The Annals of Probability
2021, Vol. 49, No. 6, 3142–3167
https://doi.org/10.1214/21-AOP1525

ON STRONG SOLUTIONS OF ITÔ'S EQUATIONS WITH $\sigma \in W^1_{\mathrm{d}}$ AND $\mathrm{b} \in L_{\mathrm{d}}$

BY N. V. KRYLOV

School of Mathematics, University of Minnesota, nkrylov@umn.edu

We consider Itô uniformly nondegenerate equations with time independent coefficients, the diffusion coefficient in $W^1_{d,\mathrm{loc}}$ and the drift in L_d. We prove the unique strong solvability for any starting point and prove that, as a function of the starting point, the solutions are Hölder continuous with any exponent < 1. We also prove that if we are given a sequence of coefficients converging in an appropriate sense to the original ones, then the solutions of approximating equations converge to the solution of the original one.

1. Introduction. Let $\mathbb{R}^d$ be a d-dimensional Euclidean space of points $x = (x^1, \dots, x^d)$ with $d \geq 3$. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space; let $\{\mathcal{F}_t\}$ be an increasing filtration of σ-fields $\mathcal{F}_t \subset \mathcal{F}$ that are complete. Let w_t be a d_1-dimensional Wiener process relative to $\{\mathcal{F}_t\}$, where $d_1 \geq d$.

Assume that, on $\mathbb{R}^d$, we are given $\mathbb{R}^d$-valued Borel functions $b, \sigma^k = (\sigma^{ik})$, $k = 1, \dots, d_1$. We are going to fix $x_0 \in \mathbb{R}^d$ and investigate the equation

$$x_t = x_0 + \int_0^t \sigma^k(x_s)\, dw^k_s + \int_0^t b(x_s)\, ds, \tag{1.1}$$

where and everywhere below the summation over repeated indices is understood.

We are interested in the so-called strong solutions, that is, solutions such that, for each $t \geq 0$, x_t is $\mathcal{F}^w_t$-measurable, where $\mathcal{F}^w_t$ is the completion of $\sigma(w_s : s \leq t)$. We present sufficient conditions for the equation to have a strong solution and also for the solution to be unique (strong uniqueness). A very rich literature on the weak uniqueness problem for (1.1) is beyond the scope of this article. For the author, who is a probabilist in the first place, the interest in stochastic equations with irregular coefficients first appeared in the study of time-homogeneous controlled diffusion processes, where, as a rule, optimal controls are achieved on diffusions with discontinuous coefficients. One of the main questions is then, "How irregular could they be to still be admissible controls?"

After the classical work by Itô showing that there exists a unique strong solution of (1.1) if σ^k and b are Lipschitz continuous (may also depend on time and ω), many efforts were applied to relax these Lipschitz conditions. In case $d = d_1 = 1$, T. Yamada and S. Watanabe [27] relaxed the Lipschitz condition on σ to the Hölder (1/2)-condition (and even slightly weaker condition) and kept b Lipschitz (slightly less restrictive). Much attention was paid to equations with continuous coefficients satisfying the so-called monotonicity conditions (see, for instance, [8] and the references therein).

T. Yamada and S. Watanabe [27] also put forward a very strong theorem, basically saying that strong uniqueness implies the existence of strong solutions. Unlike the present paper, the majority of papers on the subject after that time are using their theorem. S. Nakao ([18]) proved the strong solvability in time homogeneous case if $d = d_1 = 1$ and σ is bounded

Received July 2020; revised March 2021.
MSC2020 subject classifications. Primary 60H10; secondary 60J60.
Key words and phrases. Strong solutions, vanishing mean oscillation, singular coefficients, martingale problem.

away from zero and infinity and is locally of bounded variation. He also assumed that b is bounded, but from his arguments it is clear that the summability of $|b|$ suffices. In this respect his result shows that our results are also true if $d = 1$. However, the general case that $d = 2$ is quite open. By the way, the restriction $d \geq 3$ appears because of embedding theorems, as is explained before Theorem 4.4.

A. Veretennikov seems to be the first who in [24] not only proved the existence of strong solutions in the time inhomogeneous multidimensional case when b is bounded but also considered the case of σ^k in Sobolev class, namely, $\sigma^k_x \in L_{2d,\mathrm{loc}}$. He used A. Zvonkin's method (see [29]) of transforming the equation in such a way that the drift term disappears. X. Zhang in [28] considered time inhomogeneous equations under some conditions which, for the time homogeneous case (our case), become $\sigma^k_x, b \in L_p$ with $p > d$. For more detailed information on the time inhomogeneous case, we refer the reader to [26, 28] and the references therein.

Even the case when the σ^k's are constant and the process is nondegenerate attracted much attention, especially in the time inhomogeneous setting. We discuss some of the results in the particular case of b independent of t. M. Röckner and the author in [15] proved, among other things, the existence of strong solutions when $b \in L_p$ with $p > d$. If b is bounded, A. Shaposhnikov ([21, 22]) proved the so-called path-by-path uniqueness, which, basically, means that, for almost any trajectory w_t, there is only one solution (adapted or not). This result was already announced by A. Davie before with a very entangled proof which left many doubtful.

In a fundamental work by L. Beck, F. Flandoli, M. Gubinelli, and M. Maurelli ([2]), the authors investigate such equations from the points of view of Itô stochastic equations, stochastic transport equations and stochastic continuity equations. Their article contains an enormous amount of information and a vast references list. We compare only those of their results which have counterparts in the present article. In what concerns our situation, they require (σ^k constant and the process is nondegenerate) $b \in L_{p,\mathrm{loc}}$ with $p > d$ or $p = d$ but $\|b\|_{L_p}$ to be sufficiently small, and they prove strong solvability and strong uniqueness (actually, path-by-path-uniqueness which is stronger) only for almost all starting points x. We assume that $\sigma^k_x, b \in L_d$ and for uniformly nondegenerate and bounded σ^k prove that, for any x, equation (1.1) has a unique strong solution.

Our approach is absolutely different from all articles mentioned above and all articles which one can find in their references. We do not use Yamada–Watanabe theorem or transformations of the noise. Instead, our method is based on an analytic criterion for the existence of strong solutions which first appeared in [25], some 45 years ago, and never used since then. To make this method work, we use results of many authors, most relevant of which are the results in [4, 11–14].

Simple examples of equations for which we prove the existence of unique strong solutions are

$$dx_t = \big(2 + I_{x_t \neq 0}\zeta(x_t)\sin\big(\ln|\ln|x_t|\,|\big)\big)\,dw_t, \qquad dx_t = dw_t + \zeta(x_t)\big(|x_t|\ln|x_t|\big)^{-1} l\,dt,$$

where ζ is any smooth $d \times d$-matrix valued function vanishing for $|x| > 1/2$ satisfying $|\zeta| \leq 1$ and l is any vector in $\mathbb{R}^d$. Observe that, in the first equation, the diffusion coefficient is discontinuous at the origin. The reader can find more examples of existence and *nonexistence* of strong solution in Remark 5.10.

We conclude the Introduction by some notation. We set $u_x = Du$ to be the gradient of u, u_{xx} to be the matrix of its second-order derivatives,

$$D_{x^i}u = D_i u = u_{x^i} = \frac{\partial}{\partial x^i}u, \qquad u_{x^i\eta^j} = D_{x^i\eta^j}u = D_{x^i}D_{\eta^j}u,$$

$$\partial_t u = \frac{\partial}{\partial t}u \qquad u_{(\xi)} = \xi^i u_{x^i}.$$

If $\sigma(x) = (\sigma^i(x))$ is vector-valued (column-vector), by σ_x we mean the matrix whose ijth element is $\sigma^i_{x^j}$. If c is a matrix (in particular, vector), we set $|c|^2 = \operatorname{tr} cc^*$ ($= \operatorname{tr} c\bar{c}^*$ if c is complex-valued).

For $p \in [1, \infty)$ by L_p, we mean the space of Borel (perhaps complex vector- or matrix-valued) functions on $\mathbb{R}^d$ with finite norm given by

$$\|f\|^p_{L_p} = \int_{\mathbb{R}^d} |f(x)|^p \, dx.$$

By W^2_p we mean the space of Borel functions u on $\mathbb{R}^d$ whose Sobolev derivatives u_x and u_{xx} exist and $u, u_x, u_{xx} \in L_p$. The norm in W^2_p is given by

$$\|u\|_{W^2_p} = \|u_{xx}\|_{L_p} + \|u\|_{L_p}.$$

Similarly W^1_p is defined. As usual, we write $f \in L_{p,\mathrm{loc}}$ if $f\zeta \in L_p$ for any $\zeta \in C^\infty_0$ ($= C^\infty_0(\mathbb{R}^d)$). Similarly $W^{\cdot}_{\cdot,\mathrm{loc}}$ are defined.

If a Borel $\Gamma \subset \mathbb{R}^d$, by $|\Gamma|$ we mean its Lebesgue measure. Finally,

$$B_R(x) = \{y \in \mathbb{R}^d : |x - y| < R\}, \qquad B_R = B_R(0).$$

2. Main results. Set $a^{ij} = \sigma^{ik}\sigma^{jk}$, $a = (a^{ij})$. Fix numbers $\delta \in (0, 1)$ and $\|b\|, \|\sigma^k_x\| \in (0, \infty)$.

ASSUMPTION 2.1. We have

$$\delta^{-1}|\lambda|^2 \ge a^{ij}(x)\lambda^i\lambda^j \ge \delta|\lambda|^2 \tag{2.1}$$

for all $\lambda, x \in \mathbb{R}^d$. Also,

$$\|b\|_{L_d} \le \|b\|.$$

ASSUMPTION 2.2. For any k we have $\sigma^k \in W^1_{d,\mathrm{loc}}$ and

$$\|\sigma^k_x\|_{L_d} \le \|\sigma^k_x\|.$$

Recall that $d \ge 3$.

THEOREM 2.3. *Under the above assumptions for any $x_0 \in \mathbb{R}^d$, equation* (1.1) *has a strong solution x_t. If y_t is also a solution of* (1.1), *then with probability one $x_t = y_t$ for all t.*

THEOREM 2.4. *Under the above assumptions, suppose that we are also given sequences $\sigma^k(n)$, $b(n)$, $n = 1, 2\dots$, $k = 1, \dots, d_1$ of functions having the same meaning as σ^k, b and satisfying Assumptions 2.1 and 2.2 with the same δ, $\|b\|$ and $\|\sigma^k_x\|$. Assume that $b(n) \to b$ and $\sigma^k_x(n) \to \sigma^k_x$ in L_d as $n \to \infty$, and we are given a sequence $x(n) \to x_0$. Finally, let $\sigma^k(n) \to \sigma^k$ (a.e.) as $n \to \infty$. Then, for any $m, T \in (0, \infty)$,*

$$\lim_{n\to\infty} E \sup_{t \le T} |x_t(n, x(n)) - x_t|^m = 0,$$

where $x_t(n, x(n))$ are the solutions of (1.1) *in which x_0, σ^k and b are replaced by $x(n)$, $\sigma^k(n)$ and $b(n)$, respectively.*

THEOREM 2.5. *Under the above assumptions there is a function $x_t(x) = x_t(\omega, x)$ which is a solution of* (1.1) *with x in place of x_0 and for each $\alpha < 1$ and ω is α-Hölder continuous with respect to x and $(\alpha/2)$-Hölder continuous with respect to t on each set $[0, T] \times \bar{B}_R$, $T, R \in (0, \infty)$.*

REMARK 2.6. The main emphasis of the article is to treat the case that $b \in L_d$. It is known (see, for instance, [3] [11]) that, even if $d_1 = d$ and (σ^k) is a unit matrix, there are cases when $b \in L_{d-\varepsilon}$ for any $\varepsilon \in (0, 1)$ but not for $\varepsilon = 0$, and there are no solutions of (1.1).

However, our results are new also if b is bounded or $b \equiv 0$. In that case the arguments are not so technically involved and allow any $d \geq 1$ rather than $d \geq 3$. In Remark 5.10 we show an example with $b \equiv 0$ and $\sigma_x^k \in L_{d-\varepsilon}$ for any $\varepsilon \in (0, 1)$ but not for $\varepsilon = 0$, when there are no strong solutions. In this regard, Assumption 2.2 seems to be optimal.

The rest of the article is organized as follows. As we mentioned above, our main tool is an analytic criterion for the existence of strong solutions. To derive it, we develop necessary facts from the theory of semigroups generated by elliptic operators in Section 3. Then, in Section 4 we relate the semigroup from Section 3 to the semigroup of the corresponding Markov diffusion process. In Section 5 we derive our analytic criterion. Section 6 is devoted to some estimates of the series involved in the criterion when σ^k and b are smooth. In Sections 7, 8 and 9 we prove Theorems 2.3, 2.4 and 2.5, respectively.

3. An analytic semigroup. In this section, Assumption 2.1 is supposed to be satisfied, but Assumption 2.2 is replaced with a weaker Assumption 3.5 which comes after some discussion.

Introduce the uniformly elliptic operators

$$Lu(x) = (1/2)a^{ij}(x)u_{x^i x^j}(x) + b^i(x)u_{x^i}(x),$$
$$L_0 u(x) = (1/2)a^{ij}(x)u_{x^i x^j}(x)$$

acting on functions given on $\mathbb{R}^d$.

Denote

$$\operatorname{osc}(a, B_\rho(x)) = |B_\rho|^{-2} \int_{y,z \in B_\rho(x)} |a(y) - a(z)|\, dy\, dz,$$
$$a_r^{\#} = \sup_{x \in \mathbb{R}^d} \sup_{\rho < r} \operatorname{osc}(a, B_\rho(x)).$$

Here is a consequence of Theorem 2 of [4]. We are dealing with complex-valued functions and denote $\mathbb{R}^{d+1} = \{(x^0, x^1, \dots, x^d) : x^k \in \mathbb{R}\}$.

LEMMA 3.1. *For any $p \in (1, \infty)$ and $\varepsilon \in (0, 1]$, there exists $\theta_0 = \theta_0(d, \delta, \varepsilon, p) > 0$ such that, if there is $r_0 > 0$ for which $a_{r_0}^{\#} \leq \theta_0$, then there exist $\lambda_0 \geq 1$, N_0, depending only on d, δ, ε, p, r_0, such that, for any $u \in W_p^2$ and $\lambda \geq \lambda_0$,*

$$\sum_{r,s=0}^{d} \|D_{rs} u\|_{L_p(\mathbb{R}^{d+1})} + \lambda \|u\|_{L_p(\mathbb{R}^{d+1})} \leq N_0 \| L_0 u \pm \varepsilon i D_0^2 u - (1 \pm \varepsilon i)\lambda u \|_{L_p(\mathbb{R}^{d+1})}. \tag{3.1}$$

PROOF. As is easy to see, Theorem 1 of [4] is applicable to the operator $Mu := (1 \pm \varepsilon i)^{-1}(L_0 u \pm \varepsilon i D_0^2 u)$, and it yields an estimate for

$$u \in W_p^{1,2}((-\infty, 0) \times \mathbb{R}^{d+1}) = \{u \in L_p((-\infty, 0) \times \mathbb{R}^{d+1}) : \partial_t u, u_x, u_{xx} \in L_p((-\infty, 0) \times \mathbb{R}^{d+1})\}$$

similar to (3.1), where $\mathbb{R}^{d+1}$ is replaced with $(-\infty, 0) \times \mathbb{R}^{d+1}$ and $L_0 u \pm \varepsilon i D_0^2 u - (1 \pm \varepsilon i)\lambda u$ is replaced with $Mu - \partial u/\partial t - \lambda u$. By substituting in this estimate $u(x)e^t$, we get (3.1), and the lemma is proved. □

REMARK 3.2. Without introducing the new coordinate, Theorem 1 of [4] implies that in Lemma 3.1 one can replace (3.1) with

$$\sum_{r,s=1}^{d} \|D_{rs}u\|_{L_p} + \lambda\|u\|_{L_p} \le N_0\|L_0u - \lambda u\|_{L_p} \tag{3.2}$$

valid for any $u \in W_p^2$ and $\lambda \ge \lambda_0(d, \delta, p, r_0)$ with $N_0 = N_0(d, \delta, p, r_0)$ as long as $a_{r_0}^{\#} \le \theta_0(d, \delta, p)$. Since $\|L_0u - \lambda u\|_{L_p} \le \|L_0u - \mu u\|_{L_p} + |\lambda - \mu|\|u\|_{L_p}$, estimate (3.2) easily implies that, in the same situation,

$$\sum_{r,s=1}^{d} \|D_{rs}u\|_{L_p} + |\mu|\|u\|_{L_p} \le 2N_0\|L_0u - \mu u\|_{L_p} \tag{3.3}$$

as long as $|\mu| \ge \lambda_0$ and $|\Im\mu| \le 2\varepsilon_0\Re\mu$, where $\varepsilon_0 = \varepsilon_0(d, \delta, p, r_0) > 0$.

LEMMA 3.3. *For any $p \in (1, \infty)$, there exists $\theta_0 = \theta_0(d, \delta, p)$ such that, if there is $r_0 > 0$ for which $a_{r_0}^{\#} \le \theta_0$, then there exist $\lambda_0 \ge 1$, N_0, depending only on d, δ, p, r_0, such that, for any $u \in W_p^2$ and complex λ such that $\Re\lambda \ge \lambda_0$,*

$$\sum_{r,s=1}^{d} \|D_{rs}u\|_{L_p} + |\lambda|\|u\|_{L_p} \le N_0\|L_0u - \lambda u\|_{L_p}. \tag{3.4}$$

PROOF. We use an idea from [1]. Take a nonnegative $\zeta \in C_0^\infty(\mathbb{R})$ such that ζ^p has unit integral, $u \in W_p^2$, and plug into (3.1) the function $u(x)e^{i\mu x_0}\zeta(x_0)$ and $\varepsilon = \varepsilon_0$. Then, we get, for $\lambda \ge \lambda_0$ and $\mu \in \mathbb{R}$, that

$$\begin{aligned} &\sum_{r,s=1}^{d} \|D_{rs}u\|_{L_p} + (\lambda + \mu^2)\|u\|_{L_p} - N(1 + |\mu|)\|u\|_{L_p} \\ &\quad \le N\|L_0u - [(1 \pm \varepsilon_0 i)\lambda \pm \varepsilon_0 i\mu^2]u\|_{L_p} + N(1 + |\mu|)\|u\|_{L_p}. \end{aligned} \tag{3.5}$$

Now, take $\hat\lambda$ such that $\Re\hat\lambda \ge \lambda_0$. If $|\Im\hat\lambda| \le 2\varepsilon_0\Re\hat\lambda$, we have (3.4) with $\hat\lambda$ in place of λ thanks to (3.3).

If $\Im\hat\lambda \ge 2\varepsilon_0\Re\hat\lambda$, set $\lambda = \Re\hat\lambda$, $\varepsilon_0\mu^2 = \Im\hat\lambda - \varepsilon_0\lambda$. Then,

$$|\hat\lambda|^2 \le ((2\varepsilon_0)^{-2} + 1)(\Im\hat\lambda)^2, \qquad \mu^2 \le \varepsilon_0^{-1}\Im\hat\lambda \le \varepsilon_0^{-1}|\hat\lambda|, \qquad \lambda + \mu^2 = \varepsilon_0^{-1}\Im\hat\lambda$$

and (3.5) with upper signs yields

$$\sum_{r,s=1}^{d} \|D_{rs}u\|_{L_p} + |\hat\lambda|\|u\|_{L_p} \le N\|L_0u - \hat\lambda u\|_{L_p} + N(1 + |\hat\lambda|^{1/2})\|u\|_{L_p}.$$

By increasing λ_0, we absorb the last term on the right into the left-hand side for $\Re\hat\lambda \ge \lambda_0$, and we come to (3.4) with $\hat\lambda$ in place of λ if $\Im\hat\lambda \ge 0$. The case of $\Im\hat\lambda \le 0$ is treated by using (3.5) with lower signs. The lemma is proved. □

The argument in the second part of Remark 3.2 also allows us to deduce from Lemma 3.3 the following.

LEMMA 3.4. *Lemma 3.3 holds true if we replace the restriction $\Re\lambda \ge \lambda_0$ in it with $\lambda \in \Gamma$, where $\Gamma = \{\Re\lambda \ge \lambda_0\} \cup \{\varepsilon_0|\Im\lambda| \ge -\Re\lambda + \mu_0\}$, with $\varepsilon_0 > 0$ and $\mu_0 > \lambda_0$ which depend only on d, δ, p, r_0.*

In the rest of the section, we impose the following.

ASSUMPTION 3.5 (p, r_0). We have $a^{\#}_{r_0} \le \theta_0(d,\delta,p)$, where θ_0 is taken in a way to accommodate Lemmas 3.3 and 3.4.

REMARK 3.6. It is well known that if $a_x \in L_d$, then $a^{\#}_{r_0} \to 0$ as $r_0 \downarrow 0$ (see, for instance, Theorem 10.2.5 of [9] or Exercise 10.2.8 and the hint to it in the same book). Therefore, Assumption 3.5 is weaker than Assumption 2.2.

On the basis of Lemma 3.4, we can repeat what was done in [14] and obtain the first part of the following result about the full operator L.

THEOREM 3.7. *Let $p \in (1,d)$. Then, under Assumptions 2.1 and 3.5 there exist $\lambda_0 \ge 1$, N_0, depending only on d, δ, p, r_0 and ν_b (introduced below), such that, for any $u \in W^2_p$ and $\lambda \in \Gamma$,*

$$\sum_{r,s=1}^{d} \|D_{rs}u\|_{L_p} + |\lambda|\|u\|_{L_p} \le N\|Lu - \lambda u\|_{L_p}, \tag{3.6}$$

where ν_b is defined by the condition

$$N_1\|bI_{|b|\ge\nu_b}\|_{L_d} \le 1$$

with a constant $N_1 = N_1(d,\delta,r_0,p)$. Furthermore, for any $\lambda \in \Gamma$ and $f \in L_p$, there is a unique $u \in W^2_p$ such that $\lambda u - Lu = f$.

The "existence" part of this theorem, as usual, is proved by the method of continuity.

Denote by $R_\lambda f$ the solution u from Theorem 3.7. Then, the fact that the norm of R_λ, as an operator in L_p, decreases as $N/|\lambda|$ for $\lambda \in \Gamma$ allows us to use the well-known construction introduced by Hille ([5]). We use the following facts which the reader can find, for instance, in [19]. For complex t in the sector $S := \{|\Im t| < \varepsilon_0 \Re t\}$ with ε_0 from Lemma 3.4, set

$$\hat{T}_t = \frac{1}{2\pi i}\int_{\partial\Gamma} e^{tz} R_z\,dz, \tag{3.7}$$

where the integral is taken in a counterclockwise direction. Below in this section

$$p \in (1,d).$$

THEOREM 3.8.

(i) *Formula (3.7) defines $\hat{T}_t$ in S as an analytic semigroup of bounded operators in L_p with norms bounded by a constant, depending only on ε, d, δ, p, r_0 and ν_b as long as $t \in \{|\Im t| \le \varepsilon \Re t, |t| \le (\varepsilon_0 - \varepsilon)^{-1}\}$ for any given $\varepsilon < \varepsilon_0$;*

(ii) *The infinitesimal generator of this semigroup is L with domain W^2_p;*

(iii) *For $g \in W^2_p$, the function $\hat{T}_t g(x)$ is a unique solution of the problem*

$$\partial_t u(t,x) = Lu(t,x), \quad t > 0, \qquad \lim_{t\downarrow 0}\|u(t,\cdot) - g\|_{L_p} = 0$$

in the class of u such that $u(t,\cdot) \in W^2_p$ and (strong L_p-derivative) $\partial_t u(t,\cdot) \in L_p$ for each $t > 0$;

(iv) *For any $T \in (0,\infty)$, there is a constant N, depending only on T, d, δ, p, r_0 and ν_b such that, for each $t \in (0,T]$ and $f \in L_p$,*

$$\|\hat{T}_t f\|_{W^2_p} \le \frac{N}{t}\|f\|_{L_p}, \qquad \|D\hat{T}_t f\|_{L_p} \le \frac{N}{\sqrt{t}}\|f\|_{L_p}. \tag{3.8}$$

Actually, the second estimate in (3.8) is not to be found explicitly in [19], but it follows by interpolation from the first one and the fact that $\|\hat{T}_t f\|_{L_p} \le N\|f\|_{L_p}$.

We will also need a stability result before which we make the following.

REMARK 3.9. Let $d > p > d/2$ and $f \in W^2_p$. Then, $\hat{T}_t f \in W^2_p$, and, by embedding theorems, $\hat{T}_t f$ has a modification that is bounded and continuous in x, which we still call $\hat{T}_t f$. Also, $(\lambda - L)\hat{T}_t f = \hat{T}_t(\lambda - L)f \to (\lambda - L)\hat{T}_s f$ in L_p as $t \to s$. Hence, $\hat{T}_t f \to \hat{T}_s f$ in W^2_p which, by embedding theorems, implies that $\hat{T}_t f(x) \to \hat{T}_s f(x)$ uniformly on $\mathbb{R}^d$. Therefore, $\hat{T}_t f(x)$ is a bounded continuous function on $[0, T] \times \mathbb{R}^d$ for any $T \in (0, \infty)$.

Moreover, by embedding theorems, if $u \in W^2_p$, then, for any $x \in \mathbb{R}^d$,

$$|u(x)| \le N(\|u_{xx}\|_{L_p} + \|u\|_{L_p}),$$

where $N = N(d, p)$. By substituting here $u(cx)$ in place of $u(x)$ and taking minimum of the right-hand side with respect to $c > 0$, we come to the well-known estimate

$$|u(x)| \le N\|u_{xx}\|^{d/(2p)}_{L_p}\|u\|^{1-/(2p)}_{L_p}.$$

Applying this and (3.8) yields that, for $t \le T$ and any $x \in \mathbb{R}^d$,

$$|\hat{T}_t f(x)| \le \frac{N}{t^{d/(2p)}}\|f\|_{L_p}, \tag{3.9}$$

where N depends only on T, d, δ, p, r_0 and ν_b.

THEOREM 3.10. *Let $d > p > d/2$, and let $a_n, b_n, n = 1, 2, \dots,$ have the same meaning as a, b, respectively. Suppose that, for each n, they satisfy Assumptions 2.1(with the same $\delta, \|b\|$) and 3.5 (p, r_0) (with the same θ_0). Assume that $a_n \to a$ (a.e.) and $b_n \to b$ in L_d as $n \to \infty$. Denote by $\hat{T}^n_t$ the semigroups constructed on the basis of (3.7) when R_z is replaced with R^n_z that is the inverse operator to $z - L_n$, where $L_n = (1/2)a^{ij}_n D_{ij} + b^i_n D_i$. Then, for any $t > 0$ and $f \in L_p$ we have $\hat{T}^n_t f \to \hat{T}_t f$ in W^2_p and hence uniformly on $\mathbb{R}^d$ as $n \to \infty$.*

PROOF. It suffices to prove the convergence in W^2_p, and formula (3.7) and estimate (3.6) and the dominated convergence theorem show that it suffices to prove that $\|R^n_z f - R_z f\|_{W^2_p} \to 0$ for $z \in \Gamma$. In light of (3.6),

$$\begin{aligned}\|R^n_z f - R_z f\|_{W^2_p} &\le N\|(z - L^n)(R^n_z f - R_z f)\|_{L_p}\\ &= N\|(L - L^n)R_z f\|_{L_p}\\ &\le N\||a^n - a||(R_z f)_{xx}|\|_{L_p} + N\||b^n - b||(R_z f)_x|\|_{L_p},\end{aligned}$$

where the constants N are independent of n. In the last sum the first term tends to zero by the dominated convergence theorem. Concerning the second one, observe that, by the Hölder and Sobolev inequalities,

$$\begin{aligned}\||b^n - b||(R_z f)_x|\|_{L_p} &\le \|b^n - b\|_{L_d}\|(R_z f)_x\|_{L_{pd/(d-p)}}\\ &\le N(d, p)\|b^n - b\|_{L_d}\|(R_z f)_{xx}\|_{L_p}.\end{aligned}$$

Therefore, it also goes to zero and the theorem is proved. □

4. Relation of $\hat{T}_t$ to a diffusion process. Fix $p \in [d_0, d)$, where $d_0 = d_0(d, \delta) \in (d/2, d)$ is taken from [11], and suppose that Assumptions 2.1 and 3.5 (p, r_0) are satisfied.

Define $\Omega = C([0, \infty), \mathbb{R}^d)$, and, for $\omega = \omega_\cdot \in \Omega$, define $x_t(\omega) = \omega_t$. Also, set $\mathcal{M}_t = \mathcal{N}_t = \sigma(x_s : s \le t)$. Let $X = (x_t, \infty, \mathcal{M}_t, P_x)$ be a Markov process corresponding to the operator L constructed in [11] (we need only Assumptions 2.1 for that). We know from [11] that, for each $x_0 \in \mathbb{R}^d$, with P_{x_0}-probability one

$$x_t = x_0 + \int_0^t \sqrt{a(x_s)}\, dB_s + \int_0^t b(x_s)\, ds, \tag{4.1}$$

where B_t is a d-dimensional Wiener process on Ω relative to $\mathcal{N}_t^{x_0}$ that are the completions of $\mathcal{N}_t$ with respect to P_{x_0}.

LEMMA 4.1. *There is an extension of the probability space $(\Omega, \mathcal{N}_\infty^{x_0}, P_{x_0})$ that carries a d_1-dimensional Wiener process w_t such that the above x_t satisfies* (1.1).

PROOF. Enlarge the probability space $(\Omega, \mathcal{N}_\infty^{x_0}, P_{x_0})$ in such a way that it will carry a d_1-dimensional Wiener process $\hat{B}_t$, the first d coordinate of which coincides with those of B_t. Then, introduce $\hat{\sigma}^k = a^{-1/2}\sigma^k$, and observe that, for each x, the vectors $\xi_i(x) = (\hat{\sigma}^{i1}(x), \dots, \hat{\sigma}^{id_1}(x))$, $i = 1, \dots, d$, are orthogonal to each other and have unit length. By using the Gram–Schmidt procedure it is not hard to complement them in such a way that $\xi_i(x)$, $i = 1, \dots, d_1$, are orthogonal to each other, have unit length, and are Borel with respect to x. In that case the matrix $Q(x)$, having as rows the $\xi_i(x)$'s, is orthogonal and

$$w_t := \int_0^t Q^*(x_s)\, d\hat{B}_s$$

is a d_1-dimensional Wiener process. After that, it only remains to note that $\hat{\sigma}^k Q^{rk} = e_r I_{r \le d}$, where e_r is the rth basis vector so that

$$\sigma^k(x_s)\, dw_s^k = a^{1/2}(x_s)\hat{\sigma}^k(x_s) Q^{rk}(x_s)\, d\hat{B}_s^r = a^{1/2}(x_s)\, dB_s.$$

The lemma is proved. □

We know from [14] that, under Assumption 3.5 (p, r_0), solutions of (1.1) are weakly unique and, therefore, talking about the properties of solutions of (1.1) we may use some results from [13] about the process X.

The following result regarding X is taken from [13].

LEMMA 4.2. *Denote*

$$T_t f(x) = E_x f(x_t).$$

Then, (Theorem 4.8 of [13]) for any $q \ge d_0$ there are constants N and $\mu > 0$, depending only on d, q, δ, and $\|b\|$, such that for any Borel nonnegative f given on $\mathbb{R}^d$ and $t > 0$ we have

$$T_t f(0) \le N t^{-d/(2q)} \|\Phi_t f\|_{L_q}, \tag{4.2}$$

where $\Phi_t(x) = \exp(-\mu|x|/\sqrt{t})$. Furthermore (Corollary 4.9 of [13]), for $q \ge d_0$ such that $q > d/2 + 1$, there exists a constant $N = N(q, d, \delta, \|b\|)$ such that, for any $T \in (0, \infty)$ and nonnegative Borel $f(t, x)$ given on $[0, T] \times \mathbb{R}^d$, we have

$$E_0 \int_0^T f(t, x_t)\, dt \le N T^{(q-1)/q - d/(2q)} \|\Phi_T f\|_{L_q([0,T]\times\mathbb{R}^d)}. \tag{4.3}$$

Finally (Lemma 6.4 of [13] and (4.2)), $T_t f(x)$ is a continuous (even locally Hölder continuous) function of $(t, x) \in (0, \infty) \times \mathbb{R}^d$ if $f \in L_q$ with $q \ge d_0$.

Note that if $u \in W_q^{1,2}([0,T] \times \mathbb{R}^d)$ and $q > d/2+1$, then u has a modification which is bounded and continuous on $[0,T] \times \mathbb{R}^d$. Therefore, talking about u of class $W_q^{1,2}([0,T] \times \mathbb{R}^d)$, we will always mean this modification.

THEOREM 4.3 (Itô's formula). *Let $q \ge d_0$ and $q > d/2+1$, and let $u \in W_q^{1,2}([0,T] \times \mathbb{R}^d)$. Then, with probability one for all $t \in [0,T]$, we have*

$$u(t,x_t) = u(0,x_0) + \int_0^t (\partial_t + L)u(s,x_s)\,ds + \int_0^t \sigma^{ik} D_i u(s,x_s)\,dw_s^k, \tag{4.4}$$

where the stochastic integral is a square integrable martingale on $[0,T]$ (and x_t is a solution of (1.1)).

This theorem is proved by using (4.3) in the same way as Theorem 1.3 of [11] is proved on the basis of Theorem 2.6 of [11].

Recall that $p \in [d_0, d)$ and $d_0 \in (d/2, d)$, so that there are values of $p > d/2+1$ since $d \ge 3$.

THEOREM 4.4. *Let $p > d/2+1$, $T \in (0,\infty)$ and $f \in L_p \cap L_{2p}$. Then:*

(i) *For each $t > 0$ and $x \in \mathbb{R}^d$, we have $\hat{T}_t f(x) = T_t f(x)$;*
(ii) *For each $t > 0$, for solutions of (1.1), with probability one we have*

$$f(x_t) = T_t f(x_0) + \int_0^t \sigma^{ik} D_i T_{t-s} f(x_s)\,dw_s^k, \tag{4.5}$$

where $\sigma^{ik} D_i T_{t-s} f(x) = (\sigma^{ik} D_i T_{t-s} f)(x)$ and similar notation is also used below;
(iii) *For each $t > 0$ and $x \in \mathbb{R}^d$,*

$$T_t f^2(x) = (T_t f(x))^2 + \sum_k \int_0^t T_s\Big[\Big(\sum_i \sigma^{ik} D_i T_{t-s} f\Big)^2\Big](x)\,ds. \tag{4.6}$$

PROOF. If $f \in W_p^2$, then $u(s,x) := \hat{T}_{t-s} f(x)$, $s \le t$ satisfies the condition of Theorem 4.3, and we get (4.5) with $\hat{T}_{\cdot}$ in place of $T_{\cdot}$ by that theorem. By taking the expectations of both sides, we get that $T_t f(x_0) = \hat{T}_t f(x_0)$. This holds for any x_0 and yields (4.5) as is. By taking the expectations of the squares of both sides of (4.5), we obtain (4.6). Thus, all assertions of the theorem are true if $f \in W_p^2$.

Assertion (i) holds for any $f \in L_p$, which is seen from the fact that, according to embedding theorems and (4.2), both $\hat{T}_t f(x)$ and $T_t f(x)$ are bounded linear functionals on L_p and W_p^2 is dense in L_p.

Then, as $f^n \in W_p^2$ tend to f in $L_p \cap L_{2p}$, $T_{t-s} f^n \to T_{t-s} f$ in W_p^2 for $s < t$ (see (3.8)). By embedding theorems ($p \ge d/2$) $DT_{t-s} f^n \to DT_{t-s} f$ in L_{2p} and in light of (4.2),

$$T_s\Big[\Big(\sum_i \sigma^{ik} D_i T_{t-s} f^n\Big)^2\Big](x) \to T_s\Big[\Big(\sum_i \sigma^{ik} D_i T_{t-s} f\Big)^2\Big](x)$$

for any $0 < s < t$ and $x \in \mathbb{R}^d$. Furthermore, $(f^n)^2 \to f^2$ in L_p and, due to (4.2), $T_t (f^n)^2(x) \to T_t f^2(x)$. It follows by Fatou's lemma (and (4.6)) that

$$T_t f^2(x) \ge (T_t f(x))^2 + \sum_k \int_0^t T_s\Big[\Big(\sum_i \sigma^{ik} D_i T_{t-s} f\Big)^2\Big](x)\,ds. \tag{4.7}$$

Hence, the right-hand side of (4.5) is well defined. Furthermore,

$$
\begin{aligned}
E\left|\int_0^t \sigma^{ik} D_i T_{t-s} f(x_s)\,dw_s^k - \int_0^t \sigma^{ik} D_i T_{t-s} f^n(x_s)\,dw_s^k\right|^2 \\
&= \sum_k \int_0^t T_s\Big[\Big(\sum_i \sigma^{ik} D_i T_{t-s}(f-f^n)\Big)^2\Big](x_0)\,ds \\
&\le T_t(f-f^n)^2(x_0) - \big(T_t(f-f^n)(x_0)\big)^2 \\
&\le T_t(f-f^n)^2(x_0) = E\big|f(x_t) - f^n(x_t)\big|^2,
\end{aligned}
$$

where the first inequality is due to (4.7). The last expression tends to zero in light of (4.2) which allows us to get (4.5) by passing to the limit in its version with f^n in place of f. After that, (4.6) follows as above. The theorem is proved. □

REMARK 4.5. In light of Theorem 4.4(i), estimate (4.2) is weaker in what concerns the restriction on q than (3.9). However, (4.2) is proved for just measurable σ^k.

5. A criterion for strong solutions of Itô's equations. In this section

$$p \in (d_0, d), \qquad p > d/2 + 1$$

and we suppose that Assumptions 2.1 and 3.5 (p, r_0) are satisfied.

Recall the setting from the beginning of the article. We are given a complete probability space $(\Omega, \mathcal{F}, P)$ with an increasing filtration of σ-fields $\mathcal{F}_t \subset \mathcal{F}$ that are complete. We are also given a d_1-dimensional Wiener process w_t relative to $\{\mathcal{F}_t\}$. Finally, for an $x_0 \in \mathbb{R}^d$, we are given a solution x_t of (1.1). We know from Lemma 4.1 that such a situation is quite realistic, and we also know that the solution is weakly unique. In particular, it has the same distributions as the process x_t from Section 4 has relative to P_{x_0}.

For further discussion we need the following in which $\mathcal{P}$ is the σ-field of predictable sets and $\mathcal{B}(0,\infty)$ is the Borel σ-field in $(0,\infty)$.

LEMMA 5.1. *Assume that, for $s, r \in (0,\infty)$, $\omega \in \Omega$, we are given a real-valued function $g(s,r) = g(s,r,\omega)$, $s \in (0,\infty)$, $(r,\omega) \in (0,\infty) \times \Omega$, which is measurable with respect to $\mathcal{B}(0,\infty) \otimes \mathcal{P}$, and such that, for each s,*

$$E\int_0^\infty g^2(s,r)\,dr < \infty.$$

Then, there is a function $m_{s,t} = m(s,t,\omega)$ on $[0,\infty) \times ([0,\infty) \times \Omega)$ measurable with respect to $\mathcal{B}(0,\infty) \otimes \mathcal{P}$, continuous in t for each (s,ω) and such that for each s it is martingale starting from zero, and, moreover, for each s (a.s.) for all $t \ge 0$,

$$m_{s,t} = \int_0^t g(s,r)\,dw_r. \tag{5.1}$$

PROOF. Introduce

$$\Omega_s = \Big\{\omega : \int_0^\infty g^2(s,r)\,dr < \infty\Big\}, \qquad \hat{g}(s,r) = I_{\Omega_s} g(s,r),$$

$$B_t(s) = \int_0^t \hat{g}^2(s,r)\,dr.$$

Observe that $P(\Omega_s) = 1$ so that $\Omega_s \in \mathcal{F}_0$. Also, $B_\infty(s) < \infty$ for any s and ω.

By Lemma 2.6 of [10], there exists a function $m_{s,t}$ on $[0,\infty)^2\times\Omega$ with the properties described in the statement of the lemma but satisfying (5.1) with $\hat g$ in place of g. Since $P(\Omega_s)=1$, the integrals of $\hat g$ and g coincide with probability one, and the lemma is proved. □

REMARK 5.2. As we have noted, if $f\in L_p$, then, for any $t>0$, we have $T_t f\in W^2_p$, and, hence $(p>d/2)$, $DT_t f\in L_p\cap L_{2p}$. Therefore, we can apply Theorem 4.4 and write that, for any $s<t$ (a.s.),

$$\begin{aligned}\sigma^{ik}D_iT_{t-s}f(x_s)&=T_s\big(\sigma^{ik}D_iT_{t-s}f\big)(x_0)\\&\quad+\int_0^s\sigma^{jm}D_jT_{s-r}\big(\sigma^{ik}D_iT_{t-s}f\big)(x_r)\,dw^m_r.\end{aligned}\tag{5.2}$$

After that, we want to substitute the result into (4.5) to get

$$\begin{aligned}f(x_t)&=T_tf(x_0)+\int_0^tT_s\big(\sigma^{ik}D_iT_{t-s}f\big)(x_0)\,dw^k_s\\&\quad+\int_0^t\bigg(\int_0^s\sigma^{jm}D_jT_{s-r}\big(\sigma^{ik}D_iT_{t-s}f\big)(x_r)\,dw^m_r\bigg)\,dw^k_s.\end{aligned}\tag{5.3}$$

The formal objection to do that is that we should know that the integral in (5.2) is, for instance, predictable as a function of (ω,s), and this may not happen if we allow any version of the stochastic integral to be taken for each s. However, set $h^k(s,x)=I_{s<t}\sigma^{ik}D_iT_{t-s}f(x)$, and consider

$$I^k(s,u)=\int_0^u I_{r<s}\sigma^{jm}D_jT_{s-r}h^k(s,\cdot)(x_r)\,dw^m_r.\tag{5.4}$$

This is the sum over m of stochastic integrals, and

$$\begin{aligned}&E\int_0^\infty I_{r<s}\bigg|\sum_j\sigma^{jm}D_jT_{s-r}h^k(s,\cdot)(x_r)\bigg|^2\,dr\\&\quad=E\int_0^s\bigg|\sum_j\sigma^{jm}D_jT_{s-r}h^k(s,\cdot)(x_r)\bigg|^2\,dr\\&\quad\le T_s\big((h^k(s,\cdot))^2\big)(x_0),\end{aligned}$$

where the inequality is due to (4.6). It follows from Lemma 5.1 that $I(s,u)=I(s,u,\omega)$ has a version, which we denote again $I(s,u)$, that is continuous in u for each s,ω and measurable with respect to $\mathcal{B}(0,\infty)\otimes\mathcal{P}$. Then, $I^k(s,s)$ is predictable, and we take this modification of the right-hand side of (5.4) in the right-hand side of (5.3) thus justifying (5.3).

Then, we want to repeat this procedure. Introduce

$$Q^k_tf(x)=\sigma^{ik}(x)D_iT_tf(x).\tag{5.5}$$

In this notation (4.5) and (5.3) become, respectively,

$$f(x_t)=T_tf(x_0)+\int_0^tQ^{k_1}_{t-t_1}f(x_{t_1})\,dw^{k_1}_{t_1};$$

$$\begin{aligned}f(x_t)&=T_tf(x_0)+\int_0^tT_{t_1}Q^{k_1}_{t-t_1}(x_0)\,dw^{k_1}_{t_1}\\&\quad+\int_0^t\bigg(\int_0^{t_1}Q^{k_2}_{t_1-t_2}Q^{k_1}_{t-t_1}f(x_{t_2})\,dw^{k_2}_{t_2}\bigg)\,dw^{k_1}_{t_1}.\end{aligned}$$

By induction we obtain that, for any $n \geq 1$ (a.s.) for all $t \geq 0$ ($t_0 = t$),

$$
\begin{aligned}
f(x_t) = T_t f(x_0) &+ \sum_{m=1}^{n} \int_{t>t_1>\cdots>t_m} T_{t_m} Q^{k_m}_{t_{m-1}-t_m} \cdots Q^{k_1}_{t-t_1} f(x_0)\, dw^{k_m}_{t_m} \cdots dw^{k_1}_{t_1} \\
&+ \int_{t>t_1>\cdots>t_{n+1}} Q^{k_{n+1}}_{t_n-t_{n+1}} \cdots Q^{k_1}_{t-t_1} f(x_{t_{n+1}})\, dw^{k_{n+1}}_{t_{n+1}} \cdots dw^{k_1}_{t_1},
\end{aligned}
\tag{5.6}
$$

where by the expressions like

$$\int_{t>t_1>\cdots>t_m} ::: dw^{k_m}_{t_m} \cdots dw^{k_1}_{t_1}$$

we mean

$$\int_0^t dw^{k_1}_{t_1} \int_0^{t_1} dw^{k_2}_{t_2} \ldots \int_0^{t_{m-1}} ::: dw^{k_m}_{t_m}.$$

By taking expectations of the squares of the sides in (5.6), we conclude that

$$
\begin{aligned}
&T_t f^2(x_0) \\
&\quad = \big(T_t f(x_0)\big)^2 \\
&\qquad + \sum_{m=1}^{n} \int_{t>t_1>\cdots>t_m} \big[T_{t_m} Q^{k_m}_{t_{m-1}-t_m} \cdots Q^{k_1}_{t-t_1} f(x_0)\big]^2\, dt_m \cdots dt_1 \\
&\qquad + \int_{t>t_1>\cdots>t_{n+1}} \sum_{k_1,\ldots,k_{n+1}} T_{t_{n+1}} \big[Q^{k_n}_{t_n-t_{n+1}} \cdots Q^{k_1}_{t-t_1} f\big]^2(x_0)\, dt_{n+1} \cdots dt_1.
\end{aligned}
\tag{5.7}
$$

In particular, the sequence of

$$\int_{t>t_1>\cdots>t_n} \sum_{k_1,\ldots,k_n} T_{t_n} \big[Q^{k_n}_{t_{n-1}-t_n} \cdots Q^{k_1}_{t-t_1} f\big]^2(x_0)\, dt_n \cdots dt_1$$

is decreasing.

REMARK 5.3. It turns out that proving *directly* that each term in the right-hand side of (5.7) is finite presents significant difficulties. However, observe that, due to (3.8) and (4.2) for $p \in (d_0, d)$ and $f \in L_p$, we have

$$\big|T_{t_m} Q^{k_m}_{t_{m-1}-t_m} \cdots Q^{k_1}_{t-t_1} f(x)\big| \leq \frac{N}{t_m^{d/(2p)} (t_{m-1}-t_m)^{1/2} \cdots (t-t_1)^{1/2}} \|f\|_{L_p},$$

where N depends only on m, d, δ, $\|b\|$ and ν_b. Furthermore,

$$\int_{t>t_1>\cdots>t_m} \frac{1}{t_m^{d/(2p)} (t_{m-1}-t_m)^{1/2} \cdots (t-t_1)^{1/2}}\, dt_m \cdots dt_1 < \infty.$$

Recall that $\mathcal{F}^w_t$ is the completion of $\sigma(w_s : s \leq t)$. Remark 5.2 allows us to repeat, literally, some arguments in [25] and leads to the following results.

THEOREM 5.4. *Let $f \in L_p \cap L_{2p}$, $t > 0$. Then,*

$$
\begin{aligned}
E\big(f(x_t) \mid \mathcal{F}^w_t\big) &= T_t f(x_0) \\
&\quad + \sum_{m=1}^{\infty} \int_{t>t_1>\cdots>t_m} T_{t_m} Q^{k_m}_{t_{m-1}-t_m} \cdots Q^{k_1}_{t-t_1} f(x_0)\, dw^{k_m}_{t_m} \cdots dw^{k_1}_{t_1},
\end{aligned}
$$

where the series converges in the mean square sense.

For $n \geq 1$, $t > 0$ and $s_1, \dots, s_n > 0$, define

$$Q_{s_n,\dots,s_1} f(x) = \sum_{k_1,\dots,k_n} [Q^{k_n}_{s_n} \cdots Q^{k_1}_{s_1} f]^2(x). \tag{5.8}$$

THEOREM 5.5. *Let $f \in L_p \cap L_{2p}$, $t_0 > 0$. Then, $f(x_{t_0})$ is $\mathcal{F}^w_{t_0}$-measurable iff*

$$\lim_{n\to\infty} \int_{t_0>t_1>\cdots>t_n} T_{t_n} Q_{t_{n-1}-t_n,\dots,t_0-t_1} f(x_0)\, dt_n \cdots dt_1 = 0. \tag{5.9}$$

Furthermore, under either of the above equivalent conditions,

$$\begin{aligned} f(x_t) &= T_t f(x_0) \\ &\quad + \sum_{m=1}^{\infty} \int_{t>t_1>\cdots>t_m} T_{t_m} Q^{k_m}_{t_{m-1}-t_m} \cdots Q^{k_1}_{t-t_1} f(x_0)\, dw^{k_m}_{t_m} \cdots dw^{k_1}_{t_1}. \end{aligned} \tag{5.10}$$

THEOREM 5.6. *If equation* (1.1) *has two solutions which are not indistinguishable, then it does not have any strong solution. In particular, if* (1.1) *has at least one strong solution, then the solution is unique.*

THEOREM 5.7. *If equation* (1.1) *has a strong solution on one probability space, then it has a strong solution on any other probability space carrying a d_1-dimensional Wiener process.*

REMARK 5.8. By making the change of variables $t_k = s_k + \cdots + s_n$, $k = 1, \dots, n$, we rewrite (5.9) as

$$\lim_{n\to\infty} \int_{S_n(t_0)} T_{s_n} Q_{s_{n-1},\dots,s_1,t_0-(s_1+\cdots+s_n)} f(x_0)\, ds_n \cdots ds_1 = 0, \tag{5.11}$$

where $S_n(t_0) = \{(s_1, \dots, s_n) : s_k \geq 0, s_1 + \cdots + s_n < t_0\}$.

The sequence under the limit sign in (5.11), call it $u_n(t_0)$, is decreasing for any t_0 (and x_0, see the end of Remark 5.2). Therefore, its limit will be zero for almost any t if

$$\lim_{n\to\infty} \int_0^{\infty} u_n(t_0) e^{-\nu t_0}\, dt_0 = 0,$$

where $\nu > 0$ is any number. In that case, actually, the limit of $u_n(t_0)$ is zero for all t_0, since, in light of Theorem 5.5, $f(x_{t_0})$ is $\mathcal{F}^w_{t_0}$-measurable for almost all t_0 and, by continuity, for all t_0. In this way after simple change of variables, we come to the following.

THEOREM 5.9. *Let $f \in L_p \cap L_{2p}$. Then, $f(x_t)$ is $\mathcal{F}^w_t$-measurable for any $t > 0$ if there exists a $\nu > 0$ such that*

$$\int_0^{\infty} e^{-\nu s_n} T_{s_n} \biggl(\int_{R^n_+} e^{-\nu(s_{n-1}+\cdots+s_0)} Q_{s_{n-1},\dots,s_0} f\, ds_{n-1} \cdots ds_0 \biggr)(x_0)\, ds_n \to 0 \tag{5.12}$$

as $n \to \infty$, where $R^n_+ = (0,\infty)^n$, which in light of (4.2) *holds for any x_0 if*

$$\biggl\| \int_{R^n_+} e^{-\nu(s_{n-1}+\cdots+s_0)} Q_{s_{n-1},\dots,s_0} f\, ds_{n-1} \cdots ds_0 \biggr\|^p_{L_p} \to 0. \tag{5.13}$$

We are going to prove that (5.13) holds under Assumptions 2.1 and 2.2 by showing that the series composed of the left-hand sides of (5.13) converges.

REMARK 5.10. The criterion (5.9) is proved under Assumptions 2.1 and 3.5 (p, r_0), assumptions which involve the σ^k's only implicitly, and it turns out that for some choice of the σ^k's (5.9) may hold and for another fail to hold. To illustrate this, we take $b \equiv 0$. In that case the restriction $p < d$ disappears along with $d \geq 3$ (which is a consequence of $p < d$ and $p > d/2 + 1$). Then, we take $d_1 = d = 2$ and, following [16], set $\sigma^1(x) = x/|x|$, $\sigma^2(x) = x^*/|x|$, where $x^* = (-x^2, x^1)$ for $x \neq 0$, $\sigma^{ik}(0) = \delta^{ik}$. Then, $a^{ij}(x) = \delta^{ij}$, equation (1.1) has a solution for any x_0 (see, for instance, Lemma 4.1), and each solution is a Wiener process starting from x_0. For $x_0 \neq 0$, the solutions are strong, and, hence, (5.9) holds, because the solution never reaches the origin, the point where σ^k are not smooth. However, for $x_0 = 0$, there are no strong solutions, because, as is easy to see, rotation in x^1x^2 coordinates by any angle brings any solution it to another solution of the same equation. Therefore, for $x_0 = 0$, equation (5.9) does not hold.

Also, observe that in this example $\sigma^k \in W^1_{d-\varepsilon,\text{loc}}$ for any $\varepsilon \in (0, 1)$ but not for $\varepsilon = 0$. One can construct similar examples for $d \geq 3$ starting from the following with $d = 3$, $d_1 = 9$ and σ^k's that are the kth columns of the matrix

$$\frac{1}{|x|}\begin{pmatrix} x^1 & x^2 & x^3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & x^1 & x^2 & x^3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & x^1 & x^2 & x^3 \end{pmatrix}.$$

Again, $a^{ij} = \delta^{ij}$, $\sigma^k \in W^1_{d-\varepsilon,\text{loc}}$ for any $\varepsilon \in (0, 1)$ but not for $\varepsilon = 0$, and if x_t is a solution of (1.1) with $x_0 = 0$, then $-x_t$ is also a solution of (1.1) with $x_0 = 0$.

6. Some estimates in the case of C^∞ coefficients. We suppose that σ^k, b satisfy Assumption 2.1 and are infinitely differentiable with each derivative bounded.

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space, let $\{\mathcal{F}_t\}$ be an increasing filtration of σ-fields $\mathcal{F}_t \subset \mathcal{F}$ that are complete. Let w_t be a d_1-dimensional Wiener process relative to $\{\mathcal{F}_t\}$. We also assume that there is a $(d+1)$ independent d-dimensional Wiener, relative to $\{\mathcal{F}_t\}$, process $B_t^{(0)}, \dots, B_t^{(d)}$ independent of w_t. Take $x, \eta \in \mathbb{R}^d$, a nonnegative bounded infinitely differentiable K_0, the role of which will be emphasized later, with each derivative bounded given on $\mathbb{R}^d$, and consider the following system:

$$x_t = x + \int_0^t \sigma^k(x_s)\,dw_s^k + \int_0^t b(x_s)\,ds, \tag{6.1}$$

$$\begin{aligned}\eta_t = \eta &+ \int_0^t \sigma^k_{(\eta_s)}(x_s)\,dw_s^k + \int_0^t b_{(\eta_s)}(x_s)\,ds \\ &+ \int_0^t K_0(x_s)\,dB_s^{(0)} + \int_0^t K_0(x_s)\eta_s^k\,dB_s^{(k)}.\end{aligned} \tag{6.2}$$

As is well known, (6.1) has a unique solution which we denote by $x_t(x)$. By substituting it into (6.3), we see that the coefficients of (6.3) grow linearly in η, and, hence, (6.3) also has a unique solution, which we denote by $\eta_t(x, \eta)$. By the way, observe that equation (6.3) is linear with respect to η_t. Therefore, $\eta_t(x, \eta)$ is an affine function of η. For the uniformity of notation, we set $x_t(x, \eta) = x_t(x)$. It is also well known (see, for instance, Sections 2.7 and 2.8 of [7]) that, as a function of x and (x, η), the processes $x_t(x)$ and $\eta_t(x, \eta)$ are infinitely differentiable in an appropriate sense (specified below), their derivatives satisfy the equations which are obtained by formal differentiation of (6.1) and (6.3), respectively, and, for any $n \geq 0$, $T \in (0, \infty)$, $l_k, \xi_k \in \mathbb{R}^d$, $k = 1, \dots, n$ (if $n \geq 1$), $x, \eta \in \mathbb{R}^d$, and $q \geq 1$,

$$E \sup_{t \leq T} \left| \left(\prod_{k=1}^n (lb) D_{(l_k, \xi_k)} \right) (x_t, \eta_t)(x, \eta) \right|^q \leq N\big(1 + |\eta|^m\big), \tag{6.3}$$

where N is a certain constant independent of (x,η), $m=m(n,q)$, and, for instance, by $(lb)D_{(l,\xi)}\eta_t(x,\eta)$ we mean a process ζ_t such that, for any $q\ge 1$ and $S\in(0,\infty)$,

$$\lim_{\varepsilon\downarrow 0} E\sup_{t\le S}\big|\zeta_t-\varepsilon^{-1}\big(\eta_t(x+\varepsilon l,\eta+\varepsilon\xi)-\eta_t(x,\eta)\big)\big|^q=0.$$

LEMMA 6.1. *Let $f(x,\eta)$ be infinitely differentiable and such that each of its derivatives grows as $|x|+|\eta|\to\infty$ not faster than polynomially. Then, the function $u(t,x,\eta):=Ef((x_t,\eta_t)(x,\eta))$ is infinitely differentiable in (x,η), and each of its derivatives is continuous in t and is by absolute value bounded on each finite time interval by a constant times $(1+|x|+|\eta|)^m$ for some m. Furthermore, $u(t,x,\eta)$ is continuously differentiable in t and, for $t\ge 0$ and $(x,\eta)\in\mathbb{R}^{2d}$,*

$$\begin{aligned}\partial_t u(t,x,\eta) &= (1/2)\sigma^{ik}\sigma^{jk}(x)u_{x^ix^j}(t,x,\eta)+\sigma^{ik}\sigma^{jk}_{(\eta)}(x)u_{x^i\eta^j}(t,x,\eta)\\ &\quad+(1/2)\sigma^{ik}_{(\eta)}\sigma^{jk}_{(\eta)}(x)u_{\eta^i\eta^j}(t,x,\eta)\\ &\quad+(1/2)K_0^2(x)\big(1+|\eta|^2\big)\delta^{ij}u_{\eta^i\eta^j}(t,x,\eta)\\ &\quad+b^i(x)u_{x^i}(t,x,\eta)+b^i_{(\eta)}(x)u_{\eta^i}(t,x,\eta)\\ &=:\check{L}(x,\eta)u(t,x,\eta).\end{aligned}\tag{6.4}$$

The first assertion of this lemma follows easily from what is said before it. Then, the fact that (6.4) holds follows from the Markov property of (x_t,η_t) and from the first assertion. The claimed property of $\partial_t u$ follows from (6.4).

LEMMA 6.2. *Let $\eta\in\mathbb{R}^d$ and $\xi_t(x,\eta)=(lb)D_\eta x_t(x)$. Then*:

(i) *$\xi_t(x,\eta)$ satisfies (6.3) with $K_0\equiv 0$.*
(ii) *For any t,*

$$\xi_t(x,\eta)=E\big(\eta_t(x,\eta)\mid\mathcal{F}_t^w\big).\tag{6.5}$$

(iii) *If $f(x)$ is infinitely differentiable with bounded derivatives, then*

$$Ef_{(\eta_t(x,\eta))}\big(x_t(x)\big)\big(=E\big(f_{(\eta_t(x,\eta))}\big)\big(x_t(x)\big)\big)=\big(Ef\big(x_t(x)\big)\big)_{(\eta)}.\tag{6.6}$$

PROOF. Assertion (i) is well known (see, for instance, [7]). The right-hand side of (6.5) satisfies (6.3) with $K_0\equiv 0$ owing to the linearity of $g_{(\eta)}$ in η and independence of $B_\cdot$ and $w_\cdot$. Therefore, due to uniqueness, assertion (ii) follows from (i). Assertion (iii) follows from (ii) and the fact that (see, for instance, [7])

$$\big(Ef\big(x_t(x)\big)\big)_{(\eta)}=Ef_{(\xi_t(x,\eta))}\big(x_t(x)\big).$$

The lemma is proved. □

Now follows one of the most important computations. The idea behind it is the following. If we formally differentiate both parts of (5.10) in the direction η and then take the expectations of the squares of both sides, then we obtain an equality in (6.7) below if we also replace on the left $\eta_t(x,\eta)$ by $\xi_t(x,\eta)$. Then, the inequality follows from Lemma 6.2.

LEMMA 6.3. *Let $x, \eta \in \mathbb{R}^d$, and let $f(x)$ be infinitely differentiable with bounded derivatives. Then, for any $t \in (0, \infty)$ $(t_0 = t)$,*

$$
\begin{aligned}
&E\big[f_{(\eta_t(x,\eta))}(x_t(x))\big]^2 \\
&\quad \ge \big[\big(T_t f(x)\big)_{(\eta)}\big]^2 \\
&\qquad + \sum_{n=1}^{\infty} \sum_{k_1,\dots,k_n} \int_{t>t_1>\dots>t_n} \big[\big(T_{t_n} Q^{k_n}_{t_{n-1}-t_n} \cdots Q^{k_1}_{t-t_1} f(x)\big)_{(\eta)}\big]^2 \, dt_n \cdots dt_1.
\end{aligned} \tag{6.7}
$$

PROOF. Introduce the notation $\check{T}_t u(x, \eta) = Eu((x_t, \eta_t)(x, \eta))$. Then, similarly to (4.5), for smooth bounded $u(x, \eta)$ by dropping for simplicity the arguments x and η in $x_s(x)$ and $\eta_s(x, \eta)$, we get

$$
\begin{aligned}
u(x_t, \eta_t) &= \check{T}_t u(x, \eta) + \int_0^t K_0 D_{\eta^i} \check{T}_{t-t_1} u(x_{t_1}, \eta_{t_1}) \big(dB^{i(0)}_{t_1} + \eta^k_{t_1} \, dB^{i(k)}_{t_1}\big) \\
&\quad + \int_0^t \big[\sigma^{ik}(x_{t_1}) D_{x^i} \check{T}_{t-t_1} u(x_{t_1}, \eta_{t_1}) + \sigma^{ik}_{(\eta_{t_1})}(x_{t_1}) D_{\eta^i} \check{T}_{t-t_1} u(x_{t_1}, \eta_{t_1})\big] \, dw^k_{t_1}.
\end{aligned}
$$

It follows that

$$
\begin{aligned}
Eu^2(x_t, \eta_t) &\ge \big(\check{T}_t u(x, \eta)\big)^2 \\
&\quad + \sum_k \int_0^t E\big[\sigma^{ik}(x_{t_1}) D_{x^i} \check{T}_{t-t_1} u(x_{t_1}, \eta_{t_1}) \\
&\quad + \sigma^{ik}_{(\eta_{t_1})}(x_{t_1}) D_{\eta^i} \check{T}_{t-t_1} u(x_{t_1}, \eta_{t_1})\big]^2 \, dt_1.
\end{aligned} \tag{6.8}
$$

By using Fatou's lemma and estimates like (6.3), one easily carries (6.8) over to smooth $u(x, \eta)$ whose derivatives have no more than polynomial growth as $|x| + |\eta| \to \infty$. In particular, one can apply (6.8) to $u(x, \eta) = f_{(\eta)}(x)$. Then, after noting that in light of (6.6) in that case,

$$
\begin{aligned}
&\sigma^{ik}(x) D_{x^i} \check{T}_{t-t_1} u(x, \eta) + \sigma^{ik}_{(\eta)}(x) D_{\eta^i} \check{T}_{t-t_1} u(x, \eta) \\
&\quad = \sigma^{ik}(x) D_{x^i} \big(T_{t-t_1} f(x)\big)_{(\eta)} + \sigma^{ik}_{(\eta)}(x) D_{\eta^i} \big(T_{t-t_1} f(x)\big)_{(\eta)} \\
&\quad = \big(\sigma^{ik}(x) D_{x^i} T_{t-t_1} f(x)\big)_{(\eta)} = \big(Q^k_{t-t_1} f(x)\big)_{(\eta)},
\end{aligned}
$$

we obtain

$$
E\big[f_{(\eta_t)}(x_t)\big]^2 \ge \big[\big(T_t f(x)\big)_{(\eta)}\big]^2 + \sum_{k_1} \int_0^t E\big[\big(Q^{k_1}_{t-t_1} f\big)_{(\eta_t)}(x_{t_1})\big]^2 \, dt_1.
$$

By applying this formula to $Q^{k_1}_{t-t_1} f$ in place of f, we get

$$
\begin{aligned}
E\big[f_{(\eta_t)}(x_t)\big]^2 &\ge \big[\big(T_t f(x)\big)_{(\eta)}\big]^2 + \sum_{k_1} \int_0^t E\big[\big(T_{t_1} Q^{k_1}_{t-t_1} f(x)\big)_{(\eta)}\big]^2 \, dt_1 \\
&\quad + \sum_{k_1,k_2} \int_0^t dt_1 \int_0^{t_1} E\big[\big(Q^{k_2}_{t_1-t_2} Q^{k_1}_{t-t_1} f\big)_{(\eta_{t_2})}(x_{t_2})\big]^2 \, dt_2.
\end{aligned}
$$

Using induction yields that, for any $n \ge 1$ $(t_0 = t)$,

$$
\begin{aligned}
&E\big[f_{(\eta_t)}(x_t)\big]^2 \\
&\quad \ge \big[\big(T_t u(x)\big)_{(\eta)}\big]^2
\end{aligned}
$$

$$+\sum_{m=1}^{n}\sum_{k_1,\dots,k_m}\int_{t>t_1>\cdots>t_m}\big[\big(T_{t_m}Q^{k_m}_{t_{m-1}-t_m}\cdots Q^{k_1}_{t-t_1}f(x)\big)_{(\eta)}\big]^2\,dt_m\cdots dt_1$$
$$+\sum_{k_1,\dots,k_{n+1}}\int_{t>t_1>\cdots>t_{n+1}}E\big[\big(Q^{k_{n+1}}_{t_n-t_{n+1}}\cdots Q^{k_1}_{t-t_1}f\big)_{(\eta_{t_{n+1}})}(x_{t_{n+1}})\big]^2\,dt_{n+1}\cdots dt_1.$$

This yields (6.7) and proves the lemma. □

Next, we want to estimate the left-hand side of (6.7) which, according to Lemma 6.1, satisfies (6.4).

In the future we will be interested in estimating not only the left-hand side of (6.7) but a slightly more general quantity. Therefore, we take an infinitely differentiable $f(x,\eta)\ge 0$ such that, for an $m>0$ and a constant N,

$$\big(|f|+|f_x|+|f_\eta|+|f_{xx}|+|f_{x\eta}|+|f_{\eta\eta}|\big)(x,\eta)\le N\big(1+|\eta|\big)^m$$

for all x, η and such that $f(x,\eta)=0$ for all η if $|x|\ge R$ for some $R>0$. Then, denote $u(t,x,\eta)=\check{T}_t f(x,\eta)$. According to what was said before Lemma 6.2 and in that lemma, if we denote $(l_t,\xi_t)=(lb)D_{(l,\xi)}(x_t,\eta_t)(x,\eta)$, then

$$u_{(l,\xi)}(t,x,\eta)=Ef_{(l_t,\xi_t)}(x_t,\eta_t)(x,\eta).$$

In particular, it follows that

$$\big|u_{(l,\xi)}(t,x,\eta)\big|\le P^{1/2}\big(|x_t(x)|\le R\big)\big(E\big|f_{(l_t,\xi_t)}(x_t,\eta_t)(x,\eta)\big|^2\big)^{1/2}.$$

Here, the second factor on the right is estimated by using (6.3). The first factor is less than $2\exp(-\mu\,\mathrm{dist}^2(x,B_R)/t)$ by Theorem 2.10 of [12], where $\mu>0$ depends only on d, δ and $\|b\|$. Similar estimates are available for the second-order derivatives of $u(t,x,\eta)$. More precisely, observe that there exist constants $\mu>0$, $\kappa=\kappa(m)$ and a function $M(t)$ bounded on each time interval $[0,T]$ such that, for all t, x, η, we have

$$\begin{aligned}(6.9)\qquad &\big|u(t,x,\eta)\big|+\big|u_x(t,x,\eta)\big|+\big|u_\eta(t,x,\eta)\big|\\ &\quad+\big|u_{xx}(t,x,\eta)\big|+\big|u_{x\eta}(t,x,\eta)\big|+\big|u_{\eta\eta}(t,x,\eta)\big|\le M(t)e^{-\mu|x|}\big(1+|\eta|^2\big)^\kappa.\end{aligned}$$

This justifies the integration we perform below.

Introduce

$$h=\big(1+|\eta|^2\big)^{-\kappa-d},$$

and observe that, for a constant $N=N(d,\kappa)$, we have

$$|\eta||h_\eta|\le N\kappa h,\qquad \big|\big((1+|\eta|^2)h\big)_{\eta\eta}\big|\le Nh.$$

THEOREM 6.4. *Let $q\ge 2$, and suppose that the above $u\ge 0$. Then, there is a constant $N_0\ge 1$, depending only on d, d_1, δ, m and q, such that, for any $\lambda\ge 1$ satisfying*

$$(6.10)\qquad N_0\Big(\sum_k\|\sigma^k_x I_{|\sigma^k_x|>\lambda}\|_{L_d}+\|bI_{|b|>\lambda}\|_{L_d}\Big)\le 1,$$

there exists a constant N, depending only on λ, $\|\sigma^k_x\|$, d, δ, m and q, and there is a function K_0 such that, for $t\ge 0$,

$$(6.11)\qquad \int_{\mathbb{R}^{2d}}h(\eta)u^q(t,x,\eta)\,dx\,d\eta\le e^{Nt}\int_{\mathbb{R}^{2d}}h(\eta)f^q(x,\eta)\,dx\,d\eta.$$

The proof of this theorem proceeds as usual by multiplying (6.4) by $h(\eta)u^{q-1}(t,x,\eta)$ and integrating by parts over $[0,t]\times\mathbb{R}^{2d}$. The integral of the left-hand side is

$$q^{-1}\int_{\mathbb{R}^{2d}} h(\eta)u^q(t,x,\eta)\,dx\,d\eta - q^{-1}\int_{\mathbb{R}^{2d}} h(\eta)f^q(x,\eta)\,dx\,d\eta.$$

Therefore, in light of Gronwall's inequality, to prove the theorem it suffices to prove the following estimate.

LEMMA 6.5. *Let $q\geq 2$ and $\kappa\geq 0$. Then, there is a constant $N_0\geq 1$, depending only on d, d_1, δ, κ and q, such that for any $\lambda\geq 1$ satisfying*

$$N_0\Big(\sum_k \|\sigma^k_x I_{|\sigma^k_x|>\lambda}\|_{L_d} + \|bI_{|b|>\lambda}\|_{L_d}\Big)\leq 1, \tag{6.12}$$

there exists a constant N, depending only on λ, $\|\sigma^k_x\|$, d, d_1, δ, κ and q and there is a function K_0 such that for any smooth function $v(x,\eta)\geq 0$ (independent of t), for which condition (6.9) is satisfied with v in place of u and some M, we have

$$\int_{\mathbb{R}^{2d}} h(\eta)v^{q-1}(x,\eta)\check{L}v(x,\eta)\,dx\,d\eta \leq N\int_{\mathbb{R}^{2d}} h(\eta)v^q(x,\eta)\,dx\,d\eta. \tag{6.13}$$

PROOF. For simplicity of notation, we drop the arguments x,η. We also write $U\sim V$, if their integrals over $\mathbb{R}^{2d}$ coincide, and $U\prec V$, if the integral of U is less than or equal to that of V. Below the constants called N, sometimes with indices, depend only on d, d_1, δ, κ and q unless specifically noted otherwise.

Set $w=v^{q/2}$, and note simple formulas,

$$v^{q-1}v_x=(2/q)ww_x,\qquad v^{q-2}v_{x^i}v_{x^j}=(4/q^2)w_{x^i}w_{x^j}.$$

Then, denote by $\check{L}_1$ the sum of the first-order terms in $\check{L}$, and observe that integrating by parts shows that

$$\begin{aligned} hv^{q-1}b^i_{(\eta)}v_{\eta^i} &\sim -(1/q)h_{\eta^i}b^i_{(\eta)}v^q-(1/q)hb^i_{x^i}v^q\\ &\sim (2/q)\eta^k h_{\eta^i}b^i ww_{x^k}+(2/q)hb^i ww_{x^i}.\end{aligned}$$

Hence,

$$hv^{q-1}\check{L}_1 v\sim (2/q)\eta^k h_{\eta^i}b^i ww_{x^k}+(4/q)hb^i ww_{x^i}.$$

We take a number $\lambda\geq 1$, and write $b=\hat{b}+\check{b}$, where $\hat{b}=bI_{|b|>\lambda}$. Following [23], we observe that, by the Hölder and Sobolev inequalities ($d\geq 3$),

$$\begin{aligned}\int_{\mathbb{R}^d}|\hat{b}^i ww_{x^k}|\,dx &\leq \Big(\int_{\mathbb{R}^d}|w_x|^2\,dx\Big)^{1/2}\Big(\int_{\mathbb{R}^d}|\hat{b}|^2|w|^2\,dx\Big)^{1/2}\\ &\leq \Big(\int_{\mathbb{R}^d}|w_x|^2\,dx\Big)^{1/2}\|\hat{b}\|_{L_d}\|w\|_{L_{2d/(d-2)}}\\ &\leq N\|\hat{b}\|_{L_d}\int_{\mathbb{R}^d}|w_x|^2\,dx,\end{aligned} \tag{6.14}$$

where N depends only on d. Since $|\eta||h_\eta|\leq N(\kappa,d)h$, it follows that

$$\eta^k h_{\eta^i}\hat{b}^i ww_{x^k}\prec N\|\hat{b}\|_{L_d}h|w_x|^2.$$

Similarly, $(4/q)h\hat{b}^i ww_{x^i}\prec N\|\hat{b}\|_{L_d}h|w_x|^2$. We estimate the remaining terms in $hv^{q-1}\check{L}_1 v$ roughly like

$$|\check{b}^i ww_{x^k}|\leq\lambda|w||w_x|\leq\varepsilon|w_x|^2+\varepsilon^{-1}\lambda^2|w|^2$$

and conclude that, for any $\varepsilon > 0$,

$$h^q v^{q-1}\check{L}_1 v \prec \big(N\|\hat{b}\|_{L_d}+\varepsilon\big)h|w_x|^2+N\varepsilon^{-1}\lambda^2 h|w|^2. \tag{6.15}$$

Starting to deal with the second order derivatives, note that

$$\begin{aligned} hv^{q-1}(1/2)\sigma^{ik}\sigma^{jk}v_{x^ix^j} &\sim -\big((q-1)/2\big)v^{q-2}h\sigma^{ik}v_{x^i}\sigma^{jk}v_{x^j}\\ &\quad -(1/2)h\big[\sigma^{ik}_{x^i}\sigma^{jk}+\sigma^{ik}\sigma^{jk}_{x^i}\big]v^{q-1}v_{x^j}\\ &= -\big((2q-2)/q^2\big)h\sigma^{ik}w_{x^i}\sigma^{jk}w_{x^j}\\ &\quad -(1/q)h\big[\sigma^{ik}_{x^i}\sigma^{jk}+\sigma^{ik}\sigma^{jk}_{x^i}\big]ww_{x^j}\\ &\le -(1/q)h\sigma^{ik}w_{x^i}\sigma^{jk}w_{x^j}\\ &\quad + h\big|\big[\sigma^{ik}_{x^i}\sigma^{jk}+\sigma^{ik}\sigma^{jk}_{x^i}\big]ww_{x^j}\big|, \end{aligned}$$

where the inequality (to simplify the writing) is due to the fact that $q \ge 2$. In this inequality the first term on the right is dominated in the sense of $\prec$ by

$$-(1/q)\delta h|w_x|^2$$

(see (2.1)). The remaining term contains ww_{x^i}, and we treat it like above writing $\sigma^k_x = \hat{\sigma}^k + \check{\sigma}^k$, where $\hat{\sigma}^k = \sigma^k_x I_{|\sigma^k_x|>\lambda}$. Then, we get

$$\begin{aligned} hv^{q-1}(1/2)\sigma^{ik}\sigma^{jk}v_{x^ix^j} &\prec N\lambda^2\varepsilon^{-1}h|w|^2\\ &\quad -\Big[(1/q)\delta - N\sum_k\|\hat{\sigma}^k\|_{L_d}-\varepsilon\Big]h|w_x|^2. \end{aligned} \tag{6.16}$$

Next,

$$\begin{aligned} hv^{q-1}\sigma^{ik}\sigma^{jk}_{(\eta)}v_{x^i\eta^j} &\sim -(q-1)h\sigma^{ik}v^{q-2}v_{\eta^j}\sigma^{jk}_{(\eta)}v_{x^i}\\ &\quad - v^{q-1}v_{x^i}\big[h_{\eta^j}\sigma^{ik}\sigma^{jk}_{(\eta)}+h\sigma^{ik}\sigma^{jk}_{x^j}\big]\\ &= -\big((4q-4)/q^2\big)h\sigma^{ik}w_{\eta^j}\sigma^{jk}_{(\eta)}w_{x^i}\\ &\quad -(2/q)ww_{x^i}\big[h_{\eta^j}\sigma^{ik}\sigma^{jk}_{(\eta)}+h\sigma^{ik}\sigma^{jk}_{x^j}\big]. \end{aligned}$$

We estimate the first term on the right roughly using

$$\big|\sigma^{ik}w_{\eta^j}\sigma^{jk}_{(\eta)}w_{x^i}\big| \le \varepsilon|w_x|^2+N\varepsilon^{-1}|\eta|\sum_k|\sigma^k_x|^2|w_\eta|^2.$$

The second term contains ww_{x^i} and allows the same handling as before. Therefore,

$$\begin{aligned} hv^{q-1}\sigma^{ik}\sigma^{jk}_{(\eta)}v_{x^i\eta^j} &\prec \Big(\varepsilon+N\sum_k\|\hat{\sigma}^k\|_{L_d}\Big)h|w_x|^2\\ &\quad + N\varepsilon^{-1}h|\eta|\sum_k|\sigma^k_x|^2|w_\eta|^2+N\lambda^2\varepsilon^{-1}h|w|^2. \end{aligned} \tag{6.17}$$

The last term in $hv^{q-1}\check{L}v$ containing σ is

$$\begin{aligned} hv^{q-1}(1/2)\sigma^{ik}_{(\eta)}\sigma^{jk}_{(\eta)}v_{\eta^i\eta^j} &\sim -\big((q-1)/2\big)h\sigma^{ik}_{(\eta)}v^{q-2}v_{\eta^j}\sigma^{jk}_{(\eta)}v_{\eta^i}\\ &\quad -(1/2)v^{q-1}\sigma^{ik}_{(\eta)}v_{\eta^i}\big[h_{\eta^j}\sigma^{jk}_{(\eta)}+h\sigma^{jk}_{x^j}\big]\\ &\quad -\big(1/(2q)\big)h\big(v^q\big)_{\eta^i}\sigma^{ik}_{x^j}\sigma^{jk}_{(\eta)}\\ &\prec Nh\big(|\eta|^2|w_\eta|^2+w^2\big)\sum_k|\sigma^k_x|^2+I, \end{aligned}$$

where

$$I=-(1/(2q))h(w^2)_{\eta^i}\sigma^{ik}_{x^j}\sigma^{jk}_{(\eta)}$$
$$\sim(1/(2q))w^2\sigma^{ik}_{x^j}[h_{\eta^i}\sigma^{jk}_{(\eta)}+h\sigma^{jk}_{x^i}]\prec Nh\sum_k|\sigma^k_x|^2w^2.$$

To estimate the last term, we basically use the derivation of (6.14). We have

$$\int_{\mathbb{R}^d}|\hat\sigma^k_x|^2w^2\,dx\le\|\hat\sigma^k_x\|^2_{L_d}\|w\|^2_{L_{2d/(d-2)}}\le N\|\hat\sigma^k_x\|^2_{L_d}\int_{\mathbb{R}^d}|w_x|^2\,dx. \tag{6.18}$$

Below we show how to choose the constant N_0 in the condition (6.12) under which our assertion is true. But observe that with any such choice, $\|\hat\sigma^k_x\|^2_{L_d}\le 1$, and, therefore, (just to keep some uniformity in our estimates) (6.18) implies that

$$\int_{\mathbb{R}^d}|\hat\sigma^k_x|^2w^2\,dx\le N\|\hat\sigma^k_x\|_{L_d}\int_{\mathbb{R}^d}|w_x|^2\,dx. \tag{6.19}$$

Hence,

$$I\prec Nh\lambda^2w^2+Nh\sum_k\|\hat\sigma^k_x\|_{L_d}|w_x|^2$$

and

$$\begin{aligned}hv^{q-1}(1/2)\sigma^{ik}_{(\eta)}\sigma^{jk}_{(\eta)}v_{\eta^i\eta^j}&\prec Nh|\eta|^2|w_\eta|^2\sum_k|\sigma^k_x|^2\\&\quad+Nh\lambda^2w^2+Nh\sum_k\|\hat\sigma^k_x\|_{L_d}|w_x|^2.\end{aligned} \tag{6.20}$$

Finally,

$$\begin{aligned}&hv^{q-1}(1/2)K_0^2(1+|\eta|^2)\delta^{ij}v_{\eta^i\eta^j}\\&\quad\sim-((2q-2)/q^2)hK_0^2(1+|\eta|^2)|w_\eta|^2\\&\qquad-(2/q)K_0^2(h(1+|\eta|^2))_{\eta^i}ww_{\eta^i}\\&\quad\sim-((2q-2)/q^2)hK_0^2(1+|\eta|^2)|w_\eta|^2+(1/q)w^2K_0^2\delta^{ij}(h(1+|\eta|^2))_{\eta^i\eta^j}\\&\quad\prec-(1/q)hK_0^2(1+|\eta|^2)|w_\eta|^2+Nw^2K_0^2h.\end{aligned} \tag{6.21}$$

By combining (6.15), (6.16), (6.17), (6.20) and (6.21), we see that, for any $\varepsilon\in(0,1]$,

$$\begin{aligned}hv^{q-1}\check Lv&\prec N\varepsilon^{-1}\lambda^2h|w|^2\\&\quad-\Big[(1/q)\delta-N_1\Big(\sum_k\|\hat\sigma^k_x\|_{L_d}+\|\hat b\|_{L_d}\Big)-N_2\varepsilon\Big]h|w_x|^2\\&\quad+|w_\eta|^2h(1+|\eta|^2)\Big[N_3\varepsilon^{-1}\sum_k|\sigma^k_x|^2-(1/q)K_0^2\Big]+N_4w^2K_0^2h.\end{aligned} \tag{6.22}$$

Here, one sees clearly why introducing K_0, which in no way helped us in (6.7), is actually crucial. With $K_0\equiv 0$, we would not be able to estimate the term with $|w_\eta|^2$. Now, take and fix ε so that $N_2\varepsilon\le\delta/(2q)$. After that set

$$K_0^2=1+N_3q\varepsilon^{-1}\sum_k|\sigma^k_x|^2$$

(1 is added to guarantee the smoothness of K_0), and observe that, similarly to (6.18) and (6.19),

$$N_4 w^2 K_0^2 h = N_4 w^2 h + N h w^2 \sum_k |\sigma_x^k|^2 \prec N_5 h \sum_k \|\hat{\sigma}_x^k\|_{L_d} |w_x|^2 + N h \lambda^2 w^2.$$

Then, (6.22) becomes

$$h v^{q-1} \check{L} v \prec N \lambda^2 h |w|^2 - \left[(1/(2q))\delta - (N_1 + N_5) \left(\sum_k \|\hat{\sigma}_x^k\|_{L_d} + \|\hat{b}\|_{L_d} \right) \right] h |w_x|^2.$$

We can certainly believe that $N_1 \geq 1$, take N_0 in (6.12) to be equal to $(2q/\delta)(N_1 + N_5)$ (≥ 1) and conclude

$$h v^{q-1} \check{L} v \prec N \lambda^2 h |w|^2.$$

The lemma is proved. □

We finish the section with an approximation result.

LEMMA 6.6. *Let Assumptions 2.1 and 2.2 be satisfied. Then, there are sequences $\sigma^k(n)$, $b(n)$, $n = 1, 2, \dots$, $k = 1, \dots, d_1$, of infinitely differentiable functions with each derivative bounded having the same meanings as σ^k, b in the beginning of the article, satisfying Assumptions 2.1 and 2.2 with $\delta/2$ in place of δ and the same $\|b\|$ and $\|\sigma_x^k\|$ for sufficiently large n and such that $\sigma^k(n) \to \sigma^k$ as $n \to \infty$ (a.e.) and $\sigma_x^k(n), b(n) \to \sigma_x^k, b$ in L_d as $n \to \infty$.*

PROOF. Take a nonnegative $\zeta \in C_0^\infty$ with unit integral and support in B_1 and set $\zeta_n(x) = n^d \zeta(nx)$, $u(n, x) = u(x) * \zeta_n(x)$. Then, the well-known properties of convolutions imply all stated properties apart from what concerns (2.1).

Denote by σ the $d \times d_1$-matrix whose columns are the σ^k's, and note that

$$|\sigma^*(n, x)\lambda| \leq \zeta_n(x) * |\sigma^*(x)\lambda| \leq \delta^{-1/2}|\lambda|.$$

Therefore, we need only prove that, for sufficiently large n,

$$|\sigma^*(n, x)\lambda| \geq |\lambda|\delta^{-1/2}/\sqrt{2}. \tag{6.23}$$

For any y we have

$$\begin{aligned} |\sigma^*(n, x)\lambda| &\geq |\sigma^*(y)\lambda| - |(\sigma^*(n, x) - \sigma^*(y))\lambda| \\ &\geq |\lambda|\delta^{1/2} - |(\sigma^*(n, x) - \sigma^*(y))\lambda| \\ &\geq |\lambda|(\delta^{1/2} - |\sigma^*(n, x) - \sigma^*(y)|). \end{aligned}$$

Furthermore,

$$\begin{aligned} &\int_{\mathbb{R}^d} |\sigma^*(n, x) - \sigma^*(x - y)| \zeta_n(y)\, dy \\ &\quad \leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |\sigma^*(x - z) - \sigma^*(x - y)| \zeta_n(y) \zeta_n(z)\, dy\, dz \\ &\quad \leq N(d) \max \zeta^2 \operatorname{osc}(\sigma, B_{1/n}(x)). \end{aligned}$$

The latter tends to zero uniformly with respect to x since $\sigma_x \in L_d$ (cf. Remark 3.6). This certainly proves the lemma. □

7. Proof of Theorem 2.3. According to Theorems 5.6 and 5.7, it suffices to prove that at least one of the solutions of (1.1) is strong. We will be dealing with the solution from Lemma 4.1.

Let $f \in C_0^\infty$. First, we deal with smooth coefficients and develop necessary estimates. By Lemma 6.3 and Theorem 6.4 for $t \geq 0$ and $q \geq 2$, we have

$$\int_{\mathbb{R}^{2d}} h(\eta) u^q(t,x,\eta)\,dx\,d\eta \leq N e^{Nt}, \tag{7.1}$$

where (and below) N depends only on $f, d, d_1, \delta, m = m(f), q$ and λ, defined by (6.10) and

$$u(t,x,\eta) := \sum_{n=1}^{\infty} \sum_{k_1,\dots,k_n} \int_{t>t_1>\cdots>t_n} \big[\big(T_{t_n} Q^{k_n}_{t_{n-1}-t_n} \cdots Q^{k_1}_{t-t_1} f(x)\big)_{(\eta)}\big]^2\,dt_n \cdots dt_1.$$

Obviously, $u(t,x,\eta)$ is a quadratic function of η. Hence, (7.1) implies that, for any $R \in (0,\infty)$,

$$\int_{\mathbb{R}^d} \sup_{|\eta|\leq R} u^q(t,x,\eta)\,dx \leq N e^{Nt} R^{2q}. \tag{7.2}$$

Observe that in notations (5.5) and (5.8),

$$\begin{aligned}\sum_k u(t,x,\sigma^k) &= \sum_{n=1}^{\infty} \int_{t>t_1>\cdots>t_n} Q_{t_n,t_{n-1}-t_n,\dots,t-t_1} f(x)\,dt_n \cdots dt_1 \\ &= \sum_{n=1}^{\infty} \int_{S_n(t)} Q_{s_n,s_{n-1},\dots,s_1,t-(s_1+\cdots+s_n)} f(x)\,ds_n \cdots ds_1 \\ &=: \sum_{n=1}^{\infty} I_n(t,x)\end{aligned}$$

($S_n(t)$ is introduced in Remark 5.8). Next, for $\nu > 0$, by Jensen's inequality

$$\begin{aligned}\sum_{n=1}^{\infty} \int_{\mathbb{R}^d} \left(\int_0^\infty e^{-\nu t} I_n(t,x)\,dt\right)^q dx &\leq \nu^{1-q} \int_0^\infty e^{-\nu t} \left(\sum_{n=1}^{\infty} \int_{\mathbb{R}^d} I_n^q(t,x)\,dx\right) dt \\ &\leq \nu^{1-q} \int_0^\infty e^{-\nu t} \int_{\mathbb{R}^d} \left(\sum_k u(t,x,\sigma^k)\right)^q dx\,dt,\end{aligned}$$

which, thanks to (7.2), implies that for appropriate ν, depending only on f, d, d_1, δ, q and λ,

$$\sum_{n=1}^{\infty} \int_{\mathbb{R}^d} \left(\int_0^\infty e^{-\nu t} I_n(t,x)\,dt\right)^q dx \leq N, \tag{7.3}$$

where N depends only only on f, d, d_1, δ, q and λ.

Estimate (7.3) has been derived only for infinitely differentiable σ and b. However, using smooth approximations (Lemma 6.6), Theorem 3.10 and Fatou's lemma prove (7.3) also in our general case. Indeed, although the constant N in (7.3) for each approximation depends on λ, satisfying (6.10) for the approximating σ_x^k and b, it can be taken the same as long as the approximations are sufficiently close in L_d to the original σ_x^k and b.

Finally, by observing that

$$\int_0^\infty e^{-\nu t} I_n(t,x)\,dt = \int_{R_+^{n+1}} e^{-\nu(s_0+\cdots+s_n)} Q_{s_n,\dots,s_0} f(x)\,ds_n \cdot \cdots \cdot ds_0,$$

referring to Theorem 5.9 and taking $q = p$, we conclude that $f(x_t)$ is $\mathcal{F}_t^w$-measurable. The arbitrariness of f and t finishes the proof.

8. Proof of Theorem 2.4. Take a bounded smooth function f with compact support. By Theorems 2.3 and 5.5, for any t,

$$\begin{aligned} & f\big(x_t(n,x(n))\big) \\ & \quad = T_t(n)f\big(x(n)\big) \\ & \qquad + \sum_{m=1}^{\infty}\int_{t>t_1>\cdots>t_m} T_{t_m}(n)Q^{k_m}_{t_{m-1}-t_m}(n)\cdots Q^{k_1}_{t-t_1}(n)f\big(x(n)\big)\,dw^{k_m}_{t_m}\cdots dw^{k_1}_{t_1}, \end{aligned} \tag{8.1}$$

where $T_t(n)$ and $Q^k_t(n)$ are the operators corresponding to $\sigma^k(n)$, $b(n)$. First, we prove that $E|f(x_t(n,x(n))) - f(x_t)|^2 \to 0$ as $n \to \infty$. Since $Ef^2(x_t(n,x(n))) \to Ef^2(x_t)$ (see Theorem 3.10), it suffices to prove that $f(x_t(n,x(n))) \to f(x_t)$ weakly in $L_2(\Omega, \mathcal{F}^w_t, P)$. Furthermore, according to [6], the linear combinations of the multiple Itô integrals of the type

$$\int_{t>t_1>\cdots>t_m} \phi(t_1,\dots,t_m)\,dw_{t_m}\cdots dw_{t_1},$$

where m is arbitrary and ϕ is an arbitrary bounded (nonrandom) Borel function, are dense in $L_2(\Omega,\mathcal{F}^w_t,P)$. Therefore, it suffices to prove that, for all such m and ϕ,

$$\begin{aligned} & Ef\big(x_t(n,x(n))\big)\int_{t>t_1>\cdots>t_m}\phi(t_1,\dots,t_m)\,dw_{t_m}\cdots dw_{t_1} \\ & \qquad \to Ef(x_t)\int_{t>t_1>\cdots>t_m}\phi(t_1,\dots,t_m)\,dw_{t_m}\cdots dw_{t_1}. \end{aligned}$$

In light of (8.1), this is equivalent to proving that

$$\begin{aligned} & \int_{t>t_1>\cdots>t_m}\phi(t_1,\dots,t_m)T_{t_m}(n)Q^{k_m}_{t_{m-1}-t_m}(n)\cdots Q^{k_1}_{t-t_1}(n)f\big(x(n)\big)\,dt_m\cdots dt_1 \\ & \quad \to \int_{t>t_1>\cdots>t_m}\phi(t_1,\dots,t_m)T_{t_m}Q^{k_m}_{t_{m-1}-t_m}\cdots Q^{k_1}_{t-t_1}f(x_0)\,dt_m\cdots dt_1. \end{aligned}$$

This relation is indeed true, which follows by the dominated convergence theorem from Theorem 3.10 and Remark 5.3.

Next, observe that, for any $T \in (0,\infty)$ and bounded smooth $\mathbb{R}^d$-valued $\tilde b$ with compact support,

$$\begin{aligned} I &:= \varlimsup_{n\to\infty} E\sup_{t\le T}\left|\int_0^t b\big(n,x_s(n,x(n))\big)\,ds - \int_0^t b(x_s)\,ds\right| \\ &\le \varlimsup_{n\to\infty} E\int_0^T \big|b\big(n,x_s(n,x(n))\big) - b(x_s)\big|\,ds \\ &\le \varlimsup_{n\to\infty} E\int_0^T \big|b\big(n,x_s(n,x(n))\big) - \tilde b\big(x_s(n,x(n))\big)\big|\,ds \\ &\quad + \varlimsup_{n\to\infty}\int_0^T E\big|\tilde b\big(x_s(n,x(n))\big) - \tilde b(x_s)\big|\,ds \\ &\quad + \varlimsup_{n\to\infty} E\int_0^T \big|\tilde b(x_s) - b(x_s)\big|\,ds. \end{aligned}$$

Here, the middle term vanishes by the first part of the proof. Owing to Lemma 4.2, the two remaining terms are majorated by

$$N\Big(\lim_{n\to\infty}\|b_n - \tilde b\|_{L_d} + \|b - \tilde b\|_{L_d}\Big) = N\|b - \tilde b\|_{L_d}$$

that can be made arbitrarily small by an appropriate choice of $\tilde{b}$. Hence, $I=0$.

Similarly,

$$J := \varlimsup_{n\to\infty}\Big(E\sup_{t\le T}\Big|\int_0^t \sigma^k(n,x_s(n,x(n)))\,dw_s^k - \int_0^t \sigma^k(x_s)\,dw_s^k\Big|\Big)^2$$
$$\le N\varlimsup_{n\to\infty}\sum_k E\int_0^T |\sigma^k(n,x_s(n,x(n))) - \sigma^k(x_s)|^2\,ds$$
$$\le N\varlimsup_{n\to\infty}\sum_k E\int_0^T |\sigma^k(n,x_s(n,x(n))) - \sigma^k(x_s)|\,ds$$
$$\le N\sum_k \|\Phi_T(\sigma^k - \hat{\sigma}^k)\|_{L_d},$$

where $\hat{\sigma}^k$ are smooth functions with compact support. It follows that $J=0$, and, together with $I=0$, this implies that

$$\lim_{n\to\infty} E\sup_{t\le T}|x_t(n,x(n)) - x_t| = 0. \tag{8.2}$$

By Corollary 1.2 of [12], for any $m\ge 0$,

$$E\sup_{t\le T}|x_t(n,x(n)) - x(n)|^{2m} + E\sup_{t\le T}|x_t - x_0|^{2m} \le N(m,d,\delta,\|b\|)T^m,$$

and this along with (8.2) yields the result. The theorem is proved.

9. Proof of Theorem 2.5. First, we assume that σ^k and b are infinitely differentiable with each derivative bounded. In that case, it is known since [3] (see also [17]) that one can define $x_t(x)$ in such a way that it becomes differentiable in x for all (ω,t) and the derivative of x_t in the direction of η satisfies the same equation as $\xi_t(x,\eta)$ from Lemma 6.2, for which (6.5) holds. In particular, for any even $\kappa\ge 2$ and f with compact support ($(\cdot,\cdot)$ is the scalar product in $\mathbb{R}^d$),

$$E\big((Df)(x_t(x)),\eta_t(x,\eta)\big)^\kappa \ge E\big((Df)(x_t(x)),\xi_t(x,\eta)\big)^\kappa =: v(t,x,\eta).$$

By Theorem 6.4 with $q=2$, there is a constant $m=m(\kappa)$ such that, for any $\lambda>0$ satisfying (6.10), there exists a constant N, depending only on λ, d, δ, m, such that, for $t\ge 0$,

$$\begin{aligned}&\int_{\mathbb{R}^{2d}} h(\eta)v^2(t,x,\eta)\,dx\,d\eta\\ &\quad\le e^{Nt}\int_{\mathbb{R}^{2d}} h(\eta)|f_{(\eta)}(x)|^{2\kappa}\,dx\,d\eta = N(d,\kappa)e^{Nt}\int_{\mathbb{R}^d}|f_x(x)|^{2\kappa}\,dx =: M_t.\end{aligned} \tag{9.1}$$

Next, for any $R\in(0,\infty)$,

$$\begin{aligned}E\int_{B_R}|D(f(x_t(x)))|^\kappa\,dx &= NE\int_{B_R}\int_{\mathbb{R}^d}((f(x_t(x))_{(\eta)})^\kappa h(\eta)\,d\eta\,dx\\ &= N\int_{B_R}\int_{\mathbb{R}^d} v(t,x,\eta)h(\eta)\,d\eta\,dx.\end{aligned}$$

By using (9.1) and Hölder's inequality we obtain that

$$E\int_{B_R}|D(f(x_t(x)))|^\kappa\,dx \le N(d,\kappa)R^{d/2}M_t^{1/2}.$$

By Morrey's theorem (see, for instance, Theorem 10.2.1 of [9]) this implies that, for any $\kappa > d$,

$$E \sup_{x,y\in B_R} \frac{|f(x_t(x)) - f(x_t(y))|^\kappa}{|x-y|^{\kappa-d}} \le N(d,\kappa)R^{d/2}M_t^{1/2}. \tag{9.2}$$

Note that (9.2) is certainly applicable to vector-valued f. Fix $\rho \ge 2R$, and take a smooth f with support in $B_{4\rho}$ such that $f(x) = x$ for $|x| \le 2\rho$ and $|f_x| \le 2$. Then, $M_t \le N(T,d,\delta,\lambda,\kappa)\rho^d$ and for $x, y \in B_R$ and $t \le T$

$$\begin{aligned} E\big|x_t(x) - x_t(y)\big|^\kappa &\le N\rho^{d/2}|x-y|^{\kappa-d} \\ &\quad + N(\kappa)E\big(\big|x_t(x)\big|^\kappa + \big|x_t(y)\big|^\kappa\big)I_{|x_t(x)|+|x_t(y)|\ge 2\rho}, \end{aligned}$$

where N depends only on R, T, λ, d, δ and κ. We estimate the second term on the right by using Hölder's inequality and Theorem 2.10 of [12] and find that it is dominated by

$$\begin{aligned} &N(\kappa)P^{1/2}\big(\big|x_t(x)\big| + \big|x_t(y)\big| \ge 2\rho\big)\big(\big(E\big|x_t(x)\big|^{2\kappa}\big)^{1/2} + \big(E\big|x_t(y)\big|^{2\kappa}\big)^{1/2}\big) \\ &\quad \le Ne^{-\mu\rho^2}, \end{aligned}$$

where N depends only on R, T, λ, d, δ and κ and $\mu > 0$ depends only on T, d, δ and $\|b\|$. Thus, for $\rho \ge 2R$,

$$E\big|x_t(x) - x_t(y)\big|^\kappa \le Ne^{-\mu\rho^2} + N\rho^{d/2}|x-y|^{\kappa-d}.$$

By taking here $\kappa > 2d$ and $\mu\rho^2 = -\ln|x-y|^{\kappa-2d}$, we find that

$$E\big|x_t(x) - x_t(y)\big|^\kappa \le N|x-y|^{\kappa-2d}, \tag{9.3}$$

where N depends only on $R, T, \lambda, d, \delta, \|b\|$ and κ, provided that $-\ln|x-y|^{\kappa-2d} \ge 2R/\mu$. However, if $-\ln|x-y|^{\kappa-2d} \le 2R/\mu$, (9.3) is obvious.

Estimate (9.3) so far is proved only for infinitely differentiable coefficients, but usual approximations, Theorem 2.4, and Fatou's lemma, allow us to obtain (9.3) in our general case where, naturally, by $x_t(x)$ we mean the strong solution of (1.1) with $x_0 = x$.

In the general case we also have by Corollary 1.2 of [12] that, for any $q \ge 1$,

$$E\big|x_t(x) - x_s(x)\big|^q \le N|t-s|^{q/2}, \tag{9.4}$$

where $N = N(d, \delta, \|b\|, q)$.

Now, the arbitrariness of κ and q leads to the claimed result by a version of Kolmogorov's theorem which can be found, for instance, in [20] or simply derived by using $|x_t(x) - x_s(y)| \le |x_t(x) - x_s(x)| + |x_s(x) - x_s(y)|$. The theorem is proved.

Acknowledgments. The author is sincerely grateful to the referees for their comments which helped correct some glitches and overall improve the presentation.

REFERENCES

[1] AGMON, S. (1962). On the eigenfunctions and on the eigenvalues of general elliptic boundary value problems. *Comm. Pure Appl. Math.* **15** 119–147. MR0147774 https://doi.org/10.1002/cpa.3160150203

[2] BECK, L., FLANDOLI, F., GUBINELLI, M. and MAURELLI, M. (2019). Stochastic ODEs and stochastic linear PDEs with critical drift: Regularity, duality and uniqueness. *Electron. J. Probab.* **24** 136. MR4040996 https://doi.org/10.1214/19-ejp379

[3] BLAGOVEŠČENSKIĬ, J. N. and FREĬDLIN, M. I. (1961). Some properties of diffusion processes depending on a parameter. *Dokl. Akad. Nauk SSSR* **138** 508–511; in Russian, English translation in Soviet Math. Dokl., Vol. 2 (1961), 633–636. MR0139196

[4] DONG, H. and KIM, D. (2011). On the L_p-solvability of higher order parabolic and elliptic systems with BMO coefficients. *Arch. Ration. Mech. Anal.* **199** 889–941. MR2771670 https://doi.org/10.1007/s00205-010-0345-3

[5] HILLE, E. (1948). *Functional Analysis and Semi-Groups. American Mathematical Society Colloquium Publications* **31**. Amer. Math. Soc., New York.

[6] ITÔ, K. (1951). Multiple Wiener integral. *J. Math. Soc. Japan* **3** 157–169. MR0044064 https://doi.org/10.2969/jmsj/00310157

[7] KRYLOV, N. V. (1977). *Controlled Diffusion Processes*. Nauka, Moscow. in Russian; English translation by Springer, 1980.

[8] KRYLOV, N. V. (1984). Extremal properties of solutions of stochastic equations. *Teor. Veroyatn. Primen.* **29** 209–221. in Russian; English translation in Theory Probab. Appl., Vol. 29 (1984), No. 2, 205–217. MR0749910

[9] KRYLOV, N. V. (2008). *Lectures on Elliptic and Parabolic Equations in Sobolev Spaces. Graduate Studies in Mathematics* **96**. Amer. Math. Soc., Providence, RI. MR2435520 https://doi.org/10.1090/gsm/096

[10] KRYLOV, N. V. (2011). On the Itô–Wentzell formula for distribution-valued processes and related topics. *Probab. Theory Related Fields* **150** 295–319. MR2800911 https://doi.org/10.1007/s00440-010-0275-x

[11] KRYLOV, N. V. (2020). On stochastic equations with drift in L_d. Available at http://arxiv.org/abs/2001.04008.

[12] KRYLOV, N. V. (2020). On stochastic Itô processes with drift in L_d. http://arxiv.org/abs/2001.03660.

[13] KRYLOV, N. V. (2021). On diffusion processes with drift in L_d. *Probab. Theory Related Fields* **179** 165–199. MR4221656 https://doi.org/10.1007/s00440-020-01007-3

[14] KRYLOV, N. V. (2021). Elliptic equations with VMO A, B $\in L_d$, and C $\in L_{d/2}$. *Trans. Amer. Math. Soc.* **374** 2805–2822. MR4223034 https://doi.org/10.1090/tran/8282

[15] 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 https://doi.org/10.1007/s00440-004-0361-z

[16] KRYLOV, N. V. and ZVONKIN, A. K. (1981). On strong solutions of stochastic differential equations. *Sel. Math. Sov.* **1** 19–61.

[17] KUNITA, H. (1990). *Stochastic Flows and Stochastic Differential Equations. Cambridge Studies in Advanced Mathematics* **24**. Cambridge Univ. Press, Cambridge. MR1070361

[18] NAKAO, S. (1972). On the pathwise uniqueness of solutions of one-dimensional stochastic differential equations. *Osaka Math. J.* **9** 513–518. MR0326840

[19] PAZY, A. (1983). *Semigroups of Linear Operators and Applications to Partial Differential Equations. Applied Mathematical Sciences* **44**. Springer, New York. MR0710486 https://doi.org/10.1007/978-1-4612-5561-1

[20] POTTHOFF, J. (2009). Sample properties of random fields. II. Continuity. *Commun. Stoch. Anal.* **3** 331–348. MR2604006 https://doi.org/10.31390/cosa.3.3.02

[21] SHAPOSHNIKOV, A. (2017). Correction to the paper "Some remarks on Davie's uniqueness theorem". Available at arXiv:1703.06598.

[22] SHAPOSHNIKOV, A. V. (2016). Some remarks on Davie's uniqueness theorem. *Proc. Edinb. Math. Soc.* (2) **59** 1019–1035. MR3570126 https://doi.org/10.1017/S0013091515000589

[23] STAMPACCHIA, G. (1965). Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus. *Ann. Inst. Fourier* (*Grenoble*) **15** 189–257. MR0192177

[24] VERETENNIKOV, A. Y. (1980). Strong solutions and explicit formulas for solutions of stochastic integral equations. *Mat. Sb.* **111** 434–452. in Russian, English translation in Math. USSR-Sbornik, Vol. 39 (1981), No. 3, 387–403.

[25] VERETENNIKOV, A. Y. and KRYLOV, N. V. (1976). On explicit formulas for solutions of the stochastic equations. *Mat. Sb.* **100** 266–284. in Russian; English translation in Math. USSR Sbornik Vol. 29 (1976), No. 2, 239–256.

[26] XIE, L. and ZHANG, X. (2020). Ergodicity of stochastic differential equations with jumps and singular coefficients. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** 175–229. MR4058986 https://doi.org/10.1214/19-AIHP959

[27] YAMADA, T. and WATANABE, S. (1971). On the uniqueness of solutions of stochastic differential equations, I, II. *J. Math. Kyoto Univ.* **11** 155–167; (1971), **11** 553–563.

[28] ZHANG, X. (2011). Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients. *Electron. J. Probab.* **16** 1096–1116. MR2820071 https://doi.org/10.1214/EJP.v16-887

[29] ZVONKIN, A. K. (1974). A transformation of the phase space of a diffusion process that will remove the drift. *Mat. Sb.* **93** 129–149. In Russian, English translation in Math. USSR-Sbornik, Vol. 22 (1974), No. 1, 129–149.