Illinois Journal of Mathematics
Volume 54, Number 3, Fall 2010, Pages 1069–1114
S 0019-2082

KALMAN–BUCY FILTER AND SPDES WITH GROWING LOWER-ORDER COEFFICIENTS IN W_p^1 SPACES WITHOUT WEIGHTS

N. V. KRYLOV

Dedicated to D. L. Burkholder

ABSTRACT. We consider divergence form uniformly parabolic SPDEs with VMO bounded leading coefficients, bounded coefficients in the stochastic part, and possibly growing lower-order coefficients in the deterministic part. We look for solutions which are summable to the pth power, $p \geq 2$, with respect to the usual Lebesgue measure along with their first-order derivatives with respect to the spatial variable.

Our methods allow us to include Zakai's equation for the Kalman–Bucy filter into the general filtering theory.

1. Introduction

We consider divergence form uniformly parabolic SPDEs with bounded VMO leading coefficients, bounded coefficients in the stochastic part, and possibly growing lower-order coefficients in the deterministic part. We look for solutions which are summable to the pth power, $p \geq 2$, with respect to the usual Lebesgue measure along with their first-order derivatives with respect to the spatial variable. The present paper seems to be the first one treating the unique solvability of these equations without imposing any *special* conditions on the relations between the coefficients or on their *derivatives*.

This article in its spirit is similar to the author's recent articles [18], [12], [15], and [16] and we spare the reader the common part of the comments about the literature, which can be found in the above references. The main idea, we use, originated from [18] and [12] and relies on application of special

Received February 1, 2010; received in final form May 14, 2010.
Partially supported by NSF grant DMS-06-53121.
2010 *Mathematics Subject Classification.* 60H15, 93E11.

cut-off functions whose support evolves in time in a manner adapted to the drift terms. The paper consists of two parts: Sections 2–6 are devoted to some general issues of the theory of SPDEs with growing coefficients and in Sections 7–9 we apply the results of the previous sections to show that the filtering equations corresponding to the Kalman–Bucy filter fall into the general theory.

In a sense, the methods of the first part of the present article arose as a combination of the methods from [15] and [16] which allow us to combine the method used for PDE equations with irregular (VMO) higher-order coefficients, growing lower-order coefficients, and $p>1$ with the methods which work in similar situation for SPDEs if $p=2$. Since we are interested in higher regularity of solutions (see, for instance, Theorem 3.4), we use the power of summability $p\geq 2$ and, in contrast with [15], this forces us to require some regularity of the higher-order coefficients. Roughly speaking, we need the second-order coefficients of the deterministic part of the equation belong to VMO in x and the first-order coefficients of the stochastic part to be uniformly continuous in x. In particular, the results of the present article do not generalize those of [15].

On the other hand, if we drop all stochastic terms, then we obtain the results of [16] for $p\geq 2$, which by duality, available for deterministic equations, allows one to extend the result to full range $p>1$. Concerning the *deterministic* equations with growing coefficients in spaces with or without weights it is worth mentioning that

(i) Equations in spaces with weights are treated, for instance, in [1], [3], [5], [23], and [25] for time independent coefficients, part of the result of which are extended in [6] to time-dependent Ornstein–Uhlenbeck operators;

(ii) Equations in spaces without weights are treated, for instance, in [24], [26], [27], and [4].

Some conclusions in the above cited papers are quite similar to ours but the corresponding assumptions are not as general in what concerns the regularity of the coefficients. However, these papers contain a lot of additional important information, which is probably impossible to obtain by using our methods.

The second part of the article is devoted to the Kalman–Bucy filter. One can say that one of the sources of interest in SPDEs with growing coefficients is Zakai's equation for filtering density in the case of partially observable diffusion processes. This equation has divergence form which makes it possible to use the results of the first part of the article. In a very particular case of Gaussian processes, the filtering density is given by the Kalman–Bucy filter. Generally, part of the coefficients of filtering equations in case of Gaussian processes grow. When the coefficients of an SPDE grow, it is quite natural to consider the equations in function spaces with weights which would restrict the set of solutions in such a way that all terms in the equation will be from the same space as the free terms. There are very many articles which use this idea in $\mathcal{L}_2$- and $\mathcal{L}_p$-settings (see, for instance, [2], [9], [7], [8] and the references therein). Unfortunately, the application of the spaces with weights do not allow one to treat filtering equations corresponding to the Kalman–Bucy filter even without the so-called cross terms when the operators Λ_t^k in (7.11) are of zeroth order. The main obstacle here is that the zeroth order coefficient of Λ_t^k is a linear function of x. In the general theory, which we develop in this article, we do not allow it to grow either and we use an auxiliary function to "kill" this coefficient. The construction of this auxiliary function exploits a specific structure of the equation and allows us to transform the general filtering equation (7.11) to its "reduced" form (8.1), which does not contain the zeroth order term in the stochastic part. After that, one can use a simple change of the unknown function shifting the x variables in such a way that the stochastic part of (8.1) will disappear altogether and the equation will become a parabolic equation with time inhomogeneous and random Ornstein–Uhlenbeck operator. The fact that the operator is time inhomogeneous makes it impossible to apply any results based, for instance, on the semigroup approach and even specifically aimed at the Ornstein–Uhlenbeck operator, which one can find in the above mentioned recent articles such as [3], [5], [25], or other results on elliptic operators with unbounded coefficients such as in [27]. The results of [2] are not applicable either because in [2] the zeroth-order coefficient is assumed to grow quadratically if the firs-order coefficients grow linearly. However, the results of [9] on general SPDEs with growing coefficients are applicable to the reduced form of the SPDE for the Kalman–Bucy filter and they provide existence and uniqueness theorems in Sobolev spaces with $p=2$ and weights depending on t, x and ω. By the way, a drawback of using weights depending on t is that one cannot extract from the results for general SPDEs any result for deterministic elliptic equations.

If one concentrates on $p=2$, then one can use the results from [6] where the Ornstein–Uhlenbeck time inhomogeneous operators are investigated in Sobolev spaces with Gaussian time dependent weight. Again this would allow one to investigate (8.1) in Sobolev spaces with $p=2$ and weights depending on t, x and ω. We deal with any $p\geq 2$ and do not use weights.

The article is organized as follows. In Section 2, we introduce basic notation, function spaces, and equations. Section 3 contains our main results concerning SPDEs. Section 4 contains the proof of Theorem 3.1 concerning an a priori estimate and Theorem 3.4 about regularity properties of solutions. In Section 5, we prove the existence Theorem 3.3.

In Section 6, we prove a version of Itô's formula which allows us to use the results of the previous sections to derive the filtering equation without using anything from the filtering theory itself. We do it by following [20] and [14]. In Section 7, we state our main result about the equation corresponding to Kalman–Bucy filter. We consider the so-called conditionally Gaussian process in the spirit of [22]. However, in contrast with [22], our coefficients depend only on the current state of the two-component process under consideration and are not allowed to depend on the whole past of the observable component. In Section 8, we consider the "reduced" form (8.1) of the main filtering equation (7.11). The results of the previous sections turn out to be applicable to (8.1). In the final Section 9, we finish proving Theorems 7.1 and 7.4, part of assertions of the former being proved in Section 8.

2. General setting

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 $\mathcal{P}=\mathcal{P}(\{\mathcal{F}_t\})$ the predictable σ-field in $\Omega\times(0,\infty)$ associated with $\{\mathcal{F}_t\}$. Let w^k_t, $k=1,2,\dots$, be independent one-dimensional Wiener processes with respect to $\{\mathcal{F}_t\}$. Let τ be a stopping time.

We consider the second-order operator L_t

$$L_tu_t(x)=D_i\big(a^{ij}_t(x)D_ju_t(x)+\mathfrak{b}^i_t(x)u_t(x)\big)+b^i_t(x)D_iu_t(x)-c_t(x)u_t(x), \tag{2.1}$$

and the first-order operators

$$\Lambda^k_tu_t(x)=\sigma^{ik}_t(x)D_iu_t(x)+\nu^k_t(x)u_t(x)$$

acting on functions $u_t(x)$ defined on $\Omega\times\mathbb{R}^{d+1}_+$, where $\mathbb{R}^{d+1}_+=[0,\infty)\times\mathbb{R}^d$, and given for $k=1,2,\dots$ (the summation convention is enforced throughout the article), where

$$D_i=\frac{\partial}{\partial x^i}.$$

We set $\mathbb{R}_+=[0,\infty)$.

Our main concern in the first part of the paper is proving the unique solvability of the equation

$$du_t=(L_tu_t-\lambda u_t+D_if^i_t+f^0_t)\,dt+(\Lambda^k_tu_t+g^k_t)\,dw^k_t,\quad t\leq\tau, \tag{2.2}$$

with an appropriate initial condition at $t=0$, where $\lambda\geq 0$ is a constant. The precise assumptions on the coefficients, free terms, and initial data will be given later. First, we introduce appropriate function spaces.

Fix a number

$$p\geq 2,$$

and denote $\mathcal{L}_p=\mathcal{L}_p(\mathbb{R}^d)$. We use the same notation $\mathcal{L}_p$ for vector- and matrix-valued or else ℓ_2-valued functions such as $g_t=(g^k_t)$ in (2.2). For instance, if $u(x)=(u^1(x),u^2(x),...)$ is an ℓ_2-valued measurable function on $\mathbb{R}^d$, then

$$\|u\|^p_{\mathcal{L}_p}=\int_{\mathbb{R}^d}|u(x)|^p_{\ell_2}\,dx=\int_{\mathbb{R}^d}\left(\sum_{k=1}^\infty|u^k(x)|^2\right)^{p/2}dx.$$

As usual,

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

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

Recall that τ is a stopping time and introduce

$$\mathbb{L}_p(\tau) := \mathbb{L}_p(\{\mathcal{F}_t\},\tau) := \mathcal{L}_p((0,\tau]],\mathcal{P},\mathcal{L}_p),$$
$$\mathbb{W}^1_p(\tau) := \mathbb{W}^1_p(\{\mathcal{F}_t\},\tau) := \mathcal{L}_p((0,\tau]],\mathcal{P},W^1_p),$$
$$\mathbb{L}_p = \mathbb{L}_p(\infty), \qquad \mathbb{W}^1_p = \mathbb{W}^1_p(\infty).$$

Remember that the elements of $\mathbb{L}_p(\tau)$ need only belong to $\mathcal{L}_p$ on a predictable subset of $(0,\tau]]$ of full measure. For the sake of convenience, we will always assume that they are defined everywhere on $(0,\tau]]$ at least as generalized functions. Similar situation occurs in the case of $\mathbb{W}^1_p(\tau)$.

The following definition is most appropriate for investigating our equations if the coefficients of L_t and Λ^k_t are bounded.

DEFINITION 2.1. Introduce $\mathcal{W}^1_p(\tau)$, as the space of functions $u_t = u_t(\omega,\cdot)$ on $\{(\omega,t): 0\le t\le\tau(\omega), t<\infty\}$ with values in the space of generalized functions on $\mathbb{R}^d$ and having the following properties:

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

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

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

$$(u_{t\wedge\tau},\varphi) = (u_0,\varphi) + \sum_{k=1}^\infty \int_0^t I_{s\le\tau}(g^k_s,\varphi)\,dw^k_s + \int_0^t I_{s\le\tau}\big((f^0_s,\varphi)-(f^i_s,D_i\varphi)\big)\,ds. \tag{2.3}$$

In particular, for any $\phi\in C^\infty_0$, the process $(u_{t\wedge\tau},\phi)$ is $\mathcal{F}_t$-adapted and (a.s.) continuous. In case that property (iii) holds, we write

$$du_t = (D_i f^i_t + f^0_t)\,dt + g^k_t\,dw^k_t, \qquad t\le\tau.$$

Finally, set $\mathcal{W}^1_p = \mathcal{W}^1_p(\infty)$.

REMARK 2.1. The reader understands that if u is a generalized function on $\mathbb{R}^d$, then (u,ϕ) represents the result of the action of u on the test function $\phi\in C^\infty_0$. When u is a locally integrable function, (u,ϕ) is the integral of the product $u\phi$. According to these notation

$$(f^0_s,\varphi)-(f^i_s,D_i\varphi) = (\bar f_s,\phi),$$

where the function $\bar f_s$ with values in the space of generalized functions is defined by $\bar f_s = D_i f^i_s + f^0_s$. In the framework of Definition 2.1 we have $\bar f\in\mathcal{L}_p((0,\tau]],\mathcal{P},H^{-1}_p)$, where $H^{-1}_p = (1-\Delta)^{1/2}\mathcal{L}_p$. One also knows that any $\bar f\in\mathcal{L}_p((0,\tau]],\mathcal{P},H^{-1}_p)$ is written as $\bar f_s = D_i f^i_s + f^0_s$ with some $f^j\in\mathbb{L}_p(\tau)$.

Also introduce the spaces of initial data in the same way as in [11].

DEFINITION 2.2. Let u_0 be an $\mathcal{F}_0$-measurable function on Ω with values in the space of generalized functions on $\mathbb{R}^d$. We write $u_0 \in \operatorname{tr}\mathcal{W}_p^1 = \operatorname{tr}\mathcal{W}_p^1(\mathcal{F}_0)$ if there exists a function $v \in \mathcal{W}_p^1$ such that $dv_t = (\Delta v_t - v_t)\,dt$, $t \in \mathbb{R}_+$, and $v_0 = u_0$. In such a case, we set

$$\|u_0\|^p_{\operatorname{tr}\mathcal{W}_p^1} = E\|v\|^p_{\mathbb{W}_p^1}.$$

One knows that $\operatorname{tr}\mathcal{W}_p^1$ is a Banach space, v in the above definition is unique and $\mathcal{F}_0$-measurable.

We give the definition of solution of (2.2) adopted throughout the article and which in case the coefficients of L_t and Λ_t^k are bounded coincides with the one obtained by applying Definition 2.1.

DEFINITION 2.3. Let $f^j \in \mathbb{L}_p(\tau)$, $j = 0, \dots, d$, $g = (g^1, g^2, \dots) \in \mathbb{L}_p(\tau)$. By a solution of (2.2) (relative to $\{\mathcal{F}_t\}$) with initial condition $u_0 \in \operatorname{tr}\mathcal{W}_p^1$ we mean a function $u \in \mathbb{W}_p^1(\tau)$ (not $\mathcal{W}_p^1(\tau)$) such that

(i) For any $\phi \in C_0^\infty$, the integrals in

$$(u_{t\wedge\tau}, \phi) = (u_0, \phi) + \sum_{k=1}^\infty \int_0^t I_{s\le\tau}(\sigma_s^{ik} D_i u_s + \nu_s^k u_s + g_s^k, \phi)\,dw_s^k \tag{2.4}$$
$$+ \int_0^t I_{s\le\tau}[(b_s^i D_i u_s - (c_s + \lambda)u_s + f_s^0, \phi)$$
$$- (a_s^{ij} D_j u_s + \mathfrak{b}_s^i u_s + f_s^i, D_i\phi)]\,ds$$

are well defined and are finite for all finite $t \in \mathbb{R}_+$ and the series converges uniformly on finite subinterval of $\mathbb{R}_+$ in probability;

(ii) For any $\phi \in C_0^\infty$ with probability one, equation (2.4) holds for all $t \in \mathbb{R}_+$.

Observe that for any solution of (2.2) in the sense of the above definition and any $\phi \in C_0^\infty$ the process $(u_{t\wedge\tau}, \phi)$ is continuous (a.s.) and $\mathcal{F}_t$-adapted.

Also notice that, if the coefficients of L and Λ^k are bounded, then any $u \in \mathcal{W}_p^1(\tau)$ is a solution of (2.2) with appropriate free terms since if (2.3) holds, then (2.2) holds (always in the sense of Definition 2.3) as well with

$$f_t^i - a_t^{ij} D_j u_t - \mathfrak{b}^i u_t, \quad i = 1, \dots, d, \qquad f_t^0 + (c_t + \lambda)u_t - b_t^i D_i u_t,$$
$$g_t^k - \sigma^{ik} D_i u_t - \nu_t^k u_t$$

in place of f_t^i, $i = 1, \dots, d$, f_t^0, and g_t^k, respectively.

3. Main results for SPDEs

For $\rho > 0$, denote $B_\rho(x) = \{y \in \mathbb{R}^d : |x - y| < \rho\}$, $B_\rho = B_\rho(0)$.

ASSUMPTION 3.1. (i) The functions $a_t^{ij}(x)$, $\mathfrak{b}_t^i(x)$, $b_t^i(x)$, $c_t(x)$, $\sigma_t^{ik}(x)$, $\nu_t^k(x)$ are real valued, measurable with respect to $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}_+^{d+1})$, $\mathcal{F}_t$-adapted for any x, and $c \ge 0$.

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

$$(a^{ij} - \alpha^{ij})\xi^i \xi^j \geq \delta |\xi|^2, \qquad |a^{ij}| \leq \delta^{-1}, \qquad |\nu|_{\ell_2} \leq \delta^{-1},$$

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

(iii) For any $x \in \mathbb{R}^d$ (and ω), the function

$$\int_{B_1} \big(|\mathfrak{b}_t(x+y)| + |b_t(x+y)| + |c_t(x+y)|\big)\, dy \tag{3.1}$$

is locally integrable to the p'th power on $\mathbb{R}_+ = [0, \infty)$, where $p' = p/(p-1)$.

Notice that the matrix $a = (a^{ij})$ need not be symmetric. Also notice that in Assumption 3.1(iii) the ball B_1 can be replaced with any other ball without changing the set of admissible coefficients $\mathfrak{b}, b, c$.

Recall that as is well known if $u \in \mathbb{W}^1_p(\tau)$, then owing to the boundedness of ν and σ and the fact that $Du, u, g \in \mathbb{L}_p(\tau)$, $p \geq 2$, the first series on the right in (2.4) converges uniformly in probability and the series is a continuous local martingale. Furthermore, if we denote it by m_t, then for any $T \in \mathbb{R}_+$

$$\begin{aligned} E \sup_{t \leq T} |m_t|^p & \\ &\leq NE\left(\sum_{k=1}^{\infty} \int_0^{\tau\wedge T} (\sigma^{ik}_s D_i u_s + \nu^k_s u_s + g^k_s, \phi)^2\, ds\right)^{p/2} \\ &\leq N \|\phi\|^{p/2}_{\mathcal{L}_1} \\ &\quad \times E\left(\int_0^{\tau\wedge T} \sum_{k=1}^{\infty} (|\sigma^{ik}_s|^2 |D_i u_s|^2 + |\nu^k_s|^2 |u_s|^2 + |g^k_s|^2, |\phi|)\, ds\right)^{p/2} \\ &\leq N\big(\|u\|^p_{\mathbb{W}^1_p(\tau)} + \|g\|^p_{\mathbb{L}_p(\tau)}\big), \end{aligned} \tag{3.2}$$

where the constants N depend only on ϕ, d, p, δ, and T.

Assumption 3.2. There exists a function $\kappa(r)$, $r \in \mathbb{R}_+$, such that $\kappa(0+) = 0$ and for any $\omega \in \Omega$, $t \geq 0$, $x, y \in \mathbb{R}^d$, and $i = 1, \dots, d$ we have

$$|\sigma^{i\cdot}_t(x) - \sigma^{i\cdot}_t(y)|_{\ell_2} \leq \kappa(|x-y|).$$

The following assumptions contain parameters $\gamma_a, \gamma_b \in (0, 1]$, whose values will be specified later. They also contain constants $K \geq 0$, $\rho_0, \rho_1 \in (0, 1]$ which are fixed.

Assumption 3.3. For any $\omega \in \Omega$, $\rho \in (0, \rho_0]$, $t \geq 0$, and $i, j = 1, \dots, d$, we have

$$\rho^{-2d-2} \int_t^{t+\rho^2} \left(\sup_{x \in \mathbb{R}^d} \int_{B_\rho(x)} \int_{B_\rho(x)} |a^{ij}_s(y) - a^{ij}_s(z)|\, dy\, dz\right) ds \leq \gamma_a. \tag{3.3}$$

Obviously, the left-hand side of (3.3) is less than

$$N(d)\sup_{t\geq 0}\sup_{|x-y|\leq 2\rho}|a^{ij}_t(x)-a^{ij}_t(y)|,$$

which implies that Assumption 3.3 is satisfied with any $\gamma_a>0$ if, for instance, a is uniformly continuous in x uniformly in ω and t. Recall that if a is independent of t and for any $\gamma_a>0$ there is a $\rho_0>0$ such that Assumption 3.3 is satisfied, then one says that a is in VMO.

We take and fix a number $q=q(d,p)$ such that

$$q\geq \max(d,p)\quad \text{if } p\neq d,\qquad q>d\quad \text{if } p=d. \tag{3.4}$$

ASSUMPTION 3.4. For any $\omega\in\Omega$, $\mathfrak{b}:=(\mathfrak{b}^1,\dots,\mathfrak{b}^d)$, $b:=(b^1,\dots,b^d)$, and $(t,x)\in\mathbb{R}^{d+1}$, we have

$$\int_{B_{\rho_1}(x)}\int_{B_{\rho_1}(x)}|\mathfrak{b}_t(y)-\mathfrak{b}_t(z)|^q\,dy\,dz+\int_{B_{\rho_1}(x)}\int_{B_{\rho_1}(x)}|b_t(y)-b_t(z)|^q\,dy\,dz$$
$$+\int_{B_{\rho_1}(x)}\int_{B_{\rho_1}(x)}|c_t(y)-c_t(z)|^q\,dy\,dz\leq KI_{q>d}+\rho_1^d\gamma_b.$$

Obviously, Assumption 3.4 is satisfied if b, $\mathfrak{b}$, and c are independent of x. They also are satisfied with any $q>d$, $\gamma_b=0$, and $\rho_1=1$ on the account of choosing K appropriately if, say,

$$|\mathfrak{b}_t(x)-\mathfrak{b}_t(y)|+|b_t(x)-b_t(y)|+|c_t(x)-c_t(y)|\leq K_1$$

whenever $|x-y|\leq 1$, where K_1 is a constant. In particular, Assumption 3.4 is satisfied if $\mathfrak{b}$, b, and c are globally Lipschitz continuous:

$$\begin{aligned}|\mathfrak{b}_t(x)-\mathfrak{b}_t(y)|&+|b_t(x)-b_t(y)|+|c_t(x)-c_t(y)|\\ &\leq K_1|x-y|\quad \forall x,y\in\mathbb{R}^d, t\geq 0.\end{aligned} \tag{3.5}$$

We see that Assumption 3.4 allows b, $\mathfrak{b}$, and c growing linearly in x. Here is our result on a priori estimates of solutions of (2.2).

THEOREM 3.1. *There exist*

$$\begin{aligned}&\gamma_a=\gamma_a(d,\delta,p),\qquad \gamma_b=\gamma_b(d,\delta,p,\kappa,\rho_0)\in(0,1],\\ &N=N(d,\delta,p,\kappa,\rho_0),\qquad \lambda_0=\lambda_0(d,\delta,p,\kappa,\rho_0,\rho_1,K)\geq 1\end{aligned}$$

such that, if the above assumptions are satisfied and $\lambda\geq\lambda_0$ *and* u *is a solution of* (2.2) *with initial data* $u_0\in \operatorname{tr}\mathcal{W}^1_p$ *and some* $f^j,g\in\mathbb{L}_p(\tau)$, *then*

$$\begin{aligned}\lambda\|u\|^2_{\mathbb{L}_p(\tau)}+\|Du\|^2_{\mathbb{L}_p(\tau)}\leq N&\left(\sum_{i=1}^d\|f^i\|^2_{\mathbb{L}_p(\tau)}+\|g\|^2_{\mathbb{L}_p(\tau)}\right)\\ &+N\lambda^{-1}\|f^0\|^2_{\mathbb{L}_p(\tau)}+N\|u_0\|^2_{\operatorname{tr}\mathcal{W}^1_p}.\end{aligned} \tag{3.6}$$

REMARK 3.1. There is an unusual property of u_t, which is nontrivial even if $f_t^j = g_t^k \equiv 0$.

Namely, assume that $g \equiv 0$. Take a predictable ℓ_2-valued process ξ_t such that $(\nu_t, \xi_t)_{\ell_2} \geq 0$ and $(\nu_t, \xi_t)_{\ell_2}$ and $(\sigma_t^{i\cdot}, \xi_t)$ are independent of x (which happens, for instance, if $\nu = 0$ and σ is independent of x) and

$$\int_0^\tau |\xi_t|_{\ell_2}^2 \, dt < \infty$$

(a.s.) and assume that $E\rho_\tau(\xi) = 1$, where

$$\rho_t(\xi) = \rho_t(\xi, dw) := \exp\left(- \int_0^t \xi_s^k \, dw_s^k - \frac{1}{2} \int_0^t |\xi_s|_{\ell_2}^2 \, ds \right).$$

Then the assertion of Theorem 3.1 holds with the *same* γ_a, γ_b, λ_0, and N if we understand $\|v\|_{\mathbb{L}_p(\tau)}^p$ for all v's as

$$E\rho_\tau \int_0^\tau \|v_t\|_{\mathcal{L}_p}^p \, dt.$$

Indeed, one can change the probability measure by using Girsanov's theorem. This will add a new drift term in the deterministic part of (2.2) and this additional drift depends only on (ω, t). This will also add the term $-(\nu_t, \xi_t)_{\ell_2} u_t \, dt$, where $(\nu_t, \xi_t)_{\ell_2}$ is nonnegative and also independent of x. Then the result follows immediately from Theorem 3.1.

Theorem 3.1 admits the following version if τ is bounded.

THEOREM 3.2. *Let $T \in (0, \infty)$ be a constant and suppose that $\tau \leq T$. Assume that the above assumptions are satisfied with γ_a and γ_b from Theorem 3.1. Let $\lambda = 0$ and let u be a solution of (2.2) with initial data $u_0 \in \operatorname{tr} \mathcal{W}_p^1$ and some $f^j, g \in \mathbb{L}_p(\tau)$. Then*

$$\|u\|_{\mathbb{W}_p^1(\tau)}^2 \leq N \left(\sum_{i=0}^d \|f^i\|_{\mathbb{L}_p(\tau)}^2 + \|g\|_{\mathbb{L}_p(\tau)}^2 + \|u_0\|_{\operatorname{tr} \mathcal{W}_p^1}^2 \right), \tag{3.7}$$

where $N = N(d, \delta, p, \kappa, \rho_0, \rho_1, K, T)$.

This result is a trivial consequence of Theorem 3.1 since, for any constant μ, the function $v_t := u_t e^{-\mu t}$ satisfies (2.2) with $\lambda + \mu$, $f_t^j e^{-\mu t}$, and $g_t^k e^{-\mu t}$ in place of λ, f_t^j, and g_t^k, respectively. If μ is large enough and $\tau \leq T$, estimate (3.6) for v implies (3.7) indeed.

REMARK 3.2. Theorems 3.1 and 3.2 provide uniqueness of solutions of (2.2). The a priori estimates (3.6) and (3.7) can also be used to investigate continuous dependence of solutions on the coefficients and other data.

To prove the existence, we need stronger assumptions because, generally, Assumption 3.4 does not guarantee that

$$D_i(\mathfrak{b}_t^i u_t) + b_t^i D_i u_t - c_t u_t$$

can be written even locally as $D_i\hat f^i_t+\hat f^0_t$ with $\hat f^j\in\mathbb{L}_p(\tau)$ if we only know that $u\in\mathbb{W}^1_p(\tau)$ even if $\mathfrak{b}$, b, and c are independent of x. We can only prove our Lemma 5.2 if we have a certain control on this expression.

Assumption 3.5. For any $x\in\mathbb{R}^d$ (and ω), the function (3.1) is locally integrable to the power $p/(p-2)$ (locally bounded if $p=2$) on $\mathbb{R}_+=[0,\infty)$.

Remark 3.3. Assumptions 3.4 and 3.5 are both satisfied if the global Lipschitz condition (8.7) holds and $b_t(0)$, $\mathfrak{b}_t(0)$, and $c_t(0)$ are bounded for each ω.

Theorem 3.3. *Let the above assumptions be satisfied with γ_a and γ_b taken from Theorem 3.1. Take $\lambda\ge\lambda_0$, where λ_0 is defined in Theorem 3.1, and take $u_0\in\operatorname{tr}\mathcal{W}^1_p$. Then there exists a unique solution of (2.2) with initial condition u_0.*

Remark 3.4. If the stopping time τ is bounded, then in the above theorem one can take $\lambda_0=0$. This is shown by the same argument as after Theorem 3.2.

In general, the continuity properties in t of the solution from Theorem 3.3 are unknown. For instance, we do not know if $\|u_{t\wedge\tau}\phi\|_{\mathcal{L}_p}$ is continuous (a.s.) for any $\phi\in C^\infty_0$. However, under stronger assumptions we can say more about regularity of u. In the following theorem by H^γ_p, we mean $(1-\Delta)^{-\gamma/2}\mathcal{L}_p$.

Theorem 3.4. *Under the above assumptions suppose that for each $x\in\mathbb{R}^d$ the function (3.1) is bounded on $(0,\tau]\!]$. Then the (unique) solution u possesses the following properties:*

(i) *For any $\phi\in C^\infty_0$, we have $\phi u\in\mathcal{W}^1_p(\tau)$;*

(ii) *For any $\phi\in C^\infty_0$, the process $u_{t\wedge\tau}\phi$ is continuous on $\mathbb{R}_+$ as an $\mathcal{L}_p$-valued process (a.s.);*

(iii) *If $p>2$ and τ is bounded and we have two numbers α and β such that*

$$\frac{2}{p}<\alpha<\beta\le 1,$$

then for any $\phi\in C^\infty_0$ (a.s.)

$$u\phi\in C^{\alpha/2-1/p}([0,\tau],H^{1-\beta}_p).$$

In particular, if $p>d+2$, then

(a) *for any $\varepsilon\in(0,\varepsilon_0]$, with*

$$\varepsilon_0=1-\frac{d+2}{p},$$

(a.s.) for any $t\in[0,\tau]$ we have $u_t\phi\in C^{\varepsilon_0-\varepsilon}(\mathbb{R}^d)$ and the norm of $u_t\phi$ in this space is bounded as a function of t;

(b) *for any* ε *as in* (a) (*a.s.*) *for any* $x \in \mathbb{R}^d$ *we have* $u_\cdot(x)\phi(x) \in C^{(\varepsilon_0-\varepsilon)/2}([0,\tau])$ *and the norm of* $u_\cdot(x)\phi(x)$ *in this space is bounded as a function of* x.

Observe that assertions (ii) and (iii) of Theorem 3.4 follow from assertion (i) proved in Remark 4.1. In case of assertion (ii) this is shown in [13]. The main part of assertion (iii) follows from assertion (i) and Corollary 4.12 [10]. By applying Sobolev's embedding theorems, assertion (iii) (a) is obtained after taking α and β close to $2/p$ and (iii) (b) after taking α and β close to $1-d/p$.

REMARK 3.5. Let $p_1, p_2 \in [2,\infty)$, let τ be bounded (cf. Remark 3.4), and let the assumptions of Theorem 3.3 be satisfied for any $p \in [p_1, p_2]$ with γ_a and γ_b which are suitable for all $p \in [p_1, p_2]$. Then it turns out that the solution from Theorem 3.3 corresponding to $p = p_1$ coincides with the one obtained for $p = p_2$.

This fact is obtained in the same way as the proof of Theorem 3.4 of [16] is obtained from the proof of Theorem 3.3 [16].

Our last main result on general SPDEs bears on the measurability of u_t with respect to σ-fields which are smaller than $\mathcal{F}_t$. It will be used in Section 8 and this is the reason why we use the somewhat strange notation $\tilde{y}_t$ and $\tilde{\mathrm{B}}^k_t$ below. We suppose that all the above assumptions are satisfied with γ_a and γ_b taken from Theorem 3.1 and let $\tilde{\mathcal{F}}_t$, $t \geq 0$, be a filtration of complete with respect to $\mathcal{F}, P$ σ-fields such that $\mathcal{F}_t \supset \tilde{\mathcal{F}}_t$. Our aim is to show that sometimes u_t is $\tilde{\mathcal{F}}_t$-adapted even if some terms in (2.2) are not $\tilde{\mathcal{F}}_t$-adapted. However, the equation is assumed to have a special structure. The result is not surprising because in the notation, introduced below, the equation

$$du_t = (\Lambda^k_t u_t + g^k_t)\, dw^k_t + (L_t u_t + \hat{b}^i_t D_i u_t - \hat{c}_t u_t + D_i f^i_t + f^0_t + \hat{f}_t)\, dt, \quad t \leq \tau \tag{3.8}$$

is written as

$$du_t = (\Lambda^k_t u_t + g^k_t)\, d\tilde{y}^k_t + (L_t u_t + D_i f^i_t + f^0_t)\, dt, \quad t \leq \tau. \tag{3.9}$$

THEOREM 3.5. *Fix a number* $T \in (0,\infty)$. *Assume that we are given an* ℓ_2*-valued process* $\tilde{\mathrm{B}}_t$ *which is* $\mathcal{F}_t$*-adapted, jointly measurable with respect to* (ω, t), *and such that* $|\tilde{\mathrm{B}}_t|_{\ell_2}$ *is locally square integrable on* $\mathbb{R}_+$ *and* $E\rho_T = 1$, *where*

$$\rho_t = \rho_t(\tilde{\mathrm{B}}, dw) = \exp\left(-\int_0^t \tilde{\mathrm{B}}^k_s\, dw^k_s - \frac{1}{2}\int_0^t |\tilde{\mathrm{B}}_s|^2_{\ell_2}\, ds\right).$$

Suppose that Assumption 3.1(i) *is satisfied with* $\tilde{\mathcal{F}}_t$ *in place of* $\mathcal{F}_t$ *and the processes*

$$\tilde{y}^k_t = w^k_t + \int_0^t \tilde{\mathrm{B}}^k_s\, ds$$

are $\tilde{\mathcal{F}}_t$ adapted. Introduce

$$\bar{b}^i_t(x) = \sigma^{ik}_t(x)\tilde{\mathrm{B}}^k_t, \qquad \bar{c}_t(x) = -\nu^k_t(x)\tilde{\mathrm{B}}^k_t$$

and suppose that $b + \bar{b}$ and $c + \bar{c}$ satisfy Assumption 3.4 with γ_b from Theorem 3.1, for any $x \in \mathbb{R}^d$ (and ω) we have $\bar{c}_t(x) \le K$, and the function

$$\int_{B_1} \big(|\bar{b}_t(x+y)| + |\bar{c}_t(x+y)|\big)\,dy \tag{3.10}$$

is locally integrable to the power $p/(p-2)$ (locally bounded if $p=2$) on $\mathbb{R}_+$. Let τ be an $\tilde{\mathcal{F}}_t$-stopping time such that $\tau \le T$.

Then, for any initial data $u_0 \in \operatorname{tr}\mathcal{W}^1_p(\tilde{\mathcal{F}}_0)$ and $f^j, g \in \mathbb{L}_p(\{\tilde{\mathcal{F}}_t\}, \tau)$ such that $\tilde{f} := (g, \tilde{\mathrm{B}})_{\ell_2} \in \mathbb{L}_p(\{\mathcal{F}_t\}, \tau)$,

(i) *equation (3.8) has a unique solution u relative to $\{\mathcal{F}_t\}$ in the sense of Definition 2.3,*

(ii) *for any $\phi \in C^\infty_0$ the process $(u_{t\wedge\tau}, \phi)$ is $\tilde{\mathcal{F}}_t$-adapted.*

Proof. Owing to the argument after Theorem 3.2 allowing us to introduce as large λ as we wish, assertion (i) follow immediately from Theorem 3.3.

To prove (ii), we use a change of measure. Define $\tilde{P}(d\omega) = \rho_T(\omega)P(d\omega)$, notice that by Girsanov's theorem the processes $\tilde{y}^k_t$, $t \le T$, are independent Wiener processes with respect to $\tilde{P}, \mathcal{F}_t$. By assumption, they are $\tilde{\mathcal{F}}_t$-adapted and since $\tilde{\mathcal{F}}_t \subset \mathcal{F}_t$ the increments $\tilde{y}^k_{t+s} - \tilde{y}^k_t$ are independent of $\tilde{\mathcal{F}}_t$ if $s \ge 0$. Thus $(\tilde{y}^k_t, \tilde{\mathcal{F}}_t)$ are independent Wiener processes. Introduce $\tilde{E}$ as the expectation sign relative to $\tilde{P}$.

After rewriting (3.8) in form (3.9) and applying Theorems 3.1 and 3.3, we get that there exists a unique solution $\tilde{u}$ of (3.8) with initial data u_0 relative to $\{\tilde{\mathcal{F}}_t\}$ in the sense of Definition 2.3 on the new probability space, that is with $\mathbb{L}_p(\tau)$ and $\mathbb{W}^1_p(\tau)$ replaced with $\tilde{\mathbb{L}}_p(\{\tilde{\mathcal{F}}_t\}, \tau)$ and $\tilde{\mathbb{W}}^1_p(\{\tilde{\mathcal{F}}_t\}, \tau)$, respectively, where the norms in these spaces are defined as

$$\tilde{E}\int_0^\tau \|u_t\|^p_{\mathcal{L}_p}\,dt \quad \text{and} \quad \tilde{E}\int_0^\tau \|u_t\|^p_{W^1_p}\,dt$$

raised to the power $1/p$, respectively.

Now for $n \ge 2$, we introduce $\mathcal{F}_t$-stopping times

$$\tau_n = \tau \wedge \inf\{t \ge 0 : \rho_t \le 1/n\}$$

and observe that

$$E\int_0^{\tau_n} \|\tilde{u}_t\|^p_{\mathcal{L}_p}\,dt \le nE\rho_{\tau_n}\int_0^{\tau_n} \|\tilde{u}_t\|^p_{\mathcal{L}_p}\,dt = n\tilde{E}\int_0^{\tau_n} \|\tilde{u}_t\|^p_{\mathcal{L}_p}\,dt < \infty.$$

Similar estimates hold if we replace $\mathcal{L}_p$ with W^1_p. By recalling that $\tilde{\mathcal{F}}_t \subset \mathcal{F}_t$, we conclude that $\tilde{u}$ is a solution of (3.8) relative to $\{\mathcal{F}_t\}$ with τ_n in place of τ. By uniqueness, in the sense of distributions $\tilde{u}_t I_{t\le\tau_n} = u_t I_{t\le\tau_n}$ for almost all (ω, t), that is, $(\tilde{u}_t, \phi) I_{t\le\tau_n} = (u_t, \phi) I_{t\le\tau_n}$ for almost all (ω, t) for each fixed

$\phi\in C_0^\infty$. Then it follows from the integral form of (3.8) that for each $\phi\in C_0^\infty$ with probability one $(\tilde u_{t\wedge\tau_n},\phi)=(u_{t\wedge\tau_n},\phi)$ for all t. Upon letting $n\to\infty$, we replace τ_n with τ and it only remains to observe that $(\tilde u_{t\wedge\tau},\phi)$ is $\tilde{\mathcal F}_t$-measurable. The theorem is proved. $\square$

The following is almost identical to Remark 3.5 of [15].

REMARK 3.6. We do not use the spaces with weights. However, there is a trivial and since very long time known way how to use results like ours for treating equations in spaces with weights. For instance, let $\psi_t(x)>0$ be a nonrandom smooth function on $\mathbb{R}^{d+1}$. Introduce, $\partial_t=\partial/\partial t$,

$$\hat{\mathfrak b}_t^i=\mathfrak b_t^i-a_t^{ij}D_j\ln\psi_t,\qquad \hat b_t^i=b_t^i-a_t^{ij}D_j\ln\psi_t,$$
$$\hat c_t=c_t+(b_t^i+\mathfrak b_t^i)D_i\ln\psi_t-a_t^{ij}(D_i\ln\psi_t)D_j\ln\psi_t-\partial_t\ln\psi_t,$$
$$\hat\nu_t^k=\nu_t^k-\sigma_t^{ik}D_i\ln\psi_t,$$
$$\hat f_t^i=\psi_tf_t^i,\quad i=1,\dots,d,\qquad \hat f_t^0=f_t^0\psi_t-f_t^iD_i\psi_t,\qquad \hat g_t^k=g_t^k\psi_t.$$

Suppose that, if we replace b, $\mathfrak b$, c, and ν with $\hat b$, $\hat{\mathfrak b}$, $\hat c$, and $\hat\nu$, respectively, then Assumptions 3.1, 3.4, and 3.5 are satisfied with γ_a and γ_b from Theorem 3.1. Finally, assume that $\hat f^j,\hat g\in\mathbb{L}_2(\tau)$ and $u_0\psi_0\in\operatorname{tr}\mathcal{W}_p^1$. Then it turns out that for $\lambda\ge\lambda_0$ (λ_0 is taken from Theorem 3.1) equation (2.2) has a unique solution u such that $u\psi\in\mathbb{W}_p^1(\tau)$.

This fact is almost trivial since u satisfies (2.2) if and only if $v:=u\psi$ satisfies the version of (2.2) which is obtained as the result of the replacements described above and also the replacement of f^j,g with $\hat f^j,\hat g$, respectively. In addition, the natural estimate of the $\mathbb{W}_p^1(\tau)$-norm of v gives an estimate of u in an appropriate space with weights.

As a specification of the above, in the setting of Remark 3.3 take a $T\in(0,\infty)$, set $\tau=T$, and for $\theta\in(0,\infty)$ introduce

$$\ln\psi_t(x)=-\theta e^{\theta^2(t-T)}\sqrt{1+|x|^2}.$$

Obviously, $D_i\ln\psi$ are bounded for $t\le T$. Furthermore, it is not hard to see that if θ is large enough, then $\hat c_t\ge0$ for $t\le T$. Also, if $|x-y|\le1$, then owing to the fact that $|D_{ij}\ln\psi_t(x)|\le N(1+|x|)^{-1}$ for $t\le T$, where N is a constant, we have

$$\begin{aligned}&|b_t^i(x)D_i\ln\psi_t(x)-b_t^i(y)D_i\ln\psi_t(y)|\\&\quad\le\big|\big(b_t^i(x)-b_t^i(y)\big)D_i\ln\psi_t(x)\big|+N(1+|x|)|D\ln\psi_t(x)-D\ln\psi_t(y)|\\&\quad\le K|D\ln\psi_t(x)|+N\end{aligned}$$

for $t\le T$. Estimates similar to this one show that $\hat b$, $\hat{\mathfrak b}$, and $\hat c$ satisfy Assumption 3.4 for $t\le T$. By what is said in the beginning of the current remark, if $u_0\psi_0\in\operatorname{tr}\mathcal{W}_p^1$ (for instance, $u_0(x)=x^1$), then (2.2) has a unique solution u

such that $u\psi \in \mathbb{W}^1_p(T)$. Since $D\ln\psi$ is bounded, the inclusion $u\psi \in \mathbb{W}^1_p(T)$ is equivalent to $u\psi \in \mathbb{L}_p(T)$, $\psi Du \in \mathbb{L}_p(T)$.

To the best of the author's knowledge even in this special case the result in this generality was not known before.

4. Proof of Theorems 3.1 and 3.4

In this section, we suppose that Assumptions 3.1, 3.2, 3.3, and 3.4 are satisfied with some $\gamma_a, \gamma_b \in (0,1]$ and start by showing that the requirement (i) of Definition 2.3 is automatically satisfied for any $u \in \mathbb{W}^1_p(\tau)$. Take a nonnegative $\xi \in C_0^\infty(B_{\rho_1})$ with unit integral and define

$$\bar{b}_s(x) = \int_{B_{\rho_1}} \xi(y) b_s(x-y)\,dy, \qquad \bar{\mathfrak{b}}_s(x) = \int_{B_{\rho_1}} \xi(y)\mathfrak{b}_s(x-y)\,dy,$$
$$\bar{c}_s(x) = \int_{B_{\rho_1}} \xi(y) c_s(x-y)\,dy. \tag{4.1}$$

We may assume that $|\xi| \le N(d)\rho_1^{-d}$.

REMARK 4.1. By Corollary 5.4 of [16], for $x_0 \in \mathbb{R}^d$, $v \in \mathcal{L}_p$, $\phi \in W^1_{p'}$, and $u \in W^1_p$ we have

$$\big(|b_s - \bar{b}_s(x_0)| I_{B_{\rho_1}(x_0)} v, |\phi|\big) \le N\|v\|_{\mathcal{L}_p}\|\phi\|_{W^1_{p'}},$$
$$\big\|I_{B_{\rho_1}(x_0)}|\mathfrak{b}_s - \bar{\mathfrak{b}}_s(x_0)|u\big\|_{\mathcal{L}_p} + \big\|I_{B_{\rho_1}(x_0)}|c_s - \bar{c}_s(x_0)|u\big\|_{\mathcal{L}_p} \le N\|u\|_{W^1_p}, \tag{4.2}$$

where $N = N(d, p, \rho_1, K)$. In particular,

$$\big(|b_s| I_{B_{\rho_1}(x_0)} v, |\phi|\big) \le \big(N + |\bar{b}_s(x_0)|\big)\|v\|_{\mathcal{L}_p}\|\phi\|_{W^1_{p'}}, \tag{4.3}$$

the latter implying that $|b_s| I_{B_{\rho_1}(x_0)} v \in H^{-1}_p$. It is also seen that if $u \in \mathbb{W}^1_p(\tau)$ and $|\bar{b}_s(x_0)|$ is a *bounded* function on $(0,\tau]\!]$, then

$$I_{B_{\rho_1}}(x_0) b^i D_i u \in \mathcal{L}_p((0,\tau]\!], \mathcal{P}, H^{-1}_p).$$

Similarly,

$$\begin{aligned}&\big\|I_{B_{\rho_1}(x_0)}|\mathfrak{b}_s|u\big\|_{\mathcal{L}_p} + \big\|I_{B_{\rho_1}(x_0)}|c_s|u\big\|_{\mathcal{L}_p}\\ &\qquad \le \big(N + |\bar{\mathfrak{b}}_s(x_0)| + |\bar{c}_s(x_0)|\big)\|u\|_{W^1_p}.\end{aligned} \tag{4.4}$$

By the way, Remark 2.1 now shows that under the conditions of Theorem 3.4 for any solution u of (2.2) and $\phi \in C_0^\infty$ with support lying in a ball of radius ρ_1 we have $u\phi \in \mathcal{W}^1_p(\tau)$. Of course, the restriction on the size of support of ϕ is easily removed and this proves assertion (i) of Theorem 3.4.

LEMMA 4.1. *Let $R \in (0,\infty)$. Then there exists a sequence of bounded stopping times $\tau_n \to \infty$ such that for any $\omega \in \Omega$, $u \in \mathcal{L}_p((0,\tau), W_p^1)$, and $\phi \in C_0^\infty(B_R)$*

$$\int_0^{\tau_n \wedge \tau} \bigl(|(b_s^i D_i u_s, \phi)| + |(\mathfrak{b}_s^i u_s, D_i \phi)| + |(c_s u_s, \phi)|\bigr)\, ds \le n\|u\|_{\mathcal{L}_p((0,\tau),W_p^1)} \|\phi\|_{W_{p'}^1}, \tag{4.5}$$

so that requirement (i) *in Definition 2.3 can be dropped.*

Proof. By having in mind partitions of unity, we convince ourselves that it suffices to prove (4.5) under the assumption that ϕ has support in a ball $B_{\rho_1}(x_0)$. Observe that by (4.4) and Hölder's inequality

$$|(\mathfrak{b}_s^i u_s, D_i \phi)| + |(c_s u_s, \phi)| \le N\bigl(1 + |\bar{\mathfrak{b}}_s(x_0)| + |\bar{c}_s(x_0)|\bigr) \|u_s\|_{W_p^1} \|\phi\|_{W_{p'}^1}. \tag{4.6}$$

It follows again by Hölder's inequality that

$$\int_0^{t \wedge \tau} \bigl(|(\mathfrak{b}_s^i u_s, D_i \phi)| + |(c_s u_s, \phi)|\bigr)\, ds \le N \chi_t \|u\|_{\mathcal{L}_p((0,\tau),W_p^1)} \|\phi\|_{W_{p'}^1},$$

where

$$\chi_t = t^{1/p'} + \left(\int_0^t |\bar{\mathfrak{b}}_s(x_0)|^{p'}\, ds\right)^{1/p'} + \left(\int_0^t |\bar{c}_s(x_0)|^{p'}\, ds\right)^{1/p'}.$$

After that, in what concerns $\mathfrak{b}$ and c, it only remains to recall Assumption 3.1(iii). Similarly the integral of $|(b_s^i D_i u_s, \phi)|$ is estimated by using (4.3) and the lemma is proved. □

REMARK 4.2. Estimates (4.3) and (4.4) show that for any $u \in \mathbb{W}_p^1$ for almost all (ω, s) the functions $b_s^i D_i u_s$, $D_i(\mathfrak{b}_s^i u_s)$, and $c_s u_s$ are distributions on $\mathbb{R}^d$.

Since bounded linear operators are continuous we obtain the following.

COROLLARY 4.2. *Let R, τ_n, ϕ be as in Lemma 4.1. Then the operators*

$$u_t \to \int_0^{t\wedge\tau_n} (b_s^i D_i u_s, \phi)\, ds, \qquad u_t \to \int_0^{t\wedge\tau_n} (\mathfrak{b}_s^i u_s, D_i \phi)\, ds,$$

$$u_t \to \int_0^{t\wedge\tau_n} (c_s u_s, \phi)\, ds$$

are continuous as operators from $\mathbb{W}_p^1(\tau)$ to $\mathcal{L}_p((0,\tau_n]\!])$ for any n.

This result will be used in Section 5.

Now we prove Theorem 3.1 in a particular case.

LEMMA 4.3. *Let $\mathfrak{b}^i$, b^i, and c be independent of x and let $u_0 = 0$. Then the assertion of Theorem 3.1 holds, naturally, with $\lambda_0 = \lambda_0(d, \delta, p, \rho_0, \kappa)$ (independent of ρ_1).*

Proof. First, let $c \equiv 0$. We want to use the Itô-Wentzell formula to get rid of the first-order terms. Observe that (2.2) reads as

$$
\begin{aligned}
(4.7)\quad du_t = {} & (\Lambda^k_t u_t + g^k_t) dw^k_t \\
& + \big[D_i\big(a^{ij}_t D_j u_t + (\mathfrak{b}^i_t + b^i_t) u_t + f^i_t\big) + f^0_t - \lambda u_t\big]\, dt, \quad t \le \tau.
\end{aligned}
$$

Recall that from the start (see Definition 2.3) it is assumed that $u \in \mathbb{W}^1_p(\tau)$. Then one can find a predictable set $A \subset (0, \tau]\!]$ of full measure such that $I_A f^j$, $j = 0, 1, \dots, d$, $I_A g$, and $I_A D_i u$, $i = 1, \dots, d$, are well defined as $\mathcal{L}_p$-valued predictable functions satisfying

$$\int_0^\tau I_A \left(\sum_{j=0}^d \|f^j_t\|^p_{\mathcal{L}_p} + \|g_t\|^p_{\mathcal{L}_p} + \|Du_t\|^p_{\mathcal{L}_p} \right) dt < \infty.$$

Replacing f^j, g, and $D_i u$ in (4.7) with $I_A f^j$, $I_A g$, and $I_A D_i u$, respectively, will not affect (4.7). Similarly one can treat the term $h_t = (\mathfrak{b}^i_t + b^i_t) u_t$ for which

$$\int_0^{T \wedge \tau} \|h_t\|_{\mathcal{L}_p}\, dt < \infty$$

(a.s.) for each $T \in \mathbb{R}_+$, owing to Assumption 3.1 and the fact that $u \in \mathbb{L}_p(\tau)$.

After these replacements all terms on the right in (4.7) will be of class $\mathfrak{D}^1$ and $\mathfrak{D}^2$ as appropriate since a and σ are bounded (see the definition of $\mathfrak{D}^1$ and $\mathfrak{D}^2$ in [17]). This allows us to apply Theorem 1.1 of [17] and for

$$B^i_t = \int_0^t (\mathfrak{b}^i_s + b^i_s)\, ds, \qquad \hat{u}_t(x) = u_t(x - B_t)$$

obtain that

$$(4.8)\qquad d\hat{u}_t = [D_i(\hat{a}^{ij}_t D_j \hat{u}_t) - \lambda \hat{u}_t + D_i \hat{f}^i_t + \hat{f}^0_t]\, dt + (\hat{\Lambda}^k_t \hat{u}_t + \hat{g}^k_t)\, dw^k_t,$$

where $\hat{\Lambda}^k_t = \hat{\sigma}^{ik}_t D_i + \hat{\nu}^k_t$ and

$$(\hat{a}^{ij}_t, \hat{\sigma}^{ik}_t, \hat{\nu}^k_t, \hat{f}^j_t, \hat{g}^k_t)(x) = (a^{ij}_t, \sigma^{ik}_t, \nu^k_t, f^j_t, g^k_t)(x - B_t).$$

Obviously, $\hat{u}$ is in $\mathbb{W}^1_p(\tau)$ and its norm coincides with that of u. Equation (4.8) shows that $\hat{u} \in \mathcal{W}^1_p(\tau)$.

Next observe that owing to (3.3), for any $\omega \in \Omega, \rho \in (0, \rho_0], t \ge 0$, and $i, j = 1, \dots, d$ we have

$$\rho^{-2d-2} \int_t^{t+\rho^2} \left(\sup_{x \in \mathbb{R}^d} \int_{B_\rho(x)} \int_{B_\rho(x)} |\hat{a}^{ij}_s(y) - \hat{a}^{ij}_s(z)|\, dy\, dz \right) ds \le \gamma_a,$$

which in terms of [11] implies that the couple $(\hat{a}, \hat{\sigma})$ is $(\varepsilon, \varepsilon)$-regular at any point of $\mathbb{R}_+ \times \mathbb{R}^d$ for any $\varepsilon \in (0, \rho_0]$. Then owing to our Assumptions 3.1(ii) and 3.2 one can choose $\varepsilon = \varepsilon(\delta, \kappa) \in (0, \rho_0]$ so that Assumption 2.2 of [11] is satisfied.

By Theorem 2.2 of [11] if Assumption 3.3 is satisfied with $\gamma_a = \gamma_a(d,\delta,p) > 0$, specified in its proof, and if $\lambda \geq \lambda_0(d,\delta,p,\kappa,\rho_0) \geq 1$, then

$$\lambda\|\hat{u}\|^2_{\mathbb{L}_p(\tau)} + \|D\hat{u}\|^2_{\mathbb{L}_p(\tau)} \leq N\left(\sum_{i=1}^d \|\hat{f}^i\|^2_{\mathbb{L}_p(\tau)} + \|\hat{g}\|^2_{\mathbb{L}_p(\tau)} + \lambda^{-1}\|\hat{f}^0\|^2_{\mathbb{L}_p(\tau)}\right),$$

where $N = N(d,\delta,p,\kappa,\rho_0)$. This coincides with (3.6) and proves the lemma in case $c \equiv 0$.

In the general case, observe that owing to Assumption 3.1(iii) there exists a sequence of stopping times $\tau_n \uparrow \tau$ such that

$$\int_0^{\tau_n} c_s\,ds \leq n.$$

Clearly, if we can prove (3.6) with τ_n in place of τ, then by passing to the limit we will get (3.6) as is. Therefore, without losing generality, we assume that

$$\sup_\Omega \int_0^\infty c_s\,ds < \infty.$$

Then introduce

$$\xi_t = \exp\left(\int_0^t c_s\,ds\right).$$

By the above argument, we have $\bar{u} := \xi u \in \mathbb{W}^1_p(\tau)$ and

$$\begin{aligned} d\bar{u}_t = {} & \big[D_i(a^{ij}_t D_j\bar{u}_t + [\mathfrak{b}^i_t + b^i_t]\bar{u}_t + \xi_t f^i_t) + \xi_t f^0_t - \lambda\bar{u}_t\big]\,dt \\ & + (\Lambda^k_t\bar{u}_t + \xi_t g^k_t)\,dw^k_t, \quad t \leq \tau. \end{aligned}$$

By the above result for any stopping time $\tau' \leq \tau$

$$\begin{aligned} & \lambda^{p/2}\|\xi u\|^p_{\mathbb{L}_p(\tau')} + \|\xi Du\|^p_{\mathbb{L}_p(\tau')} \\ & \quad = \lambda^{p/2}\|\bar{u}\|^p_{\mathbb{L}_p(\tau')} + \|D\bar{u}\|^p_{\mathbb{L}_p(\tau')} \\ & \quad \leq N\left(\sum_{i=1}^d \|\xi f^i\|^p_{\mathbb{L}_p(\tau')} + \|\xi g\|^p_{\mathbb{L}_p(\tau')} + \lambda^{-p/2}\|\xi f^0\|^p_{\mathbb{L}_p(\tau')}\right). \end{aligned} \tag{4.9}$$

If needed, one can enlarge the original probability space in such a way that there will exist an exponentially distributed, with parameter one, random variable η independent of $\{\mathcal{F}_t, t \geq 0\}$. We assume that the enlargement is not needed and define

$$\phi_t = p\int_0^t c_s\,ds, \qquad \psi_s = \tau \wedge \inf\{t \geq 0 : \phi_t \geq s\}, \quad \tau' = \psi_\eta.$$

Notice that

$$\{\omega : \psi_s > t\} = \{\omega : \tau > t, \phi_t < s\}.$$

Hence,

$$\{\omega : \tau' > t\} = \{\omega : \tau > t, \phi_t < \eta\}.$$

It follows that τ' is a stopping time with respect to $\mathcal{F}_t \vee \sigma(\eta)$. Furthermore, for any nonnegative predictable (relative to the original filtration $\mathcal{F}_t$) process h_t we have

$$E\int_0^{\tau'} h_t\,dt = \int_0^\infty Eh_t E\{I_{\tau'>t} \mid \mathcal{F}_t\}\,dt$$
$$= \int_0^\infty Eh_t I_{\tau>t} e^{-\phi_t}\,dt = E\int_0^\tau h_t \xi_t^{-p}\,dt.$$

This and (4.9) immediately lead to (3.6) and the lemma is proved. □

To proceed further take $\bar{\mathfrak{b}}, \bar{b}$, and $\bar{c}$ from (4.1). From Lemma 4.2 of [12] and Assumption 3.4 it follows that, for $h_t = \bar{\mathfrak{b}}_t, \bar{b}_t, \bar{c}_t$, it holds that $|D^n h_t| \le M_n$, where $M_n = M_n(n, d, \rho_1, K) \ge 1$ and $D^n h_t$ is any derivative of h_t of order $n \ge 1$ with respect to x. By Corollary 4.3 of [12], we have $|h_t(x)| \le K(t)(1 + |x|)$, where for each ω the function $K(t) = K(\omega, t)$ is locally integrable with respect to t on $\mathbb{R}_+$. Owing to these properties the equation

$$x_t = x_0 - \int_{t_0}^t (\bar{\mathfrak{b}}_s + \bar{b}_s)(x_s)\,ds, \quad t \ge t_0, \tag{4.10}$$

for any (ω and) $(t_0, x_0) \in \mathbb{R}^{d+1}_+$ has a unique solution $x_t = x_{t_0,x_0,t}$. Obviously, the process $x_{t_0,x_0,t}$, $t \ge t_0$, is $\mathcal{F}_t$-adapted.

Next, for $i = 1, 2$ set $\chi^{(i)}(x)$ to be the indicator function of $B_{\rho_1/i}$ and introduce

$$\chi^{(i)}_{t_0,x_0,t}(x) = \chi^{(i)}(x - x_{t_0,x_0,t}) I_{t \ge t_0}.$$

By using the above results and reproducing the proofs of Lemma 5.5 of [15], where $p = 2$ and SPDEs are treated, and Lemma 5.8 of [16], where p is general but only PDEs are considered, we easily obtain the following.

LEMMA 4.4. *Suppose that Assumption 3.3 is satisfied with $\gamma_a = \gamma_a(d, \delta, p)$ taken from Lemma 4.3. Assume that we are given a function u which is a solution of (2.2) with some $f^j, g \in \mathbb{L}_p(\tau)$, and $\lambda \ge \lambda_0 = \lambda_0(d, \delta, p, \rho_0, \kappa)$, where $\lambda_0(d, \delta, p, \rho_0, \kappa)$ is taken from Lemma 4.3. Take $(t_0, x_0) \in \mathbb{R}^{d+1}_+$ and assume that $u_t = 0$ if $t \le t_0 \wedge \tau$. Then*

$$\begin{aligned}
&\lambda^{p/2}\|\chi^{(2)}_{t_0,x_0} u\|^p_{\mathbb{L}_p(\tau)} + \|\chi^{(2)}_{t_0,x_0} Du\|^p_{\mathbb{L}_p(\tau)} \\
&\quad \le N\left(\sum_{i=1}^d \|\chi^{(1)}_{t_0,x_0} f^i\|^p_{\mathbb{L}_p(\tau)} + \|\chi^{(1)}_{t_0,x_0} g\|^p_{\mathbb{L}_p(\tau)}\right) \\
&\qquad + N\lambda^{-p/2}\|\chi^{(1)}_{t_0,x_0} f^0\|^p_{\mathbb{L}_p(\tau)} \\
&\qquad + N\gamma_b^{p/q}\|\chi^{(1)}_{t_0,x_0} Du\|^p_{\mathbb{L}_p(\tau)} + N^*\lambda^{-p/2}\|\chi^{(1)}_{t_0,x_0} Du\|^p_{\mathbb{L}_p(\tau)} \\
&\qquad + N^*\|\chi^{(1)}_{t_0,x_0} u\|^p_{\mathbb{L}_p(\tau)} + N^*\lambda^{-p/2}\sum_{i=1}^d \|\chi^{(1)}_{t_0,x_0} f^i\|^p_{\mathbb{L}_p(\tau)},
\end{aligned} \tag{4.11}$$

where N is a constant depending only on d, δ, p, ρ_0, and κ and N^* depends only on the same objects, γ_b, ρ_1, *and* K.

Upon integrating through equation (4.11) with respect to x_0 and repeating the arguments in the proofs of Lemma 5.6 of [15] or Lemma 5.9 of [16], we obtain the following result in which $M_1(d, \rho_1, K)$ is the constant introduced before Lemma 4.4.

LEMMA 4.5. *Suppose that Assumption* 3.3 *is satisfied with* $\gamma_a = \gamma_a(d, \delta, p)$ *taken from Lemma* 4.3. *Assume that we are given a function u which is a solution of* (2.2) *with some* $f^j, g \in \mathbb{L}_p(\tau)$, *and* $\lambda \ge \lambda_0 = \lambda_0(d, \delta, p, \rho_0, \kappa)$, *where* $\lambda_0(d, \delta, p, \rho_0, \kappa)$ *is taken from Lemma* 4.3. *Take an* $s_0 \in \mathbb{R}_+$ *and assume that* $u_t = 0$ *if* $t \le s_0 \wedge \tau$. *Then for* $I_{s_0} := I_{(s_0, t_0)}$, *where* $t_0 = s_0 + M_1^{-1}$, *we have*

$$
\begin{aligned}
\text{(4.12)}\qquad & \lambda^{p/2} \|I_{s_0} u\|^p_{\mathbb{L}_p(\tau)} + \|I_{s_0} Du\|^p_{\mathbb{L}_p(\tau)} \\
&\le N\left(\sum_{i=1}^d \|I_{s_0} f^i\|^p_{\mathbb{L}_p(\tau)} + \|I_{s_0} g\|^p_{\mathbb{L}_p(\tau)}\right) \\
&\quad + N\lambda^{-p/2} \|I_{s_0} f^0\|^p_{\mathbb{L}_p(\tau)} + N\gamma_b^{p/q} \|I_{s_0} Du\|^p_{\mathbb{L}_p(\tau)} \\
&\quad + N^*\lambda^{-p/2} \|I_{s_0} Du\|^p_{\mathbb{L}_p(\tau)} + N^* \|I_{s_0} u\|^p_{\mathbb{L}_p(\tau)} \\
&\quad + N^*\lambda^{-p/2} \sum_{i=1}^d \|I_{s_0} f^i\|^p_{\mathbb{L}_p(\tau)},
\end{aligned}
$$

where N is a constant depending only on d, δ, p, ρ_0, and κ and N^* depends only on the same objects, γ_b, ρ_1, *and* K.

Proof of Theorem 3.1. First, we show how to choose an appropriate $\gamma_b = \gamma_b(d, \delta, p, \rho_0, \kappa)$. Call N_0 the constant factor of $\gamma_b^{p/q} \|I_{s_0} Du\|^p_{\mathbb{L}_p(\tau)}$ in (4.12) and choose a $\gamma_b \in (0, 1]$ in such a way that $N_0 \gamma_b^{p/q} \le 1/2$. Then under the assumptions of Lemma 4.5 we have

$$
\begin{aligned}
\text{(4.13)}\qquad & \lambda^{p/2} \|I_{s_0} u\|^p_{\mathbb{L}_p(\tau)} + \|I_{s_0} Du\|^p_{\mathbb{L}_p(\tau)} \\
&\le N\left(\sum_{i=1}^d \|I_{s_0} f^i\|^p_{\mathbb{L}_p(\tau)} + \|I_{s_0} g\|^p_{\mathbb{L}_p(\tau)}\right) \\
&\quad + N\lambda^{-p/2} \|I_{s_0} f^0\|^p_{\mathbb{L}_p(\tau)} + N^*\lambda^{-p/2} \|I_{s_0} Du\|^p_{\mathbb{L}_p(\tau)} \\
&\quad + N^* \|I_{s_0} u\|^p_{\mathbb{L}_p(\tau)} + N^*\lambda^{-p/2} \sum_{i=1}^d \|I_{s_0} f^i\|^p_{\mathbb{L}_p(\tau)}.
\end{aligned}
$$

To proceed further, assume that

$$
\text{(4.14)}\qquad u_0 = 0.
$$

After γ_b has been fixed, we recall that $M_1 = M_1(d, \rho_1, K)$ and we take a $\zeta \in C_0^\infty(\mathbb{R})$ with support in $(0, M_1^{-1})$ such that

$$\int_{-\infty}^{\infty} \zeta^p(t)\,dt = 1. \tag{4.15}$$

For $s \in \mathbb{R}$, define $\zeta_t^s = \zeta(t-s)$, $u_t^s(x) = u_t(x)\zeta_t^s$. Obviously, $u_t^s = 0$ if $0 \le t \le s_+ \wedge \tau$. Therefore, we can apply (4.13) to u_t^s by taking $s_0 = s_+$ and observing that

$$\begin{aligned} du_t^s = {} & \big(L_t u_t^s - \lambda u_t^s + D_i(\zeta_t^s f_t^i) + \zeta_t^s f_t^0 + u_t(\zeta_t^s)'\big)\,dt \\ & + (\Lambda_t^k u_t^s + \zeta_t^s g_t^k)\,dw_t^k, \quad t \le \tau. \end{aligned}$$

We also use the fact that for $t \ge 0$, as is easy to see, $I_{s_+}(t)\zeta_t^s = \zeta_t^s$. Then for and $\lambda \ge \lambda_0 = \lambda_0(d, \delta, p, \rho_0, \kappa)$, where $\lambda_0(d, \delta, p, \rho_0, \kappa)$ is taken from Lemma 4.3, we obtain

$$\begin{aligned} & \lambda^{p/2}\|\zeta^s u\|^p_{\mathbb{L}_p(\tau)} + \|\zeta^s Du\|^p_{\mathbb{L}_p(\tau)} \\ & \le N\left(\sum_{i=1}^d \|\zeta^s f^i\|^p_{\mathbb{L}_p(\tau)} + \|\zeta^s g\|^p_{\mathbb{L}_p(\tau)}\right) \\ & \quad + N\lambda^{-p/2}\big(\|\zeta^s f^0\|^p_{\mathbb{L}_p(\tau)} + \|(\zeta^s)' u\|^p_{\mathbb{L}_p(\tau)}\big) \\ & \quad + N^*\lambda^{-p/2}\|\zeta^s Du\|^p_{\mathbb{L}_p(\tau)} + N^*\|\zeta^s u\|^p_{\mathbb{L}_p(\tau)} \\ & \quad + N^*\lambda^{-p/2}\sum_{i=1}^d \|\zeta^s f^i\|^p_{\mathbb{L}_p(\tau)}. \end{aligned} \tag{4.16}$$

We integrate through this relation with respect to $s \in \mathbb{R}$, use (4.15) and

$$\int_{-\infty}^{\infty} |(\zeta_t^s)'|^p\,ds = \int_{-\infty}^{\infty} |\zeta'(t)|^p\,dt = N^*.$$

Then we conclude

$$\begin{aligned} \lambda^{p/2}\|u\|^p_{\mathbb{L}_p(\tau)} + \|Du\|^p_{\mathbb{L}_p(\tau)} \le {} & N_1\left(\sum_{i=1}^d \|f^i\|^p_{\mathbb{L}_p(\tau)} + \|g\|^p_{\mathbb{L}_p(\tau)}\right) \\ & + N_1\lambda^{-p/2}\|f^0\|^p_{\mathbb{L}_p(\tau)} + N_1^*\lambda^{-p/2}\|Du\|^p_{\mathbb{L}_p(\tau)} \\ & + N_1^*\|u\|^p_{\mathbb{L}_p(\tau)} + N_1^*\lambda^{-p/2}\sum_{i=1}^d \|f^i\|^p_{\mathbb{L}_p(\tau)}. \end{aligned}$$

Without losing generality, we assume that $N_1 \ge 1$ and we show how to choose $\lambda_0 = \lambda_0(d, \delta, p, \rho_0, \rho_1, \kappa, K) \ge 1$. Above, we assumed that $\lambda \ge \lambda_0(d, \delta, p, \rho_0, \kappa)$, where $\lambda_0(d, \delta, p, \rho_0, \kappa)$ is taken from Lemma 4.3. Therefore, we take

$$\lambda_0 = \lambda_0(d, \delta, p, \rho_0, \rho_1, \kappa, K) \ge \lambda_0(d, \delta, p, \rho_0, \kappa)$$

such that $\lambda_0^{p/2} \geq 2N_1^*$ (recall that $N_1^* = N_1^*(d, \delta, p, \rho_0, \rho_1, \kappa, K)$). Then we obviously come to (3.6) (with $u_0 = 0$).

A standard method to remove assumption (4.14) by subtracting from u the solution of the heat equation $dv_t = (\Delta v_t - v_t)\,dt$ with initial data u_0 does not work because it leads to subtracting the terms $D_i(\mathfrak{b}^i v) + b^i D_i v$, which one should include into the free terms $D_i f^i + f^0$ in the equation. Generally, this is impossible because we only know that $D_i v \in \mathbb{L}_p(\tau)$ and if we multiply $D_i v$ by an arbitrary function of x with linear growth, the inclusion may fail.

Therefore, we use a different method. The idea is to shift all data along the time axis by 1, consider our equations on $(1, \hat{\tau}]\!]$, where $\hat{\tau} = 1 + \tau$, and supplement this equation with an equation for $t \in [0, 1]$ with zero initial data and such that the value of its solution at time 1 would coincide with u_0. Then the two equations combined would give an equation on $(0, \hat{\tau}]\!]$ with zero initial condition, which would allow us to apply the above result.

Formally, we need to have Wiener processes on $[0, \infty)$ and after shifting they will be defined only on $[1, \infty)$ (and satisfy $w_1^k = 0$). Therefore, we augment if needed our probability space in such a way that we may assume that there are Wiener processes $\bar{w}_t^1, \bar{w}_t^2, \dots,$ $t \geq 0$, independent of $\{\mathcal{F}_s, s \geq 0\}$. Then define $\mathcal{F}_t^{\bar{w}}$ as the completion of $\sigma(\bar{w}_s : s \leq t)$,

$$\hat{\mathcal{F}}_t = \mathcal{F}_0 \vee \mathcal{F}_t^{\bar{w}}, \quad t \in [0, 1], \qquad \hat{\mathcal{F}}_t = \mathcal{F}_{t-1} \vee \mathcal{F}_1^{\bar{w}}, \quad t \geq 1,$$
$$\hat{w}_t^k = \bar{w}_t^k, \quad t \in [0, 1], \qquad \hat{w}_t^k = \bar{w}_1^k + w_{t-1}^k, \quad t \geq 1, \qquad \hat{\tau} = 1 + \tau,$$

and for $t \geq 1$ define the coefficients and the free terms by following the example $\hat{a}_t^{ij} = a_{t-1}^{ij}$.

Next, take the function v from Definition 2.2 and for $t \in [0, 1]$ set

$$\hat{a}_t^{ij} = \delta^{ij}, \qquad \hat{f}_t^i = -2t D_i v_{1-t}, \qquad \hat{f}_t^0 = (1 + t + \lambda t) v_{1-t},$$

where $\lambda \geq \lambda_0$ with λ_0 determined in the first part of the proof. We define all other coefficients with hats and the free terms $\hat{g}_t^k$ to be zero for $t \in [0, 1]$. Notice that for $\hat{u}_t = t v_{1-t}$, $t \in [0, 1]$, we have

$$d\hat{u}_t = [D_i(\hat{a}_t^{ij} D_j \hat{u}_t + \hat{f}_t^i) + \hat{f}_t^0 - \lambda \hat{u}_t]\,dt.$$

Moreover, $\hat{u}_0 = 0$, $\hat{u}_1 = u_0$, and $\hat{u}_t$ is $\hat{\mathcal{F}}_t$-adapted. Therefore, naturally we define $\hat{u}_t = u_{t-1}$ for $t \geq 1$.

It is easy to see that if we construct the operators $\hat{L}_t$ and $\hat{\Lambda}_t^k$ from the coefficients with hats, then

$$d\hat{u}_t = (\hat{L}_t \hat{u}_t - \lambda \hat{u}_t + D_i \hat{f}_t^i + \hat{f}_t^0)\,dt + (\hat{\Lambda}_t^k \hat{u}_t + \hat{g}_t^k)\,d\hat{w}_t^k, \quad t \leq \hat{\tau}.$$

By the first part of the proof,

$$\begin{aligned} \lambda \|u\|_{\mathbb{L}_p(\tau)}^2 + \|Du\|_{\mathbb{L}_p(\tau)}^2 &\leq \lambda \|\hat{u}\|_{\mathbb{L}_p(\hat{\tau})}^2 + \|D\hat{u}\|_{\mathbb{L}_p(\hat{\tau})}^2 \\ &\leq N \left(\sum_{i=1}^d \|\hat{f}^i\|_{\mathbb{L}_p(\hat{\tau})}^2 + \|\hat{g}\|_{\mathbb{L}_p(\hat{\tau})}^2 \right) + N\lambda^{-1} \|\hat{f}^0\|_{\mathbb{L}_p(\hat{\tau})}^2 \end{aligned}$$

$$\leq N\left(\sum_{i=1}^{d}\|f^i\|^2_{\mathbb{L}_p(\tau)}+\|g\|^2_{\mathbb{L}_p(\tau)}\right)+N\lambda^{-1}\|f^0\|^2_{\mathbb{L}_p(\tau)}$$
$$+N(\|v\|^2_{\mathbb{L}_p}+\|Dv\|^2_{\mathbb{L}_p}).$$

It only remains to notice that the last term is dominated by $N\|u_0\|^2_{\mathrm{tr}\,\mathcal{W}^1_p}$. The theorem is proved. □

5. Proof of Theorem 3.3

Throughout this section, we suppose that the assumptions of Theorem 3.3 are satisfied.

Owing to Theorem 3.1, implying that the solution in $\mathbb{W}^1_p(\tau)$ is unique, and having in mind setting all data equal to zero for $t>\tau$, we see that without loss of generality we may assume that $\tau=\infty$. Set

$$\mathbb{L}_p=\mathbb{L}_p(\infty),\qquad \mathbb{W}^1_p=\mathbb{W}^1_p(\infty).$$

We need two auxiliary results.

LEMMA 5.1. *For any $T,R\in(0,\infty)$ (and ω), we have*

$$\int_0^T\int_{B_R}\big(|\mathfrak{b}_s(x)|^{p'}+|b_s(x)|^{p'}+c_s^{p'}(x)\big)\,dx\,ds<\infty. \tag{5.1}$$

This lemma is proved in the same way as Lemma 6.1 of [16] on the basis of Assumptions 3.1(iii) and 3.4 and the fact that $q\geq p'$.

The solution of our equation will be obtained as the weak limit of the solutions of equations with cut-off coefficients. Therefore, the following result is relevant.

LEMMA 5.2. *Let $\phi\in C_0^\infty$, u^m, $u\in\mathbb{W}^1_p$, $m=1,2,\dots$, be such that $u^m\to u$ weakly in $\mathbb{W}^1_p$. For $m=1,2,\dots$ define $\chi_m(t)=(-m)\vee t\wedge m$, $\mathfrak{b}^i_{mt}=\chi_m(\mathfrak{b}^i_t)$, $b^i_{mt}=\chi_m(b^i_t)$, and $c_{mt}=\chi_m(c_t)$. Then there is a sequence of bounded stopping times $\tau_n\to\infty$ such that, for any n, the functions*

$$\int_0^t(b^i_{ms}D_iu^m_s,\phi)\,ds,\qquad\int_0^t(\mathfrak{b}^i_{ms}u^m_s,D_i\phi)\,ds,\qquad\int_0^t(c_{ms}u^m_s,\phi)\,ds \tag{5.2}$$

converge weakly in the space $\mathcal{L}_p((0,\tau_n]])$ as $m\to\infty$ to

$$\int_0^t(b^i_sD_iu_s,\phi)\,ds,\qquad\int_0^t(\mathfrak{b}^i_su_s,D_i\phi)\,ds,\qquad\int_0^t(c_su_s,\phi)\,ds, \tag{5.3}$$

respectively.

Proof. Let R be such that $\phi(x)=0$ for $|x|\geq R$. We take $\tau_n\to\infty$ such that each of them is bounded, they are smaller than the ones from Lemma 4.1, and are such that the left-hand side of (5.1) with $T=\tau_n$ is less than n.

By Corollary 4.2 and by the fact that (strongly) continuous operators are weakly continuous, we obtain that

$$\int_0^t (b^i_s D_i u^m_s, \phi)\, ds \to \int_0^t (b^i_s D_i u_s, \phi)\, ds$$

as $m \to \infty$ weakly in the space $\mathcal{L}_p((0, \tau_n]\!])$ for any n. Therefore, in what concerns the first function in (5.2), it suffices to show that

$$\int_0^t \bigl(D_i u^m_s, (b^i_s - b^i_{ms})\phi\bigr)\, ds \to 0$$

weakly in $\mathcal{L}_p((0, \tau_n]\!])$. In other words, it suffices to show that for any $\xi \in \mathcal{L}_{p'}((0, \tau_n]\!])$

$$E\int_0^{\tau_n} \xi_t \biggl(\int_0^t \bigl(D_i u^m_s, (b^i_s - b^i_{ms})\phi\bigr)\, ds\biggr)\, dt \to 0.$$

This relation is rewritten as

$$E\int_0^{\tau_n} \bigl(D_i u^m_s, \eta_s(b^i_s - b^i_{ms})\phi\bigr)\, ds \to 0, \tag{5.4}$$

where

$$\eta_s := \int_s^{\tau_n} \xi_t\, dt.$$

Observe that by the choice of τ_n, we have

$$\begin{aligned} E\int_0^{\tau_n} |\eta_s|^{p'} \int_{|x|\le R} |b_s(x)|^{p'}\, dx\, ds &\le E \sup_{t\le\tau_n} |\eta_s|^{p'} \int_0^{\tau_n}\int_{|x|\le R} |b_s(x)|^{p'}\, dx\, ds \\ &\le nE\biggl(\int_0^{\tau_n} |\xi_s| ds\biggr)^{p'} < \infty. \end{aligned}$$

It follows by the dominated convergence that $\eta_s(b^i_s - b^i_{ms})\phi \to 0$ as $m \to \infty$ strongly in $\mathbb{L}_{p'}(\tau_n)$. By assumption $Du^m \to Du$ weakly in $\mathbb{L}_p(\tau_n)$. This implies (5.4). Similarly, one proves our assertion about the remaining functions in (5.2). The lemma is proved. □

Proof of Theorem 3.3. Recall that we may assume that $\tau = \infty$. Since the case $p = 2$ is dealt with in [15] (under much milder assumptions), we also assume that $p > 2$. Define $\mathfrak{b}_{mt}$, b_{mt}, and c_{mt} as in Lemma 5.2 and consider equation (2.2) with $\mathfrak{b}_{mt}$, b_{mt}, and c_{mt} in place of $\mathfrak{b}_t$, b_t, and c_t, respectively. Obviously, $\mathfrak{b}_{mt}$, b_{mt}, and c_{mt} satisfy Assumption 3.4 with the same γ_b and K as $\mathfrak{b}_t$, b_t, and c_t do. By Theorem 3.1 and the method of continuity for $\lambda \ge \lambda_0(d, \delta, p, \kappa, \rho_0, \rho_1, K)$ there exists a unique solution u^m of the modified equation on $\mathbb{R}$.

By Theorem 3.1, we also have

$$\|u^m\|_{\mathbb{L}_p} + \|Du^m\|_{\mathbb{L}_p} \le N,$$

where N is independent of m. Hence, the sequence of functions u^m is bounded in the space $\mathbb{W}^1_p$ and consequently has a weak limit point $u\in\mathbb{W}^1_p$. For simplicity of presentation, we assume that the whole sequence u^m converges weakly to u.

Take a $\phi\in C_0^\infty$. Then by Lemma 5.2 for appropriate τ_n we have that the functions (5.2) converge to (5.3) weakly in $\mathcal{L}_p((0,\tau_n]])$ as $m\to\infty$ for any n. Owing to (3.2) and the fact that bounded linear operators are weakly continuous, the stochastic terms in the equations for u^m_t also converge weakly in $\mathcal{L}_p((0,\tau_n]])$ as $m\to\infty$ for any n. Obviously, the same is true for $(u^m_t,\phi)\to(u_t,\phi)$ and the remaining terms entering the equation for u^m_t. Hence, by passing to the weak limit in the equation for u^m_t we see that for any $\phi\in C_0^\infty$ equation (2.4) holds for *almost any* (ω,t).

Until this moment, Assumption 3.5 was not needed. We will need it in order to be able to apply Theorem 3.1 of [19] and find an appropriate modification of u_t.

Take a $\psi\in C_0^\infty$ and observe that $u\psi\in\mathbb{W}^1_2(T)$ and $g\psi\in\mathbb{L}_2(T)$ for any $T\in(0,\infty)$ which implies that

$$m^\psi_t:=u_0\psi+\sum_{k=1}^\infty\int_0^t\psi(\Lambda^k_su_s+g^k_s)\,dw^k_s$$

is well defined as an $\mathcal{L}_2$-valued continuous martingale such that for any $\phi\in\mathcal{L}_2$ with probability one

$$(m^\psi_t,\phi)=(u_0\psi,\phi)+\sum_{k=1}^\infty\int_0^t\big(\psi(\Lambda^k_su_s+g^k_s),\phi\big)\,dw^k_s \tag{5.5}$$

for all $t\in\mathbb{R}_+$.

Notice that for any $\phi\in C_0^\infty$

$$(u_t\psi,\phi)=\int_0^t(u^*_s\psi,\phi)\,ds+(m^\psi_t,\phi) \tag{5.6}$$

for almost all (ω,t), where u^*_s is a function with values in the space of distributions on $\mathbb{R}^d$ defined by

$$u^*_s=L_su_s-\lambda u_s+D_if^i_s+f^0_s$$

(see Remark 4.2).

Next, take an $R\in(0,\infty)$ and let $W^{-1}_{p'}(B_R)$ denote the dual space for

$$\overset{0}{W}{}^1_p(B_R):=W^1_p(B_R)\cap\{v:\,v|_{\partial B_R}=0\}.$$

Estimate (4.6) combined with the facts that, $p'<p$ and that one can cover B_R with finitely many balls of radius ρ_1 shows that for any $\phi\in C_0^\infty(B_R)$

$$|(D_i(\mathfrak{b}^i_su_s),\phi)|\le N\bigg(1+\int_{B_{R+1}}|\mathfrak{b}_s|\,dx\bigg)\|u_s\|_{W^1_p}\|\phi\|_{W^1_p},$$

where N is independent of ω, s, u_s, ϕ. Due to the arbitrariness of ϕ and the fact that $C_0^\infty(B_R)$ is dense in $\overset{0}{W}{}_p^1(B_R)$, we conclude that (for almost all (ω, s)) we have $D_i(\mathfrak{b}_s^i u_s) \in W_{p'}^{-1}(B_R)$ and

$$\|D_i(\mathfrak{b}_s^i u_s)\|_{W_{p'}^{-1}(B_R)} \le N\left(1 + \int_{B_{R+1}} |\mathfrak{b}_s|\, dx\right)\|u_s\|_{W_p^1}.$$

Here the right-hand side is locally summable on $\mathbb{R}_+$ to the power p' (a.s.) owing to Assumption 3.5, Hölder's inequality, and the fact that $u \in \mathbb{W}_p^1$. Similar statements are true for $b_s^i D_i u_s$, $c_s u_s$, and u_s^*.

Now, since $u\psi \in \mathcal{L}_p(\mathbb{R}_+, \overset{0}{W}{}_p^1(B_R))$ and $\overset{0}{W}{}_p^1(B_R)$ is dense in $\mathcal{L}_2(B_R)$, by Theorem 3.1 of [19] we get that there exist an event Ω^ψ of full probability and a continuous $\mathcal{L}_2(B_R)$-valued $\mathcal{F}_t$-adapted process u_t^ψ such that $u_t^\psi = u_t\psi$ as $\mathcal{L}_2(B_R)$-valued functions for almost all (ω, t) and for any $\omega \in \Omega^\psi$, $t \in \mathbb{R}_+$, and $\phi \in C_0^\infty(B_R)$ we have

$$(u_t^\psi, \phi) = \int_0^t (u_s^*\psi, \phi)\, ds + (m_t^\psi, \phi). \tag{5.7}$$

Take a $\psi \in C_0^\infty$ such that $\psi(x) = 1$ for $|x| \le 1$ and for $k = 1, 2, \ldots$ define $\psi_k(x) = \psi(x/k)$ and

$$\Omega' = \bigcap_{k=1}^\infty \Omega^{\psi_k}.$$

Clearly, $P(\Omega') = 1$. We will further reduce Ω' in the following way. Obviously (see (5.5)), if $\psi', \psi'' \in C_0^\infty$ and $\psi' = \psi''$ on B_R and $\phi \in \mathcal{L}_2$ is such that $\phi = 0$ outside B_R, then with probability one we have $(m_t^{\psi'}, \phi) = (m_t^{\psi''}, \phi)$ for all t.

Let Φ be the union over $n = 1, 2, \ldots$ of countable subsets of $C_0^\infty(B_n)$ each of which everywhere dense in $\mathcal{L}_2(B_n)$. For $\phi \in C_0^\infty$ denote $d(\phi)$ the smallest radius of the balls centered at the origin containing the support of ϕ. Then by the above for $\phi \in C_0^\infty$ the events

$$\Omega(\phi) = \{\omega \in \Omega : (m_t^{\psi_k}, \phi) = (m_t^{\psi_j}, \phi),\ \forall t \in \mathbb{R}_+, k, j \ge d(\phi)\},$$
$$\Omega'' = \Omega' \bigcap \bigcap_{\phi \in \Phi} \Omega(\phi)$$

have probability one. Since m_t^ψ are $\mathcal{L}_2$-valued and $\Phi \cap C_0^\infty(B_n)$ is dense in $\mathcal{L}_2(B_n)$, we have that for $\omega \in \Omega''$, $t \in \mathbb{R}_+$, and any $\phi \in C_0^\infty(B_n)$ it holds that

$$(m_t^{\psi_k}, \phi) = (m_t^{\psi_j}, \phi)$$

as long as $i, j \ge n$.

Then (5.7) implies that for any $\omega \in \Omega''$, $t \in \mathbb{R}_+$, and $\phi \in C_0^\infty(B_n)$ we have $(u_t^{\psi_j}, \phi) = (u_t^{\psi_k}, \phi)$ for all $j, k \ge n$. In particular, for any $\omega \in \Omega''$, $t \in \mathbb{R}_+$, $n = 1, 2 \ldots$ it holds that $u_t^{\psi_j} = u_t^{\psi_k}$ as distributions on B_n for $j, k \ge n$ and there exists a distribution $\bar u_t$ on $\mathbb{R}^d$ such that $\bar u_t = u_t^{\psi_k}$ on B_n for all $k \geq n$. Since $u_t^{\psi_k} = u_t\psi_k$ for almost all (ω,t), we have that $\bar u_t = u_t$ (as distributions on $\mathbb{R}^d$) for almost all (ω,t). The inclusion $u \in \mathbb{W}_p^1$ now yields $\bar u \in \mathbb{W}_p^1$.

It also follows from (5.7) that if $\omega \in \Omega''$, $t \in \mathbb{R}_+$, and $\phi \in C_0^\infty$ is such that $\phi = 0$ outside B_n, then for any $j \geq n$

$$(\bar u_t,\phi) = (u_t^{\psi_j},\phi) = \int_0^t (L_s u_s - \lambda u_s + D_i f_s^i + f_s^0, \phi)\,ds + (m_t^{\psi_j},\phi).$$

By having in mind (5.5), we conclude that for any $\phi \in C_0^\infty$ with probability one for all $t \in \mathbb{R}_+$

$$(\bar u_t,\phi) = (u_0,\phi) + \int_0^t (L_s u_s - \lambda u_s + D_i f_s^i + f_s^0, \phi)\,ds \\ + \sum_{k=1}^\infty \int_0^t (\Lambda_s^k u_s + g_s^k, \phi)\,dw_s^k.$$

Now it only remains to observe that since $\bar u_s = u_s$ for almost all (ω,s), we can replace u_s with $\bar u_s$ in the above equation. The theorem is proved. □

6. Itô's formula for the product of two processes of class $\mathcal{W}_{2,loc}^1(\tau)$

The results of this section will be used in a few places below, in particular, in the proof of Lemma 8.5. Recall that the spaces $\mathcal{W}_p^1(\tau)$ are introduced in Definition 2.1.

THEOREM 6.1. *Let τ be a stopping time and let $u, \tilde u$, f^j, $\tilde f^j$, $g = (g^1, g^2, \dots)$, $\tilde g = (\tilde g^1, \tilde g^2, \dots)$ be some functions such that for any $\phi \in C_0^\infty$ we have $\phi u, \phi\tilde u \in \mathcal{W}_2^1(\tau)$, $\phi f^j, \phi \tilde f^j \in \mathbb{L}_2(\tau)$, $j = 0, \dots, d$, and $\phi g, \phi\tilde g \in \mathbb{L}_2(\tau)$. Assume that in the sense of generalized functions*

$$du_t = (D_i f_t^i + f_t^0)\,dt + g_t^k\,dw_t^k, \qquad d\tilde u_t = (D_i \tilde f_t^i + \tilde f_t^0)\,dt + \tilde g_t^k\,dw_t^k, \quad t \leq \tau.$$

Then

$$d(u_t\tilde u_t) = [\tilde u_t(D_i f_t^i + f_t^0) + u_t(D_i \tilde f_t^i + \tilde f_t^0) + h_t]\,dt \\ + (\tilde u_t g_t^k + u_t \tilde g_t^k)\,dw_t^k, \quad t \leq \tau,$$

where $h_t := (g_t, \tilde g_t)_{\ell_2}$, in the sense of generalized functions, that is, for any $\phi \in C_0^\infty$, with probability one,

$$(u_{t\wedge\tau}\tilde u_{t\wedge\tau},\phi) = (u_0\tilde u_0,\phi) + \int_0^t I_{s\leq\tau}(\tilde u_s g_s^k + u_s \tilde g_s^k, \phi)\,dw_s^k \\ + \int_0^t I_{s\leq\tau}[(\tilde u_s f_s^0,\phi) - (f_s^i, D_i(\tilde u_s\phi)) \\ + (u_s \tilde f_s^0,\phi) - (\tilde f_s^i, D_i(u_s\phi)) + (h_s,\phi)]\,ds \tag{6.1}$$

for all t.

Proof. To prove (6.1), we only need to consider the case that $\tilde{u} = u$. Indeed, then by writing down the stochastic differential of $|u_t + \lambda \tilde{u}_t|^2$, where λ is an arbitrary constant, and comparing the coefficients of λ, we would come to (6.1). In other words, to prove (6.1), we need only prove that for any $\phi \in C_0^\infty$ with probability one

$$
\begin{aligned}
(6.2)\quad (u_{t\wedge\tau}^2, \phi) = (u_0^2, \phi) &+ 2\int_0^t I_{s\le\tau}(u_s g_s^k, \phi)\,dw_s^k \\
&+ \int_0^t I_{s\le\tau}[2(u_s f_s^0, \phi) - 2(f_s^i, D_i(u_s\phi)) + (|g_s|_{\ell_2}^2, \phi)]\,ds
\end{aligned}
$$

for all t.

Next, observe that for any $\psi, \phi \in C_0^\infty$, with probability one

$$
\begin{aligned}
(\psi u_{t\wedge\tau}, \phi) = (\psi u_0, \phi) &+ \int_0^t I_{t\le\tau}(\psi g_s^k, \phi)\,dw_s^k \\
&+ \int_0^t I_{t\le\tau}[(\psi f_s^0 - f_s^i D_i\psi, \phi) - (\psi f_s^i, D_i\phi)]\,dt
\end{aligned}
$$

for all t. This means that

$$
d(\psi u_t) = \big(\psi f_t^0 - f_t^i D_i\psi + D_i(\psi f_t^i)\big)\,dt + \psi g_t^k\,dw_t^k, \quad t \le \tau.
$$

By well-known results, in particular, by Itô's formula (see, for instance, [13]) there is a set $\Omega' \subset \Omega$ of full probability such that

(i) $\psi u_{t\wedge\tau} I_{\Omega'}$ is a continuous $\mathcal{L}_2$-valued $\mathcal{F}_t$-adapted function on $[0,\infty)$;
(ii) for all $t \in [0,\infty)$ and $\omega \in \Omega'$, Itô's formula holds:

$$
\begin{aligned}
(6.3)\quad &\int_{\mathbb{R}^d} |\psi u_{t\wedge\tau}|^2\,dx \\
&\quad = \int_{\mathbb{R}^d} |\psi u_0|^2\,dx + 2\int_0^t I_{s\le\tau}\int_{\mathbb{R}^d} \psi^2 u_s g_s^k\,dx\,dw_s^k \\
&\qquad + \int_0^t I_{s\le\tau}\left(\int_{\mathbb{R}^d}[2u_s f_s^0\psi^2 - 2f_s^i D_i(\psi^2 u_s) + \psi^2|g_s|_{\ell_2}^2]\,dx\right)ds.
\end{aligned}
$$

This proves (6.2) if we replace there ϕ with ψ^2. However, for any $\phi \in C_0^\infty$ one can find $\psi_1, \psi_2 \in C_0^\infty$ such that $\phi = \psi_1^2 - \psi_2^2$. Indeed, one can take sufficiently large $N, R > 0$ and take $\psi_1(x) = \exp(-(R^2 - |x|^2)^{-1})$ for $|x| < R$ and $\psi_1(x) = 0$ for $|x| \ge R$ and define $\psi_2 = (\psi_1^2 - \phi)^{1/2}$. This implies that (6.2) holds for any $\phi \in C_0^\infty$ with probability one for all t and proves the theorem. □

COROLLARY 6.2. *Let u, f, g be as in Theorem 6.1, let a nonrandom $\psi \in W_2^1$, and let a random process x_t be given as*

$$
x_t = \int_0^t \sigma_s^k\,dw_s^k + \int_0^t b_s\,ds
$$

for some predictable $\mathbb{R}^d$*-valued functions* σ_t^k *and* b_t *such that*

$$E\int_0^\tau \Big(\sum_k |\sigma_t^k|^2 + |b_t|\Big)\,dt < \infty.$$

Then in the sense of generalized functions

$$\begin{aligned} d(u_t\psi_t) = {} & [D_i(u_t a_t^{ij} D_j\psi_t) - a_t^{ij}(D_i u_t) D_j\psi_t + u_t b_t^i D_i\psi_t + D_i(\psi_t f_t^i) \\ & - f_t^i D_i\psi_t + \psi_t f_t^0 + g_t^k \sigma_t^{ik} D_i\psi_t]\,dt + [\psi_t g_t^k + u_t\sigma_t^{ik} D_i\psi_t]\,dw_t^k, \quad t \le \tau, \end{aligned}$$

where $\psi_t(x) = \psi(x + x_t)$ *and* $2a_t^{ij} = \sigma_t^{ik}\sigma_t^{jk}$.

Indeed, observe that by Itô's formula and the stochastic Fubini theorem, for any $\phi \in C_0^\infty$,

$$\begin{aligned} \int_{\mathbb{R}^d} \psi_{t\wedge\tau}\phi\,dx &= \int_{\mathbb{R}^d} \psi(x)\phi(x - x_{t\wedge\tau})\,dx \\ &= \int_{\mathbb{R}^d} \psi\phi\,dx + \int_0^t I_{s\le\tau}\int_{\mathbb{R}^d} \psi_s[a_s^{ij} D_{ij}\phi - b_s^i D_i\phi]\,dx\,ds \\ &\quad + \int_0^t I_{s\le\tau}\int_{\mathbb{R}^d} \sigma_s^{ik}\phi D_i\psi_s\,dx\,dw_s^k, \end{aligned}$$

where the coefficient of ds equals

$$\int_{\mathbb{R}^d} \phi[a_s^{ij} D_{ij}\psi_s + b_s^i D_i\psi_s]\,dx.$$

Furthermore, for instance,

$$\begin{aligned} E\int_0^\tau \int_{\mathbb{R}^d} \sum_k |\sigma_s^{ik} D_i\psi_s|^2\,dx\,ds &\le E\int_0^\tau \int_{\mathbb{R}^d} \sum_k |\sigma_s^k|^2 |D\psi_s|^2\,dx\,ds \\ &= \int_{\mathbb{R}^d} |D\psi|^2\,dx E\int_0^\tau \sum_k |\sigma_s^k|^2\,ds < \infty. \end{aligned}$$

It follows that $\psi_{\cdot} \in \mathcal{W}_2^1(\tau)$ and

$$d\psi_t = [D_i(a_t^{ij} D_j\psi_t) + b_t^i D_i\psi_t]\,dt + \sigma_t^{ik} D_i\psi_t\,dw_t^k$$

in the sense of generalized functions, so that the desired result follows from Theorem 6.1.

7. Kalman–Bucy filter

We take a $T \in (0,\infty)$ and on $[0,T]$ consider a d_1-dimensional two component process $z_t = (x_t, y_t)$ with x_t being d-dimensional and y_t $(d_1 - d)$-dimensional. We assume that z_t is a diffusion process defined as a solution of the system

$$\begin{aligned} dx_t &= b(t, z_t)\,dt + \theta(t, y_t)\,dw_t, \\ dy_t &= B(t, z_t)\,dt + \Theta(t, y_t)\,dw_t \end{aligned} \tag{7.1}$$

with some initial data.

ASSUMPTION 7.1. The functions b, θ, B and Θ are Borel measurable functions of (t,z) and (t,y) as appropriate and θ and Θ are bounded and satisfy the Lipschitz condition with respect to y with a constant independent of t. We have

$$b(t,z)=x^*\dot{b}(t,y)+b(t,0,y), \qquad B(t,z)=x^*\dot{B}(t,y)+B(t,0,y),$$

where $\dot{b}$ and $\dot{B}$ are bounded matrix-valued functions of appropriate dimensions, $b(t,0)$ and $B(t,0)$ are bounded, and $\dot{b}(t,y)$, $\dot{B}(t,y)$, $b(t,0,y)$, and $B(t,0,y)$ satisfy the Lipschitz condition with respect to y with a constant independent of t.

In the rest of the article, we use the notation

$$D_i=\frac{\partial}{\partial x^i}, \qquad D_{ij}=D_iD_j$$

only for $i,j=1,\dots,d$.

REMARK 7.1. Note that

$$\dot{b}^{ij}(t,y)=D_ib^j(t,z), \qquad \dot{B}^{ij}(t,y)=D_iB^j(t,z). \tag{7.2}$$

Set

$$\check{\theta}(t,y)=\begin{pmatrix}\theta(t,y)\\ \Theta(t,y)\end{pmatrix}, \qquad \check{a}(t,y)=\frac{1}{2}\check{\theta}\check{\theta}^*(t,y), \qquad \check{b}(t,z)=\begin{pmatrix}b(t,z)\\ B(t,z)\end{pmatrix}, \tag{7.3}$$

$$\check{L}(t,z)=\check{a}^{ij}(t,y)\frac{\partial^2}{\partial z^i\,\partial z^j}+\check{b}^i(t,z)\frac{\partial}{\partial z^i}, \tag{7.4}$$

where $t\in[0,T]$, $z=(x,y)\in\mathbb{R}^{d_1}$, and we use the summation convention over all "reasonable" values of repeated indices, so that the summation in (7.4) is performed for $i,j=1,\dots,d_1$.

Observe that

$$dz_t=\check{\theta}(t,z_t)\,dw_t+\check{b}(t,z_t)\,dt. \tag{7.5}$$

ASSUMPTION 7.2. The symmetric matrix $\check{a}(t,y)$ is uniformly nondegenerate. In particular, the matrix $\Theta\Theta^*$ is invertible and

$$\Psi:=(\Theta\Theta^*)^{-\frac{1}{2}}$$

is a bounded function of (t,y).

REMARK 7.2. It is well known (see, for instance, [14]) that in light of Assumption 7.2 the matrix

$$\hat{a}(t,y)=a(t,y)-\alpha(t,y)$$

is uniformly (with respect to (t,y)) nondegenerate, where

$$a=\frac{1}{2}\theta\theta^*, \qquad \alpha=\frac{1}{2}\sigma\sigma^*, \quad \sigma=\theta\Theta^*\Psi,$$

REMARK 7.3. Everywhere below we use the stipulation that if we are given a function $\xi(t,x,y)$, then we denote

$$\xi_t = \xi_t(x) = \xi(t,x,y_t) \tag{7.6}$$

unless it is explicitly specified otherwise. For instance, $\Psi_t = \Psi(t,y_t)$, $\Theta_t = \Theta(t,y_t)$, $\sigma_t = \theta_t \Theta_t^* \Psi_t$.

Next we introduce a few more notation. Let (note the size and shape of B)

$$\mathrm{B} = \Psi B, \qquad \mathrm{B}_t(x) = \Psi_t B_t(x) = \Psi(t,y_t) B(t,x,y_t)$$

and set

$$L_t(x) = a_t^{ij} D_i D_j + b_t^i(x) D_i, \tag{7.7}$$

$$\begin{aligned} L_t^*(x) u_t(x) &= D_i D_j (a_t^{ij} u_t(x)) - D_i (b_t^i(x) u_t(x)) \\ &= D_j \big(a_t^{ij} D_i u_t(x) - b_t^j(x) u_t(x)\big), \end{aligned} \tag{7.8}$$

$$\Lambda_t^k(x) u_t(x) = \sigma_t^{ik} D_i u_t(x) + \mathrm{B}_t^k(x) u_t(x), \tag{7.9}$$

$$\Lambda_t^{k*}(x) u_t(x) = -\sigma_t^{ik} D_i u_t(x) + \mathrm{B}_t^k(x) u_t(x), \tag{7.10}$$

where $t \in [0,T]$, $x \in \mathbb{R}^d$, $k = 1,\dots,d_1 - d$, and as above we use the summation convention over all "reasonable" values of repeated indices, so that the summation in (7.7), (7.8), (7.9), and (7.10) is performed for $i,j = 1,\dots,d$ (whereas in (7.4) for $i,j = 1,\dots,d_1$).

Finally, by $\mathcal{F}_t^y$ we denote the completion of $\sigma\{y_s : s \le t\}$ with respect to $P, \mathcal{F}$.

ASSUMPTION 7.3. There exists an $\varepsilon > 0$ and a function $Q(x) = Q(\omega, x)$ which is $\mathcal{F}_0^y$-measurable in ω, quadratic in x, and

(i) For all $x \in \mathbb{R}^d$ (and ω)

$$\varepsilon^{-1} |x|^2 \ge x^i x^j D_{ij} Q \ge \varepsilon |x|^2;$$

(ii) We have $\pi_0 e^Q \in \operatorname{tr} \mathcal{W}_p^1$, where π_0 is the conditional density of x_0 given y_0.

Assumption 7.3 is satisfied, for instance, in the classical setting of the Kalman–Bucy filter when π_0 is a Gaussian density.

THEOREM 7.1. *There exists a process $\bar\pi$ on $[0,T]$ such that*

(i) *$\bar\pi_t$ is $\mathcal{F}_t^y$-adapted and, for any $r \in [1,p]$, with probability one $\bar\pi_t$ is a continuous $\mathcal{L}_r$-valued process on $[0,T]$ and $\bar\pi_0 = \pi_0$;*

(ii) *There exists an increasing sequence of $\mathcal{F}_t^y$-stopping times $\tau_m \le T$ such that $P(\tau_m = T) \to 1$ and $\bar\pi \in \mathbb{W}_p^1(\tau_m)$ for any m;*

(iii) *In the sense of Definition 2.3 for any m*

$$d\bar\pi_t = \Lambda_t^{k*} \bar\pi_t \, d\tilde y_t^k + L_t^* \bar\pi_t \, dt, \quad t \le \tau_m, \tag{7.11}$$

where

$$\tilde y_t^k = \int_0^t \Psi_s^{kr} \, dy_s^r.$$

Furthermore, for any m *and* $\phi \in C_0^\infty$ *we have* $\bar{\pi}\phi \in \mathcal{W}_p^1(\tau_m)$;

(iv) *We have* $\bar{\pi}_t \geq 0$ *for all* $t \in [0,T]$ (*a.s.*),

$$0 < \int_{\mathbb{R}^d} \bar{\pi}_t(x)\,dx = (\bar{\pi}_t, 1) < \infty \tag{7.12}$$

for all $t \in [0,T]$ (*a.s.*), *and for any* $t \in [0,T]$ *and real-valued, bounded or nonnegative*, (*Borel*) *measurable function* f *given on* $\mathbb{R}^d$

$$E[f(x_t)|\mathcal{F}_t^y] = \frac{(\bar{\pi}_t, f)}{(\bar{\pi}_t, 1)} \quad (a.s.). \tag{7.13}$$

REMARK 7.4. Equation (7.13) shows (by definition) that

$$\pi_t(x) := \frac{\bar{\pi}_t(x)}{(\bar{\pi}_t, 1)}$$

is a conditional density of distribution of x_t given $y_s, s \leq t$. Since, generally, $(\bar{\pi}_t, 1) \neq 1$, one calls $\bar{\pi}_t$ an unnormalized conditional density of distribution of x_t given $y_s, s \leq t$. Thus, Theorem 7.1 allows us to characterize the conditional density and being combined with Theorem 3.4 allows us to obtain fine regularity properties of it.

The following result is obtained by repeating what is said after Theorem 3.4 and taking into account that with probability one $\tau_m = T$ for all large m.

THEOREM 7.2. (i) *For any* $\phi \in C_0^\infty$ *the process* $\bar{\pi}_t\phi$ *is continuous on* $[0,T]$ *as an* $\mathcal{L}_p$*-valued process* (*a.s.*);

(ii) *If* $p > 2$ *and we have two numbers* α *and* β *such that*

$$\frac{2}{p} < \alpha < \beta \leq 1,$$

then for any $\phi \in C_0^\infty$ (*a.s.*)

$$\bar{\pi}\phi \in C^{\alpha/2-1/p}([0,T], H_p^{1-\beta}).$$

In particular, if $p > d+2$, *then*

(a) *for any* $\varepsilon \in (0, \varepsilon_0]$, *with* $\varepsilon_0 = 1 - (d+2)/p$, (*a.s.*) *for any* $t \in [0,T]$ *we have* $\bar{\pi}_t\phi \in C^{\varepsilon_0-\varepsilon}(\mathbb{R}^d)$ *and the norm of* $\bar{\pi}_t\phi$ *in this space is bounded as a function of* t;

(b) *for any* ε *as in* (a) (*a.s.*) *for any* $x \in \mathbb{R}^d$ *we have* $\bar{\pi}_\cdot(x)\phi(x) \in C^{(\varepsilon_0-\varepsilon)/2}([0,T])$ *and the norm of* $\bar{\pi}_\cdot(x)\phi(x)$ *in this space is bounded as a function of* x.

In the general filtering theory equation (7.11) is known as Zakai's equation. From the point of view of the Sobolev space theory of SPDEs the most unpleasant feature of (7.11) in our particular case is the presence of $\mathrm{B}_t^k(x)\bar{\pi}_t$ in the stochastic term with $\mathrm{B}_t^k(x)$ which is unbounded in x. However, in the theory of linear PDEs it was observed that if an equation has a zeroth order term and we know a particular nonzero solution, then the ratio of the unknown function and this particular solution satisfies an equation without zeroth order term (cf. (8.1)).

The way to find a particular solution of (7.11) is suggested by filtering theory. Imagine that $\check{b}$ is affine with respect to z and $\check{\theta}$ is independent of z. Then as easy to see z_t is a Gaussian process and hence the conditional density of x_t given y_s, $s \leq t$, is Gaussian, that is, its logarithm is a quadratic function in x. Therefore, we were looking for a particular solution as $e^{-Q_t(x)}$, where $Q_t(x)$ is a quadratic function with respect to x, and finding the equation for $Q_t(x)$ (see (7.19)) was pretty straightforward.

After we "kill" the zeroth-order term our equation falls into the scheme of Section 3 even though it still has growing first order coefficients in the deterministic part of the equation. Finding $\bar{\pi}_t$ in the described way allows us to follow the scheme suggested in [20] thus avoiding using filtering theory. However, we still encounter an additional difficulty that certain exponential martingales may not have moments of order > 1, unlike the situation in [20], and, to prove that they are martingales indeed, we use the Liptser–Shiryaev theorem (see [22]). This way of proceeding was used by Liptser in [21] (see also [22]) while treating filtering problem for the so-called conditionally Gaussian processes.

Finding a particular solution of (7.11) is based on the following lemma which is probably well known. We give its proof in the end of Section 8 just for completeness. Set

$$\dot{\mathrm{B}}_t = \dot{B}_t \Psi_t. \tag{7.14}$$

LEMMA 7.3. *The following system of equations about $d \times d$-symmetric matrix-valued process W_t, $\mathbb{R}^d$-valued process V_t, and real-valued process U_t*

$$\frac{d}{dt} W_t = (\dot{\mathrm{B}}_t \sigma_t^* - \dot{b}_t) W_t + W_t^* (\sigma_t \dot{\mathrm{B}}_t^* - \dot{b}_t^*) - 2 W_t^* \hat{a}_t W_t + \dot{\mathrm{B}}_t \dot{\mathrm{B}}_t^*, \tag{7.15}$$

$$\begin{aligned} dV_t = {} & -(W_t \sigma_t + \dot{\mathrm{B}}_t)\, d\tilde{y}_t \\ & + \big[(\dot{\mathrm{B}}_t \sigma_t^* - \dot{b}_t) V_t - 2 W_t \hat{a}_t V_t \\ & + W_t \big(\sigma_t \mathrm{B}_t(0) - b_t(0)\big) + \dot{\mathrm{B}}_t \mathrm{B}_t(0)\big]\, dt, \end{aligned} \tag{7.16}$$

$$\begin{aligned} dU_t = {} & -\big(V_t^* \sigma_t + \mathrm{B}_t^*(0)\big)\, d\tilde{y}_t \\ & + \bigg[a_t^{ij} W_t^{ij} + V_t^* \big(\sigma_t \mathrm{B}_t(0) - b_t(0)\big) \\ & - V_t^* \hat{a}_t V_t + \frac{1}{2} |\mathrm{B}_t(0)|^2 + \operatorname{tr} \dot{b}_t \bigg]\, dt, \end{aligned} \tag{7.17}$$

has a unique $\mathcal{F}_t^y$-adapted solution with initial conditions $W_0^{ij} = D_{ij} Q$, $V_0^i = D_i Q(0)$, $U_0 = Q(0)$. Furthermore, $\varepsilon_1^{-1}(\delta^{ij}) \geq W_t \geq \varepsilon_1(\delta^{ij})$ on $[0,T]$, where $\varepsilon_1 > 0$ is a constant independent of ω and t (depending on T among other things).

Observe that the coefficients in (7.17) are independent of x.

REMARK 7.5. Set

$$Q_t(x) = \frac{1}{2} W_t^{ij} x^i x^j + V_t^i x^i + U_t. \tag{7.18}$$

Then by using Itô's formula one easily checks that for any $x \in \mathbb{R}^d$

$$\begin{aligned} dQ_t(x) = &-\big(\sigma_t^{ik} D_i Q_t(x) + \mathrm{B}_t^k(x)\big)\, d\tilde{y}_t^k \\ &+ \bigg[a_t^{ij} D_{ij} Q_t(x) + D_i b_t^i + \big(\sigma_t^{ik} \mathrm{B}_t^k(x) - b_t^i(x)\big) D_i Q_t(x) \\ &- \hat{a}_t^{ij} (D_i Q_t(x)) D_j Q_t(x) + \frac{1}{2} |\mathrm{B}_t(x)|^2\bigg]\, dt \end{aligned} \tag{7.19}$$

and $\eta_t = e^{-Q_t}$ satisfies

$$d\eta_t(x) = \Lambda_t^{r*} \eta_t(x)\, d\tilde{y}_t^r + L_t^* \eta_t(x)\, dt. \tag{7.20}$$

By the way, $Q_t(x)$ is a unique $\mathcal{F}_t^y$-adapted function depending quadratically on x, satisfying (7.19), and such that $Q_0 = Q$. Indeed, uniqueness follows from the fact that $D_{ij} Q_t$, $D_i Q_t(0)$, and $Q_t(0)$ are easily shown to satisfy (7.15), (7.16), and (7.17), respectively.

Our method also allows us to derive the classical equations for the Kalman–Bucy filter.

THEOREM 7.4. *Replace requirement* (ii) *in Assumption 7.3 with the assumption that $\pi_0 = e^{-Q}$. Then for any t (a.s.) we have $\pi_t(x) = C_t e^{-Q_t(x)}$, where C_t is a normalizing process obtained from the condition that*

$$C_t \int_{\mathbb{R}^d} e^{-Q_t(x)}\, dx = 1.$$

This theorem is proved in Section 9.

REMARK 7.6. After just completing the square and finding the stochastic differential of the remaining term, we find that

$$\begin{aligned} Q_t(x) = &\frac{1}{2} |W_t^{1/2} x + W_t^{-1/2} V_t|^2 + \int_0^t \big(V_s^* W_s^{-1} \dot{\mathrm{B}}_s - \mathrm{B}_s^*(0)\big)\, d\tilde{y}_s \\ &+ \frac{1}{2} \int_0^t |\dot{\mathrm{B}}_s^* W_s^{-1} V_s - \mathrm{B}_s(0)|^2\, ds + A_t, \end{aligned} \tag{7.21}$$

with a bounded on $\Omega \times [0, T]$ function

$$A_t := \int_0^t \bigg[a_s^{ij} W_s^{ij} + \operatorname{tr} \dot{b}_s - \frac{1}{2} \|W_s^{1/2} \sigma_s + W_s^{-1/2} \dot{\mathrm{B}}_s\|^2\bigg]\, ds,$$

where for a matrix u we use the notation $\|u\|^2 = \operatorname{tr} uu^*$. This shows that in the situation of Theorem 7.4

$$\bar{x}_t := E(x_t|\mathcal{F}_t^y) = \int_{\mathbb{R}^d} x\pi_t(x)\,dx = -W_t^{-1}V_t,$$
$$\Sigma_t := E\big((x_t - \bar{x}_t)(x_t - \bar{x}_t)^*|\mathcal{F}_t^y\big) = W_t^{-1}$$

and allows one to derive the classical Kalman–Bucy equations for $\bar{x}_t$ and Σ_t from (7.15) and (7.16).

8. An auxiliary function

The assumptions from Section 7 are supposed to hold. Set

$$\hat{b}_t^i(x) = \sigma_t^{ik}\mathrm{B}_t^k(x) - 2\hat{a}_t^{ij}D_jQ_t(x).$$

THEOREM 8.1. *The equation*

$$d\hat{\pi}_t = -\sigma_t^{ik}D_i\hat{\pi}_t d\tilde{y}_t^k + [a_t^{ij}D_{ij}\hat{\pi}_t - b_t^iD_i\hat{\pi}_t + \hat{b}_t^iD_i\hat{\pi}_t]\,dt, \quad t \le T, \tag{8.1}$$

with initial data $\hat{\pi}_0 = e^Q\pi_0$ has a unique solution in the sense of Definition 2.3.

This theorem is a direct consequence of Remark 3.4 and Theorem 3.3 since the coefficients b and $\hat{b}$ in (8.1) are affine functions of x and have *bounded* derivatives in x.

LEMMA 8.2. *Almost surely $\hat{\pi}_t$ is a continuous $\mathcal{L}_p$-valued process on $[0,T]$. Furthermore, $G_t\|\hat{\pi}_t\|_{\mathcal{L}_p}^p$ is a decreasing function of t (a.s.), where G_t is a bounded function on $\Omega \times [0,T]$ defined by*

$$G_t := \exp\int_0^t (D_i\hat{b}_s^i - D_ib_s^i)\,ds = \exp\int_0^t \operatorname{tr}(\sigma_s\dot{\mathrm{B}}_s^* - \hat{a}_sW_s - \dot{b}_s)\,ds.$$

In particular, on the set where $\tau := T \wedge \inf\{t \ge 0 : \|\hat{\pi}_t\|_{\mathcal{L}_p} = 0\} < T$ we have $\|\hat{\pi}_t\|_{\mathcal{L}_p} = 0$ for $\tau \le t \le T$ (a.s.).

Proof. Set

$$\xi_t^i = \int_0^t \sigma_s^{ik}\,d\tilde{y}_s^r, \qquad \xi_t = (\xi_t^i), \qquad \tau_m = T \wedge \inf\{t \ge 0 : |z_t| + |\xi_t| \ge m\}.$$

The purpose to stop z_t is that on $(0,\tau_m]\!]$, we have

$$|\sigma_t\mathrm{B}_t(x)| + |b_t(x)| + |\hat{b}_t(x)| \le N(1 + |x|),$$

where the constant N is independent of ω, t, x. Why we also stop ξ_t will become clear later.

Observe that for any $\psi \in C_0^\infty$ the process $\psi\hat{\pi}_t$ satisfies an equation obtained by multiplying through (8.1) by ψ. Then after writing $\psi D_i(a_t^{ij}D_j\hat{\pi}_t)$ as $D_i(\psi a_t^{ij}D_j\hat{\pi}_t) - a_t^{ij}(D_j\hat{\pi}_t)D_i\psi$ and noting that the other coefficients multiplied by ψ are bounded functions on $(0,\tau_m]\!] \times \mathbb{R}^d$ we see that

$$\psi\hat{\pi}_t \in \mathcal{W}_p^1(\tau_m) \tag{8.2}$$

for any m. It follows from [13] that with probability one $\psi\hat{\pi}_{t\wedge\tau_m}$ is a continuous $\mathcal{L}_p$-valued process and since, for each ω, $\tau_m = T$ if m is sufficiently large, with probability one $\psi\hat{\pi}_t$ is a continuous $\mathcal{L}_p$-valued process on $[0,T]$ for any $\psi \in C_0^\infty$.

Then take a nonnegative radially symmetric and radially decreasing function $\phi \in C_0^\infty$ such that $|D\phi| \le 1$, introduce $\phi^n(x) = \phi(x/n)$, $n = 1, 2, \dots,$

$$\phi_t^n(x) = \phi^n(x - \xi_t)$$

and use Corollary 6.2 with τ_m in place of τ (recall (8.2)). Then we find

$$\begin{aligned} d(\hat{\pi}_t\phi_t^n) = {} & -\sigma_t^{ik} D_i(\hat{\pi}_t\phi_t^n)\, d\tilde{y}_t^k + \phi_t^n(\hat{b}_t^i - b_t^i) D_i\hat{\pi}_t\, dt \\ & + [D_i(\phi_t^n a_t^{ij} D_j\hat{\pi}_t) - (\hat{a}_t^{ij} + a_t^{ij})(D_i\phi_t^n) D_j\hat{\pi}_t \\ & + D_i(\hat{\pi}_t\alpha_t^{ij} D_j\phi_t^n)]\, dt, \quad t \le \tau_m. \end{aligned} \tag{8.3}$$

As above we conclude that $\phi_t^n\hat{\pi} \in \mathcal{W}_p^1(\tau_m)$ and that, owing to [13], with probability one $\phi_t^n\hat{\pi}_t$ is a continuous $\mathcal{L}_p$-valued process and (a.s.)

$$\|\phi_{t\wedge\tau_m}^n\hat{\pi}_{t\wedge\tau_m}\|_{\mathcal{L}_p}^p = \|\phi^n\hat{\pi}_0\|_{\mathcal{L}_p}^p + I_t^{1n} + I_t^{2n} + I_t^{3n}$$

for all t, where

$$\begin{aligned} I_t^{1n} &= -p(p-1)\int_0^{t\wedge\tau_m} a_s^{ij} \int_{\mathbb{R}^d} |\phi_s^n|^p |\hat{\pi}_s|^{p-2}(D_i\hat{\pi}_s) D_j\hat{\pi}_s\, dx\, ds \le 0, \\ I_t^{2n} &= -\int_0^{t\wedge\tau_m} [D_i\hat{b}_s^i - D_i b_s^i] \int_{\mathbb{R}^d} |\phi_s^n|^p |\hat{\pi}_s|^p\, dx\, ds, \\ I_t^{3n} &= \int_0^{t\wedge\tau_m} \int_{\mathbb{R}^d} |\hat{\pi}_s|^p \psi_s^n\, dx\, ds, \\ \psi_s^n &= p a_s^{ij} D_{ij}|\phi_s^n|^p + (p-1)(p-2)|\phi_s^n|^{p-2}\alpha_s^{ij}(D_i\phi_s^n) D_j\phi_s^n \\ &\quad + (2-p)|\phi_s^n|^{p-1}\alpha_s^{ij} D_{ij}\phi_s^n + (b_s^i - \hat{b}_s^i) D_i|\phi_s^n|^p, \end{aligned}$$

where for simplicity of notation the argument x is dropped.

Observe that $|D\phi_s^n| \le 1/n$ and for $s \le \tau_m$ we have $|b_s - \hat{b}_s| \le N(1 + |x|)$, where N is independent of s, x, and ω. Furthermore, $D\phi_s^n \to 0$ as $n \to \infty$ and for $s < \tau_m$

$$\begin{aligned} |x||D\phi_s^n(x)| = \frac{|x|}{n}\left|(D\phi)\left(\frac{x-\xi_s}{n}\right)\right| &\le \frac{|\xi_s|}{n} + \frac{|x-\xi_s|}{n}\left|(D\phi)\left(\frac{x-\xi_s}{n}\right)\right| \\ &\le m + \sup_y |y||D\phi(y)|. \end{aligned}$$

By adding that $\hat{\pi} \in \mathbb{W}_p^1(T)$, we conclude that $I_t^{3n} \to 0$ uniformly in t (a.s.). Analyzing I_t^{1n} and I_t^{2n} is almost trivial and

$$\|\phi_{t\wedge\tau_m}^n\hat{\pi}_{t\wedge\tau_m}\|_{\mathcal{L}_p}^p \to \|\hat{\pi}_{t\wedge\tau_m}\|_{\mathcal{L}_p}^p$$

as $n\to\infty$ by the monotone convergence theorem. It follows that (a.s.) for all t

$$\|\hat{\pi}_{t\wedge\tau_m}\|^p_{\mathcal{L}_p}=\|\hat{\pi}_0\|^p_{\mathcal{L}_p}-\int_0^{t\wedge\tau_m}(D_i\hat{b}^i_s-D_ib^i_s)\|\hat{\pi}_s\|^p_{\mathcal{L}_p}\,ds$$
$$-p(p-1)\int_0^{t\wedge\tau_m}a^{ij}_s\int_{\mathbb{R}^d}|\hat{\pi}_s|^{p-2}(D_i\hat{\pi}_s)D_j\hat{\pi}_s\,dx\,ds.$$

Obviously on can drop τ_m in this formula and then obtain that (a.s.) for all $t\le T$

$$G_t\|\hat{\pi}_t\|^p_{\mathcal{L}_p}=\|\hat{\pi}_0\|^p_{\mathcal{L}_p}-p(p-1)\int_0^t G_sa^{ij}_s\int_{\mathbb{R}^d}|\hat{\pi}_s|^{p-2}(D_i\hat{\pi}_s)D_j\hat{\pi}_s\,dx\,ds,$$

which implies that $G_t\|\hat{\pi}_t\|^p_{\mathcal{L}_p}$ is decreasing and continuous (a.s.). Furthermore, since $\phi^n_t\hat{\pi}_t$ are continuous $\mathcal{L}_p$-valued processes, $\hat{\pi}_t$ is at least a weakly continuous $\mathcal{L}_p$-valued function, but since $\|\hat{\pi}_t\|^p_{\mathcal{L}_p}$ is (absolutely) continuous, $\hat{\pi}_t$ is strongly continuous. This proves the lemma. □

REMARK 8.1. After we know that $\hat{\pi}_t$ is a continuous $\mathcal{L}_p$-valued process on $[0,T]$ the last assertion of Lemma 8.2 can be also obtained from uniqueness of solutions of (8.1) because the τ in Lemma 8.2 is a stopping time and $\hat{\pi}_{t\wedge\tau}$ is obviously a solution of (8.1) implying that on the set where $\tau<T$ we have $\hat{\pi}_t=0$ for $\tau\le t\le T$.

Before stating the following lemma, we introduce a stipulation accepted throughout the rest of the paper that if we are given a function $\xi(t,x,y)$, then we denote

$$\tilde{\xi}_t=\xi_t(x_t)=\xi(t,x_t,y_t). \tag{8.4}$$

The reader encountered above already one of these abbreviated notation (see (7.6)).

LEMMA 8.3. *Introduce*

$$\tilde{w}_t=\int_0^t\Psi_s\Theta_s\,dw_s,\qquad \tilde{\mathrm{B}}_t=\mathrm{B}_t(x_t)=\Psi(t,y_t)B(t,x_t,y_t).$$

Then $\tilde{w}_t$ is a Wiener process and the process

$$\rho_t=\rho_t(\tilde{\mathrm{B}},d\tilde{w})=\exp\Big(-\int_0^t\tilde{\mathrm{B}}^k_s\,d\tilde{w}^k_s-\frac{1}{2}\int_0^t|\tilde{\mathrm{B}}_s|^2_{\ell_2}\,ds\Big)$$

is a martingale on $[0,T]$.

Proof. The first assertion follows from Lévy's theorem. To prove the second one, observe that

$$\int_0^t\tilde{\mathrm{B}}^k_s\,d\tilde{w}^k_s=\int_0^t\tilde{\mathrm{B}}^*_s\Psi_s\Theta_s\,dw_s.$$

Furthermore, the system

$$
\begin{aligned}
dx_t &= \big(b(t,z_t) - \theta(t,y_t)\Theta^*(t,y_t)\Psi^2(t,y_t)B(t,z_t)\big)\,dt + \theta(t,y_t)\,dw_t,\\
dy_t &= \Theta(t,y_t)\,dw_t,
\end{aligned}
$$

which is obtained from (7.1) by formal application of the measure change, has a unique solution with initial data z_0 since its coefficients are locally Lipschitz in z and grow as $|z|\to\infty$ not faster than linearly. In this situation by the Liptser–Shiryaev theorem, ρ is a martingale since

$$
\int_0^T |\Psi(t,y(t))B(t,x(t),y(t))|^2\,dt < \infty
$$

for any deterministic functions $x(t)$ and $y(t)$ which are continuous on $[0,T]$. The lemma is proved. □

LEMMA 8.4. *The process* $\hat\pi_t$ *is* $\mathcal{F}_t^y$*-adapted.*

Proof. Observe that in equation (8.1) we have

$$
d\tilde y_t^k = \Psi_t^{kr}\,dy_t^r = d\tilde w_t^k + \tilde{\mathrm{B}}_t^k\,dt,
$$

where, as it is pointed out above, $\tilde w_t$ is a Wiener process. Furthermore, the processes $\tilde y_t^k$ is $\mathcal{F}_t^y$-adapted since such are Ψ_t^{kr} and equation (8.1) is rewritten as

$$
\begin{aligned}
d\hat\pi_t = {}& -\sigma_t^{ik}D_i\hat\pi_t\,d\tilde w_t^k + [D_i(a_t^{ij}D_j\hat\pi_t) - b_t^iD_i\hat\pi_t\\
&+ (\hat b_t^i - \sigma_t^{ik}\tilde{\mathrm{B}}_t^k)D_i\hat\pi_t]\,dt, \quad t\le T.
\end{aligned}
\tag{8.5}
$$

Here $\sigma_t^{ik}\tilde{\mathrm{B}}_t^k$ is independent of x and for each ω the trajectories of $\sigma_t^{ik}\tilde{\mathrm{B}}_t^k$ are locally bounded on $\mathbb{R}_+$, which shows that in order to be able to apply Theorem 3.5 it only remains to refer to Lemma 8.3. The lemma is proved. □

LEMMA 8.5. *The assertions* (i)–(iii) *of Theorem* 7.1 *hold for* $\bar\pi_t := e^{-Q_t}\hat\pi_t$.

Proof. Assertion (i) of Theorem 7.1 follows immediately from Lemma 8.2, the continuity of Q_t, and the boundedness of $W_t = (D_{ij}Q_t)$ away from zero.

To prove assertion (ii), notice that $\hat\pi \in \mathbb{W}_p^1(T)$ and

$$
\int_0^t \|\bar\pi_s\|_{W_p^1}^p\,ds
$$

is an $\mathcal{F}_t^y$-adapted continuous process on $[0,T]$. Then after introducing

$$
\tau_m' = T\wedge \inf\left\{t\ge 0: \int_0^t \|\bar\pi_s\|_{W_p^1}^p\,ds \ge m\right\}
$$

we get that $\bar\pi \in \mathbb{W}_p^1(\tau_m')$ and $\tau_m' = T$ for all large m (a.s.).

We now prove that $\bar\pi$ satisfied (7.11) define $\Phi_t = \Psi_t^{-1}$ and observe that

$$
(d\tilde y_t^k)\,d\tilde y_t^r = \delta^{kr}\,dt, \qquad dy_t^k = \Phi_t^{kr}\,d\tilde w_t^r + \tilde B_t^k\,dt,
$$

($\tilde B_t = B(t,z_t)$). Recall that $\eta_t(x) = \exp(-Q_t(x))$ satisfies equation (7.20) for each x with probability one for all $t\in[0,T]$. It turns out that this equation also holds in the sense of generalized functions. Owing to the special structure of Q_t, this follows from the stochastic version of Fubini's theorem (see, for instance, Lemma 2.7 of [17]).

Next, for $m=1,2,\dots$ set

$$\tau''_m = T\wedge \inf\{t\geq 0: |z_t|+|DQ_t(0)|\geq m\}. \tag{8.6}$$

Note that for a constant N_0 independent of m for $t<\tau''_m$ we have

$$|\mathrm{B}_t(x)|+|b_t(x)|\leq N_0(1+|x|+m),\qquad |\tilde{\mathrm{B}}_t|+|\tilde{B}_t|\leq N_0(1+2m).$$

Furthermore, $D_iQ_t(x)=x^jD_{ij}Q_t+D_iQ_t(0)$, so that increasing N_0 if needed we may assume that for $t<\tau''_m$

$$|DQ_t(x)|\leq N_0(1+|x|+m).$$

Then as is easy to see (cf. (8.2)) $u_t:=\hat{\pi}_t$ and $\tilde{u}_t:=\eta_t$ satisfy the condition of Theorem 6.1 with appropriate $f,\tilde{f},g,\tilde{g}$ and τ''_m in place of τ.

By Theorem 6.1 in the sense of generalized functions

$$d(\eta_t\hat{\pi}_t)=I^r_t\,d\tilde{y}^r_t+J_t\,dt,\quad t\leq\tau''_m,$$

where

$$\begin{aligned}
I^r_t &= \hat{\pi}_t\Lambda^{r*}_t\eta_t-\eta_t\sigma^{ir}_tD_i\hat{\pi}_t=\Lambda^{r*}_t(\eta_t\hat{\pi}_t),\\
J_t &= -(\eta_t\mathrm{B}^k_t-\sigma^{ik}_tD_i\eta_t)\sigma^{jk}_tD_j\hat{\pi}_t+\hat{\pi}_tL^*_t\eta_t\\
&\qquad+\eta_t[a^{ij}_tD_{ij}\hat{\pi}_t-b^i_tD_i\hat{\pi}_t+(\sigma^{ik}_t\mathrm{B}^k_t+2\eta^{-1}\hat{a}^{ij}_tD_j\eta_t)D_i\hat{\pi}_t]\\
&\quad=\hat{\pi}_tL^*_t\eta_t+\eta_t(a^{ij}_tD_{ij}\hat{\pi}_t-b^i_tD_i\hat{\pi}_t)+2a^{ij}_t(D_i\hat{\pi}_t)D_j\eta_t=L^*_t(\eta_t\hat{\pi}_t).
\end{aligned}$$

In other words (see Theorem 6.1) for any $\phi\in C^\infty_0$ with probability one

$$(\bar{\pi}_{t\wedge\tau''_m},\phi)=(\bar{\pi}_0,\phi)+\int_0^t I_{s\leq\tau''_m}(\bar{\pi}_s,\Lambda^k_s\phi)\,d\tilde{y}^k_s+\int_0^t I_{s\leq\tau''_m}(\bar{\pi}_s,L_s\phi)\,ds$$

for all $t\geq 0$. Obviously, one can take here $\tau_m:=\tau'_m\wedge\tau''_m$ in place of τ''_m and then after recalling that $\bar{\pi}\in\mathbb{W}^1_p(\tau'_m)$ one concludes that $\bar{\pi}$ is a solution of (7.11) in the sense of Definition 2.3. The final assertion in (iii) is obtained in the same way as (8.2). The lemma is proved. $\square$

To better orient the reader, it is worth noting that in the next lemma the second factor on the left in (8.7) contains the negative of two terms in (7.21).

LEMMA 8.6. *We have*

$$\begin{aligned}
\rho_t(\tilde{\mathrm{B}},d\tilde{w})\exp\biggl(-\int_0^t\bigl(V^*_sW^{-1}_s\dot{\mathrm{B}}_s-\mathrm{B}^*_s(0)\bigr)\,d\tilde{y}_s&\\
-\frac{1}{2}\int_0^t|\dot{\mathrm{B}}^*_sW^{-1}_sV_s-\mathrm{B}_s(0)|^2\,ds\biggr)&\\
=\rho_t\bigl(\tilde{\mathrm{B}}-\mathrm{B}_\cdot(0)+\dot{\mathrm{B}}^*W^{-1}V,d\tilde{w}\bigr).&
\end{aligned} \tag{8.7}$$

Furthermore, the right-hand side is a martingale on $[0,T]$.

Proof. The equality is obtained by simple manipulations. As in the proof of Lemma 8.3, to prove that (8.7) is a martingale we are going to use the Liptser–Shiryaev theorem by considering the system consisting of (7.5), (7.15), and (7.16). We do not include (7.17) because U_t does not enter (8.7). First of all, we find a smooth bounded, uniformly nondegenerate $d\times d$-matrix-valued function $F(W)$ such that $F(W_t)=W_t$. The fact that this is possible follows from Lemma 7.3. Then set

$$A(t,z,W,V)=\Theta^*(t,y)\Psi^2(t,y)\bigl(B(t,z)-B(t,0,y)+\dot B^*(t,y)F^{-1}(W)V\bigr).$$

After changing the probability measure formally, we arrive at the system consisting of (7.15) with $\sigma_t=\sigma(t,y_t)$, $\hat a_t=\hat a(t,y_t)$, and with $F(W_t)$ in place of W_t on the right and the following two equations

$$\begin{aligned}
dz_t&=\check\theta(t,y_t)\,dw_t+[\check b(t,z_t)-\check\theta(t,y_t)A(t,z_t,W_t,V_t)]\,dt,\\
dV_t&=-\bigl(F(W_t)\sigma(t,y_t)+\dot B(t,y_t)\Psi(t,y_t)\bigr)\Psi(t,y_t)\Theta(t,y_t)\,dw_t\\
&\qquad+\bigl(F(W_t)\sigma(t,y_t)+\dot B(t,y_t)\Psi(t,y_t)\bigr)\Psi(t,y_t)\Theta(t,y_t)A(t,z_t,W_t,V_t)\,dt\\
&\qquad-\bigl(F(W_t)\sigma(t,y_t)+\dot B(t,y_t)\Psi(t,y_t)\bigr)\Psi(t,y_t)B(t,z_t)\,dt\\
&\qquad+\bigl[\bigl(\dot B(t,y_t)\Psi(t,y_t)\sigma^*(t,y_t)-\dot b(t,y_t)\bigr)V_t-2F(W_t)\hat a(t,y_t)V_t\bigr]\,dt\\
&\qquad+\bigl[F(W_t)\bigl(\sigma(t,y_t)\Psi(t,y_t)B(t,0,y_t)-b(t,0,y_t)\bigr)\\
&\qquad+\dot B(t,y_t)\Psi^2(t,y_t)B(t,0,y_t)\bigr]\,dt.
\end{aligned}$$

This system has a unique solution with prescribed initial data since its coefficients are locally Lipschitz continuous and may grow to infinity as $|z|+|W|+|V|\to\infty$ not faster than linearly. Moreover,

$$\int_0^T|A(t,z(t),W(t),V(t))|^2\,dt<\infty$$

for any functions $z(t),W(t),V(t)$ which are continuous on $[0,T]$. This implies that process (8.7) is a martingale on $[0,T]$ and the lemma is proved. $\square$

Proof of Lemma 7.3. Notice that (7.17) yields U_t once W_t and V_t are found. Equation (7.16) is linear with respect to V_t and proving the existence and uniqueness of its solution presents no difficulty if W_t is known.

Equation (7.15) can be considered for each ω separately. Then the theory of ODEs allows us to conclude that a unique solution exists until it blows up and it is $\mathcal F^y_t$-adapted. Uniqueness implies that $W_t=W^*_t$. Furthermore, at least on a small time interval $W_t>0$. It turns out that $W_t>0$ on any interval of time where W_t is bounded.

Indeed, if not, then for some $t_0>0$ we would have that $\det W_{t_0}=0$, W_t is bounded on $[0,t_0]$ and $\det W_t>0$ for $t<t_0$. However, for $t<t_0$

$$\frac{d}{dt}\det W_t=\operatorname{tr}\dot W_t W_t^{-1}\det W_t, \tag{8.8}$$

and

$$\operatorname{tr}\dot W_t W_t^{-1}=2\operatorname{tr}(\dot{\mathrm{B}}_t\sigma_t^*-\dot b_t)-2\operatorname{tr}\hat a_t W_t+\operatorname{tr}\dot{\mathrm{B}}_t\dot{\mathrm{B}}_t^* W_t^{-1},$$

where the last term is nonnegative as the trace of the product of two symmetric nonnegative matrices. It follows, that $\operatorname{tr}\dot W_t W_t^{-1}$ is bounded from below on $[0,t_0)$ and hence equation (8.8) implies that $\det W_{t_0}>0$.

Next, it turns out that the solution does not blow up on $[0,T]$. Indeed

$$\frac{d}{dt}\operatorname{tr}W_tW_t=4\operatorname{tr}(\dot{\mathrm{B}}_t\sigma_t^*-\dot b_t)W_tW_t+2\operatorname{tr}\dot{\mathrm{B}}_t\dot{\mathrm{B}}_t^*W_t-4\operatorname{tr}\hat a_tW_t^3,$$

where the last trace is nonnegative again on the interval of existence of W_t. Here

$$\operatorname{tr}\dot{\mathrm{B}}_t\dot{\mathrm{B}}_t^*W_t\le N(\operatorname{tr}W_t^2)^{1/2}\le N+\operatorname{tr}W_t^2,$$

where N is a constant. Also for two matrices A and W such that W is symmetric and nonnegative it holds that

$$(\operatorname{tr}AW^2)^2\le\|A\|\|W^2\|\le\|A\|(\operatorname{tr}W^2)^2.$$

This and Gronwall's inequality imply that W_t is bounded on $[0,T]$. Obviously the bound of W_t is uniform with respect to ω. The lower bound is also uniform since by the above $\det W_t$ is bounded away from zero on $[0,T]$ uniformly with respect to ω. The lemma is proved. □

9. Proof of Theorems 7.1 and 7.4

Take a function $\varphi\in C_0^\infty(\mathbb{R}^{d_1})$ and let $c(t,y)$ be a smooth, bounded, and nonnegative function on $[0,T]\times\mathbb{R}^{d_1-d}$. Recall that the operator $\check L$ is introduced in (7.4) and consider the following deterministic problem

$$\begin{aligned}\partial_t v(t,z)+\check Lv(t,z)-c(t,y)v(t,z)=0,\quad t\in[0,T],z\in\mathbb{R}^{d_1},\\ v(T,z)=\varphi(z),\quad z\in\mathbb{R}^{d_1}.\end{aligned}\tag{9.1}$$

REMARK 9.1. By Theorem 2.5 of [18], for any $\alpha\in(0,1)$ there exists a unique classical solution v of (9.1) such that, for any $t\in[0,T]$, $v(t,\cdot)\in C^{2+\alpha}(\mathbb{R}^{d_1})$ and the standard $C^{2+\alpha}(\mathbb{R}^{d_1})$-norms of $v(t,\cdot)$ are bounded on $[0,T]$. If we denote by $z_t(s,z)$, $t\ge s$, the solution of system (7.1) which starts at z at moment $s\le T$, then by Itô's formula we have

$$v(s,z)=E\varphi(z_T(s,z))\exp\left(-\int_s^T c_r(y_r(s,z))\,dr\right),$$
$$|v(s,z)|\le\sup|\varphi|P\{\tau(s,z)\le T\}\le\sup|\varphi|e^{N_0T}Ee^{-N_0\tau(s,z)},$$

where $N_0 > 0$ is an arbitrary constant, $\tau(s,z)$ is the first time $z_t(s,z)$ hits $\{z : |z| \le R\}$, and R is such that $\varphi(z) = 0$ for $|z| \ge R$. Take an $m \ge 0$ and introduce $\psi(z) = (1+|z|^2)^{-m}$. It is not hard to see that, if N_0 is sufficiently large, then

$$\check{L}_t\psi(z) - N_0\psi(z) \le 0.$$

By Itô's formula, for $|z| \ge R$,

$$\psi(R)Ee^{-N_0\tau(s,z)} \le \psi(z),$$

implying that for any $m \ge 0$ there is a constant N such that for all (s,z)

$$|v(s,z)| \le \frac{N}{(1+|z|^2)^m}.$$

The argument in the proof of Lemma 4.11 of [20] proves that the same estimate holds for $\partial v(s,z)/\partial z^i$ and $\partial^2 v(s,z)/\partial z^i\,\partial z^j$, $i,j = 1,\dots,d_1$.

Before we come to a crucial point, we state the following.

LEMMA 9.1. *Let ξ_t be a nonnegative continuous martingale on $[0,T]$ and let ζ_t be a continuous $\mathcal{F}_t$-adapted process given on $[0,T]$ such that $\xi_t\zeta_t$ is a local martingale on $[0,T)$. Assume that*

$$E\xi_T \sup_{[0,T]} |\zeta_t| < \infty.$$

Then $\xi_t\zeta_t$ is a martingale on $[0,T]$.

Proof. We need to prove that for any stopping time $\tau \le T$ we have $E\xi_\tau\zeta_\tau = E\xi_0\zeta_0$. Here the left-hand side equals $E\xi_T\zeta_\tau$ and we are given that there exists a sequence of stopping times $\tau_n \uparrow T$ such that $E\xi_T\zeta_{\tau\wedge\tau_n} = E\xi_0\zeta_0$. Using the dominated convergence theorem yields the desired result and proves the lemma. $\square$

LEMMA 9.2. *The process*

$$\rho_t e^{-\int_0^t c_s(y_s)\,ds} \int_{\mathbb{R}^d} v(t,x,y_t)\bar{\pi}_t(x)\,dx$$

is a martingale on $[0,T]$.

Proof. Define $(c_t = c(t,y_t),\ v_t(x) = v(t,x,y_t))$

$$D^y_k = \frac{\partial}{\partial y^k}, \qquad D^y_{kr} = D^y_k D^y_r, \qquad C_t = \exp\left(-\int_0^t c_s\,ds\right), \qquad \chi_t = C_t v_t \bar{\pi}_t.$$

We need to show that

$$\rho_t \int_{\mathbb{R}^d} \chi_t(x)\,dx \tag{9.2}$$

is a martingale.

Observe that by Itô's formula and (9.1), we have

$$
\begin{aligned}
(9.3)\qquad d[v_t(x)C_t] &= d[v_tC_t] \\
&= C_t[D^y_k v_t dy^k_t + (\partial_t v_t - c_tv_t + \check{a}^{kr}_t D^y_{kr}v_t)\,dt] \\
&= C_t\big[D^y_k v_t\Phi^{kr}_t d\tilde{w}^r_t \\
&\quad - \big(L_tv_t + 2\check{a}^{ik}_t D_iD^y_k v_t + (B^k_t - \tilde{B}^k_t)D^y_k v_t\big)\,dt\big],
\end{aligned}
$$

where we dropped the arguments x for shortness and, of course, $D^y_k v_t = (D^y_k v)(t,x,y_t)$, $D^y_{kr}v_t = (D^y_{kr}v)(t,x,y_t)$, and $D_iD^y_k v_t = (D_iD^y_k v)(t,x,y_t)$. By the way, observe that

$$
\sigma^{ir}_t\Phi^{kr}_t = 2\check{a}^{ik}_t, \qquad B^k_t = \Phi^{kr}_t \mathrm{B}^r_t.
$$

Similarly to the proof of Lemma 8.5, we conclude that (9.3) holds in the sense of distributions and that Theorem 6.1 is applicable to $v_t\bar{\pi}_t$ on the time interval $t\le\tau_m$ for any n, where τ_m are taken from Lemma 8.5. It follows that for any m for $t\le\tau_m$

$$
\begin{aligned}
d\chi_t = C_t&(\bar{\pi}_t\Phi^{kr}_t D^y_k v_t + v_t\Lambda^{r*}_t\bar{\pi}_t)\,d\tilde{w}^r_t \\
&- C_t\bar{\pi}_t\big(L_tv_t + \sigma^{ir}_t\Phi^{kr}_t D_iD^y_k v_t + \Phi^{kr}_t(\mathrm{B}^r_t - \tilde{\mathrm{B}}^r_t)D^y_k v_t\big)\,dt \\
&+ C_tv_t(L^*_t\bar{\pi}_t + \tilde{\mathrm{B}}^k_t\Lambda^{k*}_t\bar{\pi}_t)\,dt + C_t(D^y_k v_t)\Phi^{kr}_t\Lambda^{r*}_t\bar{\pi}_t\,dt.
\end{aligned}
$$

It is convenient to rearrange the above terms by using the notation

$$
\zeta^r_t = C_t(\bar{\pi}_t\Phi^{kr}_t D^y_k v_t + v_t\Lambda^{r*}_t\bar{\pi}_t).
$$

We have

$$
d\chi_t = \zeta^r_t\,d\tilde{w}^r_t + (\tilde{\mathrm{B}}^r_t\zeta^r_t + I^1_t + I^2_t)\,dt, \quad t\le\tau_m,
$$

where

$$
\begin{aligned}
I^1_t &= C_t(v_tL^*_t\bar{\pi}_t - \bar{\pi}_tL_tv_t), \\
I^2_t &= -C_t\Phi^{kr}_t\sigma^{ir}_t\big(\bar{\pi}_t\sigma^{ir}_t D_iD^y_k v_t + (D^y_k v_t)D_i\bar{\pi}_t\big) \\
&= -C_t\Phi^{kr}_t\sigma^{ir}_t D_i(\bar{\pi}_tD^y_k v_t).
\end{aligned}
$$

In the integral form this means that for any $\phi\in C^\infty_0$ with probability one

$$
\begin{aligned}
(9.4)\qquad (\chi_{t\wedge\tau_m},\phi) &= (\chi_0,\phi) + \int_0^t I_{s\le\tau_m}(\zeta^r_s,\phi)\,d\tilde{w}^r_s \\
&\quad + \int_0^t I_{s\le\tau_m}\tilde{\mathrm{B}}^r_s(\zeta^r_s,\phi)\,ds \\
&\quad + \int_0^t I_{s\le\tau_m}C_sa^{ij}_s(\bar{\pi}_sD_jv_s - v_sD_j\bar{\pi}_s, D_i\phi)\,ds \\
&\quad + \int_0^t I_{s\le\tau_m}C_s[(\bar{\pi}_sv_s, b^i_sD_i\phi) + \Phi^{kr}_s\sigma^{ir}_s(\bar{\pi}_tD^y_k v_s, D_i\phi)]\,ds.
\end{aligned}
$$

We take a ϕ such that $\phi(0)=1$ and plug ϕ_j into (9.4) in place of ϕ, where $\phi_j(x)=\phi(x/j)$, $j=1,2,\dots$.

Observe that

$$(\zeta^r_s,\phi_j)=C_s(\phi_j,\bar\pi_s\Phi^{kr}_s D^y_k v_s+v_s\Lambda^{r*}_s\bar\pi_s)$$

and for any r and k

$$\int_0^T(1,|\bar\pi_s D^y_k v_s|+|v_s\Lambda^{r*}_s\bar\pi_s|)^2\,ds$$
$$\le N\int_0^T\|\bar\pi_s\|^2_{W^1_p}\|v_s\|^2_{W^1_{p'}}\,ds\le N\|\bar\pi\|^2_{\mathbb{W}^1_p(T)}<\infty,$$

where N is independent of ω. By the dominated convergence theorem and the rules for passing to the limit under the sign of stochastic integral it follows that in probability uniformly on $[0,T]$

$$\int_0^t I_{s\le\tau_m}(\zeta^r_s,\phi_j)\,d\tilde w^r_s\to\int_0^t I_{s\le\tau_m}C_s(1,\bar\pi_s\Phi^{kr}_s D^y_k v_s+v_s\Lambda^{r*}_s\bar\pi_s)\,d\tilde w_s.$$

Similarly, and in an easier fashion one analyzes the remaining terms in (9.4) and concludes that for any m

$$d(\chi_t,1)=C_t(1,\bar\pi_t\Phi^{kr}_t D^y_k v_t+v_t\Lambda^{r*}_t\bar\pi_t)\,d\tilde y_t,\quad t\le\tau_m.$$

By using Itô's formula we then immediately obtain that the process (9.2) is at least a local martingale on $[0,T]$. We rewrite it as $\xi_t\zeta_t$, where (see Remark 7.6 and Lemma 8.6) $\xi_t=\rho_t(\tilde{\text{B}}-\text{B}.(0)+\dot{\text{B}}^*W^{-1}V,d\tilde w)$ and

$$\zeta_t=e^{-A_t-\int_0^t c_s(y_s)\,ds}\int_{\mathbb{R}^d}\hat\pi_t(x)v_t(x)\exp\biggl(-\frac12\int_0^t|W^{1/2}_s x+W^{-1/2}_s V_s|^2\,ds\biggr)\,dx.$$

Owing to Lemma 8.2 the process ζ_t is bounded on $[0,T]$ by a constant times $\|\pi_0\|_{\mathcal{L}_p}$ which along with Lemma 9.1 implies that $\xi_t\zeta_t$ is a martingale. The lemma is proved. □

Proof of Theorem 7.1. Recall that assertions (i)–(iii) are proved in Lemma 8.5. By Lemma 9.2 and Itô's formula

$$\begin{aligned}Ee^{-\int_0^T c_s(y_s)\,ds}\varphi(z_T)&=Ev(0,x_0,y_0)=E\int_{\mathbb{R}^d}v(0,x,y_0)\bar\pi_0\,dx\\&=E\rho_Te^{-\int_0^T c_s(y_s)\,ds}\int_{\mathbb{R}^d}\varphi(x,y_T)\bar\pi_T(x)\,dx\\&=E\bar{\rho_T}e^{-\int_0^T c_s(y_s)\,ds}\int_{\mathbb{R}^d}\varphi(x,y_T)\bar\pi_T(x)\,dx,\end{aligned}$$

where $\bar{\rho_T}=E(\rho_T\mid\mathcal{F}^y_T)$. Since the equality between the extreme terms holds for sufficiently wide class of functions c, we get that

$$E(\varphi(z_T)\mid\mathcal{F}^y_T)=\bar{\rho_T}\int_{\mathbb{R}^d}\varphi(x,y_T)\bar\pi_T(x)\,dx\quad\text{(a.s.)}.$$

The arbitrariness of ϕ implies that $\bar\pi_T \geq 0$ (a.s.) and

$$1 = \bar{\rho_T} \int_{\mathbb{R}^d} \bar\pi_T(x)\,dx, \qquad (1,\bar\pi_T) = \int_{\mathbb{R}^d} \bar\pi_T(x)\,dx > 0, \qquad \bar{\rho_T} = (1,\bar\pi_T)^{-1}$$

(a.s.). It follows that for any Borel $f \geq 0$ equation (7.13) holds with $t = T$.

The above argument can be repeated for any $t \leq T$ by taking t in place of T. Then we obtain (7.13) for any t. Furthermore, for any t we will have that that $\bar\pi_t \geq 0$ and $(1,\bar\pi_t) > 0$ (a.s.). Actually, the last two properties hold with probability one for all t at once since with probability one $\bar\pi_t$ is a continuous $\mathcal{L}_1$-function by Lemma 8.5 and by Lemma 8.2, on the set where $\tau = \inf\{t \geq 0 : (1,\bar\pi_t) = 0\} < T$, we have $\bar\pi_T = 0$, which only happens with probability zero. The theorem is proved. □

Proof of Theorem 7.4. We use part of notation from the proof of Lemma 9.2 but this time take $\bar\pi_t = \eta_t = e^{-Q_t}$. Then by Itô's formula and (7.20) we obtain that for each x

$$d\chi_t = \zeta_t^r(d\tilde w_t^r + \tilde{\mathrm{B}}^r\,dt) + C_t(v_t L_t^*\bar\pi_t - \bar\pi_t L_t v_t)\,dt - C_t\Phi_t^{kr}\sigma_t^{ir}D_i(\bar\pi_t D_k^y v_t)\,dt.$$

By using the stochastic Fubini theorem and integrating by parts, we see that

$$d(\chi_t,1) = (\zeta_t^r,1)(d\tilde w_t^r + \tilde{\mathrm{B}}^r\,dt)$$

which implies that process (9.2) is a local martingale on $[0,T]$. We rewrite it as $\xi_t\zeta_t$, where (see Remark 7.6 and Lemma 8.6) $\xi_t = \rho_t(\tilde{\mathrm{B}} - \mathrm{B}.(0) + \dot{\mathrm{B}}^* W^{-1}V, d\tilde w)$ and

$$\zeta_t = e^{-A_t - \int_0^t c_s(y_s)\,ds} \int_{\mathbb{R}^d} v_t(x) \exp\left(-\frac{1}{2}\int_0^t |W_s^{1/2}x + W_s^{-1/2}V_s|^2\,ds\right) dx.$$

Notice that ξ_t is a martingale and ζ_t is obviously bounded. By Lemma 9.1, we conclude that process (9.2) is a martingale.

After that it suffices to repeat the proof of Theorem 7.1 dropping unnecessary here details concerning the fact that $(1,\bar\pi_t) > 0$. The theorem is proved. □

References

[1] P. Cannarsa and V. Vespri, *Generation of analytic semigroups by elliptic operators with unbounded coefficients*, SIAM J. Math. Anal. **18** (1987), no. 3, 857–872. MR 0883572

[2] P. Cannarsa and V. Vespri, *Existence and uniqueness results for a nonlinear stochastic partial differential equation*, in Stochastic Partial Differential Equations and Applications Proceedings (G. Da Prato and L. Tubaro, eds.), Lecture Notes in Math., vol. 1236, Springer Verlag, Berlin, 1987, pp. 1–24. MR 0899293

[3] A. Chojnowska-Michalik and B. Goldys, *Generalized symmetric Ornstein–Uhlenbeck semigroups in L^p: Littlewood-Paley-Stein inequalities and domains of generators*, J. Funct. Anal. **182** (2001), 243–279. MR 1828795

[4] G. Cupini and S. Fornaro, *Maximal regularity in $L^p(\mathbb{R}^N)$ for a class of elliptic operators with unbounded coefficients*, Differential Integral Equations **17** (2004), no. 3–4, 259–296. MR 2037979

[5] B. Farkas and A. Lunardi, *Maximal regularity for Kolmogorov operators in L^2 spaces with respect to invariant measures*, J. Math. Pures Appl. **86** (2006), 310–321. MR 2257846

[6] M. Geissert and A. Lunardi, *Invariant measures and maximal L^2 regularity for nonautonomous Ornstein–Uhlenbeck equations*, J. Lond. Math. Soc. (2) **77** (2008), no. 3, 719–740. MR 2418301

[7] I. Gyöngy, *Stochastic partial differential equations on Manifolds, I*, Potential Analysis **2** (1993), 101–113. MR 1246744

[8] I. Gyöngy, *Stochastic partial differential equations manifolds II. Nonlinear filtering*, Potential Analysis **6** (1997), 39–56. MR 1436821

[9] I. Gyöngy and N. V. Krylov, *On stochastic partial differential equations with unbounded coefficients*, Potential Analysis **1** (1992), no. 3, 233–256. MR 1245229

[10] N. V. Krylov, *Some properties of traces for stochastic and deterministic parabolic weighted Sobolev spaces*, Journal of Functional Analysis **183** (2001), no. 1, 1–41. MR 1837532

[11] N. V. Krylov, *On divergence form SPDEs with VMO coefficients*, SIAM J. Math. Anal. **40** (2009), no. 6, 2262–2285. MR 2481294

[12] N. V. Krylov, *On linear elliptic and parabolic equations with growing drift in Sobolev spaces without weights*, Problemy Matemtaticheskogo Analiza **40** (2009), 77–90, in Russian; English version in Journal of Mathematical Sciences **159** (2009), no. 1, 75–90, Srpinger. MR 2544045

[13] N. V. Krylov, *Itô's formula for the L_p-norm of stochastic W^1_p-valued processes*, Probab. Theory Related Fields **147** (2010), 583–605.

[14] N. V. Krylov, *Filtering equations for partially observable diffusion processes with Lipschitz continuous coefficients*, The Oxford Handbook of Nonlinear Filtering, Oxford University Press, 2011, pp. 169–194.

[15] N. V. Krylov, *On divergence form SPDEs with growing coefficients in W^1_2 spaces without weights*, SIAM J. Math. Anal. **42** (2010), 609–633.

[16] N. V. Krylov, *On divergence form second-order PDEs with growing coefficients in W^1_p spaces without weights*, Progr. Nonlinear Differential Equations Appl. **60** (2011), 389–414.

[17] N. V. Krylov, *On the Itô-Wentzell formula for distribution-valued processes and related topics*, Probab. Theory Related Fields **150** (2011), 295–319. MR 2800911

[18] N. V. Krylov and E. Priola, *Elliptic and parabolic second-order PDEs with growing coefficients*, Comm in PDEs **35** (2010), no. 1, 1–22. MR 2748616

[19] N. V. Krylov and B. L. Rozovskii, *Stochastic evolution equations*, Itogy nauki i tekhniki **14** (1979), 71–146 in Russian; English translation: J. Soviet Math. **16** (1981), no. 4, 1233–1277. MR 0570795

[20] N. V. Krylov and A. Zatezalo, *A direct approach to deriving filtering equations for diffusion processes*, Applied Mathematics and Optimization **42** (2000), no. 3, 315–332. MR 1795614

[21] R. Sh. Liptser, *Conditionally Gaussian random processes*, Problemy Peredachi Informatsii **X** (1974), no. 2, 75–94, Russian; English translation: Problems of Information Transmission **10** (1974), no. 2, 151–167. MR 0483005

[22] R. Sh. Liptser and A. N. Shiryaev, *Statistics of Random Processes*, Nauka, Moscow, 1974 in Russian; English translation: Vols. I, II, Springer-Verlag, New York, 1977–1978, 2nd edition 2001. MR 180085 MR 1800858

[23] A. Lunardi, *Schauder estimates for a class of degenerate elliptic and parabolic operators with unbounded coefficients in $\mathbb{R}^n$*, Ann. Sc. Norm. Super Pisa, Ser. IV. **24** (1997), 133–164. MR 1475774

[24] A. Lunardi and V. Vespri, *Generation of strongly continuous semigroups by elliptic operators with unbounded coefficients in $L^p(\mathbb{R}^n)$*, Rend. Istit. Mat. Univ. Trieste **28** (1996), suppl., 251–279 (1997). MR 1602271

[25] G. Metafune, J. Prüss, A. Rhandi and R. Schnaubelt, *The domain of the Ornstein–Uhlenbeck operator on an L^p-space with invariant measure*, Ann. Sc. Norm. Super. Pisa, Cl. Sci. **5** 1 (2002), 471–485. MR 1991148

[26] G. Metafune, J. Prüss, A. Rhandi and R. Schnaubelt, *L^p-regularity for elliptic operators with unbounded coefficients*, Adv. Differential Equations **10** (2005), no. 10, 1131–1164. MR 2162364

[27] J. Prüss, A. Rhandi and R. Schnaubelt, *The domain of elliptic operators on $L^p(\mathbb{R}^d)$ with unbounded drift coefficients*, Houston J. Math. **32** (2006), no. 2, 563–576. MR 2293873

N. V. Krylov, 127 Vincent Hall, University of Minnesota, Minneapolis, MN 55455, USA

E-mail address: krylov@math.umn.edu