ON DIVERGENCE FORM SPDES WITH VMO COEFFICIENTS

N.V. KRYLOV

Abstract. We present several results on solvability in Sobolev spaces W^1_p of SPDEs in divergence form in the whole space.

1. Introduction

The theory of (usual) partial differential equations has two rather different parts depending on whether the equations are written in divergence or nondivergence form. Quite often the starting point is the same: equations with constant coefficients, and then one uses different techniques to treat different types of equations.

By now one, can say that the L_p-theory of evolutional second order SPDEs is quite well developed. The most advanced results of this theory can be found in the following papers and references therein: [2] (nondivergence type equations), [3] and [4] (divergence type equations). The results of the present paper are close to the corresponding results of [3]. However, unlike [3] we do not assume that the leading coefficients are continuous in the space variable. Instead we assume that the leading coefficients of the "deterministic" part of the equation are in VMO which is a much wider class than C. Still the leading coefficients of the "stochastic" part are assumed to be continuous in x.

The exposition in [3] and [4] is based on the theory of solvability in spaces $H^\gamma_p=(1-\Delta)^{-\gamma/2}L_p$ of SPDEs with coefficients independent of x. Then the method of "freezing" the coefficients is applied as in the general framework set out in [7]. This method does not work if the coefficients are only in VMO and we use a different technique based on recent results from [9] on deterministic parabolic equations with VMO coefficients. In addition, our technique allows us to avoid using the W^n_2-theory of SPDEs, which is a starting point in the paper [7] and subsequent articles based on it.

One more difference of our approach from the one in [3] is that we represent the free term in the deterministic part in the form $D_if^i+f^0$ with $f^j\in L_p$ (see (1.1) below) . Of course, this is just a general form of a distribution from H^{-1}_p. However, the spaces H^γ_p are most appropriate for equations

1991 *Mathematics Subject Classification.* 60H15, 35R60.

Key words and phrases. Stochastic partial differential equations, divergence equations, Sobolev spaces.

The work was partially supported by NSF Grant DMS-0653121.

in nondivergence form. One general inconvenience of these spaces is that the space or space-time dilations affect the norms in a way which is hard to control. For divergence form equations with low regularity of coefficients the most important space is H^1_p. This space coincides with the Sobolev space W^1_p and the effect of dilations on the norm or on $D_if^i+f^0$ can be easily taken into account.

The exposition here is self-contained apart from references to some very basic results of [7], [9], and [14] and is much more elementary than in [3], employing the derivatives instead of the powers of the Laplacian, and yet gives more information. In particular, the author intends to use Corollary 5.5 in order to largely simplify the theory in [3] of divergence form SPDEs in domains. It turns out that to develop this theory one need not first develop the theory of SPDEs in domains with coefficient independent of x, which in itself required quite a bit of work.

The author's interest in divergence type equations and in simplifying the theory of them appeared after he realized that the corresponding results can be applied to filtering theory of partially observable diffusion processes, given by stochastic Itô equations, and proving that, under Lipschitz and nondegeneracy conditions only, the filtering density is almost Lipschitz in x and almost Hölder $1/2$ in time. This is proved in [12] on the basis of Theorems 2.2 through 2.6 of the present article. The filtering density satisfies an SPDE usually written in terms of the operators adjoint to operators in nondivergence form with Lipschitz continuous coefficients. Writing these adjoint operators in divergence form makes perfect sense and allows us to obtain the above mentioned results (see [12]).

Our Theorem 2.2 is very close to Theorem 2.12 of [3]. Apart from weaker conditions on the coefficients, another important difference is the presence of the parameter λ in (2.7). One of differences in the proofs is that we avoid proving the solvability on small consecutive time intervals and then gluing together the results.

Let $(\Omega,\mathcal{F},P)$ be a complete probability space with an increasing filtration $\{\mathcal{F}_t, t\geq 0\}$ of complete with respect to $(\mathcal{F},P)$ σ-fields $\mathcal{F}_t\subset\mathcal{F}$. Denote by $\mathcal{P}$ the predictable σ-field in $\Omega\times(0,\infty)$ associated with $\{\mathcal{F}_t\}$. Let w^k_t, $k=1,2,...$, be independent one-dimensional Wiener processes with respect to $\{\mathcal{F}_t\}$.

We fix a stopping time τ and for $t\leq\tau$ in the Euclidean d-dimensional space $\mathbb{R}^d$ of points $x=(x^1,...,x^d)$ we are considering the following equation

$$du_t=(L_tu_t-\lambda u_t+D_if^i_t+f^0_t)\,dt+(\Lambda^k_tu_t+g^k_t)\,dw^k_t, \tag{1.1}$$

where $u_t=u_t(x)=u_t(\omega,x)$ is an unknown function,

$$L_t\psi(x)=D_j\big(a^{ij}_t(x)D_i\psi(x)+a^j_t(x)\psi(x)\big)+b^i_t(x)D_i\psi(x)+c_t(x)\psi(x),$$

$$\Lambda^k_t\psi(x)=\sigma^{ik}_t(x)D_i\psi(x)+\nu^k_t(x)\psi(x),$$

the summation convention with respect to $i,j=1,...,d$ and $k=1,2,...$ is enforced and detailed assumptions on the coefficients and the free terms will be given later.

One can rewrite (1.1) in the nondivergence form assuming that the coefficients a^{ij}_t and a^j_t are differentiable in x and then one could apply the results from [7]. It turns out that the differentiability of a^{ij}_t and a^j_t is not needed for the corresponding counterparts of the results in [7] to be true and showing this and generalizing the corresponding results of [3] is one of the main purposes of the present article.

2. Main results

Fix a number

$$p\geq 2,$$

and denote $L_p=L_p(\mathbb{R}^d)$. We use the same notation L_p for vector- and matrix-valued or else ℓ_2-valued functions such as $g_t=(g^k_t)$ in (1.1). For instance, if $u(x)=(u^1(x),u^2(x),...)$ is an ℓ_2-valued measurable function on $\mathbb{R}^d$, then

$$\|u\|^p_{L_p}=\int_{\mathbb{R}^d}|u(x)|^p_{\ell_2}\,dx=\int_{\mathbb{R}^d}\big(\sum_{k=1}^{\infty}|u^k(x)|^2\big)^{p/2}\,dx.$$

Introduce

$$D_i=\frac{\partial}{\partial x^i},\quad i=1,...,d,\quad \Delta=D_1^2+...+D_d^2.$$

By Du we mean the gradient with respect to x of a function u on $\mathbb{R}^d$.

As usual,

$$W^1_p=\{u\in L_p:Du\in L_p\},\quad \|u\|_{W^1_p}=\|u\|_{L_p}+\|Du\|_{L_p}.$$

Recall that τ is a stopping time and introduce

$$\mathbb{L}_p(\tau):=L_p((0,\tau]\!],\mathcal{P},L_p),\quad \mathbb{W}^1_p(\tau):=L_p((0,\tau]\!],\mathcal{P},W^1_p).$$

We also need the space $\mathcal{W}^1_p(\tau)$, which is the space of functions $u_t=u_t(\omega,\cdot)$ on $\{(\omega,t):0\leq t\leq\tau,t<\infty\}$ with values in the space of generalized functions on $\mathbb{R}^d$ and having the following properties:

(i) We have $u_0\in L_p(\Omega,\mathcal{F}_0,L_p)$;

(ii) We have $u\in\mathbb{W}^1_p(\tau)$;

(iii) There exist $f^i\in\mathbb{L}_p(\tau)$, $i=0,...,d$, and $g=(g^1,g^2,...)\in\mathbb{L}_p(\tau)$ such that for any $\varphi\in C^\infty_0$ with probability 1 for all $t\in[0,\infty)$ we have

$$(u_{t\wedge\tau},\varphi)=(u_0,\varphi)+\sum_{k=1}^{\infty}\int_0^t I_{s\leq\tau}(g^k_s,\varphi)\,dw^k_s$$

$$+\int_0^t I_{s\leq\tau}\big((f^0_s,\varphi)-(f^i_s,D_i\varphi)\big)\,ds. \tag{2.1}$$

In particular, for any $\phi\in C_0^\infty$, the process $(u_{t\wedge\tau},\phi)$ is $\mathcal{F}_t$-adapted and (a.s.) continuous.

The reader can find in [7] a discussion of (ii) and (iii), in particular, the fact that the series in (2.1) converges uniformly in probability on every finite subinterval of $[0,\tau]$. On the other hand, it is worth saying that the above introduced space $\mathcal{W}^1_p$ is not quite the same as $\mathcal{H}^1_p(\tau)$ in [7] or in [3]. There are three differences. One is that there is an additional restriction on u_0 in [7] and [3]. But in the main part of the article we are going to work with $\mathcal{W}^1_{p,0}(\tau)$ which is the subset of $\mathcal{W}^1_p(\tau)$ consisting of functions with $u_0=0$. Another issue is that in [7] and [3] we have $f^i=0$, $i=1,...,d$, and

$$f^0\in\mathbb{H}^{-1}_p(\tau)=L_p((0,\tau]],\mathcal{P},H^{-1}_p).$$

Actually, this difference is fictitious because one knows that any $f\in H^{-1}_p$

(a) has the form $D_if^i+f^0$ with $f^j\in L_p$ and

$$\|f\|_{H^{-1}_p}\leq N\sum_{j=0}^d\|f^j\|_{L_p},$$

where N is independent of f,f^j, and on the other hand,

(b) for any $f\in H^{-1}_p$ there exist $f^j\in L_p$ such that $f=D_if^i+f^0$ and

$$\sum_{j=0}^d\|f^j\|_{L_p}\leq N\|f\|_{H^{-1}_p},$$

where N is independent of f.

The third difference is that instead of (i) the condition $D^2u\in\mathbb{H}^{-1}_p(\tau)$ is required in in [7] and [3]. However, as it follows from Theorem 3.7 of [7] and the boundedness of the operator $D:L_p\to H^{-1}_p$, this difference disappears if τ is a bounded stopping time.

To summarize, the spaces $\mathcal{W}^1_{p,0}(\tau)$ introduced above coincide with $\mathcal{H}^1_{p,0}(\tau)$ from [7] if τ is bounded and we choose a particular representation of the deterministic part of the stochastic differential just for convenience. In the remainder of the article the spaces $\mathcal{H}^1_{p,0}(\tau)$ do not appear and none of their properties is used.

In case that property (iii) holds, we write

$$du_t=(D_if^i_t+f^0_t)\,dt+g^k_t\,dw^k_t \tag{2.2}$$

for $t\leq\tau$ and this explains the sense in which equation (1.1) is understood. Of course, we still need to specify appropriate assumptions on the coefficients and the free terms in (1.1).

Assumption 2.1. (i) The coefficients a^{ij}_t, a^i_t, b^i_t, σ^{ik}_t, c_t, and ν^k_t are measurable with respect to $\mathcal{P}\times\mathcal{B}(\mathbb{R}^d)$, where $\mathcal{B}(\mathbb{R}^d)$ is the Borel σ-field on $\mathbb{R}^d$.

(ii) There is a constant K such that for all values of indices and arguments

$$|a^i_t|+|b^i_t|+|c_t|+|\nu|_{\ell_2}\leq K,\quad c_t\leq 0.$$

(iii) There is a constant $\delta > 0$ such that for all values of the arguments and $\xi \in \mathbb{R}^d$

$$a_t^{ij}\xi^i\xi^j \leq \delta^{-1}|\xi|^2, \quad (a_t^{ij} - \alpha_t^{ij})\xi^i\xi^j \geq \delta|\xi|^2, \tag{2.3}$$

where $\alpha_t^{ij} = (1/2)(\sigma^{i\cdot}, \sigma^{j\cdot})_{\ell_2}$. Finally, the constant $\lambda \geq 0$.

It is worth emphasizing that we do not require the matrix (a^{ij}) to be symmetric.

Assumption 2.1 (i) guarantees that equation (1.1) makes perfect sense if $u \in \mathcal{W}_p^1(\tau)$. By the way, adding the term $-\lambda u_t$ with constant $\lambda \geq 0$ is one more technically convenient step. One can always introduce this term, if originally it is absent, by considering $v_t := u_t e^{\lambda t}$.

For functions $h_t(x)$ on $[0, \infty) \times \mathbb{R}^d$ and balls B in $\mathbb{R}^d$ introduce

$$h_{t(B)} = \frac{1}{|B|}\int_B h_t(x)\,dx,$$

where $|B|$ is the volume of B. Also let $\mathbb{Q}$ denote the set of all cylinders in $[0, \infty) \times \mathbb{R}^d$ of type $Q = (s, t) \times B$, where B is a ball in $\mathbb{R}^d$ and $t - s = \rho^2(Q)$, where $\rho(Q)$ is the radius of B. If $Q = (s, t) \times B \in \mathbb{Q}$, set

$$\operatorname{ocs}(h, Q) = \frac{1}{t-s}\int_s^t (|h_r - h_{r(B)}|)_{(B)}\,dr.$$

Finally, introduce the integral oscillation of h

$$\operatorname{ocs}_\varepsilon h = \sup_{Q \in \mathbb{Q}: \rho(Q) \leq \varepsilon} \operatorname{ocs}(h, Q).$$

Observe that $\operatorname{ocs}_\varepsilon u = 0$ if $u_t(x)$ is independent of x.

Assumption 2.2. For a constant $\beta > 0$, which will be specified later, there exists an $\varepsilon > 0$ such that

$$\sup_{|x-y| \leq \varepsilon} |\sigma_t^{i\cdot}(x) - \sigma_t^{i\cdot}(y)|_{\ell_2} + \operatorname{ocs}_\varepsilon a^{ij} \leq \beta$$

for all i, j, t.

Finally, we describe the space of initial data. Recall that for $p \geq 2$ the Slobodetskii space $W_p^{1-2/p} = W_p^{1-2/p}(\mathbb{R}^d)$ of functions $u_0(x)$ can be introduced as the space of traces on $t = 0$ of (deterministic) functions u such that

$$u \in L_p(\mathbb{R}_+, H_p^1), \quad \partial u/\partial t \in L_p(\mathbb{R}_+, H_p^{-1}),$$

where $\mathbb{R}_+ = (0, \infty)$. For such functions there is a (unique) modification denoted again u such that u_t is a continuous L_p-valued function on $[0, \infty)$ so that u_0 is well defined. Any such u_t is called an extension of u_0.

The norm in $W_p^{1-2/p}$ can be defined as the infimum of

$$\|u\|_{L_p(\mathbb{R}_+, H_p^1)} + \|\partial u/\partial t\|_{L_p(\mathbb{R}_+, H_p^{-1})}$$

over all extensions u_t of elements u_0. It is also well known that an equivalent norm of u_0 can be introduced as

$$\|u\|_{L_p((0,1),W^1_p)},$$

where $u = u_t$ is defined as the (unique) solution of the heat equation $\partial u_t(x)/\partial t = \Delta u_t(x)$ with initial condition $u_0(x)$.

For $s \geq 0$ we introduce

$$\mathrm{tr}_s \mathcal{W}^1_p = L_p(\Omega, \mathcal{F}_s, W^{1-2/p}_p).$$

The following auxiliary result helps understand the role of $\mathrm{tr}_s \mathcal{W}^1_p$. We use spaces $\mathcal{W}^1_p([S,T))$ and $\mathbb{W}^1_p((S,T))$, which are introduced in the same way as $\mathcal{W}^1_p(\tau)$ and $\mathbb{W}^1_p(\tau)$ but the functions are only considered on $[S,T)$ and (S,T), respectively.

Lemma 2.1. *Let $s \geq 0$ be a fixed number and let u_s be an $\mathcal{F}_s$-measurable function with values in the set of distributions over $\mathbb{R}^d$.*

(i) We have $u_s \in \mathrm{tr}_s \mathcal{W}^1_p$ if and only if there exists a $v \in \mathcal{W}^1_p([s,\infty))$ satisfying the equation

$$\partial v/\partial t = \Delta v - v, \quad t \geq s, \tag{2.4}$$

(which is a particular case of (1.1) *and is understood in the same sense) with initial data u_s. This v is unique and satisfies*

$$\|v\|_{\mathbb{W}^1_p((s,\infty))} \leq N \|u_s\|_{\mathrm{tr}_s \mathcal{W}^1_p}, \quad \|u_s\|_{\mathrm{tr}_s \mathcal{W}^1_p} \leq N \|v\|_{\mathbb{W}^1_p((s,\infty))}, \tag{2.5}$$

where the constants N are independent of s, u_s, and v.

(ii) We have $u_s \in \mathrm{tr}_s \mathcal{W}^1_p$ if and only if there exists a $v \in \mathcal{W}^1_p([s,s+1))$ such that $v_s = u_s$.

(iii) If such a v exists and $dv_t = (D_i f^i_t + f^0_t)\,dt + g^k_t\,dw^k_t$, $t \geq s$, then

$$\|u_s\|_{\mathrm{tr}_s \mathcal{W}^1_p} \leq N\big(\|v\|_{\mathbb{W}^1_p((s,s+1))} + \sum_{j=0}^d \|f^j\|_{\mathbb{L}_p((s,s+1))} + \|g\|_{\mathbb{L}_p((s,s+1))}\big), \tag{2.6}$$

where the constant N is independent of s, u_s and v.

(iv) If $s > 0$ and we have a $u \in \mathcal{W}^1_p(s)$, then $u_s \in \mathrm{tr}_s \mathcal{W}^1_p$ and

$$\|u_s\|_{\mathrm{tr}_s \mathcal{W}^1_p} \leq N\big(\|u\|_{\mathbb{W}^1_p(s)} + \sum_{j=0}^d \|f^j\|_{\mathbb{L}_p(s)} + \|g\|_{\mathbb{L}_p(s)}\big),$$

where N is independent of u, and f^j and g^k are taken from (2.2)*.*

We prove this lemma in Section 5.

Here are our main results concerning (1.1). The following theorem is very close to Theorem 2.12 of [3]. Important differences are the presence of the parameter λ in (2.7) and weaker assumptions on the coefficients of the deterministic part of the equation.

Theorem 2.2. *Let $\lambda \geq 0$, let $f^j, g \in \mathbb{L}_p(\tau)$, and let $u_0 \in \mathrm{tr}_0\mathcal{W}^1_p$. Also let $\beta \leq \beta_0$, where $\beta_0 = \beta_0(d,p,\delta) > 0$ is a (small) constant continuously depending on (p,δ) an estimate of which from below can be extracted from the proof.*

(i) Then equation (1.1) for $t \leq \tau$ has a unique solution $u \in \mathcal{W}^1_p(\tau)$ with initial data u_0.

(ii) Furthermore, if a $v \in \mathcal{W}^1_p(\infty)$ is defined by equation (2.4) with initial condition u_0, then the above solution u satisfies

$$
\begin{aligned}
&\lambda^{1/2}\|u\|_{\mathbb{L}_p(\tau)} + \|Du\|_{\mathbb{L}_p(\tau)} \\
&\leq N\big(\sum_{i=1}^d \|f^i\|_{\mathbb{L}_p(\tau)} + \|g\|_{\mathbb{L}_p(\tau)} + \|Dv\|_{\mathbb{L}_p(\tau)}\big) \\
&\quad + N\lambda^{-1/2}\|f^0\|_{\mathbb{L}_p(\tau)} + N\lambda^{1/2}\|v\|_{\mathbb{L}_p(\tau)},
\end{aligned}
\tag{2.7}
$$

provided that $\lambda \geq \lambda_0$, where the constants $N, \lambda_0 \geq 1$ depend only on d, p, K, δ, and ε.

(iii) Finally, there exists a set $\Omega' \subset \Omega$ of full probability such that $u_{t\wedge\tau}I_{\Omega'}$ is a continuous $\mathcal{F}_t$-adapted L_p-valued functions of $t \in [0,\infty)$.

Observe that estimate (2.7) shows one of good reasons for writing the free term in (1.1) in the form $D_i f^i + f^0$, because f^i, $i = 1,...,d$, and f^0 enter (2.7) differently.

Remark 2.3. As it follows from the proof of Theorem 2.2, if $p = 2$, Assumption 2.2 is not needed and mentioning β_0 and ε can be dropped in the statement. Thus we provide a new way to prove the classical result on Hilbert space solvability of SPDEs (cf., for instance, [16]).

We prove Theorem 2.2 in Section 6 after we prepare necessary tools in Sections 3-5. In Section 3 we prove uniqueness part of Theorem 2.2 on the basis of Itô's formula from [14]. Here Assumption 2.2 is not used. In Section 4 we treat the case of the heat equation with random right-hand side and present a simplified version of the corresponding result from [7]. In Section 5 we prove an auxiliary existence theorem and derive some a priori estimates.

Here is a result about continuous dependence of solutions on the data.

Theorem 2.4. *Assume that for each $n = 1, 2, ...$ we are given functions a^{ij}_{nt}, a^i_{nt}, b^i_{nt}, c_{nt}, σ^{ik}_{nt}, ν^k_{nt}, f^j_{nt}, g^k_{nt}, and u_{n0} having the same meaning and satisfying the same assumptions with the same δ, K, κ, ε, and $\beta \leq \beta_0(d,p,\delta)$ as the original ones. Assume that for $i,j = 1,...,d$ and almost all (ω,t,x) we have*

$$
\begin{gathered}
(a^{ij}_{nt}, a^i_{nt}, b^i_{nt}, c_{nt}) \to (a^{ij}_t, a^i_t, b^i_t, c_t), \\
|\sigma^{i\cdot}_{nt} - \sigma^{i\cdot}_t|_{\ell_2} + |\nu_{nt} - \nu_t|_{\ell_2} \to 0,
\end{gathered}
$$

as $n \to \infty$. Also assume that

$$
\sum_{j=0}^d (\|f^j_n - f^j\|_{\mathbb{L}_p(\tau)} + \|g_n - g\|_{\mathbb{L}_p(\tau)} + \|u_{n0} - u_0\|_{\mathrm{tr}_0\mathcal{W}^1_p} \to 0
$$

as $n \to \infty$. Take $\lambda \geq \lambda_0$, take the function u from Theorem 2.2 and let $u_n \in \mathcal{W}^1_p(\tau)$ be the unique solutions of equations (1.1) *for $t \leq \tau$ constructed from a^{ij}_{nt}, a^i_{nt}, b^i_{nt}, c_{nt}, σ^{ik}_{nt}, ν^k_{nt}, f^j_{nt}, and g^k_{nt} and having initial values u_{n0}.*

Then, as $n \to \infty$, we have $\|u_n - u\|_{\mathbb{W}^1_p(\tau)} \to 0$ and for any finite $T \in [0, \infty)$

$$E \sup_{t \leq \tau \wedge T} \|u_{nt} - u_t\|^p_{L_p} \to 0. \tag{2.8}$$

Proof. Set $v_{nt} = u_{nt} - u_t$. Then

$$dv_{nt} = (L_{nt} v_{nt} - \lambda v_{nt} + D_i \tilde{f}^i_{nt} + \tilde{f}^0_{nt})\, dt + (\Lambda^k_{nt} v_{nt} + \tilde{g}^k_{nt})\, dw^k_t,$$

where L_{nt} and Λ^k_{nt} are the operators constructed from a^{ij}_{nt}, a^i_{nt}, b^i_{nt}, c_{nt} and σ^{ik}_{nt}, ν^k_{nt}, respectively, and

$$\tilde{f}^i_{nt} = f^i_{nt} - f^i_t + (a^{ji}_{nt} - a^{ji}_t) D_j u_t + (a^i_{nt} - a^i_t) u_t,$$
$$\tilde{f}^0_{nt} = f^0_{nt} - f^0_t + (b^i_{nt} - b^i_t) D_i u_t + (c_{nt} - c_t) u_t,$$
$$\tilde{g}^k_{nt} = g^k_{nt} - g^k_t + (\sigma^{ik}_{nt} - \sigma^{ik}_t) D_i u_t + (\nu^k_{nt} - \nu^k_t) u_t.$$

By Theorem 2.2 we know that $u \in \mathbb{W}^1_p(\tau)$. This along with our assumptions and the dominated convergence theorem implies that

$$\sum_{j=0}^d \|\tilde{f}^j_n\|_{\mathbb{L}_p(\tau)} + \|\tilde{g}_n\|_{\mathbb{L}_p(\tau)} \to 0$$

as $n \to \infty$. After that by applying (2.7) to v_{nt} we immediately see that $\|u_n - u\|_{\mathbb{W}^1_p(\tau)} \to 0$.

Assertion (2.8) is, actually, a simple corollary of the above. Indeed, by introducing $\hat{f}^j_n$ and $\hat{g}^k_n$ in an obvious way, we can write

$$dv_{nt} = (D_i \hat{f}^i_{nt} + \hat{f}^0_{nt})\, dt + \hat{g}^k_{nt}\, dw^k_t, \tag{2.9}$$

and

$$\sum_{j=1}^d \|\hat{f}^j_n\|_{\mathbb{L}_p(\tau)} + \|\hat{g}_n\|_{\mathbb{L}_p(\tau)} \to 0.$$

It is standard (see, for instance, our Theorem 3.1) to derive from here the estimate

$$E \sup_{t \leq \tau \wedge T} \|u_{nt} - u_t\|^p_{L_p} \leq N\big(\sum_{j=1}^d \|\hat{f}^j_n\|_{\mathbb{L}_p(\tau \wedge T)} + \|\hat{g}_n\|_{\mathbb{L}_p(\tau \wedge T)} + E\|u_{n0} - u_0\|^p_{L_p}\big),$$

where N is independent of n. It is also well known that $W^{1-2/p}_p \subset L_p$, that is

$$\|u_{n0} - u_0\|_{L_p} \leq N \|u_{n0} - u_0\|_{W^{1-2/p}_p}.$$

By combining all this together we obtain (2.8) and the theorem is proved.

The following result could be proved on the basis of Theorem 2.4 in the same way as Corollary 5.11 of [7], where the solutions are approximated by solutions of equations with smooth coefficients and then a stopping time techniques was used. We give here a shorter proof based on a different idea.

Theorem 2.5. *Let $p_1, p_2 \in [2, \infty)$, $p_1 < p_2$, and let the assumptions of Theorem 2.2 be satisfied for $p = p_1$ and $p = p_2$. Assume that $\beta \leq \beta_0(d, p, \delta)$ for any $p \in [p_1, p_2]$. Then the solutions corresponding to $p = p_1$ and $p = p_2$ coincide, that is, there is a unique solution $u \in \mathcal{W}^1_{p_1}(\tau) \cap \mathcal{W}^1_{p_2}(\tau)$ of equation* (1.1) *with initial data u_0.*

Proof. Obviously, it suffices to concentrate on bounded τ. As is explained above in that case we may assume that λ is as large as we like. We take it so large that one could use assertion (ii) of Theorem 2.2 with any $p \in [p_1, p_2]$. Finally, we may assume that $p_1 < p_2$.

Denote by u the solution corresponding to $p = p_2$ and observe that, owing to uniqueness of solutions in $\mathcal{W}^1_{p_1}(\tau)$, we need only show that $u \in \mathcal{W}^1_{p_1}(\tau)$.

Take a $\zeta \in C_0^\infty$ such that $\zeta(0) = 1$, set $\zeta_n(x) = \zeta(x/n)$, and notice that $u^n := u\zeta_n$ satisfies

$$du^n_t = (L_t u^n_t) - \lambda u^n_t + D_i f^i_{nt} + f^0_{nt})\, dt + (\Lambda^k_t u^n_t + g^k_{nt})\, dw^k_t,$$

where

$$f^i_{nt} = f^i_t \zeta_n - u a^{ji}_t D_j \zeta_n, \quad i \geq 1,$$

$$f^0_{nt} = f^0_t \zeta_n - f^i_t D_i \zeta_n - (a^{ij}_t D_i u_t + a^j_t u) D_j \zeta_n - b^i_t u_t D_i \zeta_n,$$

$$g^k_{nt} = g^k_t \zeta_n - \sigma^{ik}_t u_t D_i \zeta_n.$$

It follows that for $p_1 \leq p \leq p_2$ we have

$$\|u^n\|_{\mathbb{W}^1_p(\tau)} \leq N\Big(\sum_{i=0}^{d} \|f^i_n\|_{\mathbb{L}_p(\tau)} + \|g_n\|_{\mathbb{L}_p(\tau)} + \|u_0\zeta_n\|_{\mathrm{tr}_0 \mathcal{W}^1_p}\Big). \tag{2.10}$$

One knows that with constants N independent of n

$$\|u_0\zeta_n\|_{\mathrm{tr}_0\mathcal{W}^1_p} \leq N\big(\|u_0\zeta_n\|_{\mathrm{tr}_0\mathcal{W}^1_{p_1}} + \|u_0\zeta_n\|_{\mathrm{tr}_0\mathcal{W}^1_{p_2}}\big) \leq N\big(\|u_0\|_{\mathrm{tr}_0\mathcal{W}^1_{p_1}} + \|u_0\|_{\mathrm{tr}_0\mathcal{W}^1_{p_2}}\big).$$

Similarly, and by Hölder's inequality

$$\|f^i_n\|_{\mathbb{L}_p(\tau)} \leq N + N\|uD\zeta_n\|_{\mathbb{L}_p(\tau)} \leq N + \|u\|_{\mathbb{L}_{p_2}(\tau)} \|D\zeta_n\|_{\mathbb{L}_q(\tau)},$$

where

$$q = \frac{pp_2}{p_2 - p}.$$

Similar estimates are available for other terms in the right-hand side of (2.10). Since

$$\|D\zeta_n\|_{\mathbb{L}_q(\tau)} = N n^{-1 + (p_2 - p)d/(p_2 p)} \to 0$$

as $n \to \infty$ if

$$\frac{1}{p} - \frac{1}{p_2} < \frac{1}{d}, \tag{2.11}$$

estimate (2.10) implies that $u \in \mathcal{W}^1_p(\tau)$.

Thus knowing that $u \in \mathcal{W}^1_{p_2}(\tau)$ allowed us to conclude that $u \in \mathcal{W}^1_p(\tau)$ as long as $p \in [p_1, p_2]$ and (2.11) holds. We can now replace p_2 with a smaller p and keep going in the same way each time increasing $1/p$ by the same amount until p reaches p_1. Then we get that $u \in \mathcal{W}^1_{p_1}(\tau)$. The theorem is proved.

In many situation the following maximum principle is useful.

Theorem 2.6. *Under the assumptions of Theorem 2.2 suppose that $\beta \leq \beta_0(d,q,\delta)$ for $q \in [2,p]$. Also suppose that $u_0 \geq 0$, $f^i = 0$, $i = 1,...,d$, $f^0 \geq 0$, $g = 0$. Then for the solution u almost surely we have $u_t \geq 0$ for all finite $t \leq \tau$.*

Proof. If $p = 2$ the result is proved in [10]. For general $p \geq 2$ take the same function ζ_n as in the preceding proof, introduce $f^i = f^{ni}\zeta_n$, $g^k_n = 0$, and call u^n the solution of (1.1) with so modified free terms and the same initial data u_0. By Theorem 2.5 we have $u^n \in \mathcal{W}^1_p(\tau) \cap \mathcal{W}^1_2(\tau)$. By the above, $u^n \geq 0$ and it only remains to use Theorem 2.4. The theorem is proved.

3. Itô's formula and uniqueness

The following two "standard" results are taken from [14].

Theorem 3.1. *Let $u \in \mathcal{W}^1_p(\tau)$, $f^j \in \mathbb{L}_p(\tau)$, $g = (g^k) \in \mathbb{L}_p(\tau)$ and assume that (2.2) holds for $t \leq \tau$ in the sense of generalized functions. Then there is a set $\Omega' \subset \Omega$ of full probability such that*

(i) $u_{t\wedge\tau} I_{\Omega'}$ is a continuous L_p-valued $\mathcal{F}_t$-adapted function on $[0,\infty)$;

(ii) for all $t \in [0,\infty)$ and $\omega \in \Omega'$ Itô's formula holds:

$$\int_{\mathbb{R}^d} |u_{t\wedge\tau}|^p \, dx = \int_{\mathbb{R}^d} |u_0|^p \, dx + p \int_0^{t\wedge\tau} \int_{\mathbb{R}^d} |u_s|^{p-2} u_s g^k_s \, dx \, dw^k_s$$

$$+ \int_0^{t\wedge\tau} \Big(\int_{\mathbb{R}^d} \big[p|u_t|^{p-2} u_t f^0_t - p(p-1)|u_t|^{p-2} f^i_t D_i u_t$$

$$+ (1/2)p(p-1)|u_t|^{p-2} |g_t|^2_{\ell_2} \big] \, dx \Big) \, dt. \tag{3.1}$$

Furthermore, for any $T \in [0,\infty)$ and

$$E \sup_{t \leq \tau\wedge T} \|u_t\|^p_{L_p} \leq 2E\|u_0\|^p_{L_p} + NT^{p-1} \|f^0\|^p_{\mathbb{L}_p(\tau)}$$

$$+ NT^{(p-2)/2} \Big(\sum_{i=1}^d \|f^i\|^p_{\mathbb{L}_p(\tau)} + \|g\|^p_{\mathbb{L}_p(\tau)} + \|Du\|^p_{\mathbb{L}_p(\tau)} \Big), \tag{3.2}$$

where $N = N(d,p)$.

Here is an "energy" identity.

Corollary 3.2. *Under the conditions of Theorem 3.1 assume that $\tau < \infty$ (a.s.). Then*

$$E \int_{\mathbb{R}^d} |u_0|^p \, dx + E \int_0^\tau \Big(\int_{\mathbb{R}^d} \big[p|u_t|^{p-2} u_t f^0_t - p(p-1)|u_t|^{p-2} f^i_t D_i u_t$$

$$+ (1/2)p(p-1)|u_t|^{p-2} |g_t|^2_{\ell_2} \big] \, dx \Big) \, dt \geq E I_{\tau<\infty} \int_{\mathbb{R}^d} |u_\tau|^p \, dx. \tag{3.3}$$

Furthermore, if τ is bounded then there is an equality instead of inequality in (3.3).

Next result implies, in particular, uniqueness in Theorem 2.2.

Lemma 3.3. *Under Assumption 2.1 there exist $\lambda_0 \geq 0$ and N depending only on d, p, K, and δ such that for any $\lambda \geq \lambda_0$ and any solution $u \in \mathcal{W}^1_{p,0}(\tau)$ of* (1.1) *we have*

$$\lambda \|u\|_{\mathbb{L}_p(\tau)} \leq N\lambda^{1/2}\big(\sum_{j=1}^d \|f^j\|_{\mathbb{L}_p(\tau)} + \|g\|_{\mathbb{L}_p(\tau)}\big) + N\|f^0\|_{\mathbb{L}_p(\tau)}. \tag{3.4}$$

Furthermore, if $a^i = b^i = \nu^k \equiv 0$, then one can take $\lambda_0 = 0$.

Proof. If (3.4) is true for $\tau \wedge T$ in place of τ and any $T \in (0, \infty)$, then it is obviously also true as is. Therefore, we may assume that τ is finite. An advantage of this assumption is that we can use Corollary 3.2. Write (3.3) with $\hat{f}^i_t$, $\hat{f}^0_t$, and $\hat{g}^k_t$ in place of f^i_t, f^0_t, and g^k_t, respectively, where

$$\hat{f}^i_t = a^{ji}_t D_j u_t + a^i_t u_t + f^i_t,$$

$$\hat{f}^0_t = b^i_t D_i u_t + (c_t - \lambda)u_t + f^0_t, \quad \hat{g}^k_t = \sigma^{ik}_t D_i u_t + \nu^k_t u_t + g^k_t.$$

Then observe that inequalities like $(a+b)^2 \leq (1+\varepsilon)a^2 + (1+\varepsilon^{-1})b^2$ show that for any $\varepsilon \in (0,1]$ we have

$$|\hat{g}_t|^2_{\ell_2} \leq (1+\varepsilon)\big|\sum_{i=1}^d \sigma^{i\cdot}_t D_i u_t\big|^2_{\ell_2} + 2\varepsilon^{-1}|\nu_t u_t + g_t|^2_{\ell_2}$$

$$\leq 2(1+\varepsilon)\alpha^{ij}_t (D_i u_t) D_j u_t + N\varepsilon^{-1}(|u_t|^2 + |g_t|^2_{\ell_2}).$$

Owing to (2.3), for $\varepsilon = \varepsilon(\delta, K, d) > 0$ small enough

$$I := (1/2)|u_t|^{p-2}|\hat{g}_t|^2_{\ell_2} - |u_t|^{p-2}\hat{f}^i_t D_i u_t + (p-1)^{-1}|u_t|^{p-2}u_t b^i_t D_i u_t$$

$$\leq -(\delta/2)|u_t|^{p-2}|Du_t|^2 + N|u_t|^{p-2}(|u_t|^2 + |g_t|^2_{\ell_2} + |Du_t|\,|u_t| + |Du_t|\sum_{i=1}^d |f^i_t|). \tag{3.5}$$

Next we use that for any $\gamma > 0$

$$|u_t|^{p-1}|Du_t| = (|u_t|^{(p-2)/2}|Du_t|)|u_t|^{p/2} \leq \gamma |u_t|^{p-2}|Du_t|^2 + \gamma^{-1}|u_t|^p,$$

$$|u_t|^{p-2}|Du_t|\,|f^i_t| \leq \gamma|u_t|^{p-2}|Du_t|^2 + \gamma^{-1}|u_t|^{p-2}|f^i_t|^2,$$

and by choosing γ appropriately find from (3.5) that

$$I \leq N|u_t|^p + N|u_t|^{p-2}\big(\sum_{i=1}^d |f^i_t|^2 + |g_t|^2_{\ell_2}\big). \tag{3.6}$$

After that Hölder's inequality and (3.3), where the right-hand side is nonnegative, immediately lead to

$$(\lambda - N_1)\|u\|^p_{\mathbb{L}_p(\tau)} \leq N\|u\|^{p-2}_{\mathbb{L}_p(\tau)}\big(\sum_{i=1}^d \|f^i\|^2_{\mathbb{L}_p(\tau)} + \|g\|^2_{\mathbb{L}_p(\tau)}\big) + N\|u\|^{p-1}_{L_p(\tau)}\|f^0\|_{L_p(\tau)}.$$

Furthermore, simple inspection of the above argument shows that, if $a^i = b^i = \nu^k \equiv 0$, then the terms with $|u_t|^2$ and $|u_t|\,|Du_t|$ in (3.5) and the term with $|u_t|^p$ in (3.6) disappear, so that we can take $N_1 = 0$ in this case (recall that $c \leq 0$). Generally, for $\lambda \geq 2N_1$ we have $\lambda - N_1 \geq (1/2)\lambda$ and

$$\bar{U}^p \leq N\bar{U}^{p-2}\bar{G}^2 + N\bar{U}^{p-1}\bar{F},$$

where

$$\bar{U} = \lambda\|u\|_{\mathbb{L}_p(\tau)}, \quad \bar{G} = \lambda^{1/2}(\|f^i\|_{\mathbb{L}_p(\tau)} + \|g\|_{\mathbb{L}_p(\tau)}), \quad \bar{F} = \|f^0\|_{L_p(\tau)}.$$

It follows that $\bar{U} \leq N(\bar{G} + \bar{F})$, which is (3.4) and the lemma is proved.

4. Case of the heat equation

To move further we need the following analytic fact established in [5] (see also [8] for a complete proof).

Lemma 4.1. *Denote by T_t the heat semigroup in $\mathbb{R}^d$ and let $p \geq 2$, $-\infty \leq a < b \leq \infty$, $g \in L_p((a,b) \times \mathbb{R}^d, \ell_2)$. Then*

$$\int_{\mathbb{R}^d}\int_a^b \Big[\int_a^t |DT_{t-s}g_s(x)|_{\ell_2}^2\,ds\Big]\,dtdx \leq N(d,p)\int_{\mathbb{R}^d}\int_a^b |g_t(x)|_{\ell_2}^2\,dtdx.$$

In this section we deal with the following model equation

$$du_t = \Delta u_t\,dt + g_t^k\,dw_t^k. \tag{4.1}$$

Lemma 4.2. *Assume that $\tau \leq T$, where the constant $T \in [0,\infty)$. Then for any $g = (g^1, g^2, ...) \in \mathbb{L}_p(\tau)$ there exists a unique $u \in \mathcal{W}_{p,0}^1(\tau)$ satisfying (4.1) for $t \leq \tau$. Furthermore, for this solution we have*

$$E\sup_{t\leq\tau}\|u_t\|_{L_p}^p \leq N(d,p)T^{(p-2)/2}\|g\|_{\mathbb{L}_p(\tau)}^p, \tag{4.2}$$

$$\|Du\|_{\mathbb{L}_p(\tau)} \leq N(d,p)\|g\|_{\mathbb{L}_p(\tau)}. \tag{4.3}$$

Proof. By replacing the unknown function u_t with $v_te^{\lambda t}$ we see that v_t satisfies

$$dv_t = (\Delta v_t - \lambda v)\,dt + e^{-\lambda t}g_t^k\,dw_t^k.$$

Since τ is bounded the inclusions $u \in \mathcal{W}_{p,0}^1(\tau)$ and $v \in \mathcal{W}_{p,0}^1(\tau)$ are equivalent and our assertion about uniqueness follows from Lemma 3.3.

In the proof of existence we borrow part of the proof of Lemma 4.1 of [7]. As we have pointed out in the Introduction the beginning of the theory of divergence and nondivergence type equations is the same. The only difference with that proof is that here we take $f \equiv 0$.

We take an integer $m \geq 1$, some bounded stopping times $\tau_0 \leq \tau_1 \leq ... \leq \tau_m \leq T$ and some (nonrandom) functions $g^{ij} \in C_0^\infty$, $i, j = 1, ..., m$. Then we define

$$g_t^k(x) = \sum_{i=1}^m g^{ik}(x)I_{(\tau_{i-1},\tau_i]}(t),$$

$$v_t(x)=\sum_{k=1}^m\int_0^t g^k_s(x)\,dw^k_s=\sum_{i,k=1}^m g^{ik}(x)(w^k_{t\wedge\tau_i}-w^k_{t\wedge\tau_{i-1}}),\quad t\geq 0.$$

Obviously, for any ω, the function $v_t(x)$ is continuous and bounded in (t,x) along with any derivative in x. Furthermore, the function and its derivative in x are Hölder $1/3$ continuous in t uniformly with respect to x (for almost any ω). Also $v_t(x)$ has compact support in x.

These properties of $v_t(x)$ imply that for any ω there exists a unique classical solution of the heat equation

$$\frac{\partial}{\partial t}\bar u_t=\Delta\bar u_t+\Delta v_t,\quad t>0,$$

with zero initial data. Furthermore,

$$\bar u_t(x)=\int_0^t T_{t-s}\Delta v_s(x)\,ds. \tag{4.4}$$

This formula shows, in particular, that $\bar u_t(x)$ is $\mathcal{F}_t$-adapted. Adding the fact that $\bar u_t$ is continuous in t proves that $\bar u_t(x)$ is predictable. The same holds for

$$(\bar u_t,\phi)=\int_0^t(T_{t-s}\Delta v_s,\phi)\,ds$$

with any $\phi\in C_0^\infty$. The following corollary of Minkowski's inequality

$$\|\bar u_t\|_{L_p}\leq\int_0^t\|\Delta v_s\|_{L_p}\,ds \tag{4.5}$$

shows that $\bar u_t$ is L_p-valued. Since $(\bar u_t,\phi)$ is predictable for any $\phi\in C_0^\infty$, $\bar u_t$ is weakly and hence strongly predictable as an L_p-valued process.

One can differentiate (4.4) with respect to x as many times as one wants and get similar statements about the derivatives of $\bar u_t$. In particular, (4.5) implies that for any multi-index α

$$E\int_0^T\int_{\mathbb{R}^d}|D^\alpha\bar u_t|^p\,dxdt\leq T^pE\int_0^T\int_{\mathbb{R}^d}|D^\alpha\Delta v_t|^p\,dxdt<\infty,$$

so that $\bar u_t\in\mathcal{W}^1_{p,0}(T)$.

Now, it is easily seen that

$$u_t(x):=\bar u_t(x)+v_t(x)$$

satisfies (4.1) pointwisely and by the above $u_t\in\mathcal{W}^1_{p,0}(T)$. The (deterministic) Fubini's theorem also shows that u_t satisfies (4.1) in the sense of distributions.

Next, we use the same simple transformation as in the proof of Lemma 4.1 of [7] and conclude that for any t and x almost surely

$$Du_t(x)=\sum_{k=1}^m\int_0^t T_{t-s}Dg^k_s(x)\,dw^k_s.$$

Hence by Burkholder-Davis-Gundy inequality

$$E|Du_t(x)|^p \le NE\big[\int_0^t |T_{t-s}Dg_s(x)|^2_{\ell_2}\,ds\big]^{p/2},$$

which along with Lemma 4.1 proves (4.3) for our particular g. Theorem 3.1 shows that (4.2) follows from (4.3) and (4.1).

The rest is trivial since the set of g's like the one above is dense in $\mathbb{L}_p(T)$ by Theorem 3.10 of [7]. The lemma is proved.

Next we introduce the parameter λ into (4.1).

Lemma 4.3. *Assume that $\tau \le T$, where the constant $T \in [0,\infty)$. Let $\lambda > 0$. Then for any $g = (g^1, g^2, ...) \in \mathbb{L}_p(\tau)$ there exists a unique $u \in \mathcal{W}^1_{p,0}(\tau)$ satisfying*

$$du_t = (\Delta u_t - \lambda u_t)\,dt + g^k_t\,dw^k_t. \tag{4.6}$$

for $t \le \tau$. Furthermore, for this solution we have

$$\lambda^{p/2}\|u\|^p_{\mathbb{L}_p(\tau)} \le N(d,p)\|g\|^p_{\mathbb{L}_p(\tau)}, \tag{4.7}$$

$$\|Du\|_{\mathbb{L}_p(\tau)} \le N(d,p)\|g\|_{\mathbb{L}_p(\tau)}. \tag{4.8}$$

Proof. Uniqueness and estimate (4.7) follow from Lemma 3.3. The existence immediately follows from Lemma 4.2 and the result of transformation described in the beginning of its proof. To establish (4.8) consider the heat equation

$$\frac{\partial}{\partial t}v_t = \Delta v_t - \lambda u_t. \tag{4.9}$$

Since $u \in \mathbb{L}_p(\tau)$, for almost any ω we have $u \in L_p((0,\tau)\times\mathbb{R}^d)$ and by by a classical result (see, for instance, [13]) for almost any ω equation (4.9) with zero initial data has a unique solution in the class of functions such that along with derivatives in x up to the second order they belong to $L_p((0,\tau)\times\mathbb{R}^d)$. Furthermore,

$$\|D^2v\|^p_{L_p((0,\tau)\times\mathbb{R}^d)} + \lambda^{p/2}\|Dv\|^p_{L_p((0,\tau)\times\mathbb{R}^d)}$$
$$+ \lambda^p\|v\|^p_{L_p((0,\tau)\times\mathbb{R}^d)} \le N\|\lambda u\|^p_{L_p((0,\tau)\times\mathbb{R}^d)}. \tag{4.10}$$

The solution v_t can be given by an integral formula, which implies that v_t is $\mathcal{F}_t$-adapted. It is also continuous as an L_p-valued process, hence, is a predictable L_p-valued process. Taking expectations of both parts of (4.10) shows that $v \in \mathcal{W}^1_p(\tau)$.

Now observe that

$$d(u_t - v_t) = \Delta(u_t - v_t)\,dt + g^k_t\,dw^k_t,$$

which by Lemma 4.2 implies that

$$\|D(u-v)\|^p_{\mathbb{L}_p(\tau)} \le N\|g\|^p_{\mathbb{L}_p(\tau)}.$$

Upon combining this with (4.10) we obtain

$$\|Du\|^p_{\mathbb{L}_p(\tau)} \le N(\|g\|^p_{\mathbb{L}_p(\tau)} + \lambda^{p/2}\|u\|^p_{\mathbb{L}_p(\tau)}),$$

which along with (4.7) yields (4.8). The lemma is proved.

5. A priori estimates in the general case

First we deal with the case when $\sigma = \nu = 0$.

Lemma 5.1. *Suppose that Assumptions 2.1 and 2.2 are satisfied. Also suppose that $\sigma^{ik} \equiv \nu^k \equiv 0$. Let $f^j \in \mathbb{L}_p(\tau)$ and $g \in \mathbb{L}_p(\tau)$. Also let $\beta \leq \beta_0$, where the way to estimate the constant $\beta_0(d,p,\delta) > 0$ is described in the proof.*

Then there exist constants $\lambda_0 \geq 1$ and N, depending only on d, p, δ, K, and ε, such that for any $\lambda \geq \lambda_0$ there exists a unique $u \in \mathcal{W}^1_{p,0}(\tau)$ satisfying (1.1) for $t \leq \tau$. Furthermore, this solution satisfies the estimate

$$\lambda^{1/2}\|u\|_{\mathbb{L}_p(\tau)} + \|Du\|_{\mathbb{L}_p(\tau)} \leq N\big(\sum_{i=1}^{d}\|f^i\|_{\mathbb{L}_p(\tau)} + \|g\|_{\mathbb{L}_p(\tau)}\big) + N\lambda^{-1/2}\|f^0\|_{\mathbb{L}_p(\tau)}. \tag{5.1}$$

Proof. Uniqueness and part of estimate (5.1) follow from Lemma 3.3. In the rest of the proof we may assume that τ is bounded and split our argument into two parts.

Case $g^k \equiv 0$. First assume that the coefficients and f^j are nonrandom. We extend the coefficients of L following the example $a^{ij}_t(x) = a^{ij}_0(x)$, $t \geq 0$, and extend f^j_t beyond $(0,\tau)$ arbitrary only requiring $f^j \in L_p(\mathbb{R}^{d+1})$.

Then by Theorem 4.5 and Remark 2.4 of [9] the equation

$$\frac{\partial}{\partial t}u_t = L_t u_t - \lambda u_t + D_i f^i_t + f^0_t \tag{5.2}$$

in $\mathbb{R}^{d+1}$ has a unique solution with finite norms

$$\|u\|_{L_p(\mathbb{R}^{d+1})} \quad \text{and} \quad \|Du\|_{L_p(\mathbb{R}^{d+1})}$$

provided that $\lambda \geq \lambda_0$. By Theorem 4.4 of [9]

$$\lambda^{1/2}\|u\|_{L_p(\mathbb{R}^{d+1})} + \|Du\|_{L_p(\mathbb{R}^{d+1})} \leq N(\sum_{i=1}^{d}\|f^i\|_{L_p(\mathbb{R}^{d+1})} + \lambda^{-1/2}\|f^0\|_{L_p(\mathbb{R}^{d+1})}). \tag{5.3}$$

By Theorem 3.1 the function u_t is a continuous L_p-valued function.

The proof of Theorem 4.4 of [9] is achieved on the basis of the a priori estimate (5.3) and the method of continuity by considering the family of equations

$$\frac{\partial}{\partial t}u_t = (\theta L_t + (1-\theta)\Delta)u_t - \lambda u_t + D_i f^i_t + f^0_t, \tag{5.4}$$

where the parameter θ changes in $[0,1]$. We remind briefly the method of continuity because we want to show that certain properties of equation (5.4) which we know for $\theta = 0$ propagate from $\theta = 0$ to $\theta = 1$.

We fix a $\theta_0 \in [0,1]$ and to solve (5.4) for given f^j define a sequence of $u^n \in L_p(\mathbb{R}, W^1_p)$ by solving the equation

$$\frac{\partial}{\partial t}u^{n+1}_t = (\theta_0 L_t + (1-\theta_0)\Delta)u^{n+1}_t - \lambda u^{n+1}_t$$

$$+ D_i f^i_t + f^0_t + (\theta - \theta_0)(L_t - \Delta)u^n, \quad n \geq 1, \quad u^0 = 0. \tag{5.5}$$

If we know that equation (5.4) is uniquely solvable with θ_0 in place of θ for arbitrary $f^j \in L_p(\mathbb{R}^{d+1})$, then the sequence u^n is well defined. Furthermore, estimate (5.3) easily shows that for θ sufficiently close to θ_0 the $L_p(\mathbb{R}, W^1_p)$ norm of $u^{n+1} - u^n$ goes to zero geometrically as $n \to \infty$. In this way passing to the limit in (5.5) we obtain the solution of (5.4) for θ close to θ_0. Then we can repeat the procedure and starting from $\theta = 0$ and moving step by step eventually reach $\theta = 1$.

For $\theta = 0$ we are dealing with solvability of the heat equation which is proved by giving the solution explicitly by means of the heat semigroup. This representation formula has two important implications

(i) For any constant $T \in \mathbb{R}$, changing f^j_t for $t \geq T$ does not affect u_t for $t \leq T$;

(ii) If f^j are $L_p(\mathbb{R}^{d+1})$-valued functions of a parameter, say ω from a measurable space, say $(\Omega, \mathcal{F}_T)$, then the solution $u \in L_p(\mathbb{R}, W^1_p)$, which now depends on ω is also $\mathcal{F}_T$-measurable.

Property (i) is obtained by inspecting the representation formula. Property (ii) is true because the mapping $L_p(\mathbb{R}^{d+1}) \ni f^j \to u \in L_p(\mathbb{R}, W^1_p)$ is continuous and hence Borel measurable.

Obviously, both properties propagate from $\theta = 0$ to $\theta = 1$ by the above method of continuity. In particular, solutions of (5.2) on the time interval $(-\infty, T]$ depend only on the values of f^j_t for $t \in (-\infty, T]$. It follows that with the same λ and N, for any $T \in \mathbb{R}$,

$$\lambda^{1/2}\|u\|_{L_p((-\infty,T),L_p)} + \|Du\|_{L_p((-\infty,T),L_p)}$$

$$\leq N(\sum_{i=1}^d \|f^i\|_{L_p((-\infty,T),L_p)} + \lambda^{-1/2}\|f^0\|_{L_p((-\infty,T),L_p)}. \tag{5.6}$$

From now on we allow the coefficients and f^j to be random, continue f^j as zero for $t < 0$ and solve (5.2) for each ω. By (5.6) with $T = 0$ we have that $u_t = 0$ for $t \leq 0$ and it makes sense considering equation (5.2) on $(0, T)$ for each $T \in (0, \infty)$ with zero initial condition. In such situation properties (i) and (ii) still hold.

In particular, if f^j are measurable $L_p((0,T), L_p)$-valued functions of a parameter, say ω from a measurable space, say $(\Omega, \mathcal{F}_T)$, then the solution $u \in L_p((0,T), W^1_p)$ is also $\mathcal{F}_T$-measurable. Then from the equation itself it follows that (u_T, ϕ) is $\mathcal{F}_T$-measurable for any $\phi \in C^\infty_0$. Since u_T takes values in L_p, it is an L_p-valued $\mathcal{F}_T$-measurable function.

If f^i_t are predictable L_p-valued function, the above conclusions are valid for any $T \in [0, \infty)$. In particular, u_t is $\mathcal{F}_t$-adapted as an L_p-valued function and since it is continuous, u_t is a predictable L_p-valued function.

These properties and the fact that (5.6) holds for any $T \in (0, \infty)$ and ω prove the lemma in the particular case under consideration.

General case. By Lemma 4.3 there is a unique solution $v \in \mathcal{W}^1_{p,0}(\tau)$ of (4.6). Observe that

$$(L_t - \Delta)v_t = D_i \hat{f}^i_t + \hat{f}^0_t,$$

where $\hat{f}^j_t$ are function of class $\mathbb{L}_p(\tau)$ defined by

$$\hat{f}^j_t = (a^{ij}_t - \delta^{ij})D_i v_t + a^j_t v_t, \quad j = 1, ..., d,$$
$$\hat{f}^0_t = b^i_t D_i v_t + c_t v_t.$$

By the above there is a unique solution $u \in \mathcal{W}^1_{p,0}(\tau)$ of

$$\frac{\partial}{\partial t} u_t = L_t u_t - \lambda u_t + (L_t - \Delta)v_t + D_i f^i_t + f^0_t.$$

Obviously, $v_t + u_t$ is a solution of class $\mathcal{W}^1_{p,0}(\tau)$ of equation (1.1). By the particular case

$$\lambda^{1/2}\|u\|_{\mathbb{L}_p(\tau)} + \|Du\|_{\mathbb{L}_p(\tau)} \le N\big(\sum_{i=1}^d (\|f^i\|_{\mathbb{L}_p(\tau)} + \|\hat{f}^i\|_{\mathbb{L}_p(\tau)})\big)$$
$$+N\lambda^{-1/2}(\|f^0\|_{\mathbb{L}_p(\tau)} + \|\hat{f}^0\|_{\mathbb{L}_p(\tau)})$$

and to obtain (5.1) it only remains to use the estimates of v_t provided by Lemma 4.3. The lemma is proved.

Now we allow $\sigma \neq 0$.

Lemma 5.2. *In the assumptions of Lemma 5.1 drop the condition that $\sigma^{ik} \equiv 0$ and instead suppose that σ^{ik}_t depends only on ω and t. Then the assertion of Lemma 5.1 is still true.*

Proof. As in the proof of Lemma 5.1 we only need to prove existence and estimate (5.1). From this lemma we know that there exists a (unique) solution in $\mathcal{W}^1_{p,0}(\tau)$ of the equation

$$dv_t(x) = (\bar{L}_t v_t(x) - \lambda v_t + D_i \bar{f}^i_t + \bar{f}^0_t)\,dt + \bar{g}^k_t(x - x_t)\,dw^k_t,$$

where the operator $\bar{L}_t$ is constructed from

$$a^{ij}_t(x - x_t) - \alpha^{ij}_t, \quad (a^j_t, b^i_t, c_t)(x - x_t),$$
$$\bar{f}^i_t(x) := f^i_t(x - x_t) - \sigma^{ik}_t g^k_t(x - x_t), \quad i = 1, ..., d,$$
$$\bar{f}^0_t(x) := f^0_t(x - x_t), \quad \bar{g}^k_t(x) = g^k_t(x - x_t),$$

and the process $x_t = (x^1_t, ..., x^d_t)$ is defined by

$$x^i_t = \int_0^t \sigma^{ik}_s\,dw^k_s.$$

By Lemma 4.7 of [7] (Itô-Wentzell formula) the function $u_t(x) := v_t(x + x_t)$ satisfies (1.1). From Lemma 5.1 we also know an estimate of v, which easily translates into (5.1). The lemma is proved.

Remark 5.3. If under the conditions of Lemma 5.2 also a^{ij}_t depend only on ω and t and $a^i = b^i = c \equiv 0$, another way of proving its assertions can be obtained by following the arguments in Section 4.3 of [7]. Even though those arguments are much longer, they allow one to prove a very general result saying roughly speaking that "whatever estimate can be established for solutions of the heat equation in Banach function spaces with norms, that are invariant under time dependent shifting the x coordinate, the same estimate with the same constant also holds for solutions of the parabolic equation with no lower order terms and with the matrix of the second order coefficients depending only on t and dominating (in the matrix sense) the unit matrix" (see [6]).

Next step is to consider equations with lower order terms in front of dw^k_t and σ also depending on x. The following lemma and its corollary are stated in a slightly more general form than it is needed in the present article. The point is that we intend to use them in a subsequent article about equations in half spaces.

Lemma 5.4. *Let $G \subset \mathbb{R}^d$ be a domain (perhaps, $G = \mathbb{R}^d$).*

(i) Suppose that Assumption 2.1 is satisfied and suppose that there is a constant $\varepsilon > 0$ such that for all $r \in (0, \varepsilon]$, $x \in G$, $y \in \mathbb{R}^d$, such that $|x - y| \leq \varepsilon$, cylinders $Q = (s,t) \times \{z : |z - y| \leq r\} \in \mathbb{Q}$, and all values of indices we have

$$|\sigma^{i\cdot}_t(x) - \sigma^{i\cdot}_t(y)|_{\ell_2} + \operatorname{ocs}(a^{ij}, Q) \leq \beta_0, \tag{5.7}$$

where $\beta_0 = \beta_0(d,p,\delta) \in (0,1]$ is a constant an estimate from below for which can be obtained from the proof.

(ii) Let $f^j, g \in \mathbb{L}_p(\tau)$ and let $u \in \mathcal{W}^1_{p,0}(\tau)$ satisfy (1.1)*.*

(iii) Assume that $u_t(x) = 0$ if $x \notin G$.

Then there exist constants $N, \lambda_0 \geq 0$ depending only on d, p, K, δ, and ε, such that estimate (5.1) *holds true whenever $\lambda \geq \lambda_0$.*

Proof. We need to know how to extend a^{ij} beyond G and preserve the smallness of the integral oscillation. The natural way would be just to refer to the known result of [1]. However, the situation here is slightly different because there is no t in [1] and in addition it seems to the author that the proof in [1] contains an error related to the fact that the integral oscillations of the product of a C^∞_0 function and a function $a \in VMO$ cannot be estimated through the integral oscillations of a alone. This is seen, for instance if $a \equiv 1$. This is why we prefer to use a different way. Without losing generality we assume that $\varepsilon \in (0, 1/2)$.

Take a nonnegative $\xi \in C^\infty_0(\mathbb{R}^d)$ with support lying in the ball of radius $\varepsilon/2$ centered at the origin and such that $\xi(x) = 1$ for $|x| \leq \varepsilon/4$ and $0 \leq \xi \leq 1$.

Assume that $0 \in G$ and set

$$\hat{a}^{ij}_t := \xi a^{ij}_t + \delta^{-1}(1-\xi)\delta^{ij}.$$

Take a $Q=(s,t)\times B \in \mathbb{Q}$ with $\rho(Q) \leq \varepsilon^2$ $(\leq \varepsilon/2)$ and let z be the center of B. Denote by N^* various (large) constants depending only on d, p, and δ (for a while they will be independent of p) and observe that $|D\xi| \leq N^*\varepsilon^{-1}$. It follows easily that if $|z| \geq \varepsilon$ then

$$\operatorname{ocs}(\hat{a}^{ij}, Q) = \delta^{-1}\delta^{ij}\operatorname{ocs}(\xi, Q) \leq N^*\varepsilon^{-1}\rho(Q) \leq N^*\varepsilon.$$

However, if $|z| \leq \varepsilon$, then by assumption (recall that $0 \in G$) and the estimates

$$(|h_r - h_{r(B)}|)_{(B)} \leq \frac{1}{|B|^2}\int_B\int_B |h_r(y_1) - h_r(y_2)|\,dy_1dy_2 \leq 2(|h_r - h_{r(B)}|)_{(B)},$$

$$|\xi(y_1)a^{ij}_r(y_1) - \xi(y_2)a^{ij}_r(y_2)| \leq \xi(y_1)|a^{ij}_r(y_1) - a^{ij}_r(y_2)| + N^*|\xi(y_1) - \xi(y_2)|$$

we have

$$\operatorname{ocs}(\hat{a}^{ij}, Q) \leq 2\operatorname{ocs}(a^{ij}, Q) + N^*\varepsilon^{-1}\rho(Q) \leq 2\beta_0 + N^*\varepsilon.$$

The latter can be made less than any prescribed constant $\hat{\beta}_0(d,p,\delta) > 0$ on the account of taking appropriately $\beta_0(d,p,\delta)$ and reducing ε.

Next, if $|y| \leq \varepsilon/2$, then

$$\begin{aligned}(\hat{a}^{ij}(x) - \alpha^{ij}(y))\eta^i\eta^j &= \xi(x)(a^{ij}(x) - \alpha^{ij}(y))\eta^i\eta^j \\ &+(1-\xi(x))(\delta^{-1}\delta^{ij} - \alpha^{ij}(y))\eta^i\eta^j,\end{aligned}$$

where, due to the fact that $|x| \leq \varepsilon/2$ if $\xi(x) \neq 0$, we have

$$\begin{aligned}\xi(x)(a^{ij}(x) - \alpha^{ij}(y))\eta^i\eta^j &= \xi(x)(a^{ij}(x) - \alpha^{ij}(x))\eta^i\eta^j + \xi(x)(\alpha^{ij}(x) - \alpha^{ij}(y))\eta^i\eta^j \\ &\geq \xi(x)\delta|\eta|^2 - \xi(x)N^*\beta_0|\eta|^2 \geq (1/2)\xi(x)\delta|\eta|^2\end{aligned}$$

if $\beta_0(d,\delta,p) > 0$ is sufficiently small. Furthermore, by assumption

$$\alpha^{ij}(y)\eta^i\eta^j \leq a^{ij}(y))\eta^i\eta^j - \delta|\eta|^2 \leq (\delta^{-1} - \delta),|\eta|^2.$$

It follows that

$$(1-\xi(x))(\delta^{-1}\delta^{ij} - \alpha^{ij}(y))\eta^i\eta^j \geq (1-\xi(x))\delta|\eta|^2,$$

$$(\hat{a}^{ij}(x) - \alpha^{ij}(y))\eta^i\eta^j \geq (1/2)\delta|\eta|^2$$

if $|y| \leq \varepsilon/2$ and β_0 is chosen appropriately.

This and Lemma 5.2 allow us to assert that

$$\lambda^{1/2}\|v\|_{\mathbb{L}_p(\tau)} + \|Dv\|_{\mathbb{L}_p(\tau)} \leq N\big(\sum_{i=1}^d \|f^i\|_{\mathbb{L}_p(\tau)} + \|g\|_{\mathbb{L}_p(\tau)}\big) + N\lambda^{-1/2}\|f^0\|_{\mathbb{L}_p(\tau)}, \tag{5.8}$$

provided that $\lambda \geq \lambda_0$, if $v \in \mathcal{W}^1_{p,0}(\tau)$ is a solution of equation (1.1) in which $\nu^k \equiv 0$, a^{ij}_t are replaced with $\hat{a}^{ij}_t$, and $\sigma^{ik}_t = \sigma^{ik}_t(y)$, where $y \in \mathbb{R}^d$ is any fixed point with $|y| \leq \varepsilon/2$. In addition the dependence of λ_0 and N on the data is the same as in Lemma 5.1.

Now take a $\zeta\in C_0^\infty(\mathbb{R}^d)$ with unit L_p-norm and support lying in the ball of radius $\varepsilon/8$ centered at the origin. For $y\in\mathbb{R}^d$ introduce

$$\zeta^y(x)=\zeta(x-y),\quad u^y_t(x)=\zeta^y(x)u_t(x)$$

so that, in particular,

$$|u_t(x)|^p=\int_{\mathbb{R}^d}|u^y_t(x)|^p\,dy. \tag{5.9}$$

Observe that for each y

$$du^y_t=\big(L_tu^y_t-\lambda u^y_t+D_j\hat f^{yj}_t+\hat f^{y0}_t\big)\,dt+\big(\sigma^{ik}_t(y)D_iu^y_t+\hat g^{ky}_t\big)\,dw^k_t, \tag{5.10}$$

where we dropped the argument x (and ω) and

$$\hat f^{yj}_t=-a^{ij}_tu_tD_i\zeta^y+\zeta^yf^j_t,\quad j=1,...,d,$$
$$\hat f^{y0}_t=-a^{ij}_t(D_iu_t)D_j\zeta^y-u_ta^j_tD_j\zeta^y-u_tb^i_tD_i\zeta^y+\zeta^yf^0_t-f^i_tD_i\zeta^y,$$
$$\hat g^{ky}_t=\zeta^y\big(\nu^k_tu_t+g^k_t+(\sigma^{ik}_t-\sigma^{ik}_t(y))D_iu_t\big)-u_t\sigma^{ik}_t(y)D_i\zeta^y.$$

Next,

$$|\hat f^{yj}_t|\le N\eta^y|u_t|+\zeta^y|f^j_t|,\quad j=1,...,d,$$
$$|\hat f^{y0}_t|\le N\eta^y(|Du_t|+|u_t|)+N\eta^y\sum_{j=0}^d|f^j_t|,$$
$$|\hat g^y_t|_{\ell_2}\le N^*\zeta^y\beta_0|Du_t|+N\eta^y|u_t|+\zeta^y|g_t|_{\ell_2},$$

where $\eta=\zeta+|D\zeta|$, $\eta^y(x)=\eta(x-y)$, and the constants N depend only on d,p,δ, and K.

Obviously, if $u^y\not\equiv0$, then, owing to the fact that $u_t(x)=0$ for $x\notin G$, y lies within the distance $\le\varepsilon/8$ from a point in G. If this point is the origin, then the support of ζ^y is in the ball of radius $\varepsilon/4$ centered at the origin in which $\xi=1$ and consequently in (5.10) one can replace a^{ij}_t with the above $\hat a^{ij}_t$ since they coincide on the support of u^y. In that case estimate (5.8) is applicable to u^y. Naturally, the same is true for any y such that $u^y\not\equiv0$ (and certainly for any y such that $u^y\equiv0$ as well) and we conclude that for each y

$$\lambda^{1/2}\|\zeta^yu\|_{\mathbb{L}_p(\tau)}+\|D(\zeta^yu)\|_{\mathbb{L}_p(\tau)}$$
$$\le N\Big(\sum_{i=1}^d\|\eta^yf^i\|_{\mathbb{L}_p(\tau)}+\|\zeta^yg\|_{\mathbb{L}_p(\tau)}\Big)+N\lambda^{-1/2}\|\eta^yf^0\|_{\mathbb{L}_p(\tau)}$$
$$+N^*\beta_0\|\eta^yDu\|_{\mathbb{L}_p(\tau)}+N\|\eta^yu\|_{\mathbb{L}_p(\tau)}+N\lambda^{-1/2}\|\eta^yDu\|_{\mathbb{L}_p(\tau)}.$$

We raise both parts of this inequality to the pth power and notice that $|\zeta^yDu_t|\le|D(\zeta^yu_t)|+\eta^y|u_t|$. Then we find that

$$\lambda^{p/2}\|\zeta^yu\|^p_{\mathbb{L}_p(\tau)}+\|\zeta^yDu\|^p_{\mathbb{L}_p(\tau)}$$
$$\le N\Big(\sum_{i=1}^d\|\eta^yf^i\|^p_{\mathbb{L}_p(\tau)}+\|\zeta^yg\|^p_{\mathbb{L}_p(\tau)}\Big)+N\lambda^{-p/2}\|\eta^yf^0\|^p_{\mathbb{L}_p(\tau)}$$

$$+N^*\beta_0^p\|\eta^y Du\|^p_{\mathbb{L}_p(\tau)}+N\|\eta^y u\|^p_{\mathbb{L}_p(\tau)}+N\lambda^{-p/2}\|\eta^y Du\|^p_{\mathbb{L}_p(\tau)}.$$

We integrate through this estimate and use formulas like (5.9). Then we obtain

$$\lambda^{p/2}\|u\|^p_{\mathbb{L}_p(\tau)}+\|Du\|^p_{\mathbb{L}_p(\tau)}$$

$$\leq N\big(\sum_{i=1}^d\|f^i\|^p_{\mathbb{L}_p(\tau)}+\|g\|^p_{\mathbb{L}_p(\tau)}\big)+N\lambda^{-p/2}\|f^0\|^p_{\mathbb{L}_p(\tau)}$$

$$+N_1^*\beta_0^p\|Du\|^p_{\mathbb{L}_p(\tau)}+N_1\|u\|^p_{\mathbb{L}_p(\tau)}+N_2\lambda^{-p/2}\|Du\|^p_{\mathbb{L}_p(\tau)},$$

where the constant N_1^* depends only on d,p, and δ and the constants N also depend on K and ε. Finally, we first reduce $\beta_0=\beta_0(d,p,\delta)$ and then increase $\lambda_0\geq 0$, if necessary, in such a way that

$$N_1^*\beta_0^p\|\ Du\|^p_{\mathbb{L}_p(\tau)}+N_1\|u\|^p_{\mathbb{L}_p(\tau)}+N_2\lambda^{-p/2}\|Du\|^p_{\mathbb{L}_p(\tau)}$$

$$\leq(1/2)\lambda^{p/2}\|u\|^p_{\mathbb{L}_p(\tau)}+(1/2)\|u_x\|^p_{\mathbb{L}_p(\tau)},$$

provided that $\lambda\geq\lambda_0$. Then we obviously arrive at (5.1). The lemma is proved.

To the best of the author's knowledge the following multiplicative estimate is new even in the deterministic case.

Corollary 5.5. *Let $\lambda=0$. Then under the assumptions of Lemma 5.4 we have*

$$\|Du\|_{\mathbb{L}_p(\tau)}\leq N\big(\sum_{i=1}^d\|f^i\|_{\mathbb{L}_p(\tau)}+\|g\|_{\mathbb{L}_p(\tau)}+\|f^0\|^{1/2}_{\mathbb{L}_p(\tau)}\|u\|^{1/2}_{\mathbb{L}_p(\tau)}+\|u\|_{\mathbb{L}_p(\tau)}\big),$$

where N depends only on $d,p,K,\delta,$ and ε.

Indeed, take a $\lambda>0$ and add and subtract the term $(\lambda_0+\lambda)u_t\,dt$ on the right in (1.1), thus introducing λ into the equation and modifying f^0_t by including into it one of $(\lambda_0+\lambda)u_t$. Then after applying (5.1), we see that

$$\|Du\|_{\mathbb{L}_p(\tau)}\leq N\big(\sum_{i=1}^d\|f^i\|_{\mathbb{L}_p(\tau)}+\|g\|_{\mathbb{L}_p(\tau)}$$

$$+(\lambda_0+\lambda)^{-1/2}\|f^0\|_{\mathbb{L}_p(\tau)}+(\lambda_0+\lambda)^{1/2}\|u\|_{\mathbb{L}_p(\tau)}\big).$$

Now it only remains to take the inf with respect to $\lambda>0$.

Proof of Lemma 2.1. By bearing in mind an obvious shifting of time we see that in the proof of assertions (i)-(iii) we may assume that $s=0$.

(i) First of all observe that uniqueness of solution of (2.4) is well known even in a much wider class than $\mathcal{W}^1_p(\infty)$.

Let $u_0\in\mathrm{tr}_0\mathcal{W}^1_p$, then $u_0\in W^{1-2/p}$ for almost each ω and there is a unique solution of the heat equation

$$dv_t=\Delta v_t\,dt$$

of class $L_p((0,1),W^1_p)$ with initial condition u_0. Furthermore,

$$\|v\|_{L_p((0,1),W^1_p)} \sim \|u_0\|_{W^{1-2/p}_p}.$$

Next take a $\zeta \in C^\infty_0(\mathbb{R})$ such that $\zeta_0 = 1$ and $\zeta_t = 0$ for $t \geq 1/2$ and define $\psi_t(x) = e^{-t}v_t(x)\zeta_t$ for $t \in [0,1]$ and as zero if $t \geq 1/2$. Notice that (a.s.)

$$\psi \in L_p(\mathbb{R}_+, W^1_p),$$

and

$$\frac{\partial}{\partial t}\psi_t = \Delta\psi_t - \psi_t + e^{-t}\zeta'_t v_t$$

Then it is a classical result that there exists a unique $\phi \in L_p(\mathbb{R}_+, W^2_p)$ which solves the equation

$$d\phi_t = (\Delta\phi_t - \phi_t + e^{-t}\zeta'_t v_t)\,dt$$

with *zero* initial condition. In addition,

$$\|\phi\|_{L_p(\mathbb{R}_+,W^2_p)} \leq N\|\zeta'_t v_t\|_{L_p(\mathbb{R}_+,L_p)} \leq N\|u_0\|_{W^{1-2/p}_p},$$

where the constants N depend only on d and p. Owing to these estimates and uniqueness, the operators mapping u_0 into v and ϕ are continuous (and nonrandom). Since u_0 is $\mathcal{F}_0$-measurable, the same is true for ψ, ϕ, and $u = \psi - \phi$, which is of class $L_p((0,1),W^1_p)$, satisfies (2.4) and equals u_0 for $t = 0$. Also for each ω

$$\|u\|_{L_p(\mathbb{R}_+,W^1_p)} \leq \|\psi\|_{L_p(\mathbb{R}_+,W^1_p)} + \|\phi\|_{L_p(\mathbb{R}_+,W^1_p)} \leq N\|u_0\|_{W^{1-2/p}_p},$$

where N depends only on d and p. By raising the extreme terms to the pth power and taking expectations we get the first inequality in (2.5) and also finish proving the "only if" part of (i).

To prove the "if" part assume that we have a $v \in \mathcal{W}^1_p(\infty)$ satisfying (2.4) and equal u_0 at $t = 0$. Then $u_t = v_t e^t$ satisfies $\partial u_t/\partial t = \Delta u_t$ and is of class $\mathcal{W}^1_p(1)$. It follows that almost all ω we have $u \in L_p((0,1),W^1_p)$, $u_0 \in W^{1-2/p}_p$, and

$$\|u_0\|_{W^{1-2/p}_p} \leq N\|u\|_{L_p((0,1),W^1_p)} \leq N\|v\|_{L_p(\mathbb{R}_+,W^1_p)}.$$

By raising all expressions to the power p and taking expectations we arrive at the second estimate in (2.5). Assertion (i) is proved.

The "only if" part in (ii) is, actually, proved above. To prove the "if" part write

$$dv_t = (D_i f^i_t + f^0_t)\,dt + g^k_t\,dw^k_t = (\Delta v_t - \lambda v_t + D_i\hat{f}^i_t + \hat{f}^0_t)\,dt + g^k_t\,dw^k_t,$$

where the constant $\lambda > 0$ will be chosen later, $\hat{f}^i_t = f^i_t - D_i v_t$, $i = 1,...,d$, $\hat{f}^0_t = f^0_t + \lambda v_t$, and $\hat{f}^j, g \in \mathbb{L}_p(1)$. Next, take the function ζ as above, set $u = v\zeta$, and observe that

$$du_t = (\Delta u_t - \lambda u_t + D_i\check{f}^i_t + \check{f}^0_t)\,dt + \check{g}^k_t\,dw^k_t, \tag{5.11}$$

where $\check{f}^0 = \zeta\hat{f}^0 + v\zeta'$, $\check{f}^i_t = \zeta\hat{f}^i_t$, $i = 1,...,d$, $\check{g}^k = \zeta g^k$ and $\check{f}^j, \check{g} \in \mathbb{L}_p(\infty)$ and $u \in \mathcal{W}^1_p(\infty)$.

By Lemma 5.1, for λ fixed and large enough (actually, one can take $\lambda = 1$, which is shown by using dilations), equation (5.11) with zero initial condition admits a unique solution $\psi \in \mathcal{W}^1_p(\infty)$ and

$$\|\psi\|_{\mathbb{W}^1_p(\infty)} \le N(\sum_{j=0}^{d} \|\check{f}^j\|_{\mathbb{L}_p(\infty)} + \|\check{g}\|_{\mathbb{L}_p(\infty)})$$

$$\le N(\sum_{j=0}^{d} \|f^j\|_{\mathbb{L}_p(1)} + \|g\|_{\mathbb{L}_p(1)} + \|v\|_{\mathbb{W}^1_p(1)}).$$

Then the difference $\phi = u - \psi$ satisfies (2.4), is of class $\mathcal{W}^1_p(\infty)$, and $\phi_0 = u_0$. By assertion (i) we have $u_0 \in \mathrm{tr}_0\mathcal{W}^1_p$, which proves the "if" part in (ii). Furthermore,

$$\|u_0\|_{\mathrm{tr}_0\mathcal{W}^1_p} \le N\|\phi\|_{\mathbb{W}^1_p(\infty)} \le N\|u\|_{\mathbb{W}^1_p(\infty)} + N\|\psi\|_{\mathbb{W}^1_p(\infty)}$$

$$\le N\|v\|_{\mathbb{W}^1_p(1)} + N\|\psi\|_{\mathbb{W}^1_p(\infty)} \le N(\sum_{j=0}^{d} \|f^j\|_{\mathbb{L}_p(1)} + \|g\|_{\mathbb{L}_p(1)} + \|v\|_{\mathbb{W}^1_p(1)}).$$

This proves assertion (iii).

To prove (iv) observe that obvious dilations of the t axis allow us to assume that $s = 1$. Then write (2.2) for $t \in [0, 1]$ and notice that tu_t admits representation (2.2) with new f^j and g^k having simple relations with u_t and the original f^j and g^k. It follows that in the rest of the proof we may assume that $u_0 = 0$.

In that case take a sufficiently large $\lambda > 0$ and consider the equation

$$dv_t = (\Delta v_t - \lambda v_t + D_i\bar{f}^i_t + \bar{f}^0_t)\,dt + \bar{g}^k_t\,dw^k_t$$

for $t \ge 0$ with zero initial condition, where

$$\bar{f}^i_t = f^i_t I_{(0,1)}(t) - D_i u_t I_{(0,1)}(t), \quad i = 1, ..., d,$$

$$\bar{f}^0_t = (f^0_t + \lambda u_t)I_{(0,1)}(t), \quad \bar{g}^k_t = g^k_t I_{(0,1)}(t).$$

By uniqueness, $v_t = u_t$ for $t \in [0, 1]$ and by assertion (iii) we have $v_1 \in \mathrm{tr}_1\mathcal{W}^1_p$. This fact combined with already known estimates of v proves assertion (iv). The lemma is proved.

6. Proof of Theorem 2.2

Owing to Lemma 2.1 we may assume that we are given a v as in assertion (i) of the lemma. By introducing a new unknown function $\bar{u} = u - v$ we see that u satisfies (1.1) and $u_0 = v_0$ if and only if $\bar{u}_0 = 0$ and

$$d\bar{u}_t = (L_t\bar{u}_t - \lambda\bar{u}_t + D_j\bar{f}^j_t + \bar{f}^0_t)\,dt + (\Lambda^k_t\bar{u}_t + \bar{g}^k_t)\,dw^k_t,$$

where

$$\bar{f}^j_t = f^j_t - D_j v_t + a^{ij}_t D_i v_t + a^j_t v_t, \quad j = 1, ..., d,$$

$$\bar{f}^0_t = f^0_t + b^i_t D_i v_t + (c_t - \lambda + 1)v_t,$$

$$\bar{g}^k_t = g^k_t + \sigma^{ik}_t D_i v_t + \nu^k_t v_t.$$

By Lemma 2.1 we have $\bar f^j, \bar g \in \mathbb{L}_p(\tau)$ and the problem of finding solutions of (1.1) with initial data u_0 is thus reduced to the same problem but with zero initial data.

Furthermore, if estimate (2.7) holds for solutions with zero initial condition, then (for $\lambda \geq \lambda_0$)

$$
\begin{aligned}
\lambda^{1/2}\|u\|_{\mathbb{L}_p(\tau)} + \|Du\|_{\mathbb{L}_p(\tau)} - \lambda^{1/2}\|v\|_{\mathbb{L}_p(\tau)} - \|Dv\|_{\mathbb{L}_p(\tau)} \\
\leq \lambda^{1/2}\|\bar u\|_{\mathbb{L}_p(\tau)} + \|D\bar u\|_{\mathbb{L}_p(\tau)} \\
\leq N\big(\sum_{i=1}^d \|\bar f^i\|_{\mathbb{L}_p(\tau)} + \|\bar g\|_{\mathbb{L}_p(\tau)}\big) + N\lambda^{-1/2}\|\bar f^0\|_{\mathbb{L}_p(\tau)} \\
\leq N\big(\sum_{i=1}^d \|f^i\|_{\mathbb{L}_p(\tau)} + \|g\|_{\mathbb{L}_p(\tau)} + \|v\|_{\mathbb{W}^1_p(\tau)}\big) \\
+N\lambda^{-1/2}(\|f^0\|_{\mathbb{L}_p(\tau)} + \|v\|_{\mathbb{W}^1_p(\tau)}) + N\lambda^{1/2}\|v\|_{\mathbb{L}_p(\tau)},
\end{aligned}
$$

which yields (2.7) in full generality.

It follows that while proving (2.7) we may also assume that $u_0 = 0$. Therefore, in the rest of the proof of assertions (i) and (ii) we assume that $u_0 = 0$.

Now we suppose that Assumption 2.2 is satisfied with $\beta \leq \beta_0(d,p,\delta)$, where $\beta_0(d,p,\delta)$ is taken from Lemma 5.4. Then we take λ_0 larger than the one in Lemma 3.3 and the one in Lemma 5.4. In that case uniqueness follows from Lemma 3.3. In the proof of existence we will rely on the method of continuity and the a priori estimate (5.1) which is established in Lemma 5.4. For $\lambda \geq \lambda_0$ and $\theta \in [0,1]$ we consider the equation

$$
du_t = [(\theta L_t + (1-\theta)\Delta)u_t - \lambda u_t + D_i f^i_t + f^0_t)\,dt + (\theta\Lambda^k_t u_t + g^k_t)\,dw^k_t. \quad (6.1)
$$

We call a $\theta \in [0,1]$ "good" if the assertion of the theorem holds for equation (6.1). Observe that 0 is a "good" point by Lemma 5.1. Now to prove the theorem it suffices to show that there exists a $\gamma > 0$ such that if θ_0 is a good point then all points of the interval $[\theta_0 - \gamma, \theta_0 + \gamma] \cap [0,1]$ are "good". So fix a "good" θ_0 and for any $v \in \mathbb{W}^1_p(\tau)$ consider the equation

$$
\begin{aligned}
du_t = [(\theta_0 L_t + (1-\theta_0)\Delta)u_t - \lambda u_t + (\theta-\theta_0)(L_t - \Delta)v_t + D_i f^i_t + f^0_t)\,dt \\
+ (\theta_0\Lambda^k_t u_t + (\theta-\theta_0)\Lambda^k v_t + g^k_t)\,dw^k_t. \quad (6.2)
\end{aligned}
$$

Observe that

$$
(L_t - \Delta)v_t = D_j\big((a^{ij} - \delta^{ij})D_i v_t + a^j_t v_t\big) + b^i_t D_i v_t + c v_t
$$

and recall that $v \in \mathbb{W}^1_p(\tau)$. It follows by assumption that equation (6.2) has a unique solution $u \in \mathcal{W}^1_{p,0}(\tau)$ $(\subset \mathbb{W}^1_p(\tau))$.

In this way, for f^j and g being fixed, we define a mapping $v \to u$ in the space $\mathbb{W}^1_p(\tau)$. It is important to keep in mind that the image u of

$v \in \mathbb{W}_p^1(\tau)$ is always in $\mathcal{W}_{p,0}^1(\tau)$. Take $v', v'' \in \mathbb{W}_p^1(\tau)$ and let u', u'' be their corresponding images. Then $u := u' - u''$ satisfies

$$du_t = [(\theta_0 L_t + (1-\theta_0)\Delta)u_t - \lambda u_t + (\theta - \theta_0)(L_t - \Delta)v_t)\,dt$$
$$+(\theta_0 \Lambda_t^k u_t + (\theta - \theta_0)\Lambda^k v_t)\,dw_t^k,$$

where $v = v' - v''$. It follows by Lemma 5.4 that

$$\|u\|_{\mathbb{W}_p^1(\tau)} \leq N|\theta - \theta_0|\,\|v\|_{\mathbb{W}_p^1(\tau)}$$

with a constant N independent of v', v'', θ_0, and θ. For θ sufficiently close to θ_0, our mapping is a contraction and, since $\mathbb{W}_p^1(\tau)$ is a Banach space, it has a fixed point. This fixed point is in $\mathcal{W}_{p,0}^1(\tau)$ and, obviously, satisfies (6.1). This proves assertion (i) of the theorem.

Estimate (2.7) is proved above in Lemma 5.4 and assertion (iii) follows from Theorem 3.1. The theorem is proved.

References

[1] P. Acquistapace, *On BMO regularity for elliptic systems*, Ann. Mat. Pura e Appl. (IV), Vol. 161 (1992), 231-269.

[2] Kyeong-Hun Kim, *On stochastic partial differential equations with variable coefficients in C^1 domains*, Stochastic Process. Appl., Vol. 112 (2004), No. 2, 261-283.

[3] Kyeong-Hun Kim, *On L_p-theory of stochastic partial differential equations of divergence form in C^1 domains*, Probab. Theory Related Fields, Vol. 130 (2004), No. 4, 473-492.

[4] Kyeong-Hun Kim, *L_p estimates for SPDE with discontinuous coefficients in domains*, Electron. J. Probab., Vol. 10 (2005), No. 1, 1-20.

[5] N.V. Krylov, *A generalization of the Littlewood-Paley inequality and some other results related to stochastic partial differential equations*, Ulam Quarterly, Vol. 2 (1994), No. 4, 16–26, http://www.ulam.usm.edu/VIEW2.4/krylov.ps

[6] N.V. Krylov, *A parabolic Littlewood-Paley inequality with applications to parabolic equations*, Topological Methods in Nonlinear Analysis, Journal of the Juliusz Schauder Center, Vol. 4 (1994), No. 2, 355-364.

[7] N.V. Krylov, *An analytic approach to SPDEs*, pp. 185-242 in Stochastic Partial Differential Equations: Six Perspectives, Mathematical Surveys and Monographs, Vol. 64, AMS, Providence, RI, 1999.

[8] N.V. Krylov, *On the foundation of the L_p-theory of SPDEs*, pp. 179-191 in Stochastic Partial Differential Equations and Applications-VII, G. Da Prato, L. Tubaro eds., A Series of Lecture Notes in Pure and Applied Math., Chapman & Hall/CRC, 2006.

[9] N.V. Krylov, *Parabolic equations with VMO coefficients in Sobolev spaces with mixed norms*, Journal of Functional Analysis, Vol. 250 (2007), 521-558.

[10] N.V. Krylov, *Maximum principle for SPDEs and its applications*, pp. 311-338 in "Stochastic Differential Equations: Theory and Applications, A Volume in Honor of Professor Boris L. Rozovskii", P.H. Baxendale, S.V. Lototsky eds., Interdisciplinary Mathematical Sciences, Vol. 2, World Scientific, 2007.

[11] N.V. Krylov, *On parabolic PDEs and SPDEs in Sobolev spaces W_p^2 without and with weights*, pp 151-198 in Topics in Stochastic Analysis and Nonparametric Estimation, P.-L. Chow, B. Mordukhovich, G. Yin eds., IMA Volumes in Mathematics and its Applications, Vol. 145 (2008), Springer, New York.

[12] N.V. Krylov, *Filtering equations for partially observable diffusion processes with Lipschitz continuous coefficients*, to appear in "The Oxford Handbook of Nonlinear Filtering", Oxford University Press.

[13] N.V. Krylov, "Lectures on elliptic and parabolic equations in Sobolev spaces", Amer. Math. Soc., Graduate Studies in Math., Vol. 96, Providence, RI, 2008.

[14] N.V. Krylov, *Itô's formula for the L_p-norm of stochastic W^1_p-valued processes*, http://arxiv.org/abs/0806.1557

[15] N.V. Krylov and B.L. Rozovskii, *On conditional distributions of diffusion processes*, Izvestiya Akademii Nauk SSSR, seriya matematicheskaya, Vol. 42 (1978), No. 2 , 356-378 in Russian; English translation in Math. USSR Izvestija, Vol. 12 (1978), No. 2, 336-356.

[16] B.L. Rozovskii, "Stochastic evolution systems", Kluwer, Dordrecht, 1990.

127 Vincent Hall, University of Minnesota, Minneapolis, MN, 55455
E-mail address: krylov@math.umn.edu