \le d} \frac{\partial u(t, x)}{\partial x_i} \frac{\partial a_{ij}(t, x)}{\partial x_j} \, dx \le \|\nabla u^{\alpha/2}\|_{L^2} \|u^{\alpha/2} \nabla a\|_{L^2}$$

$$\le \|\nabla u^{\alpha/2}\|_{L^2} \|\nabla a\|_{L^p} \|u^{\alpha/2}\|_{L^r},$$

with $1/2 = 1/p + 1/r$, which can be done since $p > d \ge 2$. Now by Sobolev embedding

$$\|u^{\alpha/2}\|_{L^r} \le \left(\int u^\alpha \, dx \right)^{\theta/2} \|\nabla u^{\alpha/2}\|_{L^2}^{1-\theta},$$

for some $\theta \in (0, 1]$, precisely $1/r = 1/2 - (1-\theta)/d$ or $(1-\theta)/d = 1/p$, provided that $p > d$. In that case, we immediately deduce that

$$\frac{d}{dt} \int u^\alpha(t, x) \, dx + \frac{C}{2} \int |\nabla u^{\alpha/2}|^2 \, dx \le C''(1 + \|\nabla a\|_{L^p}^{2/\theta}) \int u^\alpha \, dx.$$

This concludes the bound provided that

$$\int_0^T \|\nabla a\|_{L^p}^{2/\theta} < \infty,$$

which means that $\nabla a \in L^q_{t,\text{loc}}(L^p_x)$ with $1/q = \theta/2 = 1/2 - d/2p$. This exactly corresponds to the condition $2/q + d/p = 1$ with $p > d$.

Note that $p = d$ is critical here in the sense that the result could still hold in that case provided that the norm of ∇a is small enough with respect to the constant of ellipticity.

Finally, we hence deduce that for any t and any $\alpha < \infty$

$$\|u(t, \cdot)\|_{L^\alpha} \le \|u(t = 0, \cdot)\|_{L^\alpha} \le C,$$

with C independent of α since $u_0 \in L^1 \cap L^\infty$. This implies that $\|u(t, \cdot)\|_{L^\infty} \le C$ and completes the proof.

7. Proof of Proposition 2.24. We are going to prove this result under the assumptions of Corollary 2.21. The other cases are similar.

Fix a complete filtered probability space $(\Omega, (\mathcal{F}_t)_{t\ge 0}, \mathbb{P})$ equipped with a r-dimensional standard Brownian motion W. Fix also $u_0 > 0$ in $L^1 \cap L^\infty$ such that $\int_{\mathbb{R}^d} u_0(x)\, dx = 1$. Then, by Corollary 2.21, on the probability space $(\mathbb{R}^d \times \Omega, (\mathcal{B}(\mathbb{R}^d) \otimes \mathcal{F}_t)_{t\ge 0}, u_0(x)\, dx \times \mathbb{P}(d\omega))$, there is strong existence of a process $(X_t(x, \omega), t \ge 0)$ solution of (1.1) with $\xi(x, \omega) = x$ and pathwise uniqueness holds. We deduce that strong existence for almost every deterministic initial condition x holds for (1.1) on $(\Omega, (\mathcal{F}_t)_{t\ge 0}, (W_t)_{t\ge 0}, \mathbb{P})$. In addition, the family of laws $\mathbb{P}_x$ of the process $\omega \mapsto X(x, \omega)$ for $x \in \mathbb{R}^d$ forms a regular conditional probability of the law of X given ξ.

For uniqueness, the two key points are:

- first, that we are always in cases where uniqueness in law is known for *all* initial conditions in (1.1), and in particular for all deterministic initial conditions;
- second, that $u \in L^\infty$ by Corollary 2.20 (or is bounded by an explicit function in the case of Corollaries 2.16 and 2.22), $\sigma \in H^1(u)$ and $F \in W_T^{\phi, \mathrm{weak}}(u)$ (this is implied by Corollaries 2.4 and 2.12).

For all x such that strong existence holds for (1.1) with $\xi = x$, let X_t^x and $\hat{X}_t^x$ be two strong solutions of (1.1) such that $X_0^x = \hat{X}_0^x = x$ a.s. Repeating the proof of Lemma 4.1, we have

$$\mathbb{E} \log\left(1 + \frac{|X_t^x - \hat{X}_t^x|^2}{\varepsilon^2}\right)$$

$$\le C \int_0^t \mathbb{E}\big[M|\nabla\sigma|(s, X_s^x)^2 + M|\nabla\sigma|(s, \hat{X}_s^x)^2\big]\, ds$$

$$+ C \int_0^t \mathbb{E}\big[(|F| + M_{1/\varepsilon}\nabla F)(s, X_s^x) + (|F| + M_{1/\varepsilon}\nabla F)(s, \hat{X}_s^x)\big]\, ds.$$

By uniqueness in law, the two processes X^x and $\hat{X}^x$ have the same distribution $\mathbb{P}_x$, and so

$$\mathbb{E} \log\left(1 + \frac{|X_t^x - \hat{X}_t^x|^2}{\varepsilon^2}\right) \le C \int_0^t \mathbb{E}_x\big[\big((M|\nabla\sigma|)^2 + |F| + M_{1/\varepsilon}\nabla F\big)(s, X_s)\big]\, ds.$$

Let us denote by $M_t^\varepsilon(x)$ the integral in the right-hand side. Note that the left-hand side may not be a measurable function of x, but $M_t^\varepsilon(x)$ is, because $(\mathbb{P}_x)_{x \in \mathbb{R}^d}$ is a regular conditional probability of the law of X given ξ. Choosing ϕ as in the proof of Theorem 1.1 (see the Appendix),

$$\int_{\mathbb{R}^d} M_t^\varepsilon(x) u_0(x)\, ds = \int_0^t \int_{\mathbb{R}^d} ((M|\nabla\sigma|)^2 + |F| + M_{1/\varepsilon}\nabla F)(s,x) u(s,dx)\, ds$$
$$\leq C\left(1 + \frac{|\log\varepsilon|}{\varepsilon\phi(\varepsilon^{-1})}\right).$$

Now, copying the proof of Lemma 4.2,

$$\mathbb{E}(|X_t^x - \hat{X}_t^x|) \leq C\left[\sqrt{\varepsilon} + \frac{1}{L} + \frac{LM_t^\varepsilon(x)}{|\log\varepsilon|}\right].$$

Let us denote by $N_t^\varepsilon(x)$ the r.h.s. Choosing $L = (\frac{1}{|\log\varepsilon|} + \tilde{\eta}(\varepsilon))^{-1}$ with $\eta(\varepsilon) = (\varepsilon\phi(\varepsilon^{-1}))^{-1}$, we obtain

$$\int_{\mathbb{R}^d} N_t^\varepsilon(x) u_0(x)\, ds \leq C\left(\sqrt{\varepsilon} + \sqrt{\frac{1}{|\log\varepsilon|} + \tilde{\eta}(\varepsilon)}\right).$$

Since the right-hand side converges to 0 when $\varepsilon \to 0$, there exists a sequence $\varepsilon_k \to 0$ such that $N_t^{\varepsilon_k}(x) \to 0$ for almost all x. The diagonal procedure then shows the existence of a subsequence $\varepsilon_k' \to 0$ such that $N_t^{\varepsilon_k'}(x) \to 0$ for almost all x and for all t in a dense countable subset of $[0, T]$. Since the paths of X^x and $\hat{X}^x$ are continuous, we deduce that pathwise uniqueness holds for almost all $x \in \mathbb{R}^d$.

APPENDIX: SKETCH OF THE PROOF OF THEOREM 1.1

The only thing left to prove after Theorem 2.13 is: Assume $u \in L_{t,\mathrm{loc}}^{q'}(L_x^{p'}(\mathbb{R}^d))$ then show that, for some superlinear ϕ,

$$\|\sigma\|_{H_T^1(u)} \leq C\|\sigma\|_{L_t^{2q}([0,T], W_x^{1,2p})}, \qquad \|F\|_{W_T^{\phi,\mathrm{weak}}(u)} \leq C\|F\|_{L_t^q([0,T], W_x^{1,p})}.$$

From the fact that the maximal operator M is bounded on L^p, $p > 1$, this is straightforward for σ (as $2p \geq 2 > 1$).

Therefore, the key point is how to prove that for F when $p \geq 1$. We give the proof for $p = 1$, the case $p > 1$ can be treated following the same lines.

Now fix $L \geq 1$ and denote

$$h(t,x) = M_L\nabla F = \sqrt{\log L} + \int_{\mathbb{R}^d} \frac{|\nabla F(t,z)|\mathbb{1}_{|\nabla F| \geq \sqrt{\log L}}\, dz}{(L^{-1} + |x-z|)|x-z|^{d-1}}.$$

As $p' = \infty$, for almost any fixed t, $u(t, \cdot) \in L^{q'} \cap L^\infty$, and hence

$$\int h(t, x) u(t, x) \, dx$$

$$\leq \sqrt{\log L}$$

$$+ \max\left(1, \|u(t, \cdot)\|_{L^\infty}\right) \iint \min(1, u(t, x)) \frac{|\nabla F(t, z)| \mathbb{1}_{|\nabla F| \geq \sqrt{\log L}} \, dz}{(L^{-1} + |x - z|)|x - z|^{d-1}} \, dx$$

$$\leq \sqrt{\log L} + C \log L \left(\|u(t, \cdot)\|_{L^\infty} + \|u(t, \cdot)\|_{L^1}\right) \|\nabla F(t, \cdot) \mathbb{1}_{|\nabla F| \geq \sqrt{\log L}}\|_{L^1},$$

by Fubini's theorem. Note that the term $\min(1, u(t, x))$ was kept in the integral because the function $x \mapsto (L^{-1} + |x - z|)^{-1}|x - z|^{-(d-1)}$ is not integrable on $\{|x| > 1\}$.

Therefore, integrating now in time, by Hölder's estimates

$$\int_0^T \int h(t, x) u(t, x) \, dx \, dt \leq \sqrt{\log L} T + C \log L \|\nabla F \mathbb{1}_{|\nabla F| \geq \sqrt{\log L}}\|_{L_t^q(L_x^1)}.$$

Now, if $\nabla F \in L_t^q([0, T], L_x^1)$, then de la Vallée Poussin classical integrability result means that there exists a superlinear ψ s.t.

$$\|\psi(\nabla F)\|_{L_t^q([0,T], L_x^1)} < \infty.$$

Consequently,

$$\int_0^T \int h(t, x) u(t, x) \, dx \, dt \leq T \sqrt{\log L} + C \frac{(\log L)^{3/2}}{\psi(\sqrt{\log L})}.$$

We conclude that $\|\nabla F\|_{W_T^{\phi, \text{weak}}(u)}$ is bounded for ϕ defined by

$$\frac{L}{\phi(L)} = \frac{C\sqrt{\log L}}{\log L} + \frac{C\sqrt{\log L}}{\psi(\sqrt{\log L})},$$

which is hence also superlinear.

REFERENCES

[1] ALEKSANDROV, A. D. (1963). Uniqueness conditions and bounds for the solution of the Dirichlet problem. *Vestn. Leningr. Univ., Mat. Meh. Astron.* **18** 5–29. MR0164135

[2] AMBROSIO, L. (2004). Transport equation and Cauchy problem for *BV* vector fields. *Invent. Math.* **158** 227–260. MR2096794

[3] AMBROSIO, L., FUSCO, N. and PALLARA, D. (2000). *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford Univ. Press, New York. MR1857292

[4] CHAMPAGNAT, N. and JABIN, P.-E. (2010). Well posedness in any dimension for Hamiltonian flows with non BV force terms. *Comm. Partial Differential Equations* **35** 786–816. MR2753620

[5] CRIPPA, G. and DE LELLIS, C. (2008). Estimates and regularity results for the DiPerna–Lions flow. *J. Reine Angew. Math.* **616** 15–46. MR2369485

[6] DiPERNA, R. J. and LIONS, P.-L. (1989). Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.* **98** 511–547. MR1022305

[7] ENGELBERT, H. J. and KURENOK, V. P. (2000). On multidimensional SDEs without drift and with a time-dependent diffusion matrix. *Georgian Math. J.* **7** 643–664. MR1811920

[8] ENGELBERT, H. J. and SCHMIDT, W. (1989). Strong Markov continuous local martingales and solutions of one-dimensional stochastic differential equations. I. *Math. Nachr.* **143** 167–184. MR1018242

[9] ENGELBERT, H. J. and SCHMIDT, W. (1989). Strong Markov continuous local martingales and solutions of one-dimensional stochastic differential equations. II. *Math. Nachr.* **144** 241–281. MR1037172

[10] ENGELBERT, H. J. and SCHMIDT, W. (1991). Strong Markov continuous local martingales and solutions of one-dimensional stochastic differential equations. III. *Math. Nachr.* **151** 149–197. MR1121203

[11] FEDRIZZI, E. and FLANDOLI, F. (2011). Pathwise uniqueness and continuous dependence of SDEs with non-regular drift. *Stochastics* **83** 241–257. MR2810591

[12] FEDRIZZI, E. and FLANDOLI, F. (2013). Hölder flow and differentiability for SDEs with non-regular drift. *Stoch. Anal. Appl.* **31** 708–736. MR3175794

[13] FIGALLI, A. (2008). Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients. *J. Funct. Anal.* **254** 109–153. MR2375067

[14] FLANDOLI, F., GUBINELLI, M. and PRIOLA, E. (2010). Well-posedness of the transport equation by stochastic perturbation. *Invent. Math.* **180** 1–53. MR2593276

[15] ITO, K. (1951). *On Stochastic Differential Equations*. *Mem. Amer. Math. Soc.* **4**. Amer. Math. Soc., New York. MR0040618

[16] JABIN, P.-E. (2010). Differential equations with singular fields. *J. Math. Pures Appl.* (9) **94** 597–621. MR2737390

[17] KRYLOV, N. V. (1971). A certain estimate from the theory of stochastic integrals. *Teor. Verojatnost. i Primenen.* **16** 446–457. MR0298792

[18] KRYLOV, N. V. (1974). Some estimates of the probability density of a stochastic integral. *Math. USSR, Izv.* **8** 233–254.

[19] KRYLOV, N. V. (1980). *Controlled Diffusion Processes. Applications of Mathematics* **14**. Springer, New York. Translated from the Russian by A. B. Aries. MR0601776

[20] KRYLOV, N. V. (2001). The heat equation in $L_q((0, T), L_p)$-spaces with weights. *SIAM J. Math. Anal.* **32** 1117–1141. MR1828321

[21] KRYLOV, N. V. and RÖCKNER, M. (2005). Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields* **131** 154–196. MR2117951

[22] KURENOK, V. P. and LEPEYEV, A. N. (2008). On multi-dimensional SDEs with locally integrable coefficients. *Rocky Mountain J. Math.* **38** 139–174. MR2397030

[23] LE BRIS, C. and LIONS, P.-L. (2004). Renormalized solutions of some transport equations with partially $W^{1,1}$ velocities and applications. *Ann. Mat. Pura Appl.* (4) **183** 97–130. MR2044334

[24] LE BRIS, C. and LIONS, P.-L. (2008). Existence and uniqueness of solutions to Fokker–Planck type equations with irregular coefficients. *Comm. Partial Differential Equations* **33** 1272–1317. MR2450159

[25] LE BRIS, C. and LIONS, P.-L. (2012). Parabolic partial differential equations with irregular data. Related issues. Applications to stochastic differential equations. Preprint, based on the lectures of P.-L. Lions at Collège de France. Video available at http://www.college-de-france.fr/site/pierre-louis-lions/.

[26] LE GALL, J.-F. (1983). Applications du temps local aux équations différentielles stochastiques unidimensionnelles. In *Seminar on Probability, XVII. Lecture Notes in Math.* **986** 15–31. Springer, Berlin. MR0770393

[27] MENOUKEU-PAMEN, O., MEYER-BRANDIS, T., NILSSEN, T., PROSKE, F. and ZHANG, T. (2013). A variational approach to the construction and Malliavin differentiability of strong solutions of SDE's. *Math. Ann.* **357** 761–799. MR3096525

[28] RÖCKNER, M. and ZHANG, X. (2010). Weak uniqueness of Fokker–Planck equations with degenerate and bounded coefficients. *C. R. Math. Acad. Sci. Paris* **348** 435–438. MR2607035

[29] ROZKOSZ, A. and SŁOMIŃSKI, L. (1991). On existence and stability of weak solutions of multidimensional stochastic differential equations with measurable coefficients. *Stochastic Process. Appl.* **37** 187–197. MR1102869

[30] STEIN, E. M. (1970). *Singular Integrals and Differentiability Properties of Functions. Princeton Mathematical Series* **30**. Princeton Univ. Press, Princeton, NJ. MR0290095

[31] STROOCK, D. W. and VARADHAN, S. R. S. (1979). *Multidimensional Diffusion Processes. Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **233**. Springer, Berlin. MR0532498

[32] VERETENNIKOV, A. Y. (1981). On strong solutions and explicit formulas for solutions of stochastic integral equations. *Math. USSR, Sb.* **39** 387–403.

[33] VERETENNIKOV, A. Y. and KRYLOV, N. V. (1976). On explicit formulas for solutions of stochastic equations. *Math. USSR, Sb.* **29** 239–256.

[34] WATANABE, S. and YAMADA, T. (1971). On the uniqueness of solutions of stochastic differential equations. II. *J. Math. Kyoto Univ.* **11** 553–563. MR0288876

[35] YAMADA, T. and WATANABE, S. (1971). On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.* **11** 155–167. MR0278420

[36] ZHANG, X. (2005). Strong solutions of SDES with singular drift and Sobolev diffusion coefficients. *Stochastic Process. Appl.* **115** 1805–1818. MR2172887

[37] ZHANG, X. (2010). Stochastic flows of SDEs with irregular coefficients and stochastic transport equations. *Bull. Sci. Math.* **134** 340–378. MR2651896

[38] ZHANG, X. (2011). Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients. *Electron. J. Probab.* **16** 1096–1116. MR2820071

[39] ZHANG, X. (2013). Well-posedness and large deviation for degenerate SDEs with Sobolev coefficients. *Rev. Mat. Iberoam.* **29** 25–52. MR3010120

[40] ZVONKIN, A. K. (1974). A transformation of the phase space of a diffusion process that will remove the drift. *Mat. Sb.* (*N.S.*) **93(135)** 129–149, 152. MR0336813

INSTITUT ELIE CARTAN DE LORRAINE
  UMR CNRS 7502
UNIVERSITÉ DE LORRAINE, SITE DE NANCY
B.P. 70239
F-54506 VANDOEUVRE-LES-NANCY CEDEX
FRANCE
AND
INRIA NANCY-GRAND EST
TEAM TOSCA
F-54600 VILLERS-LÈS-NANCY
FRANCE
E-MAIL: nicolas.champagnat@inria.fr

CSCAMM AND
  DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND 20742
USA
E-MAIL: pjabin@umd.edu