KALMAN-BUCY FILTER AND SPDES WITH GROWING LOWER-ORDER COEFFICIENTS IN $W^1_p$ SPACES WITHOUT WEIGHTS

N.V. KRYLOV

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 $p$th 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 $p$th 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 cut-off functions whose support evolves in time in a manner adapted to the drift terms. The paper consists of two parts: Sections 2 through 6 are devoted to some general issues of the theory of SPDEs with growing coefficients and in Sections 7 through 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.

2000 Mathematics Subject Classification. 60H15, 93E11.

Key words and phrases. Stochastic partial differential equations, Kalman-Bucy filter, Sobolev spaces, growing coefficients.

The work was partially supported by NSF grant DMS-0653121.
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 \( \Lambda_k^2 \) in (7.11) are of zeroth order. The main obstacle here is that the zeroth order coefficient of \( \Lambda_k^2 \) 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 first-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 \( \omega \). 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 \( \omega \). 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 apriori 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 \((\mathcal{F}, P)\) \(\sigma\)-fields \(\mathcal{F}_t \subset \mathcal{F}\). Denote by \(\mathcal{P} = \mathcal{P}(\{\mathcal{F}_t\})\) the predictable \(\sigma\)-field in \(\Omega \times (0, \infty)\) associated with \(\{\mathcal{F}_t\}\). Let \(w^k_t, k = 1, 2, \ldots\), be independent one-dimensional Wiener processes with respect to \(\{\mathcal{F}_t\}\). Let \(\tau\) be a stopping time.

We consider the second-order operator \(L_t\)

\[
L_t u_t(x) = D_t \left( a^{ij}_t(x) D_{ij} u_t(x) + b^i_t(x) u_t(x) \right) + b^i_t(x) D_i u_t(x) - c_t(x) u_t(x),
\]

and the first-order operators

\[
\Lambda^k_t u_t(x) = \sigma^{ik}_t(x) D_i u_t(x) + \nu^k_t(x) u_t(x)
\]

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 \(k = 1, 2, \ldots\) (the summation convention is enforced throughout the article), where

\[
D_i = \frac{\partial}{\partial x^i}.
\]

We set \(\mathbb{R}^+_1 = [0, \infty)\).

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

\[
du = (L_t u_t - \lambda u_t + D_i f^i_t + f^0_t) dt + (\Lambda^k_t u_t + g^k_t) dw^k_t, \quad t \leq \tau,
\]

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 \(\ell_2\)-valued functions such as \(g_t = (g^k_t)\) in (2.2). For instance, if \(u(x) = (u^1(x), u^2(x), \ldots)\) is an \(\ell_2\)-valued measurable function on \(\mathbb{R}^d\), then

\[
\|u\|_{L_p}^p = \int_{\mathbb{R}^d} |u(x)|^p \, dx = \int_{\mathbb{R}^d} \left( \sum_{k=1}^{\infty} |u^k(x)|^2 \right)^{p/2} \, dx.
\]

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 \(\tau\) is a stopping time and introduce

\[
\mathbb{L}_p(\tau) := \mathbb{L}_p(\{\mathcal{F}_t\}, \tau) := \mathcal{L}_p([0, \tau], \mathcal{P}, \mathcal{L}_p),
\]

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

\[
\mathbb{L}_p = \mathbb{L}_p(\infty), \quad \mathbb{W}^1_p = \mathbb{W}^1_p(\infty).
\]
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 \( \mathbb{W}_p^1(\tau) \).

The following definition is most appropriate for investigating our equations if the coefficients of \( L_t \) and \( \Lambda_t^p \) are bounded.

**Definition 2.1.** Introduce \( \mathcal{W}^1_p(\tau) \), as the space of functions \( u_t = u_t(\omega, \cdot) \) on \( \{(\omega, t) : 0 \leq t \leq \tau(\omega), t < \infty\} \) 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 \( u \in \mathbb{W}_p^1(\tau) \);

(iii) There exist \( f^i \in \mathbb{L}_p(\tau), i = 0, \ldots, d, \) and \( g = (g^1, g^2, \ldots) \in \mathbb{L}_p(\tau) \) such that for any \( \varphi \in C_0^\infty = C_0^\infty(\mathbb{R}^d) \) with probability 1 for all \( t \in [0, \infty) \) we have

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

\[
+ \int_0^t I_{s \leq \tau}(f^{0}_s, \varphi) - (f^{i}_s, D_i \varphi) \, ds.
\]

(2.3)

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

\[
du_t = (D_i f^i_t + f^{0}_t) \, dt + g^k_t \, dw^k_t, \quad t \leq \tau.
\]

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 \( (u, \phi) \) represents the result of the action of \( u \) on the test function \( \phi \in C_0^\infty \). When \( u \) is a locally integrable function, \( (u, \phi) \) 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 \( \bar{f}_s = D_i f^i + f^0_s \). In the framework of Definition 2.1 we have \( \bar{f} \in \mathcal{L}_p((0, \tau], \mathcal{P}, H_p^{-1}) \), where \( H_p^{-1} = (1 - \Delta)^{1/2} \mathcal{L}_p \). One also knows that any \( \bar{f} \in \mathcal{L}_p((0, \tau], \mathcal{P}, H_p^{-1}) \) is written as \( \bar{f}_s = D_i f^i_s + f^0_s \) with some \( f^i \in \mathbb{L}_p(\tau) \).

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 \( \Omega \) with values in the space of generalized functions on \( \mathbb{R}^d \). We write \( u_0 \in \text{tr} \mathcal{W}^1_p = \text{tr} \mathcal{W}^1_p(\mathcal{F}_0) \) if there exists a function \( v \in \mathcal{W}^1_p \) 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

\[
\|u_0\|_{\text{tr} \mathcal{W}^1_p} = E\|v\|_{\mathcal{W}^1_p}.
\]
One knows that \( \text{tr} \mathcal{W}_p^1 \) 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 \( \Lambda^k_t \) are bounded coincides with the one obtained by applying Definition 2.1.

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

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

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

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 \( \Lambda^k \) 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

\[
f_t^i - a_t^{ij} D_j u_t - b_t^i u_t, \quad i = 1, \ldots, d, \quad f_t^0 + (c_t + \lambda) u_t - b_t^0 D_i u_t, \\
g_t^k - a_t^{ik} D_i u_t - \nu_t^k u_t
\]
in place of \( f_t^i, i = 1, \ldots, d, f_t^0, \) 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 \( a_t^{ij}(x), b_t^i(x), b_t^j(x), c_t(x), \sigma_t^{ik}(x), \nu_t^i(x) \) are real valued, measurable with respect to \( \mathcal{F} \otimes \mathcal{B}(\mathbb{R}^{d+1}) \), \( \mathcal{F}_t \)-adapted for any \( x, c \geq 0 \).

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

\[
(a^{ij} - \alpha^{ij}) \xi^i \xi^j \geq \delta |\xi|^2, \quad |a^{ij}| \leq \delta^{-1}, \quad |\nu|_{L^2} \leq \delta^{-1},
\]

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

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

\[
\int_{B_{1}} \left( |b_t(x + y)| + |b_t(x + y)| + |c_t(x + y)| \right) dy
\]

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 $b, b, c$.

Recall that as is well known if $u \in \mathbb{W}^1_p(\tau)$, then owing to the boundedness of $\nu$ and $\sigma$ 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 $T \in \mathbb{R}_+$

$$E \sup_{t \leq T} |m_t|^p \leq NE\left( \sum_{k=1}^{\infty} \int_0^{T \wedge r} (\sigma^i_k D_t u_s + \nu^k u_s + g^k_s, \phi)^2 ds \right)^{p/2} \leq N\|\phi\|_{\mathbb{L}^p}^{p/2} \left( \int_0^{T \wedge r} \sum_{k=1}^{\infty} (|\sigma^i_k|^2 |D_t u_s|^2 + |\nu^k|^2 |u_s|^2 + |g^k_s|^2, |\phi|) ds \right)^{p/2} \leq N\|\phi\|_{\mathbb{L}^p}^{p/2} + \|g\|_{\mathbb{L}_p(\tau)}^{p}, \tag{3.2}$$

where the constants $N$ depend only on $\phi, d, p, \delta, \tau$.

**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 $i = 1, \ldots, d$ we have

$$|\sigma^i_t(x) - \sigma^i_t(y)|_{\mathbb{L}_2} \leq \kappa(|x - y|). \tag{3.3}$$

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 $i, j = 1, \ldots, d$ we have

$$\rho^{-2d-2} \int_t^{t+\rho^2} \left( \sup_{x \in \mathbb{R}^d} \int_{B_\rho(x)} \int_{B_\rho(x)} |a_{ij}^t(y) - a_{ij}^t(z)| dy dz \right) ds \leq \gamma_a. \tag{3.4}$$

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 $\omega$ 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

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

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

$$\int_{B_{\rho_1}(x)} \int_{B_{\rho_1}(x)} |b_t(y) - b_t(z)|^q dy dz + \int_{B_{\rho_1}(x)} \int_{B_{\rho_1}(x)} |b_t(y) - b_t(z)|^q dy dz$$

$$+ \int_{B_{\rho_1}(x)} \int_{B_{\rho_1}(x)} |c_t(y) - c_t(z)|^q dy dz \leq K_I_{q > a} + \rho_1^q \gamma_b. \tag{3.5}$$
Then the result follows immediately from Theorem 3.3. They also are satisfied with any $q > d$, $\gamma_b = 0$, and $\rho_1 = 1$ on the account of choosing $K$ appropriately if, say,

$$|b_t(x) - b_t(y)| + |b_t(x) - b_t(y)| + |c_t(x) - c_t(y)| \leq K_1$$

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

$$|b_t(x) - b_t(y)| + |b_t(x) - b_t(y)| + |c_t(x) - c_t(y)| \leq K_1|x - y|, \quad \forall x, y \in \mathbb{R}^d, t \geq 0. \quad (3.5)$$

We see that Assumption 3.4 allows $b$, $b$, and $c$ growing linearly in $x$. Here is our result on apriori estimates of solutions of (2.2).

**Theorem 3.1.** There exist

$$\gamma_a = \gamma_a(d, \delta, p), \quad \gamma_b = \gamma_b(d, \delta, p, \kappa, \rho_0) \in (0, 1],
$$

$$N = N(d, \delta, p, \kappa, \rho_0), \quad \lambda_0 = \lambda_0(d, \delta, p, \kappa, \rho_0, \rho_1, K) \geq 1$$

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

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

$$+ N\lambda^{-1}\|f^0\|_{\mathbb{L}_p(\tau)}^2 + N\|u_0\|_{\mathcal{W}_p^1}^2. \quad (3.6)$$

**Remark 3.1.** There is an unusual property of $u_t$, which is nontrivial even if $f^i = g_i \equiv 0$.

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

$$\int_0^T |\xi_t|_{\ell_2}^2 \, dt < \infty \quad (a.s.)$$

and assume that $E\rho_{\tau}(\xi) = 1$, where

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

Then the assertion of Theorem 3.1 holds with the same $\gamma_a, \gamma_b, \lambda_0$, and $N$ if we understand $\|v\|_{\mathbb{L}_p(\tau)}^p$ for all $v$'s as

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

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 $(\omega, t)$. 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 3.3.
Theorem 3.1 admits the following version if \( \tau \) 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 \( \gamma_a \) and \( \gamma_b \) from Theorem 3.1. Let \( \lambda = 0 \) and let \( u \) be a solution of (2.2) with initial data \( u_0 \in \text{tr} \mathcal{W}^{1}_p \) and some \( f^j, g \in \mathbb{L}^p(\tau) \). Then

\[
\|u\|^2_{\mathcal{W}^{1}_p(\tau)} \leq N \left( \sum_{i=0}^{d} \|f^i\|^2_{\mathcal{L}^p(\tau)} + \|g\|^2_{\mathcal{L}^p(\tau)} + \|u_0\|^2_{\text{tr} \mathcal{W}^{1}_p} \right),
\]

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 \( \mu \), the function \( v_t := u_t e^{-\mu t} \) satisfies (2.2) with \( \lambda + \mu, f^j t e^{-\mu t}, \) and \( g^k t e^{-\mu t} \) in place of \( \lambda, f^j t, \) and \( g^k \), respectively. If \( \mu \) 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 apriori 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

\[
D_i (b^i_t u_t) + b_t^i D_i u_t - c_t u_t
\]

can be written even locally as \( D_i \hat{f}_i + \hat{f}_0 \) with \( \hat{f}_i^j \in \mathbb{L}^p(\tau) \) if we only know that \( u \in \mathcal{W}^{1}_p(\tau) \) even if \( b, b, \) 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 \( \omega \)) 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 (3.5) holds and \( b_t(0), b_t(0), \) and \( c_t(0) \) are bounded for each \( \omega \).

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

**Remark 3.4.** If the stopping time \( \tau \) 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 \( \phi \in C_0^\infty \). However, under stronger assumptions we can say more about regularity of \( u \). In the following theorem by \( H^p_\gamma \) 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 \( \phi \in C_0^\infty \) we have \( \phi u \in W^{1, p}_u(\tau) \);

(ii) For any \( \phi \in C_0^\infty \) the process \( u_{t \wedge \tau} \phi \) is continuous on \( \mathbb{R}^+ \) as an \( L_p \)-valued process (a.s.);

(iii) If \( p > 2 \) and \( \tau \) is bounded and we have two numbers \( \alpha \) and \( \beta \) such that

\[
2p < \alpha < \beta \leq 1
\]

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

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

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 \( u_{t \wedge \tau} \phi \in C^{\varepsilon_0 - \varepsilon}(\mathbb{R}^d) \) and the norm of \( u_{t \wedge \tau} \phi \) in this space is bounded as a function of \( t \);

(b) for any \( \varepsilon \) as in (a) (a.s.) for any \( x \in \mathbb{R}^d \) we have \( u(x) \phi(x) \in C^{(\varepsilon_0 - \varepsilon)/2}[0, \tau] \) and the norm of \( u(x) \phi(x) \) 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 \( \alpha \) and \( \beta \) close to \( 2/p \) and (iii) (b) after taking \( \alpha \) and \( \beta \) close to \( 1 - d/p \).

Remark 3.5. Let \( p_1, p_2 \in [2, \infty) \), let \( \tau \) 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 \( \gamma_a \) and \( \gamma_b \) 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 \( \sigma \)-fields which are smaller than \( F_t \). It will be used in Section 8 and this is the reason why we use the somewhat strange notation \( \tilde{y}_t \) and \( \tilde{b}_t^k \) below. We suppose that all the above assumptions are satisfied with \( \gamma_a \) and \( \gamma_b \) taken from Theorem 3.1 and let \( \tilde{F}_t, t \geq 0 \), be a filtration of complete with respect to \( \mathcal{F}, P \) \( \sigma \)-fields such that \( \mathcal{F}_t \supset \tilde{F}_t \). Our aim is to show that sometimes \( u_t \) is \( \tilde{F}_t \)-adapted even if some terms in (2.2) are not \( \tilde{F}_t \)-adapted.
However, the equation is assumed to have a special structure. The result is not surprising because in the notation, introduced below, equation

\[ du_t = (\Lambda_t^k u_t + g_t^k) dw_t^k + (L_t u_t + \hat{b}_t D_t u_t - \hat{c}_t u_t + D_t f_t^i + f_t^0 + \hat{f}_t) dt, \quad t \leq \tau \]

is written as

\[ du_t = (\Lambda_t^k u_t + g_t^k) d\tilde{y}_t^k + (L_t u_t + D_t f_t^i + f_t^0) dt, \quad t \leq \tau, \]

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

\[ \rho_t = \rho_t(\tilde{B}, dw) = \exp(-\int_0^t \tilde{b}_s^k dw_s^k - \frac{1}{2} \int_0^t |\tilde{B}_s|_{\ell_2}^2 ds). \]

Suppose that Assumption 3.4 (i) is satisfied with \( \tilde{\mathcal{F}}_t \) in place of \( \mathcal{F}_t \) and the processes

\[ \tilde{y}_t^k = u_t^k + \int_0^t \tilde{B}_s^k ds \]

are \( \tilde{\mathcal{F}}_t \) adapted. Introduce

\[ \tilde{b}_t^i(x) = \sigma_t^{ik}(x)\tilde{b}_t^k, \quad \tilde{c}_t(x) = -u_t^k(x)\tilde{b}_t^k \]

and suppose that \( b + \tilde{b} \) and \( c + \tilde{c} \) satisfy Assumption 3.4 with \( \gamma_b \) from Theorem 3.1, for any \( x \in \mathbb{R}^d \) (and \( \omega \)) we have \( \tilde{c}_t(x) \leq K \), and the function

\[ \int_{B_1} (|\tilde{b}_t(x+y)| + |\tilde{c}_t(x+y)|) dy \]

is locally integrable to the power \( p/(p-2) \) (locally bounded if \( p = 2 \)) on \( \mathbb{R}^+ \).

Let \( \tau \) be an \( \tilde{\mathcal{F}}_t \)-stopping time such that \( \tau \leq T \).

Then, for any initial data \( u_0 \in \text{tr} \mathcal{W}_1^1(\tilde{\mathcal{F}}_0) \) and \( f^j, g \in L_p(\{\tilde{\mathcal{F}}_t\}, \tau) \) such that \( \hat{f} := (g, \tilde{B})_{\ell_2} \in L_p(\{\mathcal{F}_t\}, \tau) \)

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

(ii) for any \( \phi \in C_0^\infty \) 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 \( \lambda \) as we wish, assertion (i) follow immediately from Theorem 3.3.

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

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

$$\tilde{E} \int_0^{\tau_n} \|u_t\|_{L_p}^p dt \quad \text{and} \quad \tilde{E} \int_0^{\tau_n} \|u_t\|_{W_p^1}^p dt$$

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

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

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

and observe that

$$E \int_0^{\tau_n} \|\bar{u}_t\|_{L_p}^p dt \leq nE\rho_{\tau_n} \int_0^{\tau_n} \|\bar{u}_t\|_{L_p}^p dt = n\tilde{E} \int_0^{\tau_n} \|\bar{u}_t\|_{L_p}^p dt < \infty.$$ 

Similar estimates hold if we replace $L_p$ with $W_p^1$. By recalling that $\tilde{F}_t \subset \mathcal{F}_t$, we conclude that $\tilde{u}$ is a solution of (3.3) relative to $\{\mathcal{F}_t\}$ with $\tau_n$ in place of $\tau$. By uniqueness, in the sense of distributions $\tilde{u}_t I_{\tau_n} = u_t I_{\tau_n}$ for almost all $(\omega, t)$, that is, $(\tilde{u}_t, \phi) I_{\tau_n} = (u_t, \phi) I_{\tau_n}$ for almost all $(\omega, t)$ for each fixed $\phi \in C^\infty_0$. Then it follows from the integral form of (3.3) that for each $\phi \in C^\infty_0$ 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 $\tau_n$ with $\tau$ and it only remains to observe that $(\tilde{u}_{t\wedge\tau}, \phi)$ is $\mathcal{F}_t$-measurable. The theorem is proved.

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, $\partial_t = \partial/\partial t$,

$$\hat{b}_i^t = b_i^t - a_{ij}^t D_j \ln \psi_t, \quad \hat{b}_i = b_i - a_{ij} D_j \ln \psi_t,$$

$$\hat{c}_t = c_t + (b_i^t + b_i^0) D_i \ln \psi_t - a_{ij}^t (D_i \ln \psi_t) D_j \ln \psi_t - \partial_t \ln \psi_t,$$

$$\hat{\nu}_k^t = \nu_k^t - a_{ik}^t D_i \ln \psi_t,$$

$$\hat{f}_t^i = f_i^t \psi_t, \quad i = 1, ..., d, \quad \hat{f}_t^0 = f^0_t \psi_t - f_t^1 D_1 \psi_t, \quad \hat{g}_k^t = g_k^t \psi_t.$$

Suppose that, if we replace $b$, $\hat{b}$, $c$, and $\nu$ with $\hat{b}$, $\hat{b}$, $\hat{c}$, and $\hat{\nu}$, respectively, then Assumptions 3.1, 3.4, and 3.5 are satisfied with $\gamma_0$ and $\gamma_b$ from Theorem 3.1. Finally, assume that $\hat{f}_t, \hat{g}_t \in L_2(\tau)$ and $u_0 \psi_0 \in \operatorname{tr} W_p^1(\tau)$. Then it turns out that for $\lambda \geq \lambda_0$ ($\lambda_0$ is taken from Theorem 3.1) equation (2.2) has a unique solution $u$ such that $w \psi \in W_p^1(\tau)$.

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 $f_i^t, g_t$ with $\hat{f}_t^i, \hat{g}_t$, respectively. In addition, the natural estimate of the $W_p^1(\tau)$-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 3.5 take a $T \in (0, \infty)$, set $\tau = T$, and for $0 \leq \theta \leq (0, \infty)$ introduce

$$\ln \psi_t(x) = -\theta e^\theta (t-T) \sqrt{1 + |x|^2}.$$
Obviously, $D_{t} \ln \psi$ are bounded for $t \leq T$. Furthermore, it is not hard to see that if $\theta$ is large enough, then $\tilde{c}_{t} \geq 0$ for $t \leq T$. Also, if $|x - y| \leq 1$, then owing to the fact that $|D_{t} \ln \psi(x)| \leq N(1 + |x|)^{-1}$ for $t \leq T$, where $N$ is a constant, we have
\[
|b_{t}^{i}(x)D_{t} \ln \psi_{t}(x) - b_{t}^{i}(y)D_{t} \ln \psi_{t}(y)| \leq \|(b_{t}^{i}(x) - b_{t}^{i}(y))D_{t} \ln \psi_{t}(x) + N(1 + |x|)|D_{t} \ln \psi_{t}(x) - D_{t} \ln \psi_{t}(y)| \leq K|D_{t} \ln \psi_{t}(x)| + N
\]
for $t \leq T$. Estimates similar to this one show that $\hat{b}$, $\bar{b}$, and $\bar{c}$ satisfy Assumption 3.3 for $t \leq T$. By what is said in the beginning of the current remark, if $u_{0}\psi_{0} \in \text{tr}W_{p}^{1}$ (for instance, $u_{0}(x) = x^{1}$), then (2.2) has a unique solution $u$ such that $u\psi \in W_{p}^{1}(T)$. Since $\text{D} \ln \psi$ is bounded, the inclusion $u\psi \in W_{p}^{1}(T)$ is equivalent to $u\psi \in L_{p}(T)$, $\psi Du \in L_{p}(T)$.

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 $u \in W_{p}^{1}(\tau)$. Take a nonnegative $\xi \in C_{0}^{\infty}(B_{p_{1}})$ with unit integral and define
\[
\bar{b}_{s}(x) = \int_{B_{p_{1}}}^{\xi(y)b_{s}(x - y)} dy, \quad \bar{b}_{s}(x) = \int_{B_{p_{1}}}^{\xi(y)b_{s}(x - y)} dy, \quad \bar{c}_{s}(x) = \int_{B_{p_{1}}}^{\xi(y)c_{s}(x - y)} dy.
\]
We may assume that $|\xi| \leq N(d, p, \rho_{1}, K)$. In particular,
\[
\|I_{B_{p_{1}}(x_{0})b_{s} - \bar{b}_{s}(x_{0})u\|_{L_{p}} + \|I_{B_{p_{1}}(x_{0})c_{s} - \bar{c}_{s}(x_{0})u\|_{L_{p}} \leq N\|u\|_{W_{p}^{1}}, \quad (4.2)
\]
where $N = N(d, p, \rho_{1})$. In particular,
\[
\|I_{B_{p_{1}}(x_{0})b_{s} - \bar{b}_{s}(x_{0})v\|_{L_{p}} \leq (N + |\bar{b}_{s}(x_{0})|)\|v\|_{L_{p}}\|\phi\|_{W_{p}^{1}}, \quad (4.3)
\]
the latter implying that $|b_{s}|I_{B_{p_{1}}(x_{0})v \in H_{p}^{-1}$.

Similarly, $\|I_{B_{p_{1}}(x_{0})b_{s} - \bar{b}_{s}(x_{0})u\|_{L_{p}} + \|I_{B_{p_{1}}(x_{0})c_{s} - \bar{c}_{s}(x_{0})u\|_{L_{p}} \leq (N + |\bar{b}_{s}(x_{0})| + |\bar{c}_{s}(x_{0})|)\|u\|_{W_{p}^{1}}. \quad (4.4)$
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 $\rho_1$ we have $u\phi \in W^1_p(\tau)$. Of course, the restriction on the size of support of $\phi$ 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 $\omega \in \Omega$, $u \in L_p((0,\tau), W^1_p)$, and $\phi \in C_0^\infty(B_R)$

$$
\int_{\tau_n \wedge \tau}^t \left( |\langle b_s^i D_i u_s, \phi \rangle| + |\langle c_s u_s, \phi \rangle| \right) ds 
\leq n\|u\|_{L_p((0,\tau),W^1_p)} \|\phi\|_{W^1_p},
$$

(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 $\phi$ has support in a ball $B_{\rho_1}(x_0)$. Observe that by (4.4) and Hölder’s inequality

$$
|\langle b_s^i u_s, D_i \phi \rangle| + |\langle c_s u_s, \phi \rangle| \leq N(1 + |\bar{b}_s(x_0)| + |\bar{c}_s(x_0)|) \|u_s\|_{W^1_p} \|\phi\|_{W^1_p}.
$$

(4.6)

It follows again by Hölder’s inequality that

$$
\int_{\tau_0}^{t \wedge \tau} \left( |\langle b_s^i u_s, D_i \phi \rangle| + |\langle c_s u_s, \phi \rangle| \right) ds \leq N \chi t \|u\|_{L_p((0,\tau),W^1_p)} \|\phi\|_{W^1_p},
$$

where

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

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

**Remark 4.2.** Estimates (4.3) and (4.4) show that for any $u \in W^1_p$ for almost all $(\omega,s)$ the functions $b_s^i D_i u_s$, $D_i \langle b_s^i u_s \rangle$, 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 $R, \tau_n, \phi$ be as in Lemma 4.1. Then the operators

$$
u_t \to \int_0^{t \wedge \tau_n} \langle b_s^i D_i u_s, \phi \rangle ds, \quad u_t \to \int_0^{t \wedge \tau_n} \langle b_s^i u_s, D_i \phi \rangle ds,
$$

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

are continuous as operators from $W^1_p(\tau)$ to $L_p((0,\tau_n)]$ 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 $b^i$, $b^j$, 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 $\rho_1$).

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

$$du_t = (\Lambda^k_t u_t + g^k_t) \, dw^k_t$$

$$+ \left[ D_i (a^i_t + b^i_t) u_t + (b^i_t + b^j_t) u_t + f^0_t + f^j_t - \lambda u_t \right] \, dt, \quad t \leq \tau. \quad (4.7)$$

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

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

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 $h_t = (b^i_t + b^j_t) u_t$ 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 $u \in \mathbb{W}^1_p(\tau)$.

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

$$B_t = \int_0^t (b^i_s + b^j_s) \, ds, \quad \hat{u}_t(x) = u_t(x - B_t)$$

obtain that

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

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

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

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

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

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

which in terms of [11] implies that the couple $(\hat{a}, \hat{\sigma})$ is $(\epsilon, \epsilon)$-regular at any point of $\mathbb{R}_+ \times \mathbb{R}^d$ for any $\epsilon \in (0, \rho_0]$. Then owing to our Assumptions 3.1 (ii) and 3.2 one can choose $\epsilon = \epsilon(\delta, \kappa) \in (0, \rho_0]$ 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
\[
\lambda \| \dot{u} \|_{L_p(\tau)}^2 + \| D\dot{u} \|_{L_p(\tau)}^2 \leq N\left( \sum_{i=1}^d \| f_i \|_{L_p(\tau)}^2 + \| g \|_{L_p(\tau)}^2 + \lambda^{-1} \| \check{f}^0 \|_{L_p(\tau)}^2 \right),
\]
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 \( \tau_n \) in place of \( \tau \), 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
\[
\xi_t = \exp(\int_0^t c_s \, ds).
\]
By the above argument we have \( \ddot{u} := \xi u \in \mathbb{W}_p^1(\tau) \) and
\[
d\ddot{u}_t = \left[ D_t (a^2 D_j \ddot{u}_t + [b^2 + b_1^2] \ddot{u}_t + \xi_t f^0_t - \lambda \ddot{u}_t) \right] dt + (\Lambda_t \ddot{u}_t + \xi_t \check{f}^k_t) \, dw_t^k, \quad t \leq \tau.
\]
By the above result for any stopping time \( \tau' \leq \tau \)
\[
\lambda^{p/2} \| \xi u \|_{L_p(\tau')}^p + \| \xi D u \|_{L_p(\tau')}^p = \lambda^{p/2} \| \ddot{u} \|_{L_p(\tau')}^p + \| D\ddot{u} \|_{L_p(\tau')}^p
\]
\[
\leq N \left( \sum_{i=1}^d \| f^i \|_{L_p(\tau')}^p + \| g \|_{L_p(\tau')}^p + \lambda^{-p/2} \| \check{f}^0 \|_{L_p(\tau')}^p \right).
\] (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 \( \eta \) independent of \( \{ F_t, t \geq 0 \} \). We assume that the enlargement is not needed and define
\[
\phi_t = p \int_0^t c_s \, ds, \quad \psi = \tau \wedge \inf\{ t \geq 0 : \phi_t \geq s \}, \quad \tau' = \psi, \eta.
\]
Notice that
\[
\{ \omega : \psi > t \} = \{ \omega : \tau > t, \phi_t < s \}, \quad \tau' = \psi, \eta.
\]
Hence
\[
\{ \omega : \tau' > t \} = \{ \omega : \tau > t, \phi_t < \eta \}.
\]
It follows that \( \tau' \) is a stopping time with respect to \( F_t \sigma(\eta) \). Furthermore, for any nonnegative predictable (relative to the original filtration \( F_t \)) process \( h_t \) we have
\[
E \int_0^{\tau'} h_t \, dt = \int_0^{\infty} Eh_t E \{ I_{\tau' > t} \mid F_t \} \, dt.
\]
$$= \int_{0}^{\infty} E h_{t} I_{t} e^{-\phi_{t}} dt = E \int_{0}^{\tau} h_{t} \xi_{t}^{p} dt.$$ 

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

To proceed further take \( \bar{b}, \bar{c}, \) and \( \bar{\gamma} \) from (4.11). From Lemma 4.2 of \([12]\) and Assumption 3.3 it follows that, for \( h_{t} = b_{t}, c_{t} \), it holds that \( |D^{n}h_{t}| \leq M_{n} \), where \( M_{n} = M_{n}(n,d,\rho_{1},K) \geq 1 \) and \( D^{n}h_{t} \) is any derivative of \( h_{t} \) of order \( n \geq 1 \) with respect to \( x \). By Corollary 4.3 of \([12]\) we have \( |h_{t}(x)| \leq K(t)(1 + |x|) \), where for each \( \omega \) the function \( K(t) = K(\omega,t) \) is locally integrable with respect to \( t \) on \( \mathbb{R}_{+} \).

Owing to these properties the equation

$$x_{t} = x_{0} - \int_{t_{0}}^{t} (\bar{b}_{s} + \bar{\gamma}_{s})(x_{s}) ds, \quad t \geq t_{0}, \quad (4.10)$$

for any \((\omega \) 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 \( x_{t_{0}, x_{0}, t} \), \( t \geq t_{0} \), 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

$$\chi_{t_{0}, x_{0}, t}^{(i)}(x) = \chi^{(i)}(x - x_{t_{0}, x_{0}, t}) I_{t \geq t_{0}}.$$ 

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 \( f_{j}, g \in L_{p}(\tau) \), and \( \lambda \geq \lambda_{0} = \lambda_{0}(d,\delta,p,\rho_{0},\kappa) \), 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 \( t \leq t_{0} \wedge \tau \). Then

$$\lambda^{p/2} \|\chi_{t_{0}, x_{0}}^{(2)} u\|_{L_{p}(\tau)}^{p} + \|\chi_{t_{0}, x_{0}}^{(2)} Du\|_{L_{p}(\tau)}^{p}$$

$$\leq N\left( \sum_{i=1}^{d} \|\chi_{t_{0}, x_{0}}^{(1)} f_{i}\|_{L_{p}(\tau)}^{p} + \|\chi_{t_{0}, x_{0}}^{(1)} g\|_{L_{p}(\tau)}^{p} \right) + \lambda^{p/2} \|\chi_{t_{0}, x_{0}}^{(1)} f_{0}\|_{L_{p}(\tau)}^{p}$$

$$+ N^{\gamma_{b}/q} \|\chi_{t_{0}, x_{0}}^{(1)} Du\|_{L_{p}(\tau)}^{p} + N^{\gamma_{b}/q} \|\chi_{t_{0}, x_{0}}^{(1)} Du\|_{L_{p}(\tau)}^{p}$$

$$+ N^{\gamma_{b}/q} \|\chi_{t_{0}, x_{0}}^{(1)} u\|_{L_{p}(\tau)}^{p} + \lambda^{p/2} \sum_{i=1}^{d} \|\chi_{t_{0}, x_{0}}^{(1)} f_{i}\|_{L_{p}(\tau)}^{p},$$

where \( N \) is a constant depending only on \( d,\delta,p,\rho_{0} \), and \( \kappa \) and \( N^{*} \) depends only on the same objects, \( \gamma_{b}, \rho_{1} \), 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 4.3.
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 $f^j, g \in L_p(\tau)$, and $\lambda \geq \lambda_0 = \lambda_0(d, \delta, p, \rho_0, \kappa)$, 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 $t \leq s_0 \wedge \tau$. Then for $I_{s_0} := I(s_0, t_0)$, where $t_0 = s_0 + M_1^{-1}$, we have

$$\lambda^{p/2} \| I_{s_0} u \|^p_{L_p(\tau)} + \| I_{s_0} Du \|^p_{L_p(\tau)} \leq N \left( \sum_{i=1}^{d} \| I_{s_0} f^i \|^p_{L_p(\tau)} + \| I_{s_0} g \|^p_{L_p(\tau)} \right)$$

$$+ N\lambda^{-p/2} \| I_{s_0} f^0 \|^p_{L_p(\tau)} + N\gamma_b \| I_{s_0} Du \|^p_{L_p(\tau)}$$

$$+ N^* \lambda^{-p/2} \| I_{s_0} Du \|^p_{L_p(\tau)} + N^* \| I_{s_0} u \|^p_{L_p(\tau)} + N^* \lambda^{-p/2} \| I_{s_0} f^i \|^p_{L_p(\tau)},$$

(4.12)

where $N$ is a constant depending only on $d, \delta, p, \rho_0$, and $K$ and $N^*$ depends only on the same objects, $\gamma_b, \rho_1$, 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 $\gamma_b^{p/q} \| I_{s_0} Du \|^p_{L_p(\tau)}$ in (4.12) and choose a $\gamma_b \in (0, 1]$ in such a way that $N_0 \gamma_b^{p/q} \leq 1/2$. Then under the assumptions of Lemma 4.3 we have

$$\lambda^{p/2} \| I_{s_0} u \|^p_{L_p(\tau)} + \| I_{s_0} Du \|^p_{L_p(\tau)} \leq N \left( \sum_{i=1}^{d} \| I_{s_0} f^i \|^p_{L_p(\tau)} + \| I_{s_0} g \|^p_{L_p(\tau)} \right)$$

$$+ N\lambda^{-p/2} \| I_{s_0} f^0 \|^p_{L_p(\tau)} + N^* \lambda^{-p/2} \| I_{s_0} Du \|^p_{L_p(\tau)}$$

$$+ N^* \| I_{s_0} u \|^p_{L_p(\tau)} + N^* \lambda^{-p/2} \sum_{i=1}^{d} \| I_{s_0} f^i \|^p_{L_p(\tau)}.$$

(4.13)

To proceed further assume that

$$u_0 = 0.$$  

(4.14)

After $\gamma_b$ 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

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

(4.15)

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

$$du_t^s = (I_t u_t^s - \lambda u_t^s + D_t (\zeta_s^t f^i_t) + \zeta_s^t f^0_t + u_t (\zeta_s^t)' dt + (\Lambda^k_t u_t^s + \zeta_s^t g^k_t) dw^k_t, \ t \leq \tau.$$
We also use the fact that for $t \geq 0$, as is easy to see, $I_{x_0}(t) = \zeta^4$. Then for and $\lambda \geq \lambda_0 = \lambda_0(d, \delta, p, \rho_0, \kappa)$, where $\lambda_0(d, \delta, p, \rho_0, \kappa)$ is taken from Lemma 4.3, we obtain

$$\lambda^{p/2} \| \zeta^s u \|_{L_p(\tau)}^p + \| \zeta^s Du \|_{L_p(\tau)}^p \leq N \left( \sum_{i=1}^d \| \zeta^s f_i \|_{L_p(\tau)}^p + \| \zeta^s g \|_{L_p(\tau)}^p \right)$$

$$+ N \lambda^{-p/2} \left( \| \zeta^s f^0 \|_{L_p(\tau)}^p + \| (\zeta^s)' u \|_{L_p(\tau)}^p \right)$$

$$+ N^* \lambda^{-p/2} \| \zeta^s Du \|_{L_p(\tau)}^p + N^* \| \zeta^s u \|_{L_p(\tau)}^p + N^* \lambda^{-p/2} \sum_{i=1}^d \| \zeta^s f_i \|_{L_p(\tau)}^p. \quad (4.16)$$

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

$$\int_{-\infty}^{\infty} \left| (\zeta^s)^t \right|^p ds = \int_{-\infty}^{\infty} \left| \zeta^t \right|^p dt = N^*.$$

Then we conclude

$$\lambda^{p/2} \| u \|_{L_p(\tau)}^p + \| Du \|_{L_p(\tau)}^p \leq N_1 \left( \sum_{i=1}^d \| f_i \|_{L_p(\tau)}^p + \| g \|_{L_p(\tau)}^p \right)$$

$$+ N_1 \lambda^{-p/2} \| f^0 \|_{L_p(\tau)}^p + N_1^* \lambda^{-p/2} \| Du \|_{L_p(\tau)}^p$$

$$+ N_1^* \| u \|_{L_p(\tau)}^p + N_1^* \lambda^{-p/2} \sum_{i=1}^d \| f_i \|_{L_p(\tau)}^p.$$

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

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

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 (3.33) (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 $D_i(b^i v) + b^i D_i v$, 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 $D_i v \in L_p(\tau)$ 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 $(1, \tilde{\tau}]$, where $\tilde{\tau} = 1 + \tau$, 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 $(0, \tilde{\tau}]$ 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 $w^k_1 = 0$). Therefore, we
augment if needed our probability space in such a way that we may assume
that there are Wiener processes \( \tilde{w}_1^t, \tilde{w}_2^t, \ldots, t \geq 0 \), independent of \( \{F_s, s \geq 0\} \). Then define \( F^\tilde{w}_t \) as the completion of \( \sigma(\tilde{w}_s : s \leq t) \),

\[
\tilde{F}_t = F_0 \lor F^\tilde{w}_t, \quad t \in [0, 1], \quad \tilde{F}_t = F_{t-1} \lor F^\tilde{w}_t, \quad t \geq 1,
\]

\[
\tilde{w}_t^k = \tilde{w}_t^k, \quad t \in [0, 1], \quad \tilde{w}_t^k = \tilde{w}_1^k + \tilde{w}_{t-1}^k \quad t \geq 1, \quad \tilde{\tau} = 1 + \tau,
\]

and for \( t \geq 1 \) define the coefficients and the free terms by following the example \( \tilde{a}_{ij}^k = a_{ij}^k \).

Next, take the function \( v \) from Definition \ref{dfn:2.2} and for \( t \in [0, 1] \) set

\[
\tilde{a}_{ij}^k = \delta_{ij}, \quad \tilde{f}_t^i = -2tD_i v_{1-t}, \quad \tilde{f}_t^0 = (1 + t + \lambda t) v_{1-t},
\]

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

\[
d\hat{u}_t = \left[ D_i (\tilde{a}_{ij}^k \tilde{D}_j \hat{u}_t + \tilde{f}_t^i) + \tilde{f}_t^0 - \lambda \hat{u}_t \right] dt.
\]

Moreover, \( \hat{u}_0 = 0 \), \( \hat{u}_1 = u_0 \), and \( \hat{u}_t \) is \( \tilde{F}_t \)-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 \( \tilde{L}_t \) and \( \tilde{\Lambda}_t^k \) from the coefficients with hats, then

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

By the first part of the proof

\[
\lambda \|u\|^2_{L_p(\tau)} + \|Du\|^2_{L_p(\tau)} \leq \lambda \|\hat{u}\|^2_{L_p(\tilde{\tau})} + \|D\hat{u}\|^2_{L_p(\tilde{\tau})}
\]

\[
\leq N \left( \sum_{i=1}^d \|\tilde{f}_t^i\|^2_{L_p(\tilde{\tau})} + \|\tilde{g}_t\|^2_{L_p(\tilde{\tau})} \right) + N \lambda^{-1} \|f_t^0\|^2_{L_p(\tilde{\tau})}
\]

\[
\leq N \left( \sum_{i=1}^d \|f_t^i\|^2_{L_p(\tau)} + \|g_t\|^2_{L_p(\tau)} \right) + N \lambda^{-1} \|f_t^0\|^2_{L_p(\tau)}
\]

\[
\text{+} N(\|v\|^2_{L_p} + \|Dv\|^2_{L_p}).
\]

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

5. Proof of Theorem \ref{thm:3.3}

Throughout this section we suppose that the assumptions of Theorem \ref{thm:3.3} are satisfied.

Owing to Theorem \ref{thm:3.1} implying that the solution in \( W^1_p(\tau) \) 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

\[
L_p = L_p(\infty), \quad W^1_p = W^1_p(\infty).
\]

We need two auxiliary results.
Lemma 5.1. For any $T, R \in (0, \infty)$ (and $\omega$), we have
\[ \int_0^T \int_{B_R} (|b_s(x)|^{p'} + |b_s(x)|^{p'} + c_s'(x)) \, dx \, ds < \infty. \tag{5.1} \]

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 with cut-off coefficients. Therefore, the following result is relevant.

Lemma 5.2. Let $\phi \in C_0^\infty$, $u^m, u \in \mathbb{W}_p^1$, $m = 1, 2, \ldots$, be such that $u^m \to u$ weakly in $\mathbb{W}_p^1$. For $m = 1, 2, \ldots$ define $\chi_m(t) = (-m) \vee t \wedge m$, $b^i_{mt} = \chi_m(b^i_t)$, $b^i_{mt} = \chi_m(b^i_t)$, 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
\[ \int_0^t (b^i_{ms} D_i u^m_s, \phi) \, ds, \quad \int_0^t (b^i_{ms} u^m_s, D_i \phi) \, ds, \quad \int_0^t (c_{ms} u^m_s, \phi) \, ds \tag{5.2} \]
converge weakly in the space $\mathcal{L}_p(\{0, \tau_n\})$ as $m \to \infty$ to
\[ \int_0^t (b^i_s D_i u_s, \phi) \, ds, \quad \int_0^t (b^i_s u_s, D_i \phi) \, ds, \quad \int_0^t (c_s u_s, \phi) \, ds, \tag{5.3} \]
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
\[ \int_0^t (b^i_{ms} D_i u^m_s, \phi) \, ds \to \int_0^t (b^i_s D_i u_s, \phi) \, ds \]
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
\[ \int_0^t (D_i u^m_s, (b^i_{ms} - b^i_{ms}) \phi) \, ds \to 0 \]
weakly in $\mathcal{L}_p(\{0, \tau_n\})$. In other words, it suffices to show that for any $\xi \in \mathcal{L}_{p'}(\{0, \tau_n\})$
\[ E \int_0^{\tau_n} \xi_t \left( \int_0^t (D_i u^m_s, (b^i_s - b^i_{ms}) \phi) \, ds \right) \, dt \to 0. \]
This relation is rewritten as
\[ E \int_0^{\tau_n} (D_i u^m_s, \eta_s (b^i_{ms} - b^i_{ms}) \phi) \, ds \to 0, \tag{5.4} \]
where
\[ \eta_s := \int_s^{\tau_n} \xi_t \, dt. \]
Observe that by the choice of $\tau_n$ we have
\[
E \int_0^{\tau_n} |\eta_s| |\eta_s|^{p'} \, dx \, ds \leq E \sup_{t \leq \tau_n} |\eta_s| |\eta_s|^{p'} \int_0^{\tau_n} \int_0^{\tau_n} |x| \leq R |b_s(x)| |b_s(x)|^{p'} \, dx \, ds
\]
\[
\leq nE \left( \int_0^{\tau_n} |\xi_s| \, ds \right)^{p'} < \infty.
\]
It follows by the dominated convergence that $\eta_s(b_s^t - b_s^m) \phi \to 0$ as $m \to \infty$ strongly in $L_{p'}(\tau_n)$. By assumption $Du^m \to Du$ weakly in $L_p(\tau_n)$. 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 $b_{mt}$, $b_{mt}$, and $c_{mt}$ as in Lemma 5.2 and consider equation (2.2) with $b_{mt}$, $b_{mt}$, and $c_{mt}$ in place of $b_t$, $b_t$, and $c_t$, respectively. Obviously, $b_{mt}$, $b_{mt}$, and $c_{mt}$ satisfy Assumption 3.4 with the same $\gamma_b$ and $K$ as $b_t$, $b_t$, and $c_t$ do. By Theorem 3.1 and the method of continuity for $\lambda \geq \lambda_0(d, \delta, p, \kappa, \rho_0, \rho_1, K)$ there exists a unique solution $u^m$ of the modified equation on $\mathbb{R}$.

By Theorem 3.1 we also have
\[
\|u^m\|_{L_p} + \|Du^m\|_{L_p} \leq N,
\]
where $N$ is independent of $m$. Hence the sequence of functions $u^m$ is bounded in the space $W^1_p$ and consequently has a weak limit point $u \in W^1_p$. For simplicity of presentation we assume that the whole sequence $u^m$ converges weakly to $u$.

Take a $\phi \in C^\infty_0$. Then by Lemma 5.2 for appropriate $\tau_n$ we have that the functions (5.2) converge to (5.3) weakly in $L_p([0, \tau_n])$ 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 $L_p([0, \tau_n])$ 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 $\phi \in C^\infty_0$ equation (2.4) holds for almost any $(\omega, t)$.

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 $\psi \in C^\infty_0$ and observe that $u \psi \in W^1_2(T)$ and $g \psi \in L_2(T)$ for any $T \in (0, \infty)$ which implies that
\[
m^\psi_t := u_0 \psi + \sum_{k=1}^\infty \int_0^t \psi(A_s^ku_s + g_s^k) \, dw_s^k
\]
is well defined as an $L_2$-valued continuous martingale such that for any
\[ (m^\psi_t, \phi) = (u_0\psi, \phi) + \sum_{k=1}^\infty \int_0^t (\psi(\Lambda^k_s u_s + g^k_s), \phi) \, dw^k_s \] (5.5)
for all $t \in \mathbb{R}_+$.

Notice that for any $\phi \in C_0^\infty$
\[ (u_t\psi, \phi) = \int_0^t (u^*_s\psi, \phi) \, ds + (m^\psi_t, \phi) \] (5.6)
for almost all $(\omega, t)$, where $u^*_s$ is a function with values in the space of distributions on $\mathbb{R}^d$ defined by
\[ u^*_s = L_s u_s - \lambda u_s + D_s f^s_s + f^0_s \]
(see Remark 4.2).

Next, take an $R \in (0, \infty)$ and let $W^{-1}_{\rho'}(B_R)$ denote the dual space for
\[ W^1_\rho(B_R) := W^1_\rho(B_R) \cap \{v : v|_{\partial B_R} = 0\}. \]
Estimate (4.6) combined with the facts that, $p' < p$ and that one can cover $B_R$ with finitely many balls of radius $\rho_1$ shows that for any $\phi \in C_0^\infty(B_R)$
\[ |(D_t(b^\psi_s u_s), \phi)| \leq N(1 + \int_{B_{\rho_1}} |b_s| \, dx)\|u_s\|_{W^1_\rho} \|\phi\|_{W^1_\rho}, \]
where $N$ is independent of $\omega, s, u_s, \phi$. Due to the arbitrariness of $\phi$ and the fact that $C_0^\infty(B_R)$ is dense in $W^1_\rho(B_R)$ we conclude that (for almost all $(\omega, s)$) we have $D_t(b^\psi_s u_s) \in W^{-1}_{\rho'}(B_R)$ and
\[ \|D_t(b^\psi_s u_s)\|_{W^{-1}_{\rho'}(B_R)} \leq N(1 + \int_{B_{\rho_1}} |b_s| \, dx)\|u_s\|_{W^1_\rho}. \]
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 $u \in W^1_\rho$. Similar statements are true for $b^\psi_s D_t u_s$, $c_s u_s$, and $u_s^*$.

Now, since $u_\psi \in \mathcal{L}_p(\mathbb{R}_+, W^1_\rho(B_R))$ and $W^1_\rho(B_R)$ is dense in $\mathcal{L}_2(B_R)$, by Theorem 3.1 of [19] we get that there exist an event $\Omega^\psi$ of full probability and a continuous $\mathcal{L}_2(B_R)$-valued $\mathcal{F}_t$-adapted process $u^\psi_t$ such that $u^\psi_t = u_t\psi$ as $\mathcal{L}_2(B_R)$-valued functions for almost all $(\omega, t)$ and for any $\omega \in \Omega^\psi$, $t \in \mathbb{R}_+$, and $\phi \in C_0^\infty(B_R)$ we have
\[ (u^\psi_t, \phi) = \int_0^t (u^*_s\psi, \phi) \, ds + (m^\psi_t, \phi). \] (5.7)
Take a \( \psi \in C_0^\infty \) such that \( \psi(x) = 1 \) for \( |x| \leq 1 \) 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 \( \Omega' \) in the following way. Obviously (see (5.5)), if \( \psi', \psi'' \in C_0^\infty \) and \( \psi' = \psi'' \) on \( B_R \) and \( \phi \in L_2 \) 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 \( \Phi \) be the union over \( n = 1, 2, \ldots \) of countable subsets of \( C_0^\infty(B_n) \) each of which everywhere dense in \( L_2(B_n) \). For \( \phi \in C_0^\infty \) denote \( d(\phi) \) the smallest radius of the balls centered at the origin containing the support of \( \phi \). Then by the above for \( \phi \in C_0^\infty \) the events
\[
\Omega(\phi) = \{ \omega \in \Omega : (m_t^\psi^i, \phi) = (m_t^\psi^j, \phi), \quad \forall t \in \mathbb{R}_+, k, j \geq d(\phi) \},
\]
\[
\Omega'' = \Omega' \bigcap \bigcap_{\phi \in \Phi} \Omega(\phi)
\]
have probability one. Since \( m_t^\psi \) are \( L_2 \)-valued and \( \Phi \cap C_0^\infty(B_n) \) is dense in \( L_2(B_n) \), we have that for \( \omega \in \Omega'', t \in \mathbb{R}_+, \) and any \( \phi \in C_0^\infty(B_n) \) it holds that
\[
(m_t^\psi^i, \phi) = (m_t^\psi^j, \phi)
\]
as long as \( i, j \geq n \).

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 \( j, k \geq n \). In particular, for any \( \omega \in \Omega'', t \in \mathbb{R}_+, n = 1, 2, \ldots \) it holds that \( u_t^\psi = u_t^\psi^k \) as distributions on \( B_n \) for \( j, k \geq n \) 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 \( \omega, t \), we have that \( \bar{u}_t = u_t \) (as distributions on \( \mathbb{R}^d \)) for almost all \( (\omega, t) \). 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 \( j \geq n \)
\[
(\bar{u}_t, \phi) = (u_t^\psi^j, \phi) = \int_0^t \left( L_su_s - \lambda u_s + D_if_s^i + f_s^0, \phi \right) ds + (m_t^\psi^j, \phi).
\]

By having in mind (5.5) we conclude that for any \( \phi \in C_0^\infty \) with probability one for all \( t \in \mathbb{R}_+ \)
\[
(\bar{u}_t, \phi) = (u_0, \phi) + \int_0^t \left( L_su_s - \lambda u_s + D_if_s^i + f_s^0, \phi \right) ds
\]
\[
+ \sum_{k=1}^{\infty} \int_0^t \left( \Lambda_k u_s + g_k^s, \phi \right) du_s^k.
\]

Now it only remains to observe that since \( \bar{u}_s = u_s \) for almost all \( (\omega, s) \), 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 $\mathcal{W}^1_{2,loc}(\tau)$

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 $\mathcal{W}^1_{2}(\tau)$ are introduced in Definition 2.1.

**Theorem 6.1.** Let $\tau$ be a stopping time and let $u, \tilde{u}, f^i, \tilde{f}^i, g = (g^1, g^2, \ldots), \tilde{g} = (\tilde{g}^1, \tilde{g}^2, \ldots)$ be some functions such that for any $\phi \in C_0^\infty$ we have $\phi u, \phi \tilde{u} \in \mathcal{W}^1_{2}(\tau)$, $\phi f^j, \phi \tilde{f}^j \in L_2(\tau)$, $j = 0, \ldots, d$, and $\phi g, \phi \tilde{g} \in L_2(\tau)$. Assume that in the sense of generalized functions

$$du_t = (D_i f^i_t + f^0_t) \, dt + g^k_t \, dw^k_t, \quad d\tilde{u}_t = (D_i \tilde{f}^i_t + \tilde{f}^0_t) \, dt + \tilde{g}^k_t \, dw^k_t, \quad t \leq \tau.$$  

Then

$$d(u_t \tilde{u}_t) = \left[ \tilde{u}_t(D_i f^i_t + f^0_t) + u_t(D_i \tilde{f}^i_t + \tilde{f}^0_t) + h_t \right] \, dt$$

$$+ (\tilde{u}_t g^k_t + u_t \tilde{g}^k_t) \, dw^k_t, \quad t \leq \tau,$$

where $h_t := (g_t, \tilde{g}_t)_{l_2}$, in the sense of generalized functions, that is, for any $\phi \in C_0^\infty$, with probability one,

$$(u_t \tilde{u}_t, \phi) = (u_0 \tilde{u}_0, \phi) + \int_0^t I_{s \leq \tau} (\tilde{u}_s g^k_s + u_s \tilde{g}^k_s, \phi) \, dw^k_s$$

$$+ \int_0^t I_{s \leq \tau} \left[ (\tilde{u}_s f^0_s, \phi) - (f^i_s, D_i (\tilde{u}_s \phi)) \right] ds$$

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 $\lambda$ is an arbitrary constant, and comparing the coefficients of $\lambda$, 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

$$(u^2_t, \phi) = (u_0^2, \phi) + 2 \int_0^t I_{s \leq \tau} (u_s g^k_s, \phi) \, dw^k_s$$

$$+ \int_0^t I_{s \leq \tau} \left[ 2(u_s f^0_s, \phi) - 2(f^i_s, D_i (u_s \phi)) \right] ds$$

for all $t$.

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

$$(\psi u_t, \phi) = (\psi u_0, \phi) + \int_0^t I_{t \leq \tau} (\psi g^k_s, \phi) \, dw^k_s$$

$$+ \int_0^t I_{t \leq \tau} \left[ (\psi f^0_s - f^i_s D_i \psi, \phi) - (\psi f^i_s, D_i \phi) \right] dt.$$

for all $t$. This means that

$$d(\psi u_t) = (\psi f^0_t - f^0_t D_i \psi + D_i (\psi f^i_t)) \, dt + \psi g^k_t \, dw^k_t, \quad t \leq \tau.$$
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) $\psi u_{t,\tau} I_{\Omega'}$ is a continuous $L_2$-valued $\mathcal{F}_\tau$-adapted function on $[0, \infty)$;  

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

$$
\int_{\mathbb{R}^d} |\psi u_{t,\tau}|^2 \, dx = \int_{\mathbb{R}^d} |\psi u_0|^2 \, dx + 2 \int_0^t I_{s \leq \tau} \int_{\mathbb{R}^d} \psi^2 u_s g_s^k \, dx \, dw_s^k 
+ \int_0^t I_{s \leq \tau} \left( \int_{\mathbb{R}^d} \left[ 2 u_s f_s^0 \psi^2 - 2 f_s^i D_i (\psi^2 u_s) + \psi^2 |g_s|_2^2 \right] \, dx \right) ds. \quad (6.3)
$$

This proves (6.2) if we replace there $\phi$ with $\psi^2$. 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 $|x| \geq R$ 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  

$$
x_t = \int_0^t \sigma_i^k \, dw_s^k + \int_0^t b_i \, ds
$$

for some predictable $\mathbb{R}^d$-valued functions $\sigma_i^k$ and $b_i$ such that  

$$
E \int_0^\tau \left( \sum_k |\sigma_i^k|^2 + |b_i| \right) \, dt < \infty.
$$

Then in the sense of generalized functions  

$$
d(x_t \psi_t) = \left[ D_i (u_t a_i^{ij} D_j \psi_t) - a_i^{ij} (D_j u_t) D_j \psi_t + u_t b_i^j D_i \psi_t + D_i (\psi_t f_i^j) \right] 
- f_i^j D_i \psi_t + \psi_t f_i^0 + g_k^i \sigma_i^{ik} D_i \psi_t \right) \, dt + \left[ \psi_t g_i^k + u_t \sigma_i^{ik} D_i \psi_t \right] \, dw_s^k, \quad t \leq \tau,
$$

where $\psi_t(x) = \psi(x + x_t)$ and $2a_i^{ij} = \sigma_i^k \sigma_i^{jk}$.  

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

$$
\int_{\mathbb{R}^d} \psi_t \phi \, dx = \int_{\mathbb{R}^d} \psi(x) \phi(x - x_t \tau) \, dx = \int_{\mathbb{R}^d} \psi \phi \, dx
+ \int_0^t I_{s \leq \tau} \int_{\mathbb{R}^d} \psi_s \left[ a_s^{ij} D_i \phi - b_s^j D_i \phi \right] \, dx \, ds
+ \int_0^t I_{s \leq \tau} \int_{\mathbb{R}^d} \sigma_s^{ik} \phi D_i \psi_s \, dx \, dw_s^k,
$$

where the coefficient of $ds$ equals  

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

Furthermore, for instance,  

$$
E \int_0^\tau \int_{\mathbb{R}^d} \sum_k |\sigma_s^{ik} D_i \psi_s|^2 \, dx \, ds \leq E \int_0^\tau \int_{\mathbb{R}^d} \sum_k |\sigma_s^k|^2 |D \psi_s|^2 \, dx \, ds
$$
\[
= \int_{\mathbb{R}^d} |D\psi|^2 \, dx \int_0^T \sum_k |\sigma^k_s|^2 \, ds < \infty.
\]
It follows that \( \psi \in W_2^1(\tau) \) and
\[
d\psi_t = \left[ D_i(a^i_j D_j \psi_t) + b^i_j D_i \psi_t \right] dt + \sigma^i_k D_i \psi_t dw^k_t
\]
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
\[
\begin{align*}
dx_t &= b(t, z_t) dt + \theta(t, y_t) dw_t, \\
dy_t &= B(t, z_t) dt + \Theta(t, y_t) dw_t
\end{align*}
\]
with some initial data.

Assumption 7.1. The functions \( b, \theta, B \) and \( \Theta \) are Borel measurable functions of \((t, z)\) and \((t, y)\) as appropriate and \( \theta \) and \( \Theta \) are bounded and satisfy the Lipschitz condition with respect to \( y \) with a constant independent of \( t \). We have
\[
b(t, z) = x^* \dot{b}(t, y) + b(t, 0, y), \quad B(t, z) = x^* \dot{B}(t, y) + B(t, 0, y),
\]
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
\[
D_i = \frac{\partial}{\partial x^i}, \quad D_{ij} = D_i D_j
\]
only for \( i, j = 1, \ldots, d \).

Remark 7.1. Note that
\[
\dot{b}^{ij}(t, y) = D_i \dot{b}^j(t, z), \quad \dot{B}^{ij}(t, y) = D_i B^j(t, z).
\]

Set
\[
\dot{\theta}(t, y) = \begin{pmatrix} \theta(t, y) \\ \Theta(t, y) \end{pmatrix}, \quad \dot{a}(t, y) = \frac{1}{2} \partial \dot{\theta}^*(t, y), \quad \ddot{b}(t, z) = \begin{pmatrix} b(t, z) \\ B(t, z) \end{pmatrix},
\]
\[
\dddot{\mathcal{L}}(t, z) = \dot{a}^{ij}(t, y) \frac{\partial^2}{\partial z^i \partial z^j} + \ddot{b}^i(t, z) \frac{\partial}{\partial z^i},
\]
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 \( i, j = 1, \ldots, d_1 \).
Observe that
\[ dz_t = \tilde{\theta}(t, z_t) dw_t + \tilde{b}(t, z_t) dt. \tag{7.5} \]

**Assumption 7.2.** The symmetric matrix \( \tilde{a}(t, y) \) 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
\[ a = \frac{1}{2} \theta \theta^*, \quad \alpha = \frac{1}{2} \sigma \sigma^*, \quad \sigma = \theta \Theta^* \Psi, \]

**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 \( b \))
\[ B = \Psi B, \quad B_t(x) = \Psi_t B_t(x) = \Psi(t, y_t) B(t, x, y_t) \]
and set
\[ L_t(x) = a_t^{ij} D_i D_j + b_t^i(x) D_i, \tag{7.7} \]
\[ L_t^*(x) u_t(x) = e_t^{ij} D_i u_t(x) - D_t^i(b_t^i(x) u_t(x)) = D_j(a_t^{ij} D_i u_t(x) - b_t^i(x) u_t(x)), \tag{7.8} \]
\[ \Lambda_t^k(x) u_t(x) = \sigma_t^{ik} D_i u_t(x) + b_t^k(x) u_t(x), \tag{7.9} \]
\[ \Lambda_t^{k*}(x) u_t(x) = -\sigma_t^{ik} D_i u_t(x) + b_t^k(x) u_t(x), \tag{7.10} \]
where \( t \in [0, T], x \in \mathbb{R}^d, k = 1, \ldots, d_1 - d, \) 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 \( i, j = 1, \ldots, d \) (whereas in (7.4) for \( i, j = 1, \ldots, d_1 \)).

Finally, by \( \mathcal{F}_t^y \) we denote the completion of \( \sigma \{ y_s : s \leq t \} \) 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 \( \omega \), quadratic in \( x \), and
(i) For all \( x \in \mathbb{R}^d \) (and \( \omega \))
\[ \varepsilon^{-1} |x|^2 \geq x^i x^j D_{ij} Q \geq \varepsilon |x|^2; \]
(ii) We have \( \pi_0 e^Q \in \text{tr} \mathcal{W}_p^1 \), where \( \pi_0 \) 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 $\pi_0$ is a Gaussian density.

**Theorem 7.1.** There exists a process $\tilde{\pi}$ on $[0,T]$ such that

(i) $\tilde{\pi}_t$ is $\mathcal{F}^y_t$-adapted and, for any $r \in [1,p]$, with probability one $\tilde{\pi}_t$ is a continuous $L_r$-valued process on $[0,T]$ and $\tilde{\pi}_0 = \pi_0$;

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

(iii) In the sense of Definition 2.3

$$d\tilde{\pi}_t = \Lambda^k_t \tilde{\pi}_t dy^k_t + L^k_t \tilde{\pi}_t \, dt, \quad t \leq \tau_m, \quad (7.11)$$

where

$$\tilde{y}^k_t = \int_0^t \Psi^{kr}_s \, dy^k_s.$$ 

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

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

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

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 $\mathbb{R}^d$

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

**Remark 7.4.** Equation (7.13) shows (by definition) that

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

is a conditional density of distribution of $x_t$ given $y_s, s \leq t$. Since, generally, $(\tilde{\pi}_t,1) \neq 1$, one calls $\tilde{\pi}_t$ 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.** (ii) For any $\phi \in C_0^\infty$ the process $\tilde{\pi}_t \phi$ is continuous on $[0,T]$ as an $L_p$-valued process (a.s.);

(ii) If $p > 2$ and we have two numbers $\alpha$ and $\beta$ such that

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

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

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

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 \( \tilde{\pi}_t \phi \in C^{0-\varepsilon}(\mathbb{R}^d) \) and the norm of \( \tilde{\pi}_t \phi \) in this space is bounded as a function of \( t \);

(b) for any \( \varepsilon \) as in (a) (a.s.) for any \( x \in \mathbb{R}^d \) we have \( \tilde{\pi}_t(x)\phi(x) \in C^{(0-\varepsilon)/2}(x) \) and the norm of \( \tilde{\pi}_t(x)\phi(x) \) 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 most unpleasant feature of (7.11) in our particular case is the presence of \( \tilde{\pi}_t(x) \) in the stochastic term with \( \tilde{\pi}_t(x) \) 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 \( \tilde{b} \) is affine with respect to \( z \) and \( \tilde{\theta} \) 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 \( \tilde{\pi}_t \) 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

\[
\hat{b}_t = \hat{B}_t \hat{\Psi}_t. \tag{7.14}
\]

**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 \( U_t \)

\[
\frac{d}{dt} W_t = (\hat{b}_t \sigma^*_t - \hat{\theta}_t) W_t + W_t^* (\sigma_t B^*_t - \hat{b}^*_t) - 2W_t^* \hat{a}_t W_t + \hat{b}_t \hat{B}^*_t, \tag{7.15}
\]

\[
dV_t = -(W_t \sigma_t + \hat{b}_t) d\tilde{y}_t + [(\hat{b}_t \sigma^*_t - \hat{\theta}_t)V_t - 2W_t \hat{a}_t V_t + W_t (\sigma_t B_t(0) - b_t(0)) + \hat{b}_t B_t(0)] dt, \tag{7.16}
\]

\[
dU_t = -(V_t^* \sigma_t + B^*_t(0)) d\tilde{y}_t + [a^*_t W^*_t + \hat{V}_t^* (\sigma_t B_t(0) - b_t(0)) - V_t^* \hat{a}_t V_t + \frac{1}{2} |B_t(0)|^2 + tr \hat{b}_t] dt. \tag{7.17}
\]
has a unique $\mathcal{F}_t^y$-adapted solution with initial conditions $W_{00}^{ij} = D_{ij}Q$, $V_0^i = D_iQ(0)$, $U_0 = Q(0)$. 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 $\omega$ and $t$ (depending on $T$ among other things).

Remark 7.5. Set

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

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

$$dQ_t(x) = -(\sigma_t^{ik}D_iQ_t(x) + b_t^k(x))d\tilde{y}_t^k + \left[a_t^{ij}D_{ij}Q_t(x) + D_i b_t^i\right.\
\left.+(\sigma_t^{ik}b_t^k(x) - b_t^i(x))D_iQ_t(x) - \tilde{a}_t^{ij}(D_iQ_t(x))D_jQ_t(x) + \frac{1}{2}|b_t(x)|^2\right] dt \quad (7.19)$$

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

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

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 $D_{ij}Q_t$, $D_iQ_t(0)$, 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 $\pi_t(x) = C_te^{-Q_t(x)}$, where $c_t$ is a normalizing process obtained from the condition that

$$C_t \int_{\mathbb{R}^d} e^{-Q_t(x)} dx = 1. \quad \text{This theorem is proved in Section 9}$$

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

$$Q_t(x) = \frac{1}{2}|W_t^{1/2}x + W_t^{-1/2}V_t|^2 + \int_0^t (V_s^*W_s^{-1}\dot{b}_s - b_s^*0) d\tilde{y}_s\
+ \frac{1}{2} \int_0^t |\dot{b}_s^*W_s^{-1}V_s - b_s^*0|^2 ds + A_t, \quad (7.21)$$

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

$$A_t := \int_0^t [a^{ij}W_t^{ij} + tr \dot{b}_s - \frac{1}{2}||W_t^{1/2}\sigma_s + W_t^{-1/2}\dot{b}_s||^2] ds,$$

where for a matrix $u$ we use the notation $||u||^2 = tr uu^*$. This shows that in the situation of Theorem 7.3

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

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

8. An auxiliary function

The assumptions from Section 7 are supposed to hold. Set

\[ \hat{b}_t(x) = \sigma_t^{ik} b_t^k(x) - 2\dot{a}_t^{ij} D_j Q_t(x). \]

**Theorem 8.1.** The equation

\[ d\hat{\pi}_t = -\sigma_t^{ik} D_i \hat{\pi}_t d\hat{\pi}_t^k + \left[ a_t^{ij} D_j \hat{\pi}_t - \hat{b}_t^i D_i \hat{\pi}_t + \dot{\hat{b}}_t^i D_i \hat{\pi}_t \right] dt, \quad t \leq T, \quad (8.1) \]

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 5.3 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

\[ G_t := \exp \int_0^t (D_i \hat{b}_s^i - D_i \hat{\pi}_s) ds = \exp \int_0^t \tr (\sigma_s \hat{b}_s^* - \hat{\pi}_s) ds. \]

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

Proof. Set

\[ \xi^*_t = \int_0^t \sigma_s^{ik} d\hat{\pi}_s^k, \quad \xi_t = (\xi_t^i), \quad \tau_m = T \land \inf \{ t \geq 0 : |z_t| + |\xi_t| \geq m \}. \]

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

\[ |\sigma_t B_t(x)| + |b_t(x)| + |\hat{b}_t(x)| \leq N(1 + |x|), \]

where the constant \( N \) is independent of \( \omega, t, x \). Why we also stop \( \xi_t \) will become clear later.

Observe that for any \( \psi \in C^\infty_0 \) the process \( \psi \hat{\pi}_t \) satisfies an equation obtained by multiplying through (8.1) by \( \psi \). Then after writing \( \psi D_t (a_t^{ij} D_j \hat{\pi}_t) \) as \( D_t (\psi a_t^{ij} D_j \hat{\pi}_t) = a_t^{ij} (D_j \hat{\pi}_t) D_t \psi \) and noting that the other coefficients multiplied by \( \psi \) are bounded functions on \([0, \tau_m]\) \times \mathbb{R}^d \) we see that

\[ \psi \hat{\pi}_t \in W^1_{\mathcal{L}_p}(\tau_m) \quad (8.2) \]

for any \( m \). It follows from (13) that with probability one \( \psi \hat{\pi}_t \land \tau_m \) is a continuous \( \mathcal{L}_p \)-valued process and since, for each \( \omega, \tau_m = T \) 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 \( \psi \in C^\infty_0 \).
Then take a nonnegative radially symmetric and radially decreasing function \( \phi \in C_0^\infty \) such that \( |D \phi| \leq 1 \), introduce \( \phi^n(x) = \phi(x/n) \), \( n = 1, 2, \ldots \), and use Corollary 6.2 with \( \tau_m \) in place of \( \tau \) (recall (8.2)). Then we find

\[
d(\hat{\pi}_t \phi^n_t) = -\sigma_{t}^{ik} D_i(\hat{\pi}_t \phi^n_t) \, dy_t^{k} + \phi^n_t (\hat{b}_t^{i} - b_t^{i}) D_i \hat{\pi}_t \, dt + [D_i(\phi^n_t a_t^{ij} D_j \hat{\pi}_t) - (\hat{a}_t^{ij} + a_t^{ij}) (D_i \phi^n_t) D_j \hat{\pi}_t + D_i(\pi_t \hat{a}_t^{ij} D_j \phi^n_t)] \, dt, \quad t \leq \tau_m.
\]

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

\[
\|\phi^n_t \hat{\pi}_t \|_{\mathcal{L}_p} = \|\phi^n_0 \|_{\mathcal{L}_p} + I_{t}^{1n} + I_{t}^{2n} + I_{t}^{3n}
\]

for all \( t \), where

\[
I_{t}^{1n} = -p(p - 1) \int_0^{t \wedge \tau_m} \int_{\mathbb{R}^d} a^{ij}_s \, |\phi^n_s | p \pi^n_s | p - 2 (D_i \pi^n_s) D_j \pi^n_s \, dx \, ds \leq 0,
\]

\[
I_{t}^{2n} = -\int_0^{t \wedge \tau_m} \left[ D_i \hat{b}_s^{i} - D_i b_s^{i} \right] \int_{\mathbb{R}^d} |\phi^n_s | p \pi^n_s | p \, dx \, ds,
\]

\[
I_{t}^{3n} = \int_0^{t \wedge \tau_m} \int_{\mathbb{R}^d} |\pi^n_s | p \pi^n_s \, dx \, ds,
\]

\[
\psi^n_s = pa_s^{ij} D_i |\phi^n_s | p + (p - 1) (p - 2) |\phi^n_s | p - 2 a_s^{ij} (D_i \phi^n_s) D_j \phi^n_s
\]

\[
+ (2 - p) |\phi^n_s | p - 1 a_s^{ij} D_i \phi^n_s \hat{\pi}_s (\pi^n_s - 1) + (\hat{b}_s^{i} - b_s^{i}) D_i |\phi^n_s | p,
\]

where for simplicity of notation the argument \( x \) is dropped.

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

\[
|x| |D \phi^n_s (x)| \leq \frac{|x|}{n} ((D \phi)(x - \xi_s/n) \leq \frac{\xi_s}{n} + \frac{|x - \xi_s|}{n} ((D \phi)(x - \xi_s/n)
\]

\[
\leq m + \sup_y |x| |D \phi(y)|.
\]

By adding that \( \hat{\pi} \in \mathcal{W}_p^1(T) \), 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

\[
\|\phi^n_t \hat{\pi}_t \|_{\mathcal{L}_p} \to \|\hat{\pi}_t \hat{\pi}_t \|_{\mathcal{L}_p}
\]

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

\[
\|\hat{\pi}_t \hat{\pi}_t \|_{\mathcal{L}_p} = \|\hat{\pi}_0 \|_{\mathcal{L}_p} - \int_0^{t \wedge \tau_m} (D_i \hat{b}_s^{i} - D_i b_s^{i}) \|\hat{\pi}_s \|_{\mathcal{L}_p} \, ds
\]

\[
- p(p - 1) \int_0^{t \wedge \tau_m} a^{ij}_s \int_{\mathbb{R}^d} |\pi^n_s | p - 2 (D_i \pi^n_s) D_j \pi^n_s \, dx \, ds.
\]
Obviously one can drop $\tau_m$ in this formula and then obtain that (a.s.) for all $t \leq T$

$$G_t\|\hat{\pi}_t\|_{L_p}^p = \|\hat{\pi}_0\|_{L_p}^p - p(p-1)\int_0^t G_s a^2_s \int_{\mathbb{R}^d} |\pi_s|^{p-2}(D_i\pi_s)D_j\pi_s dx ds,$$

which implies that $G_t\|\hat{\pi}_t\|_{L_p}^p$ is decreasing and continuous (a.s.). Furthermore, since $\phi^p_t \hat{\pi}_t$ are continuous $L_p$-valued processes, $\hat{\pi}_t$ is at least a weakly continuous $L_p$-valued function, but since $\|\hat{\pi}_t\|_{L_p}^p$ 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 $L_p$-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 $\tau$ 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 $\tau \leq t \leq T$.

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

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

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

Lemma 8.3. Introduce

$$\hat{w}_t = \int_0^t \Psi_s \Theta_s dw_s, \quad \hat{b}_t = b_t(x_t) = \Psi(t,y_t)B(t,x_t,y_t).$$

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

$$\rho_t = \rho_t(\hat{w}, dw) = \exp(-\int_0^t \hat{b}_s^k dw_s^k - \frac{1}{2} \int_0^t |\hat{b}_s|_2^2 ds)$$

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

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

$$\int_0^t \hat{b}_s^k dw_s^k = \int_0^t \hat{b}_s^* \Psi_s dw_s.$$

Furthermore, the system

$$dx_t = (b(t, z_t) - \theta(t, y_t)\Theta^*(t, y_t)\Psi^2(t, y_t)B(t, z_t) )\, dt + \theta(t, y_t)\, dw_t,$$

$$dy_t = \Theta(t, y_t)\, dw_t,$$

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-Shiryaevas theorem $\rho$ 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 \( \mathcal{F}^y_t \)-adapted.

Proof. Observe that in equation (8.1) we have
\[
d\hat{y}^k_t = \Psi_t^{kr} d\hat{y}^r_t = d\hat{w}^k_t + \hat{B}_t^k dt,
\]
where, as it is pointed out above, \( \hat{w}_t \) is a Wiener process. Furthermore, the processes \( \hat{y}^k_t \) are \( \mathcal{F}^y_t \)-adapted since such are \( \Psi_t^{kr} \) and equation (8.1) is rewritten as
\[
d\hat{\pi}_t = -\sigma_t^{ik} D_t \hat{\pi}_t \, d\hat{w}^k_t + \left[ D_t (a_t^{ij} D_j \hat{\pi}_t) - b_t^i D_t \hat{\pi}_t \right] dt,
\quad t \leq T.
\] (8.5)

Here \( \sigma_t^{ik} \hat{B}_t^k \) is independent of \( x \) and for each \( \omega \) the trajectories of \( \sigma_t^{ik} \hat{B}_t^k \) 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 \( \hat{\pi}_t := e^{-Q_t \hat{\pi}_t} \).

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 \( \tilde{\pi} \in \mathbb{W}^1_p (T) \) and
\[
\int_0^t \| \tilde{\pi}_s \|_{\mathbb{W}^1_p}^p ds
\]
is an \( \mathcal{F}^y_t \)-adapted continuous process on \([0, T]\). Then after introducing
\[
\tau'_{m} = T \wedge \inf\{ t \geq 0 : \int_0^t \| \tilde{\pi}_s \|_{\mathbb{W}^1_p}^p ds \geq m \}
\]
we get that \( \tilde{\pi} \in \mathbb{W}^1_p (\tau'_{m}) \) and \( \tau'_{m} = T \) for all large \( m \) (a.s.).

We now prove that \( \tilde{\pi} \) satisfied (7.11) define \( \Phi_t = \Psi_t^{-1} \) and observe that
\[
(d \hat{y}^k_t) d\hat{u}^r_t = \delta^{kr} dt, \quad d\hat{w}^k_t + \hat{B}_t^k dt,
\]
(\( \hat{B}_t = B(t, z_t) \)). 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 \( m = 1, 2, \ldots \) set
\[
\tau''_{m} = T \wedge \inf\{ t \geq 0 : |z_t| + |DQ_t(0)| \geq m \}.
\] (8.6)

Note that for a constant \( N_0 \) independent of \( m \) for \( t \leq \tau''_{m} \) we have
\[
|b_t(x)| + |\hat{b}_t(x)| \leq N_0(1 + |x| + m), \quad |\hat{\eta}_t| + |\hat{B}_t| \leq N_0(1 + 2m).
\]
Furthermore, \( D_{ij} Q_t(x) = x^j D_{ij} Q_t + D_{ij} Q_t(0) \), so that increasing \( N_0 \) if needed we may assume that for \( t \leq \tau''_{m} \)
\[
|DQ_t(x)| \leq N_0(1 + |x| + m).
\]
Then as is easy to see (cf. (8.2)) \( u_t := \tilde{\pi}_t \) and \( \tilde{u}_t := \eta_t \) satisfy the condition of Theorem 6.1 with appropriate \( f, \tilde{f}, g, \tilde{g} \) and \( \tau''_m \) in place of \( \tau \).

By Theorem 6.1 in the sense of generalized functions
\[
d(\eta_t \tilde{\pi}_t) = \nu_t^0 d\tilde{y}_t + J_t dt, \quad t \leq \tau''_m,
\]
where
\[
\nu_t^0 = \tilde{\pi}_t \Lambda^{\nu*}_t \eta_t - \eta_t \sigma^{ij}_t D_i \tilde{\pi}_t = \Lambda^{\nu*}_t (\eta_t \tilde{\pi}_t),
\]
\[
J_t = -(\eta_t B^k_t - \sigma^{ij}_t D_i \eta_t) \sigma^{jk}_t D_j \tilde{\pi}_t + \tilde{\pi}_t L^*_t \eta_t + \eta_t \left[ a^{ij}_t D_i \tilde{\pi}_t - b^i_t D_i \tilde{\pi}_t + (\sigma^{ik}_t B^k_t + 2\eta_t^{-1} a^{ij}_t D_j \eta_t) D_i \tilde{\pi}_t \right]
\]
\[
= \tilde{\pi}_t L^*_t \eta_t + \eta_t (a^{ij}_t D_i \tilde{\pi}_t - b^i_t D_i \tilde{\pi}_t) + 2a^{ij}_t (D_i \tilde{\pi}_t) D_j \eta_t = L^*_t (\eta_t \tilde{\pi}_t).
\]
In other words (see Theorem 6.1) for any \( \phi \in \mathcal{C}_0^\infty \) with probability one
\[
(\tilde{\pi}_t \wedge \tau''_m, \phi) = (\tilde{\pi}_0, \phi) + \int_0^t I_s \leq \tau''_m (\tilde{\pi}_s, \Lambda^k_s \phi) d\tilde{y}_s + \int_0^t I_s \leq \tau''_m (\tilde{\pi}_s, L_s \phi) ds
\]
for all \( t \geq 0 \). Obviously, one can take here \( \tau_m := \tau''_m \wedge \tau''_m \) in place of \( \tau''_m \) and then after recalling that \( \tilde{\pi} \in \mathbb{W}^1_\mu(\tau''_m) \) one concludes that \( \tilde{\pi} \) 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.

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
\[
\rho_t(\tilde{b}, d\tilde{w}) \exp \left( - \int_0^t (V^*_s W^{-1}_s B_s - B^*_s(0)) d\tilde{y}_s - \frac{1}{2} \int_0^t |\tilde{b}^*_s W^{-1}_s V_s - B_s(0)|^2 ds \right)
\]
\[
= \rho_t(\tilde{b} - B(0) + B^* W^{-1} V, d\tilde{w}). \tag{8.7}
\]
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 (7.15), (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
\[
A(t, z, W, V) = \Theta^*(t, y) \Psi^2(t, y)(B(t, z) - B(t, 0, y) + \tilde{B}(t, y))F^{-1}(W)V.
\]
After changing the probability measure formally we arrive at the system consisting of (7.15) with \( \sigma_t = \sigma(t, y_t), \tilde{a}_t = \tilde{a}(t, y_t) \), and with \( F(W_t) \) in place of \( W_t \) on the right and the following two equations
\[
dz_t = \tilde{b}(t, y_t) dw_t + [\tilde{b}(t, z_t) - \tilde{b}(t, y_t) A(t, z_t, W_t, V_t)] dt,
\]
\[
dV_t = -(F(W_t)\sigma(t, y_t) + \tilde{B}(t, y_t)\Psi(t, y_t))\Psi(t, y_t)\Theta(t, y_t) dw_t + (F(W_t)\sigma(t, y_t) + \tilde{B}(t, y_t)\Psi(t, y_t))\Psi(t, y_t)\Theta(t, y_t) A(t, z_t, W_t, V_t) dt
\]
\[
- (F(W_t)\sigma(t, y_t) + \tilde{B}(t, y_t)\Psi(t, y_t))\Psi(t, y_t)B(t, z_t) dt.
\]
is bounded on \([0, t]\) interval of time where \(W\) is at least on a small time interval up and it is of ODEs allows us to conclude that a unique solution exists until it blows where the last trace is nonnegative again on the interval of existence of \(W\) and uniqueness of its solution presents no difficulty if \(W\) is known. Equation (7.16) is linear with respect to \(V\) 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.

**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 \(\omega\) separately. Then the theory of ODEs allows us to conclude that a unique solution exists until it blows up and it is \(\mathcal{F}_t^\omega\)-adapted. Uniqueness implies that \(W_t = W_t^\omega\). 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 \(t < t_0\)

\[
\frac{d}{dt} \det W_t = \text{tr} \dot{W}_t W_t^{-1} \det W_t,
\]

and

\[
\text{tr} \dot{W}_t W_t^{-1} = 2 \text{tr} (\dot{a}_t \sigma_t^* - \dot{b}_t) - 2 \text{tr} a_t W_t + \text{tr} \dot{a}_t \dot{b}_t^* W_t^{-1},
\]

where the last term is nonnegative as the trace of the product of two symmetric nonnegative matrices. It follows, that \(\text{tr} \dot{W}_t W_t^{-1}\) 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

\[
\frac{d}{dt} \text{tr} W_t W_t = 4 \text{tr} (\dot{a}_t \sigma_t^* - \dot{b}_t) W_t W_t + 2 \text{tr} \dot{a}_t \dot{b}_t^* W_t - 4 \text{tr} a_t W_t^3,
\]

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

\[
\text{tr} \dot{a}_t \dot{b}_t^* W_t \leq N (\text{tr} W_t^2)^{1/2} \leq N + \text{tr} W_t^2,
\]

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

\[
(\text{tr} AW^2)^2 \leq \|A\| \|W^2\| \leq \|A\| (\text{tr} W^2)^2.
\]

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 \(\omega\). 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 \(\omega\). 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 \( \mathcal{L} \) is introduced in (7.21) and consider the following deterministic problem

\[
\partial_t v(t,z) + \mathcal{L}v(t,z) - c(t,y)v(t,z) = 0, \quad t \in [0,T], z \in \mathbb{R}^{d_1},
\]

\[
v(T,z) = \varphi(z), \quad z \in \mathbb{R}^{d_1}. \tag{9.1}
\]

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 \( z_t(s,z) \), \( t \geq s \), the solution of system (7.1) which starts at \( z \) at moment \( s \leq T \), then by Itô’s formula we have

\[
v(s,z) = E\varphi(z_T(s,z))\exp\left(-\int_s^T c_r(y_r(s,z))\,dr\right),
\]

\[
|v(s,z)| \leq \sup |\varphi| P\{\tau(s,z) \leq T\} \leq \sup |\varphi| e^{N_0 T} e^{-N_0 \tau(s,z)},
\]

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

\[
\mathcal{L} \psi(z) - N_0 \psi(z) \leq 0.
\]

By Itô’s formula, for \( |z| \geq R \),

\[
\psi(R) e^{-N_0 \tau(s,z)} \leq \psi(z),
\]

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

\[
|v(s,z)| \leq \frac{N}{(1 + |z|^2)^m}.
\]

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 \( \partial^2 v(s,z)/\partial z^i \partial z^j \), \( i,j = 1, \ldots, d_1 \).

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

Lemma 9.1. Let \( \xi_t \) be a nonnegative continuous martingale on \([0,T] \) and let \( \zeta_t \) be a continuous \( F_t \)-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 \( \tau \leq T \) 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 \( E \xi_T \zeta_{\tau_n} = E \xi_0 \zeta_0 \). Using the dominated convergence theorem yields the desired result and proves the lemma.
Lemma 9.2. The process
\[ \rho_t e^{-\int_0^t c_s(y_s) \, ds} \int_{\mathbb{R}^d} v(t, x, y_t) \pi_t(x) \, dx \]
is a martingale on \([0, T]\).

Proof. Define \((c_t = c(t, y_t), v_t(x) = v(t, x, y_t))\)
\[ D_k^y = \frac{\partial}{\partial y_k}, \quad D_{kr}^y = D_k^y D_r^y, \quad C_t = \exp(-\int_0^t c_s \, ds), \quad \chi_t = C_t v_t \pi_t. \]
We need to show that
\[ \rho_t \int_{\mathbb{R}^d} \chi_t(x) \, dx \quad (9.2) \]
is a martingale.

Observe that by Itô’s formula and \( (9.1) \) we have
\[ d[v_t(x)C_t] = d[v_t C_t] = C_t [D_k^y v_t \, dy_k^t + (\partial_t v_t - c_t v_t + \tilde{a}_t^k D_k^y v_t) \, dt] \]
\[ = C_t [D_k^y v_t \Phi_t^{kr} \, d\tilde{w}_t^r - (L_t v_t + 2\tilde{a}_t^k D_t^y v_t + (B_t^k - \tilde{B}_t^k) D_k^y v_t) \, dt], \quad (9.3) \]
where we dropped the arguments \( x \) for shortness and, of course, \( D_k^y v_t = (D_k^y v)(t, x, y_t), D_{kr}^y v_t = (D_k^y v)(t, x, y_t) \), and \( D_t^y v_t = (D_t^y v)(t, x, y_t) \).

By the way, observe that
\[ \sigma_t^{ir} \Phi_t^{kr} = 2\tilde{a}_t^k, \quad B_t^k = \Phi_t^{kr} B_t^r. \]

Similarly to the proof of Lemma \( \S 5.5 \) we conclude that \( (9.3) \) holds in the sense of distributions and that Theorem \( 6.1 \) is applicable to \( v_t \pi_t \) on the time interval \( t \leq \tau_m \) for any \( n \), where \( \tau_m \) are taken from Lemma \( \S 5.5 \).
It follows that for any \( m \) for \( t \leq \tau_m \)
\[ d\chi_t = C_t (\pi_t \Phi_t^{kr} D_k^y v_t + v_t \Lambda_t^{rs} \pi_t) \, d\tilde{w}_t^r \]
\[ = -C_t \pi_t (L_t v_t + \sigma_t^{ir} \Phi_t^{kr} D_t^y v_t + \Phi_t^{kr} (B_t^r - \tilde{B}_t^r) D_k^y v_t) \, dt \]
\[ + C_t v_t (L_t \pi_t + \tilde{B}_t^k \Lambda_t^{rs} \pi_t) \, dt + C_t (D_k^y v_t) \Phi_t^{kr} \Lambda_t^{rs} \pi_t \, dt. \]

It is convenient to rearrange the above terms by using the notation
\[ \zeta_t^r = C_t (\pi_t \Phi_t^{kr} D_k^y v_t + v_t \Lambda_t^{rs} \pi_t). \]
We have
\[ d\chi_t = \zeta_t^r \, d\tilde{w}_t^r + (\tilde{B}_t^k \zeta_t^r + I_1^1 + I_2^2) \, dt, \quad t \leq \tau_m, \]
where
\[ I_1^1 = C_t (v_t L_t^s \pi_t - \tilde{\pi}_L v_t), \quad I_2^2 = -C_t \Phi_t^{kr} \sigma_t^{ir} (\pi_t \sigma_t^{ir} D_t^y v_t + (D_k^y v_t) D_t \pi_t) \]
\[ = -C_t \Phi_t^{kr} \sigma_t^{ir} D_t (\pi_t D_k^y v_t). \]
In the integral form this means that for any \( \phi \in C_0^\infty \) with probability one
\[ (\chi_t \wedge \tau_m, \phi) = (\chi_0, \phi) + \int_0^t I_s \leq \tau_m (c_s^r, \phi) \, d\tilde{w}_s^r + \int_0^t I_s \leq \tau_m \tilde{B}_s^k (\zeta_s^r, \phi) \, ds \]
\[ + \int_0^t I_s \leq \tau_m C_s \sigma_t^{ij} (\pi_s D_j v_s - v_s D_j \pi_s, D_t \phi) \, ds \]
\[ + \int_0^t I_{S \leq \tau_m} C_S\left[ (\bar{\pi}_s v_s, b^i_s D_t \phi) + \Phi^{kr}_s \phi^{ir}_s (\bar{\pi}_t D^y_k \bar{v}_s, D_t \phi) \right] ds. \] (9.4)

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

Observe that
\[ (\xi^r_s, \phi_j) = C_s(\phi_j, \bar{\pi}_s \Phi^{kr}_s D^y_k \bar{v}_s + v_s \Lambda^{rs}_s \bar{\pi}_s) \]
and for any \( r \) and \( k \):
\[ \int_0^T (1, |\bar{\pi}_s D^y_k \bar{v}_s| + |v_s \Lambda^{rs}_s \bar{\pi}_s|)^2 ds \]
\[ \leq N \int_0^T \|\bar{\pi}_s\|^2_{W^1_p} \|v_s\|^2_{W^1_p} ds \leq N \|\pi\|^2_{W^1_p(T)} < \infty, \]
where \( N \) is independent of \( \omega \). 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 \([0,T]\)
\[ \int_0^t I_{S \leq \tau_m} (\xi^r_s, \phi_j) d\tilde{w}_s \rightarrow \int_0^t I_{S \leq \tau_m} C_s(1, \bar{\pi}_s \Phi^{kr}_s D^y_k \bar{v}_s + v_s \Lambda^{rs}_s \bar{\pi}_s) d\tilde{w}_s. \]

Similarly, and in an easier fashion one analyzes the remaining terms in (9.4) and concludes that for any \( m \)
\[ d(\chi_t, 1) = C_t(1, \bar{\pi}_t \Phi^{kr}_t D^y_k \bar{v}_t + v_t \Lambda^{rs}_t \bar{\pi}_t) d\tilde{\gamma}_t, \; t \leq \tau_m. \]

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)) \( \xi_t = \mu_t(\bar{b} - b_t(0) + \bar{b}^* W^{-1} V, d\tilde{w}) \) and
\[ \zeta_t = e^{-\int_0^t c_s(y_s) ds} \int_{\mathbb{R}^d} \bar{\pi}(x) v_t(x) \exp \left( -\frac{1}{2} \int_0^t |W^{1/2}_s x + W^{1/2}_s v_1^{1/2}|^2 ds \right) dx. \]

Owing to Lemma (8.2) the process \( \zeta_t \) is bounded on \([0,T]\) by a constant times \( \|\pi_0\|_{L_p} \) 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
\[ E e^{-\int_0^T c_s(y_s) ds} \varphi(z_T) = E v(0, x_0, y_0) = E \int_{\mathbb{R}^d} v(0, x, y_0) \bar{\pi}_0 dx \]
\[ = E \rho_T e^{-\int_0^T c_s(y_s) ds} \int_{\mathbb{R}^d} \varphi(x, y_T) \bar{\pi}_T(x) dx \]
\[ = E \rho_T e^{-\int_0^T c_s(y_s) ds} \int_{\mathbb{R}^d} \varphi(x, y_T) \bar{\pi}_T(x) dx, \]
where \( \rho_T = E(\rho_T | F^*_T) \). Since the equality between the extreme terms holds for sufficiently wide class of functions \( c \), we get that
\[ E(\varphi(z_T) | F^*_T) = \rho_T \int_{\mathbb{R}^d} \varphi(x, y_T) \bar{\pi}_T(x) dx \quad (a.s.). \]
The arbitrariness of \( \phi \) implies that \( \bar{\pi}_T \geq 0 \) (a.s.) and

\[
1 = \rho_T \int_{\mathbb{R}^d} \bar{\pi}_T(x) \, dx, \quad (1, \bar{\pi}_T) = \int_{\mathbb{R}^d} \bar{\pi}_T(x) \, dx > 0, \quad \rho_T = (1, \bar{\pi}_T)^{-1}
\]

(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 \( \bar{\pi}_t \geq 0 \) 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 \( L_1 \)-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 \( x \)

\[
d\xi_t = \zeta_t' \left( d\bar{\mu}_t^\tau + \bar{B}' \, dt \right) + C_t(v_t L_t \bar{\pi}_t - \bar{\pi}_t L_t v_t) \, dt - C_t \Phi_t^{k^r} \sigma_t^{i^r} D_t(\bar{\pi}_t D_t^k v_t) \, dt.
\]

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

\[
d\langle \xi_t, 1 \rangle = (\zeta_t', 1) \left( d\bar{\mu}_t^\tau + \bar{B}' \, dt \right)
\]

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) \( \xi_t = \rho_t (\bar{B} - B(0) + \bar{\mu}^* W^{-1} V, dw) \)

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

Notice that \( \xi_t \) is a martingale and \( \zeta_t \) 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., Vol. 18 (1987), No. 3, 857-872.

[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, pp. 1-24, Springer Verlag, 1987.

[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., Vol. 182 (2001), 243-279.

[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, Vol. 17 (2004), No. 3-4, 259-296.

[5] B. Farkas and A. Lunardi, *Maximal regularity for Kolmogorov operators in \( L^2 \) spaces with respect to invariant measures*, J. Math. Pures Appl., Vol. 86 (2006), 310-321.
[6] M. Geissert and A. Lunardi, Invariant measures and maximal $L^2$ regularity for nonautonomous Ornstein-Uhlenbeck equations, J. Lond. Math. Soc. (2), Vol. 77 (2008), No. 3, 719-740.

[7] I. Gyöngy, Stochastic partial differential equations on Manifolds, I, Potential Analysis, Vol. 2 (1993), 101-113.

[8] I. Gyöngy, Stochastic partial differential equations manifolds II. Nonlinear filtering Potential Analysis, Vol. 6 (1997), 39-56.

[9] I. Gyöngy and N.V. Krylov, On stochastic partial differential equations with unbounded coefficients, Potential Analysis, Vol. 1 (1992), No. 3, 233-256.

[10] N.V. Krylov, Some properties of traces for stochastic and deterministic parabolic weighted Sobolev spaces, Journal of Functional Analysis, Vol. 183 (2001), No. 1, 1-41.

[11] N.V. Krylov, On divergence form SPDEs with VMO coefficients, SIAM J. Math. Anal. Vol. 40 (2009), No. 6, 2262-2285.

[12] N.V. Krylov, On linear elliptic and parabolic equations with growing drift in Sobolev spaces without weights, Problemy Matematicheskogo Analiza, Vol. 40 (2009), 77-90, in Russian; English version in Journal of Mathematical Sciences, Vol. 159 (2009), No. 1, 75-90, Srping.

[13] N.V. Krylov, Itô’s formula for the $L_p$-norm of stochastic $W^{1,p}$-valued processes, to appear in PTRF.

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

[15] N.V. Krylov, On divergence form SPDEs with growing coefficients in $W^{1,1}$ spaces without weights, to appear in SIMA.

[16] N.V. Krylov, On divergence form second-order PDEs with growing coefficients in $W^{1,1}$ spaces without weights, submitted to “Parabolic Problems: Herbert Amann Festschrift”, Birkhäuser.

[17] N.V. Krylov, On the Itô-Wentzell formula for distribution-valued processes and related topics, to appear in Probab. Theory Related Fields, [http://arxiv.org/abs/0904.2752](http://arxiv.org/abs/0904.2752)

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

[19] N.V. Krylov and B.L. Rozovskii, Stochastic evolution equations, pp. 71-146 in “Izvy nauki i techniki”, Vol. 14, VINITI, Moscow, 1979, in Russian; English translation: J. Soviet Math., Vol. 16(1981), No. 4, 1233-1277.

[20] N.V. Krylov and A. Zatezalo, A direct approach to deriving filtering equations for diffusion processes, Applied Mathematics and Optimization, Vol. 42, No. 3 (2000), 315-332.

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

[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, Second edition 2001.

[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., Vol. 24 (1997), 133-164.

[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).
[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.

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

[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., Vol. 32 (2006), No. 2, 563-576.

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