HYPOELLIPTICITY FOR FILTERING PROBLEMS OF PARTIALLY OBSERVABLE DIFFUSION PROCESSES

N.V. KRYLOV

Abstract. We prove that under Hörmander’s type conditions on the coefficients of the unobservable component of a partially observable diffusion process the filtering density is infinitely differentiable and can be represented as the integral of an infinitely differentiable kernel against the prior initial distribution. These results are derived from more general results obtained for SPDEs. One the main novelty of the paper is the existence and smoothness of the kernel, another that we allow the coefficients of our partially observable process to be just measurable with respect to the time variable.

1. Introduction

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 \(\mathcal{F}_t \subset \mathcal{F}\). Let \(w^k_t, k = 1, 2, \ldots, d_1\), be independent one-dimensional Wiener processes with respect to \(\mathcal{F}_t\), where \(d_1 \geq 1\) is an integer.

Let \(d, d' \geq 1\) be integers. Consider a \(d + d'\)-dimensional two-component process \(z_t = (x_t, y_t)\) with \(x_t\) being \(d\)-dimensional and \(y_t\) \(d'\)-dimensional. We assume that \(z_t\) is a diffusion process defined as a solution of the system

\[
\begin{align*}
    dx_t &= b(t, z_t)dt + \theta^k(t, z_t)dw^k_t, \\
    dy_t &= B(t, z_t)dt + \Theta^k(t, y_t)dw^k_t
\end{align*}
\]

(1.1)

with some initial data independent of the process \(w_t\). The coefficients of (1.1) are assumed to be vector-valued functions of appropriate dimensions defined on \([0, \infty) \times \mathbb{R}^{d+d'}\). Actually \(\Theta^k(t, y)\) are assumed to be independent of \(x\), so that they are functions on \([0, \infty) \times \mathbb{R}^{d'}\) rather than \([0, \infty) \times \mathbb{R}^{d+d'}\) but as always we may think of \(\Theta^k(t, y)\) as functions of \((t, z)\) as well.

One of the main goals of the paper is to show that under Hörmander’s type conditions satisfied for \(x\) lying in a ball \(B\), in some sense uniformly with respect to \(t\) and \(y\), there exists a function \(p(t, y, x) = p(\omega, t, y, x) \geq 0\), which is infinitely differentiable with respect to \((y, x) \in B^2\) for any \(t > 0\).
and $\omega$, such that for any $f \in C_0^\infty(B)$ and $t > 0$ with probability one

$$E\{f(x_t) \mid F^y_t\} = \int_B \int_B f(x)p(t, y, x) P_0(dy)dx,$$

where $P_0$ is the conditional distribution of $x_0$ given $y_0$ and $F^y_t = \sigma\{y_s, s \leq t\}$. Naturally,

$$\int_B p(t, y, x) P_0(dy) \quad (1.2)$$

turns out to be infinitely differentiable with respect to $x \in B$ and represent the conditional density $\pi_t(x)$ of $x_t$ given $F^y_t$.

In the literature two approaches to prove infinite differentiability of $\pi_t(x)$ for degenerate processes under Hörmander’s type conditions are known. The first one is based on filtering equations for $\pi_t$, which are stochastic partial differential equations (SPDEs). This approach was initiated by Wentzell [18] and in a more general and time inhomogeneous case outlined by Kunita in [13] and [14]. It is worth noting that in [18] the coefficient $B$ is supposed to be independent of $x$ and in [13] the functions $b$ and $\theta$ are independent of $y$. Equations in [14] seem not to cover general filtering equations either. In [18], [13], and [14] the SPDE is reduced to an ordinary parabolic equation with random coefficients by using a random change of coordinates. Without this reduction Chaleyat-Maurel and Michel in [4] achieve the goal in the time homogeneous case by mimicking some steps which are used in the proof of the deterministic Hörmander theorem. In their case as well as in [3] the matrix $(\Theta^1, \ldots, \Theta^{d'})$ is assumed to have form $(0, I)$ where $I$ is $d' \times d'$ identity matrix.

However, some of the arguments in [13] and [14] are based on the claim that Hörmander’s type theorem holds and can be proved by using Malliavin calculus for equations whose coefficients are only continuous with respect to $t$. Such a proof is unknown even now. It also looks like in [4] there is a gap at the point where the authors claim without proof that, roughly speaking, what holds for the unknown function also holds for its fractional derivatives.

Our approach is also based on using filtering equations but since we allow the coefficients of (1.1) to be just measurable with respect to $t$ our type of Hörmander’s condition is more restrictive than in [4] where the coefficients are time independent. In contrast with [13] and [14] we do not appeal to Malliavin calculus and instead rely on some analytical facts which we prove for more general SPDEs.

The second approach to proving infinite differentiability of $\pi_t(x)$ almost completely ignores filtering equations and is based on Malliavin calculus and first appeared in the paper by Bismut and Michel [1]. In their model the coefficients are time independent, but the Hörmander type condition imposed, albeit global, is much weaker than ours. Kusuoka and Stroock [16] further relax the Hörmander type condition in [1] again in time independent case but in what concerns filtering they assume that $B(t, x, y)$ is independent of $x$, so that the problem becomes a problem in the theory of conditional Markov
processes rather then a more or less general filtering problem, because the coefficients of the equation for the observation process \( y_t \) are supposed to be independent of the signal process \( x_t \). This result can also be found in [17]. In the recent publication by Chaleyat-Maurel [3] and references therein one can find a detailed account of the progress concerning Malliavin calculus and filtering equations. In particular, in [3] local versions of Hörmander’s type condition from [4] is used to obtain the local regularity of solutions. It seems that these methods are not applicable in our case of coefficients only measurable with respect to \( t \).

Apart from this novelty concerning time dependence, the fact that the conditional density \( \pi_t(x) \) is represented as (1.2) with infinitely differentiable kernel seems to be new for degenerate diffusions \( z_t \) under Hörmander’s type condition.

We derive our results about filtering densities in Section 4 from results of Sections 2, 3, and 5. In these sections we treat more general SDEs than the filtering equation.

The reader understands that \( \mathbb{R}^d \) is a Euclidean space of column-vectors (written in a common abuse of notation as) \( x = (x^1, ..., x^d) \). Denote

\[
D_i = \partial / \partial x^i, \quad D_{ij} = D_i D_j,
\]

and for an \( \mathbb{R}^d \)-valued function \( \sigma_t(x) = \sigma_t(\omega, x) \) on \( \Omega \times [0, \infty) \times \mathbb{R}^d \) and functions \( u_t(x) = u_t(\omega, x) \) on \( \Omega \times [0, \infty) \times \mathbb{R}^d \) set

\[
L_{\sigma_t} u_t(x) = [D_i u_t(x)] \sigma_i^t(x).
\]

Next take an integer \( d_2 \geq 1 \) and assume that we are given \( \mathbb{R}^d \)-valued functions \( \sigma_k^t = (\sigma_k^t)^i \), \( k = 0, ..., d_2 + d_1 \), on \( \Omega \times [0, \infty) \times \mathbb{R}^d \), which are infinitely differentiable with respect to \( x \) for any \( (\omega, t) \), and define the operator

\[
L_t = (1/2) \sum_{k=1}^{d_2 + d_1} L_{\sigma_k^t}^2 + L_{\sigma_0^t}.
\]

Assume that on \( \Omega \times [0, \infty) \times \mathbb{R}^d \) we are also given certain real-valued infinitely differentiable functions \( c_t(x) \) and \( \nu_k^t(x) \), \( k = 1, ..., d_1 \), and that on \( \Omega \times [0, \infty) \times \mathbb{R}^d \) we are given real-valued functions \( f_t \) and \( g_k^t \), \( k = 1, ..., d_1 \). Then under natural additional assumptions which will be specified later the SPDE

\[
du_t = (L_t u_t + c_t u_t + f_t) \, dt + (L_{\sigma_k^t} u_t + \nu_k^t u_t + g_k^t) \, dw_t^k
\]

makes sense (where and below the summation convention over repeated indices is enforced regardless of whether they stand at the same level or at different ones).

One of the main results of this paper is Theorem 2.4 saying that if the initial condition is a generalized functions of class \( H^{2-n} \), then (1.4) has a unique solution with this initial data without any nondegeneracy or Hörmander’s type condition. Before this result was known only if \( n \geq 1 \) is an integer (see [12]). The result is important because it allows one to take a \( \delta \)-function as the initial condition.
After the existence of solutions is secured we continue our investigation under Hörmander’s type condition and in Section 3 prove, roughly speaking, that, if \((s_1, s_2) \in (0, T)\) and for any \(\omega \in \Omega\) and \(t \in (s_1, s_2)\) the Lie algebra generated by the vector-fields \(\sigma_d^k t^{d_1+k}, k = 1, ..., d_2\), has dimension \(d\) everywhere in \(B_R\) and \(f_i\) and \(g_k^i\) are infinitely differentiable in \(B_R\) for any \(\omega \in \Omega\) and \(t \in (s_1, s_2)\), then the generalized function \(u_t\) satisfying (1.4) coincides on \((s_1, s_2)\times B_R\) with a function which is infinitely differentiable with respect to \(x\). In Section 4 we apply this result to filtering problems. In the same section we apply the results of Section 5 to derive the existence of smooth filtering kernels. The results of Section 5 bear on kernels (or fundamental solutions) for more general SPDEs.

In the whole article \(T\) is a fixed number from \((0, \infty)\).

2. An existence theorem for SPDEs

Denote by \(D\) the space of generalized functions on \(\mathbb{R}^d\), and as usual introduce \(\Lambda = (1 - \Delta)^{1/2}\), \(H^n_2 = \Lambda^{-n}L_2\), where \(L_2\) is the Hilbert space of real-valued square integrable functions on \(\mathbb{R}^d\) with usual norm. The scalar product in \(H^n_2\) will be denoted by \((\cdot, \cdot)_n\).

Denote by \(H^n_2(T)\) the set of \(D\)-valued functions \(u_t\ on \Omega \times [0, T]\ such that \((u_t, \phi)\) is predictable and \((u_0, \phi)\) is \(F_0\)-measurable for any \(\phi \in C_0^\infty(\mathbb{R}^d)\) and

\[
\int_0^T \|u_t\|^2_n dt < \infty \text{ (a.s.).}
\]

Sometimes it is necessary to indicate which filtration of \(\sigma\)-fields is involved in the definition of predictable functions. In these cases we write \(H^n_2(T) = H^n_2(T, \mathcal{F})\).

Introduce \(\mathcal{H}^n_2\) as the set of \(F_0\)-measurable \(H^n_2\)-valued function on \(\Omega\). For an open ball \(B\) by \(\mathcal{H}^n_2(B)\) we mean the subset of \(\mathcal{H}^n_2\) consisting of generalized functions with (closed) support in \(B\). Define

\[
\mathcal{H}^n_2(T) = \{ u \in H^n_2(T) : E\int_0^T \|u_t\|^2_n dt < \infty \}.
\]

**Assumption 2.1.** (i) The functions \(\sigma_d^k(x), k = 0, ..., d_2 + d_1, c_t, \nu_d^k, k = 1, ..., d_1\), are infinitely differentiable with respect to \(x\) and each of their derivatives of any order is bounded on \(\Omega \times [0, T] \times \mathbb{R}^d\). These functions are predictable with respect to \((\omega, t)\) for any \(x \in \mathbb{R}^d\);

(ii) For an \(n \in \mathbb{R}\), we have that \(f \in H^n_2(T), g^k \in H^{n+1}_2(T), k = 1, ..., d_1\), and

(iii) \(u_0 \in \mathcal{H}^n_2\).

**Definition 2.1.** By a *normal* solution of (1.4) of class \(H^n_2(T)\) with initial condition \(u_0\) we mean a function \(u\) which belongs to \(H^n_2(T)\), such that (a.s.)
Let $u_t$ be a continuous $H^{n-1}_2$-valued function and with probability one

$$u_t = u_0 + \int_0^t (L_s u_t + c_s u_s + f_s) \, ds + \int_0^t (L_{\sigma_s^k} u_s + \nu_s^k u_s + g_s^k) \, dw_s^k$$  \hspace{1cm} (2.1)

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

**Remark 2.1.** The usual and stochastic integrals of Hilbert space valued functions are well defined, so that the right-hand side of (2.1) is a continuous $H^{n-2}_2$-valued process.

**Remark 2.2.** We say that a function $u$ of class $H^n_2(T)$ is a *generalized* solution of (1.4) with initial condition $u_0$ if for any $\phi \in C_0^\infty(\mathbb{R}^d)$

$$(u_t, \phi) = (u_0, \phi) + \int_0^t (L_s u_t + c_s u_s + f_s, \phi) \, ds$$

$$+ \int_0^t (L_{\sigma_s^k} u_s + \nu_s^k u_s + g_s^k, \phi) \, dw_s^k$$  \hspace{1cm} (2.2)

for almost all $(\omega, t) \in \Omega \times [0, T]$, where by $(\cdot, \cdot)$ we mean the pairing between test functions and generalized ones.

By the way, recall that if $u \in H^n_2$ and $\phi \in C_0^\infty(\mathbb{R}^d)$, then

$$(u, \phi) = (\Lambda^k u, \Lambda^m \phi)_0,$$

as long as $k \leq n$ and $k + m = 2n$.

It is a well-known result (see, for instance, [11]) that if a function $u$ of class $H^n_2(T)$ is a generalized solution of (1.4) with initial condition $u_0$, then there exists a normal solution $\hat{u}$ of (1.4) of class $H^n_2(T)$ with initial condition $u_0$ such that $\hat{u}_t$ and $u_t$ coincide as generalized functions for almost all $(\omega, t)$.

This result implies, in particular, that if a generalized solution of class $H^n_2(T)$ is such that $u_t$ is a continuous $H^m_2$-valued function for some $m$, then $\hat{u}_t = u_t$ (a.s.) for all $t \in [0, T]$, so that $u_t$ itself is a continuous $H^{n-1}_2$-valued function (a.s.) and thus a normal solution of class $H^n_2(T)$.

Next we need the following technical lemma which enables us to integrate by parts in $H^n_2$-spaces.

**Lemma 2.3.** Let $n \in \mathbb{R}$ and $\nu \in C_0^\infty(\mathbb{R}^d)$. Then there exists a constant $N$ such that for any $u \in H^{n+1}_2$, $k = 0, 1, ..., d_1 + d_2$, $t \in [0, T]$ and $\omega \in \Omega$ we have

$$|(\Lambda^n L_{\sigma_t^k} u, \Lambda^n (\nu u))_0| \leq N\|u\|_n^2,$$  \hspace{1cm} (2.3)

$$\langle \Lambda^n u, \Lambda^n (L_{\sigma_t^k} u) \rangle + \|\Lambda^n L_{\sigma_t^k} u\|_n^2 \leq N\|u\|_n^2,$$  \hspace{1cm} (2.4)

where $(\cdot, \cdot)$ is the natural pairing between $H^n_2$ and $H^{n-1}_2$.

**Proof.** By obvious reasons we may assume that $u \in C_0^\infty(\mathbb{R}^d)$ and we drop the indices $k$ and $t$ to simplify notation.

We are going to rely on some well-known properties of pseudo-differential operators. The order of a pseudo-differential operator $S$ is a number $n \in \mathbb{R}$
such that $\Lambda^{-n}S$ and $S\Lambda^{-n}$ are bounded operators in $L_2$. If the orders of two operators $S_1$ and $S_2$ are $n_1$ and $n_2$, respectively, then the order of $[S_1, S_2] = S_1S_2 - S_2S_1$ is at most $n_1 + n_2 - 1$. One also knows that the first order linear differential operators with coefficients whose every derivative of any order is bounded are pseudodifferential operator of order one.

Observe that if a pseudo-differential operator $S$ is self adjoint, then for any $u \in C_0^\infty(\mathbb{R}^d)$
\[
(L_\sigma u, S(\nu u))_0 = (L_\sigma Su, \nu u)_0 + ([S, L_\sigma]u, \nu u)_0 \\
= -(Su, \nu L_\sigma u)_0 + ([S, L_\sigma]u, \nu u)_0
\]
where $a$ is a smooth bounded function. It follows that
\[
(L_\sigma u, S(\nu u))_0 = (1/2)\left([aSu, u]_0 + ([S, L_\sigma]u, \nu u)_0 - ([\nu, S]u, L_\sigma u)_0\right).
\]
(2.5)

It is important to note for the future that if the order of $S$ is $2n$, then the order of $[S, L_\sigma]$ is at most $2n$, the order of $[\nu, S]$ is at most $2n - 1$, and consequently
\[
|(L_\sigma u, S(\nu u))_0| \leq N\|u\|_n^2.
\]
(2.6)

This with $S = \Lambda^{2n}$ yields (2.3).

By applying (2.5) with $S = \Lambda^{2n}$ and $\nu \equiv 1$, we get
\[
(\Lambda^n L_\sigma u, \Lambda^n u)_0 = (1/2)(a\Lambda^{2n} u, u)_0 + (1/2)([\Lambda^{2n}, L_\sigma]u, u)_0,
\]
which after being polarized yields that
\[
(\Lambda^n L_\sigma v, \Lambda^n u)_0 + (\Lambda^n L_\sigma u, \Lambda^n v)_0 = (a\Lambda^{2n} u, v)_0 \\
+(a\Lambda^{2n} v, u)_0 + ([\Lambda^{2n}, L_\sigma]u, v)_0 + ([\Lambda^{2n}, L_\sigma]v, u)_0
\]
if $u, v \in C_0^\infty(\mathbb{R}^d)$. We plug in here $v = L_\sigma u$ and obtain
\[
(L_\sigma^2 u, u)_n + \|L_\sigma u\|_n^2 = (a\Lambda^{2n} u, L_\sigma u)_0 \\
+(a\Lambda^{2n} L_\sigma u, u)_0 + ([\Lambda^{2n}, L_\sigma]u, L_\sigma u)_0 + ([\Lambda^{2n}, L_\sigma]L_\sigma u, u)_0.
\]

After introducing the self adjoint operators
\[
S_1 = a\Lambda^{2n} + (a\Lambda^{2n})^*, \quad S_2 = [\Lambda^{2n}, L_\sigma] + ([\Lambda^{2n}, L_\sigma])^*
\]
we rewrite the last equation as
\[
(L_\sigma^2 u, u)_n + \|L_\sigma u\|_n^2 = (L_\sigma u, S_1 u)_0 + (L_\sigma u, S_2 u)_0
\]
and obtain (2.4) owing to (2.6). The lemma is proved.

**Theorem 2.4.** In class $H_2^0(T)$ there exists an (a.s.) unique normal solution $u$ of (1.4) on $[0, T]$ with initial condition $u_0$. Furthermore, there exists a constant $N$ independent of $u, f, g$ such that
\[
E \sup_{t \leq T} \|u_t\|^2 \leq N \|u_0\|^2 + N E \int_0^T (\|f_t\|_n^2 + \sum_{k=1}^{d_1} \|g_t^{k+1}\|_n^2) \, dt.
\]
(2.7)
Proof. Step 1. First we want to derive an a priori estimate assuming that we are given a normal solution of (1.4) of class $H^{n+1}(T)$. We apply the operator $\Lambda^n$ to both sides of (1.4) written in the integral form and observe that after that the stochastic integral will belong to $H^{n-1}_2$, whereas the deterministic integral will belong to $H^n$. This allows us to apply Itô’s formula for Banach space valued processes and shows that

$$d\|u_t\|^2_t = d\|\Lambda^n u_t\|^2_0 = \int_I^n dt + 2(\Lambda^n u_t, \Lambda^n \nu_t u_t + \nu^k_t u_t + g^k_t)\bigg|_0^\infty dw^k_t,$$

where

$$I_t = \langle \Lambda^n u_t, 2\Lambda^n \nu_t \sigma^n_t u_t + f_t \rangle = \sum_{k=1}^{d_1} \|\Lambda^n \nu_t \sigma^n_t u_t + g^k_t\|^2_0,$$

and

$$I^1_t = \sum_{k=1}^{d_1} \left[ \langle \Lambda^n u_t, \Lambda^n (\sigma^n_t) u_t \rangle + \|\Lambda^n \nu_t \sigma^n_t u_t\|^2_0 \right],$$

$$I^2_t = \langle \Lambda^n u_t, 2\Lambda^n \nu_t \sigma^n_t u_t \rangle_0,$$

$$I^3_t = \langle \Lambda^n u_t, 2\Lambda^n \nu_t \sigma^n_t u_t + f_t \rangle_0,$$

$$I^4_t = \langle \Lambda^n \nu_t \sigma^n_t u_t, \Lambda^n \nu_t u_t \rangle_0,$$

$$I^5_t = \langle \Lambda^n \nu_t \sigma^n_t u_t, \Lambda^n g^k_t \rangle_0,$$

$$I^6_t = \sum_{k=1}^{d_1} \|\Lambda^n (\nu^k_t u_t + g^k_t)\|^2_0.$$

The term $I^1_t$ is estimated in (2.4) and $I^2_t$ in (2.3) (with $\nu \equiv 1$), which also provides an estimate for $I^4_t$. Almost obviously

$$|I^5_t| + |I^6_t| \leq N\|u_t\|^2_n + N\|f_t\|^2_n + \sum_{k=1}^{d_1} \|g^k_t\|^2_n,$$

where and below we denote by $N$ various constants independent of $u, f, g^k, t,$ and $\omega$. Finally,

$$I^5_t = \langle \Lambda^n u_t, \Lambda^n \nu_t \sigma^n_t u_t \rangle = \langle \Lambda^n u_t, (\Lambda^n \nu_t \sigma^n_t u_t) \rangle_0$$

and, since the order of the operator $[\Lambda^n, \nu_t \sigma^n_t]$ is at most $n$, we have

$$|I^5_t| \leq N \sum_{k=1}^{d_1} \|g^k_t\|^2_{n+1}.$$

Upon collecting our estimates we conclude that

$$d\|u_t\|^2_t \leq N \left(\|u_t\|^2_n + \|f_t\|^2_n + \sum_{k=1}^{d_1} \|g^k_t\|^2_{n+1}\right) dt.$$
By the above what is inside the square by magnitude is dominated by 
\[ +2\left(\Lambda^n u_t, \Lambda^n(L_{\sigma_t}^k u_t + \nu_t^k u_t + g_t^k)\right)_0\, du_t^k. \] (2.8)

**Step 2. Uniqueness.** Now assume that we are given two natural solutions of (1.4) of class \(H_2^2(T)\) with the same initial condition. Then for their difference, say \(m_t\) we have
\[ d\|u_t\|^2_{n-1} \leq N_1\|u_t\|^2_{n-1}\, dt + dm_t, \]
where \(m_t\) is a local martingale. Next, comparing the differentials we obtain
\[ \|u_t\|^2_{n-1} e^{-N_1 t} \leq \int_0^t e^{-N_1 s}\, dm_s. \]
Since the right-hand side is a local martingale starting at zero and the left-hand side is nonnegative, it follows, that the right-hand side is zero as is the left-hand side, which proves uniqueness.

**Step 3.** Here we prove (2.7) as an a priori estimate under the assumptions of Step 1. We follow by now an absolutely standard and well-known way. With \(N\) from (2.8) we have
\[ d(e^{-N t}\|u_t\|^2_n) \leq N\left(\|f_t\|^2_n + \sum_{k=1}^{d_1} \|g_t^k\|^2_{n+1}\right)\, dt + dm_t, \]
where \(m_t\) is a local martingale. Since the left-hand side is nonnegative, for any \(t \in [0, T]\),
\[ e^{-N t}E\|u_t\|^2_n \leq E\|u_0\|^2_n + NE \int_0^t \left(\|f_s\|^2_n + \sum_{k=1}^{d_1} \|g_s^k\|^2_{n+1}\right)\, ds, \]
\[ \sup_{t \in [0,T]} E\|u_t\|^2_n \leq NE\|u_0\|^2_n + NE \int_0^T \left(\|f_s\|^2_n + \sum_{k=1}^{d_1} \|g_s^k\|^2_{n+1}\right)\, ds. \] (2.9)

Next, notice that since \(\|u_t\|_n\) is continuous
\[ \tau_m := T \wedge \inf\{t \geq 0 : \|u_t\|_n \geq m\} \]
are stopping times and \(\tau_m \uparrow T\) as \(m \to \infty\). By Davis’s inequality (2.8) and (2.9) imply that
\[ E \sup_{t \leq \tau_m} \|u_t\|^2_n \leq NE\|u_0\|^2_n + NE \int_0^T \left(\|f_s\|^2_n + \sum_{k=1}^{d_1} \|g_s^k\|^2_{n+1}\right)\, ds \]
\[ +6E\left(\int_0^{\tau_m} \sum_{k=1}^{d_1} \left|\left(\Lambda^n u_t, \Lambda^n(L_{\sigma_t}^k u_t + \nu_t^k u_t + g_t^k)\right)_0\right|^2\, dt\right)^{1/2}. \] (2.10)

By the above what is inside the square by magnitude is dominated by
\[ N\|u_t\|_n(\|u_t\|_n + \|g_t^k\|_n). \]
Hence the last term in (2.10) is less than
\[ NE\left(\int_0^{\tau_m} \|u_t\|^2_n(\|u_t\|^2_n + \sum_{k=1}^{d_1} \|g_t^k\|^2_n)\, dt\right)^{1/2}. \]
\[ \leq NE \left( \sup_{t \leq \tau_m} \| u_t \|_n \int_0^{\tau_m} \left( \| u_t \|_n^2 + \sum_{k=1}^{d_1} \| g_k^t \|_n^2 \right) dt \right)^{1/2} \]

\[ \leq (1/2) E \sup_{t \leq \tau_m} \| u_t \|_n^2 + NE \int_0^{\tau_m} \left( \| u_t \|_n^2 + \sum_{k=1}^{d_1} \| g_k^t \|_n^2 \right) dt, \]

which after coming back to (2.10), using again (2.9), and sending \( m \to \infty \), by Fatou’s lemma yields the a priori estimate (2.7).

Step 4. Existence in a particular case. If the norms on the right in (2.7) are finite, \( u_0 \in L_2(\Omega, F_0, H_2^{n+1}) \), and our equation is uniformly nondegenerate, then (see, for instance, [7]) there exists a unique normal solution of our problem of class \( H_2^{n+2}(T) \). For this solution (2.7) is valid.

If the norms on the right in (2.7) are finite, but \( u_0 \in L_2(\Omega, F_0, H_2^n) \), and there is no nondegeneracy assumption, we approximate \( u_0 \) in the \( L_2(\Omega, F_0, H_2^n) \)-norm by a sequence \( u_0^m \in L_2(\Omega, F_0, H_2^{n+1}) \), \( m = 1, 2, \ldots \), and add into the right-hand side of (1.4) the term \( (1/m) \Delta u_t dt \) to make the equation uniformly nondegenerate. Denote by \( u_t^m \) the solutions of the so modified problems. Then (2.7) will hold with \( N \) independent of \( m \) because no constant of nondegeneracy was involved in the derivation of (2.7).

According to (2.7) the sequence \( u_t^m \) is bounded in \( H_2^n(T) \). In particular, \( (1/m) \Delta u_t \to 0 \) in \( H_2^{n-2}(T) \). Having in mind his fact and applying (2.7) to the difference \( u_t^m - u_t^k \) and \( n-2 \) in place of \( n \) we see that the sequence \( u_t^m \) is Cauchy in the space with norm, whose square is given by

\[ E \sup_{t \leq T} \| u_t \|_n^2, \]

in particular, in \( H_2^{n-2}(T) \). Let \( u_t \) be its limit in \( H_2^{n-2}(T) \) such that

\[ E \sup_{t \leq T} \| u_t - u_t^m \|_n^2 \to 0 \]

as \( n \to \infty \). Then equation (1.4) in the integral form holds in \( H_2^{n-4}(T) \). Now, since the sequence \( u_t^m \) is bounded in \( H_2^n(T) \) and converges to \( u_t \) in \( H_2^{n-2}(T) \), \( u_t \in H_2^n(T) \). After that we apply a classical result saying that if \( u_t \in H_2^n(T) \) satisfies (2.7) in generalized sense with initial condition \( u_0 \in H_2^{n-1} \), then (a.s.) \( u_t \) is a continuous \( H_2^{n-1} \)-valued function (see Remark 2.2).

To establish (2.7) for thus found normal solution take a sequence \( \phi^r \in C_0^\infty(\mathbb{R}^d) \) such that it is dense in the unit ball of \( H_2^n \). Then owing to (2.7) write for any \( j = 1, 2, \ldots \)

\[ E \sup_{t \leq T} \max_{i=1,\ldots,j} \left( \Lambda^n u_t^m, \Lambda^n \phi^r \right)_0 \leq NE \| u_0^m \|_n^2 + I, \]

where I is the second term on the right in (2.7). Since

\[ \left( \Lambda^n (u_t^m - u_t), \Lambda^n \phi^r \right)_0 \leq \| \Lambda^{n-2}(u_t^m - u_t) \|_0^2 \| \Lambda^{n+2}\phi^r \|_0^2, \]

\[ \leq \sup_{t \leq T} \max_{i=1,\ldots,j} \left( \Lambda^n u_t^m, \Lambda^n \phi^r \right)_0 \leq NE \| u_0^m \|_n^2 + I. \]
estimate (2.11) allows us to conclude from (2.12) that
\[ E \sup_{t \leq T} \max_{r=1,\ldots,j} (\Lambda^n u_t, \Lambda^n \phi^r)_0^2 \leq NE\|u_0\|_n^2 + I. \]
By letting \( j \to \infty \) and using the monotone convergence theorem and the
fact that \( \phi^r \) are dense in the unit ball of \( H^n_2 \), we get (2.7).

**Step 5. Existence in the general case.** The first assertion of the theorem
in the general case is proved as always by using stopping time \( \gamma_m \)
\[ \gamma_m = \inf\{t \geq 0 : \|u_0\|_n + \int_0^t (\|f_s\|_n^2 + \sum_{k=1}^{d_1} \|g^k_s\|_{n+1}^2) ds \geq m \}. \]
The theorem is proved.

In the remaining part of this section by \( u_t \) we mean the normal solution
from Theorem 2.4. We remind the reader that the common way of saying
that a generalized function in a domain is smooth means that there is
a smooth function which, as a as generalized function, coincides with
the given generalized one in the domain under consideration.

Theorem 2.4 and Sobolev embedding theorems immediately imply the
following.

**Corollary 2.5.** Suppose that Assumption 2.1 is satisfied for all \( n \). Then
(a.s.) the solution \( u_t \) is infinitely differentiable with respect to \( x \) and every
its derivative is a bounded continuous function of \( (t, x) \).

**Corollary 2.6.** Suppose that Assumption 2.1 is satisfied for all \( n \). Let \( D \)
be a domain in \( \mathbb{R}^d \) with \( \partial D \neq \emptyset \) and assume that for \( x \in D, t \in [0, T] \), and
\( \omega \in \Omega \) we have
\[ u_0(x), c_t(x), f_t(x) \leq 0, \quad \nu^k_t(x) = g^k_t(x) = 0, \quad k = 1, \ldots, d. \]
Then (a.s.) for all \( t \in [0, T] \) on \( D \) we have
\[ u_t \leq \max_{s \leq t} \max_{\partial D} (u_s)_+. \quad (2.13) \]
This result follows from Theorem 1.2 of [8] in which one takes \( \xi = 0, \)
\( \tilde{u} \equiv 1, \tilde{f} \equiv -c, \tilde{f}^1 \equiv 0, g \equiv 0, \) and \( \rho_s \) equal the right-hand side of (2.13) plus
a constant \( \varepsilon > 0. \) One adds \( \varepsilon \) to be sure that
\[(u_t - \rho_t \bar{u}_t)_+ = (u_t - \rho_t)_+ \]
vanishes near the boundary of \( D \) and hence belongs to \( \bar{W}^1_2(D) \). Then after
applying Theorem 1.2 of [8] one sets \( \varepsilon \downarrow 0. \)

**Theorem 2.7.** Take an \( R \in [0, \infty) \) and suppose that \( \sigma^k_t(x) = 0 \) and \( \nu^k_t(x) = 0 \)
for \( k = 1, \ldots, d_1, t \in [0, T] \), and \( \omega \in \Omega \) as long as \( |x| > R. \) Also assume
that \( \bar{g}^k \equiv 0 \) for \( k = 1, \ldots, d_1. \) Then (a.s.) there exists a (random finite)
constant \( N \) independent of \( f \) and \( u_0 \) such that
\[ \sup_{t \leq T} \|u_t\|_n^2 \leq N\|u_0\|_n^2 + N \int_0^T \|f_t\|_n^2 dt. \quad (2.14) \]
Proof. For smooth $\mathbb{R}^d$-valued functions $\sigma(x)$ on $\mathbb{R}^d$ (whose points are always considered as column vectors) by $D\sigma$ we mean a matrix with entries $(D\sigma)^{ij} = D_j \sigma^i$ and if we are given two such functions $\sigma$ and $\gamma$, then we set
\[
D\sigma\gamma := \left[ D\sigma \right]\gamma.
\] (2.15)

Consider the equation
\[
x_t = x - \int_0^t \sigma^k_s(x_s) \, du^k_s - \int_0^t b_t(x_s) \, ds,
\] (2.16)

where
\[
b_t(x) = \sigma^0_t(x) - (1/2) \sum_{k=1}^{d_1} D\sigma^k_t(x) \sigma^k_t(x).
\]

As is well known (see, for instance, [15] for more advanced treatment of the subject or see [9]), there exists a function $X_t(x)$ on $\Omega \times [0, T] \times \mathbb{R}^d$, such that
(i) it is continuous in $(t, x)$ for any $\omega$ along with each derivative of $X_t(x)$ of any order with respect to $x$,
(ii) for each $\omega$ and $t$ the mapping $x \rightarrow X_t(x)$ of $\mathbb{R}^d$ to $\mathbb{R}^d$ is one-to-one and onto and its inverse mapping $X^{-1}_t(x)$ has bounded and continuous in $(t, x)$ derivatives of any order with respect to $x$ for any $\omega$.
(iii) it is $\mathcal{F}_t$-adapted for any $x$,
(iv) for each $x$ with probability one it satisfies (2.16) for all $t \in [0, T]$.

Observe that $X_t(x) = x$ for $|x| \geq R$, and $X^{-1}_t(x) = x$ for all $t \in [0, T]$ if $|x|$ is large enough (depending on $\omega$).

Next, define the operations “hat” and “check” which transform any function $\phi_t(x)$ into
\[
\hat{\phi}_t(x) := \phi_t(X_t(x)), \quad \check{\phi} = \phi_t(X^{-1}_t(x)).
\]

Also define $\rho_t(x)$ from the equation
\[
\rho_t(X_t(y)) \det DX_t(y) = 1
\]
and observe that by the change of variables formula
\[
\int_{\mathbb{R}^d} F(X_t(y)) \phi(y) \, dy = \int_{\mathbb{R}^d} F(x) \check{\phi}_t(x) \rho_t(x) \, dx,
\] (2.17)

whenever at least one side of the equation makes sense. Finally, define the mapping “bar” which transforms any $\mathbb{R}^d$-valued function $\sigma_t(x)$ into
\[
\bar{\sigma}_t(x) = Y_t(x) \bar{\sigma}_t(x) = Y_t(x) \sigma_t(X_t(x)),
\] (2.18)

where
\[
Y = (DX)^{-1}.
\]

By Corollary 6.5 of [9] (also see Remark 2.2) the function $\hat{u}_t$ is well defined and is a normal solution of class $H^2_2(T)$ of the equation
\[
d\hat{u}_t = \left[ \sum_{k=1}^{d_2} L^k_{\bar{\sigma}_t + \bar{\hat{f}}_t} \hat{u}_t + \bar{\hat{c}}_t \hat{u}_t + \hat{\bar{f}}_t \right] dt + \hat{u}_t \hat{d}_t \, dw^k_t.
\]
By using Kolmogorov’s continuity theorem for random fields one easily shows that there exists a function $I_t(x) = I_t(\omega, x)$ which along with each its derivative of any order with respect to $x$ is continuous with respect to $(t, x) \in [0, T] \times \mathbb{R}^d$ for each $\omega$ and such that for each $(t, x)$ with probability one

$$I_t(x) = \int_0^t \dot{\phi}_s^k(x) dw_s^k.$$ 

Then define

$$\gamma_t(x) = \exp \left[ - I_t(x) - (1/2) \sum_{k=1}^d \int_0^t |\dot{\phi}_s^k(x)|^2 ds \right], \quad v_t = \dot{u}_t \gamma_t.$$ 

Of course, $\gamma_t(x) = 1$ if $|x| \geq R$. We want to apply Itô’s formula to write an equation for $v_t$, that is, for any $\phi \in C^\infty_0(\mathbb{R}^d)$ find the stochastic differential of

$$I_t(\phi) := (v_t, \phi) = (\dot{u}_t, \gamma_t \phi) = (\Lambda^{-(n-1)} \dot{u}_t, \Lambda^{-1}(\gamma_t \phi))_0.$$ 

Here $\xi_t := \Lambda^{-(n-1)} \dot{u}_t$ and $\eta_t := \Lambda^{-1}(\gamma_t \phi)$ are continuous $L_2$-valued processes, admitting stochastic differentials such that the classic formula for the squared norm is applicable. Then this formula is also applicable to $\xi_t + \lambda \eta_t$ for any number $\lambda$. By comparing the coefficients of $\lambda$ in $\|\xi_t + \lambda \eta_t\|^2_0$ and in the formula we obtain the stochastic differential of $(\xi_t, \eta_t)_0$ that is of $I_t(\phi)$.

In this way we get that with probability one for all $t \in [0, T]$

$$v_t = v_0 + \int_0^t \left[ \sum_{k=1}^{d_2} \gamma_s \lambda^{d_s \cdot d_{1+k}} (\gamma_s^{-1} v_s) + \dot{c}_s v_s + \gamma_s \dot{f}_s \right] ds. \quad (2.19)$$

Fix an $\omega$ such that (2.19) holds for all $t \in [0, T]$,

$$\int_0^T \|u_t\|^2_{H_0^n} dt < \infty,$$

$u_0(\omega, \cdot) \in H_0^n$, and $u_t(\omega, \cdot)$ is an $H_2^n - 1$-valued continuous function. Then (2.19) becomes a deterministic equation, to which Theorem 2.4 is applicable because the differential operators in (2.19) can be rewritten in a canonical form as in (1.4) with coefficients satisfying Assumption 2.1. Then applying Theorem 2.4 to each particular $\omega$, the set of which has full probability, we obtain (2.14). The theorem is proved.

**Lemma 2.8.** Take

$$f_t(x) = u_0(x) = (1 + |x|^2)^{-d}, \quad g_t^k(x) = 0$$

and call $v_t$ the normal solution of (1.4) on $[0, T]$ with so prescribed data. By Corollary 2.5 the function $v_t$ is (a.s.) infinitely differentiable with respect to $x$ and every its derivative is a bounded continuous function of $(t, x)$. We assert that with probability one for every $r \in (0, \infty)$ there exists a (random constant) $\varepsilon > 0$ such that $v_t(x) \geq \varepsilon$ for $t \in [0, T]$ and $x \in \bar{B}_r$. 
Proof. First observe that \( v_t \geq 0 \) by the maximum principle. We take an \( R > r \) and concentrate on equation (1.4) only for \( x \in B_r \). Then equation in \( B_r \) will still hold if we cut off all \( \sigma^k \) and \( \nu^k \) outside \( B_r \) so that they will vanish outside \( B_R \).

For simplicity of notation we assume that \( \sigma^k = 0 \) and \( \nu^k = 0 \) outside \( B_R \) for \( k = 1, \ldots, d_1 \) for the original coefficients.

Then making the same transformations as in the proof of Theorem 2.7 we come to the conclusion that for almost any \( \omega \)

(i) equation (2.19) holds on \( \{(t, x) : t \in [0, T], x \in X_t^{-1}(B_R)\} \) with \( \hat{v}_t \gamma_t \) in place of \( v_t \).

For each \( \omega \) this is a deterministic (degenerate) parabolic equation with bounded coefficients. Furthermore, for almost any \( \omega \)

(ii) \( \gamma_s \hat{f}_s > 0 \) and \( \hat{v}_t \gamma_t \geq 0 \) on \( \{(t, x) : t \in [0, T], x \in X_t^{-1}(B_R)\} \).

Now we want to use the maximum principle to show that \( v_t(x) \) cannot take zero value in \([0, T] \times B_r \) whenever \( \omega \) is such that (i) and (ii) hold. Were the coefficients of (2.19) continuous in \( t \), this would be just a trivial matter. In our case we still need a little argument. Assume the contrary: there is a point \((t_0, x_0) \in [0, T] \times B_r \) such that \( v_{t_0}(x_0) = 0 \). Obviously, \( t_0 > 0 \). Then, since \( v_t \geq 0 \), for \( t \in (0, t_0) \) we get

\[
0 = \hat{v}_t(x_0)\gamma_t(x_0) + \int_t^{t_0} \left[ \sum_{k=1}^{d_2} \gamma_s L^2_{\sigma^k_s \hat{f}_s} \hat{v}_s + \hat{c}_s v_s + \gamma_s \hat{f}_s \right](x_0) \, ds
\]

\[
\geq \int_t^{t_0} \left[ \sum_{k=1}^{d_2} \gamma_s L^2_{\sigma^k_s \hat{f}_s} \hat{v}_s + \hat{c}_s v_s + \gamma_s \hat{f}_s \right](x_0) \, ds + I_t,
\]

where \( I_t \) is defined as the difference of the above two integrals. Since \( v_t(x) \) and its derivatives with respect to \( x \) are continuous with respect to \( t \), we have

\[
\lim_{t \uparrow t_0} \frac{1}{t_0 - t} I_t = 0.
\]

By taking into account that \( v_{t_0}(x_0) = 0 \), the first order derivatives of \( v_{t_0}(x) \) vanish at \( x_0 \), and the Hessian is nonnegative at \( x_0 \) we conclude that

\[
0 \geq \lim_{t \uparrow t_0} \frac{1}{t_0 - t} \int_t^{t_0} \gamma_s \hat{f}_s(x_0) \, ds,
\]

which is impossible because \( \gamma_s \hat{f}_s(x_0) \) is strictly positive on \([0, t_0] \). The lemma is proved.

**Theorem 2.9.** Suppose that Assumption 2.1 is satisfied for all \( n \). Also assume that there is an \( r \in (0, \infty) \) such that \( u_0(x), f_t(x) \leq 0 \), and \( g^k_t(x) = 0 \) for \( k = 1, \ldots, d_1 \) and \( x \notin B_r \). Then (a.s.) there exists a (random finite) constant \( N \) independent of \( f, g^k \), and \( u_0 \) such that (a.s) for \( t \in [0, T] \) and \( |x| \geq r \) we have

\[
u_t(x) \leq N \max_{s \leq t} \max_{\partial B_r}(u_s) + .
\]  
\[ (2.20) \]
Proof. Take $v_t$ from Lemma 2.8 and set

$$\rho_t = \max_{s \leq t} \max_{|x|=r} (u_t(x)/v_t(x))_+.$$ 

By Lemma 2.8 the process $\rho_t$ is finite, nonnegative, increasing, and continuous with probability one. Furthermore, $(u_t - \rho_t v_t)_+$ vanishes on $\partial B_r$. This along with the fact that $u_0(x) = f_t(x) \leq 0$ and $g_k^b(x) = 0$ for $k = 1, \ldots, d_1$ and $x \notin B_r$ by Theorem 1.2 of [8] implies that $u_t - \rho_t v_t \leq 0$ in $[0, T] \times B_R^c$ (a.s.), which obviously proves the theorem.

3. Hypoellipticity

Recall the notation associated with (2.15) and for two smooth $\mathbb{R}^d$-valued functions $\sigma$ and $\gamma$ on $\mathbb{R}^d$ set, as usual,

$$[\sigma, \gamma] = D\gamma\sigma - D\sigma\gamma.$$ 

Assumption 3.1. Assumption 2.1 (i) is satisfied, Assumption 2.1 (ii) is satisfied for all $n$ and, for an $n$, the function $u_0$ is an $\mathcal{F}_0$-measurable $H^n_\mathbb{R}$-valued function on $\Omega$.

Fix an $R_0 \in (0, \infty)$ and introduce collections of $\mathbb{R}^d$-valued functions defined on $\Omega \times [0, T] \times B_{R_0}$ inductively as $\mathbb{L}_0 = \{\sigma^{d_1+1}, \ldots, \sigma^{d_1+d_2}\}$,

$$\mathbb{L}_{n+1} = \mathbb{L}_n \cup \{[\sigma^{d_1+k}, M] : k = 1, \ldots, d_2, M \in \mathbb{L}_n\}, \quad n \geq 0.$$ 

For any multi-index $\alpha = (\alpha_1, \ldots, \alpha_d)$, $\alpha_i \in \{0, 1, \ldots\}$, introduce as usual

$$D\alpha = D^{\alpha_1}_1 \cdots D^{\alpha_d}_d, \quad |\alpha| = \alpha_1 + \ldots + \alpha_d.$$ 

Also define $BC^{\infty}_b$ as the set of real-valued measurable functions $a$ on $\Omega \times [0, T] \times \mathbb{R}^d$ such that for each $t \in [0, T]$ and $\omega \in \Omega$, $a_t(x)$ is infinitely differentiable with respect to $x$, and for any $\omega \in \Omega$ and multi-index $\alpha$ we have

$$\sup_{t, x \in [0, T] \times \mathbb{R}^d} |D\alpha a_t(x)| < \infty.$$ 

Finally we denote by Lie$_n$ the set of (finite) linear combinations of elements of $\mathbb{L}_n$ with coefficients which are of class $BC^{\infty}_b$. Observe that the vector-field $\sigma^0$ is not explicitly included into Lie$_n$. Finally, fix an $S \in [0, T)$ and introduce

$$G = (S, T) \times B_{R_0}.$$ 

Assumption 3.2. For every $\omega \in \Omega$ and $\zeta \in C_0^\infty(B_{R_0})$ there exists an $n \in \mathbb{R}$ such that we have $\xi I_{[S,T]}(\zeta) \in \text{Lie}_n$ for any $\xi \in \mathbb{R}^d$.

The following result will be used quite often. It is a particular case of Theorem 2.3 of [9]. By $u$ in this theorem and everywhere below in this section we mean the normal solution which exists due to Theorem 2.4.
Theorem 3.1. Take \( s_0 \in (S, T) \), \( r \in (0, R_0) \) and take a \( \zeta \in C_0^\infty(B_{R_0}) \) such that \( \zeta = 1 \) on a neighborhood of \( B_r \). Then

(i) with probability one \( u_t(x) \) is infinitely differentiable with respect to \( x \) for \( (t, x) \in (S, T) \times B_{R_0} \) and each derivative is a continuous function in \( (S, T) \times B_{R_0} \).

(ii) for any multi-index \( \alpha \) and \( l \) such that

\[
2(l - |\alpha| - 2) > d + 1, \tag{3.1}
\]

with probability one there exists a (random, finite) constant \( N \), independent of \( u, f \), and \( g^k \), such that

\[
\sup_{(t,x) \in [s_0,T] \times B_r} |D^\alpha u_t(x)|^2 \leq N \int_S^T \left[ \|f_t \zeta\|_{H^{l+1}_2}^2 + \|u_t \zeta\|_{H^l_2}^2 \right] dt, \tag{3.2}
\]

provided that \( g^k \zeta \equiv 0 \), \( k = 1, ..., d_1 \).

If we additionally assume that \( u_S \) is infinitely differentiable in \( B_{R_0} \) for every \( \omega \), then assertion (i) holds true with \( [S, T] \times B_{R_0} \) in place of \( (S, T) \times B_{R_0} \), and assertion (ii) with \( s_0 = S \) if we add to the right-hand side of (3.2) a constant (independent of \( u \)) times \( \|\zeta u_S\|_{H^{l+1}_2}^2 \).

Here is a generalization of the corresponding results of paper [5], where there is no stochastic terms in the equation. This is a generalization because no continuity hypothesis in time on the coefficients is imposed. The types of Hörmander’s condition imposed in [5] and here coincide. It is worth noting, however, that the result of [5] bears on the equation formally adjoint to the one we consider when there is no stochastic terms and no dependence on \( \omega \). Such equations have the same form as ours and have the same Lie \( \mathfrak{a} \). Recall that \( H^0_2(B_r) \) is introduced before Assumption 2.1.

Theorem 3.2. Suppose that \( S = 0 \), take \( s_0 \in (0, T) \), \( r \in (0, R_0) \), and suppose that \( u_0 \in H^0_2(B_r) \). Then \( u \) is (a.s.) infinitely differentiable with respect to \( x \) for \( (t, x) \in (\{s_0, T\} \times \mathbb{R}^d) \cup ([0, T] \times B_{R_0}^c) \) and each derivative is a bounded continuous function in \( ([s_0, T] \times \mathbb{R}^d) \cup ([0, T] \times B_{R_0}^c) \).

Proof. Take \( r < r_1 < r_2 < R \) and observe that \( B_{r_2} \setminus B_{r_1} \) can be covered by a finite number of balls lying inside \( B_{R_0} \setminus B_r \), where \( u_0 = 0 \). By applying Theorem 3.1 to each such ball we conclude that \( u_t \) is infinitely differentiable with respect \( x \) in \([0, T] \times (B_{r_2} \setminus B_{r_1}) \) and each its derivatives is bounded and continuous in \([0, T] \times (B_{r_2} \setminus B_{r_1}) \).

Take a \( \zeta \in C_0^\infty(B_{r_2}) \) such that \( \zeta = 1 \) in a neighborhood of \( B_{r_1} \) and set \( \eta = 1 - \zeta \). Then \( u_t \eta \) satisfies an equation similar to (1.4) but with different \( f \) and \( g^k \) which are obtained by adding to the original ones \( u_t \) or its first-order derivatives multiplied by \( C_0^\infty(\mathbb{R}^d) \) functions which vanish outside \( B_{r_2} \setminus B_{r_1} \). The initial condition for \( u_t \eta \) is obviously zero. By Theorem 2.4 and embedding theorems we conclude that (a.s.) \( u_t \eta \) is infinitely differentiable
with respect to $x$ for $(t, x) \in [0, T] \times \mathbb{R}^d$ and each derivative is a bounded continuous function in $[0, T] \times \mathbb{R}^d$.

The function $u_t \zeta$ satisfies an equation with the properties similar to those of the equation for $u_t \eta$ and by Theorem 3.1 (a.s.) is infinitely differentiable with respect to $x$ for $(t, x) \in [s_0, T] \times B_{R_0}$ and each derivative is a continuous function in $[s_0, T] \times B_{R_0}$. This proves the present theorem since $u_t = u_t \eta + u_t \zeta$ and $\zeta = 0$ outside $B_{R_0}$.

4. APPLICATIONS TO FILTERING PROBLEMS

Here we come back to system (1.1). Let $K, \delta > 0$ be some fixed constants. We denote by $\theta$ and $\Theta$ the matrix-valued functions having $\theta^k$ and $\Theta^k$, respectively, as their $k$th columns.

**Assumption 4.1.** The functions $b, \theta, B,$ and $\Theta$ are Borel measurable and bounded functions of their arguments. Each of them satisfies the Lipschitz condition in $z$ with constant $K$. These functions are infinitely differentiable with respect to $x$ and each derivative of any order is a bounded function of $(t, z)$.

**Assumption 4.2.** For all $\xi \in \mathbb{R}^{d'}$ and $(t, y)$

$$|\Theta^*(t, y)\xi| \geq \delta|\xi|.$$  

Notice that in [4] and [3] this condition is satisfied since there $\Theta = (0, I)$, where $I$ is the unit $d' \times d'$-matrix. In [16] there is no such condition because of a very peculiar filtering problem considered when the coefficients of the equation defining $y_t$ are independent of $x_t$.

**Remark 4.1.** Owing to Assumption 4.2, $d' \leq d_1$, the symmetric matrix $\Theta \Theta^*$ is invertible, and

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

is a bounded function of $(t, y)$.

**Assumption 4.3.** The random vector $z_0 = (x_0, y_0)$ is independent of the process $w_t$.

Next, we introduce a few more notation. Let

$$\Psi_t = \Psi(t, y_t), \quad \Theta_t = \Theta(t, y_t),$$

$$a_t(x) = \frac{1}{2} \theta \theta^*(t, x, y_t), \quad b_t(x) = b(t, x, y_t), \quad B_t(x) = B(t, x, y_t),$$

$$\theta_t(x) = \theta(t, x, y_t), \quad \sigma_t(x) = -\theta_t(x) \Theta_t^* \Psi_t, \quad \beta_t(x) = \Psi_t B_t(x).$$

The columns $\sigma^k_t(x), k = 1, ..., d_1$, of the matrix-valued function $\sigma_t(x)$ will play the role of $\sigma^k_t(x), k = 1, ..., d_1$, in the setting of Section 2. For $i = 1, ..., d_1$ the vector-functions $\sigma_t^{d_1+i}(x)$ are defined as the $i$th columns of $\theta_t(x) = \theta_t(x) \Theta_t^* \Psi_t \Theta_t$.
so that $d_2 = d_1$ in the notation of Section 2, $\sigma^0_t$ will be introduced later. Observe that if $d' = d_1$, then $\Theta_t$ takes values in the set of square $d_1 \times d_1$ matrices and owing to Assumption 4.2 is nondegenerate. It follows that $\Theta^* \Psi^2 \Theta$ is the identity operator for all $(t, y)$. In that case $\sigma^{d_1+i}_t \equiv 0, i = 1, \ldots, d_1$, and there is no hope to get any smoothness of the posterior distribution of $x_t$ unless the initial distribution has a smooth density. Therefore, we impose the following.

**Assumption 4.4.** We have $d' < d_1$ and, for an $S \in [0, T)$ and $R_0 \in (0, \infty)$, Assumption 3.2 is satisfied.

In a very popular so-called triangular scheme in which $\Theta = (0, \hat{\Theta})$, where $\hat{\Theta}$ is a nondegenerate square $d' \times d'$-matrix valued function, one can easily check that for $i = 1, \ldots, d_1 - d'$

$$\sigma^{d_1+i}_t(x) = \theta_i(x)$$

and $\sigma^{d_1+i}_t(x) \equiv 0$ for $i = d_1 - d' + 1, \ldots, d_1$.

Introduce $F^y_t$ as the completion of $\sigma\{y_s : s \leq t\}$ and denote by $P_0$ the regular version of the conditional distribution of $x_0$ given $y_0$.

**Theorem 4.2.** Take an $s_0 \in (S, T]$. Let $n$ be a negative integer such that $n < -1 - d/2$. Then there exists a function $\pi$ of class $H^{n+1}_2(T, F^y)$, such that

(i) $\pi_t$ is a continuous $H^n_2$-valued function on $[0, T]$,
(ii) (a.s.) $\pi_t(x)$ is infinitely differentiable with respect to $x$ for $(t, x) \in [s_0, T] \times B_{R_0}$ and each derivative is a continuous function in $[s_0, T] \times B_{R_0}$,
(iii) if $S = 0$ and the closed support of $P_0$ is a subset of $B_{R_0}$, then (a.s.) $\pi_t(x)$ is infinitely differentiable with respect to $x$ for $(t, x) \in ([s_0, T] \times \mathbb{R}^d) \cup ([0, T] \times B_{R_0}^c)$ and each derivative is a bounded continuous function in $([s_0, T] \times \mathbb{R}^d) \cup ([0, T] \times B_{R_0}^c)$,
(iv) for any $f \in C_0^\infty(\mathbb{R}^d)$ and $t \in [0, T]$ with probability one

$$\pi_t, f = E\{f(x_t) \mid F^y_t\}.$$

Before proving the theorem we prove the following.

**Lemma 4.3.** There exists an $H^{n+1}_2$-valued weakly continuous $F^y_t$-adapted process $\pi_t$ such that assertion (iv) of Theorem 4.2 holds.

Proof. Take an $f \in C_0^\infty(\mathbb{R}^d)$. By the famous Fujisaki-Kallianpur-Kunita theorem the process $E\{f(x_t) \mid F^y_t\}$ has a continuous modification, which we denote by $P_t(f)$. Then a well-known procedure (see, for instance, Chapter 5, §3.3 [17] or the Appendix in [10]) allows us to further modify, if necessary, $P_t(f)$, so that the new modification for which we use the same notation

(i) is continuous in $t$ and $F^y_t$-adapted,
(ii) for any $\omega \in \Omega, t \geq 0$, and any nonnegative $f \in C_0^\infty(\mathbb{R}^d)$ we have

$$0 \leq P_t(f) \leq \sup f,$$
(iii) for any \( \omega \in \Omega, t \geq 0, f, g \in C^\infty_0(\mathbb{R}^d) \), and numbers \( \alpha, \beta \) we have

\[
P_t(\alpha f + \beta g) = \alpha P_t(f) + \beta P_t(g).
\]

Then by Riesz-Markov theorem, there exists a measure-valued function \( P_t(dx) \) with \( P_t(\mathbb{R}^d) \leq 1 \) such that

\[
P_t(f) = \int_{\mathbb{R}^d} f(x) P_t(dx)
\]

for any \( \omega \in \Omega, t \geq 0, \) and \( f \in C^\infty_0(\mathbb{R}^d) \).

By recalling that finite measures on \( \mathbb{R}^d \) belong to \( H^{n+1}_2 \), we identify \( P_t(dx) \) with a generalized function \( \pi_t \in H^{n+1}_2 \). Observe that

\[
\| \pi_t \|_{H^{n+1}_2} \leq N P_t(\mathbb{R}^d),
\]

where \( N \) is the embedding constant. This and the continuity of \( (\pi_t, f) \) for \( f \in C^\infty_0(\mathbb{R}^d) \) shows that \( \pi_t \) is weakly continuous as a \( H^{n+1}_2 \)-valued function. All other assertions of the lemma follow from the above. The lemma is proved.

Remark 4.4. If \( f(t, x, y) \) is a Borel bounded function, then for any \( t \geq 0 \) with probability one

\[
E\{f(t, x, y) \mid \mathcal{F}^y_t\} = \int_{\mathbb{R}^d} f(t, x, y) P_t(dx), \tag{4.1}
\]

where the right-hand side is a predictable function with respect to \( \{\mathcal{F}^y_t\} \).

Indeed, we have seen that

\[
E\{f(x) \mid \mathcal{F}^y_t\} = \int_{\mathbb{R}^d} f(x) P_t(dx)
\]

(a.s.) for any \( t \geq 0 \) and \( f \in C^\infty_0(\mathbb{R}^d) \), where the right-hand side is \( \mathcal{F}^y_t \)-predictable. This implies our claim in an absolutely standard way.

**Proof of Theorem 4.2.** Set

\[
L_t(x) = a^{ij}_t(x) D_{ij} + b^i_t(x) D_i, \tag{4.2}
\]

\[
L^*_t(x) u_t(x) = D_{ij}(a^{ij}_t(x) u_t(x)) - D_i(b^i_t(x) u_t(x)),
\]

\[
\Lambda^k_t(x) u_t(x) = \beta^k_t(x) u_t(x) + \sigma^k_t(x) D_i u_t(x), \tag{4.3}
\]

\[
\Lambda^k_t(x) u_t(x) = \beta^k_t(x) u_t(x) - D_i(\sigma^k_t(x) u_t(x))
\]

where \( t \in [0, \infty), x \in \mathbb{R}^d, k = 1, \ldots, d_1 \), and as above we use the summation convention over all “reasonable” values of repeated indices, so that the summation in (4.2) and (4.3) is done for \( i, j = 1, \ldots, d \) (whereas in (4.4) for \( k = 1, \ldots, d_1 \)).

By the Fujisaki-Kallianpur-Kunita theorem, Lemma 4.3, and Remark 4.4 for any \( f \in C^\infty_0(\mathbb{R}^d) \) with probability one for all \( t \geq 0 \)

\[
(\pi_t, f) = (\pi_0, f) + \int_0^t (\pi_s, L_s f) ds + \int_0^t (\pi_s, \Lambda^k_s f - \beta^k_s f) d\vec{w}^k_s, \tag{4.4}
\]
where \((\hat{w}^1_t, \ldots, \hat{w}^{d_1}_t)\) is a standard Wiener process with respect to the filtration \(\{\mathcal{F}^y_t\}\) and
\[
\tilde{\beta}_t = \int_{\mathbb{R}^d} \beta(t, x, y) P_t(dx).
\]

By a classical result about the Itô formula for the squared norm of Banach space-valued processes, equation (4.4) implies that (a.s.) \(\pi_t\) is a continuous \(H^n\)-valued process (see Remark 2.2). Next, elementary computations show that
\[
a_t = \frac{1}{2} \sum_{k=1}^{2d_1} \sigma^k_t \sigma^{k*}_t, \quad L^*_t = \frac{1}{2} \sum_{k=1}^{2d_1} L^2_{\sigma^k_t} + L_{\sigma^0_t} + c_t,
\]
where
\[
\sigma^0_t = \frac{1}{2} L_{\sigma^k_t} \sigma^k_t + (\text{div} \sigma^k_t) \sigma^k_t - b_t, \quad c_t = D_{ij} a^{ij}_t - \text{div} b_t
\]
and for \(k = 1, \ldots, d_1\)
\[
\Lambda^k - \tilde{\beta}^k = L_{\sigma^k_t} + \nu^k_t,
\]
where
\[
\nu^k_t = \beta^k_t - \tilde{\beta}^k_t - \text{div} \sigma^k_t.
\]

After that the remaining assertion (ii) and (iii) of the present theorem follows directly from Theorems 3.1 and 3.2. The theorem is proved.

In the following two theorems we, actually, speak about \(\pi_t(x)\) which is defined as \(\pi_t(x)\) when \(P_0\) is the \(\delta\)-function concentrated at \(y\). These theorems are just direct consequences of Theorems 5.4 and 5.5 and of what was said in the proof of Theorem 4.2.

**Theorem 4.5.** Assume that \(S = 0\) and take \(r \in (0, R_0)\) and \(s_0 \in (0, T)\). Then there exists a nonnegative function \(p_t(x, y) = p_t(\omega, y, x)\) defined for
\[
(\omega, t, y, x) \in \Omega \times (0, T] \times B_r \times \mathbb{R}^d
\]
such that

(i) it is infinitely differentiable with respect to \(y\) in \(B_r\), each of its \(y\)-derivative of any order is a bounded function of \((t, y, x) \in [s_0, T] \times B_r \times \mathbb{R}^d\), and the functions
\[
\int_{B_r} |D^\alpha_y p_t(y, x)|^2 dy
\]
are bounded on \([s_0, T] \times B_r\) for any \(\omega \in \Omega\) and multi-index \(\alpha\),

(ii) for any probability distribution \(P_0\), which is concentrated in \(B_r\), with probability one we have
\[
\pi_t(x) = \int_{B_r} p_t(y, x) P_0(dy)
\]
for all \((t, x) \in (0, T] \times \mathbb{R}^d\),

(iii) for any \(y \in B_r\) with probability one for any \(t \in (0, T]\) and multi-index \(\alpha\) the function \(D^\alpha_y p_t(y, x)\) is infinitely differentiable with respect to \(x\) and each of its \(x\)-derivative of any order is a bounded function of \((t, x) \in [s_0, T] \times \mathbb{R}^d\),
(iv) the function \( p_t(y, x) \) is infinitely differentiable with respect to \((x, y) \in B_r \times B_r\) and each its derivative of any order is a bounded function of \((t, y, x) \in [s_0, T] \times B_r^2\) for any \( \omega \).

We have somewhat better properties of \( p_t(x, y) \) for a special class of filtering problems.

**Theorem 4.6.** Assume that \( S > 0 \), take \( r \in (0, R_0) \), \( s_0 \in (S, T) \) and assume that there exists an \( R \in (0, \infty) \) such that, if \( |x| \geq R \), then \( \theta(t, x, y) = 0 \) and \( B(t, x, y) \) is independent of \( x \). Then there exists a nonnegative function \( p_t(x, y) = p_t(\omega, y, x) \) defined for \((\omega, t, y, x) \in \Omega \times [s_0, T] \times \mathbb{R}^d \times B_r\) such that

- (i) it is infinitely differentiable with respect to \((x, y) \in \mathbb{R}^d \times B_r\), each its derivative of any order is a bounded function of \((t, y, x) \in [s_0, T] \times \mathbb{R}^d \times B_r\) and the functions \( \|p_t(\cdot, x)\|_l \) are bounded on \([s_0, T] \times B_r\) for any \( \omega \in \Omega \) and \( l \geq 0 \),

- (ii) with probability one we have

\[
\pi_t(x) = \int_{\mathbb{R}^d} p_t(y, x) P_0(dy)
\]

for all \((t, x) \in (S, T] \times B_r\).

### 5. Fundamental solutions of SPDEs

Here we continue our investigation of solutions of general equations (1.4) under the assumptions stated in Section 3. We also assume that

\[
g^k \equiv 0, \quad k = 1, \ldots, d_1.
\]

Here again \( u \) is the normal solution from Theorem 2.4 and \( n \) is from Assumption 2.1.

**Lemma 5.1.** Take an \( R \in [0, \infty) \), \( s_0 \in (S, T) \), \( r \in (0, R_0) \), and assume that \( \sigma_i^k = 0 \) and \( \nu_i^k = 0 \) outside \( B_R \) for any \( k = 1, \ldots, d_1, t, \) and \( \omega \). Then, for any \( l \) and multi-index \( \alpha \), such that

\[
2(l - |\alpha| - 2) > d + 1,
\]

with probability one there exists a (random, finite) constant \( N \) independent of \( u \) such that

\[
\sup_{(t, x) \in [s_0, T] \times B_r} |D^\alpha u_t(x)|^2 \leq N \|u_0\|^2_n + N \int_0^T \|f_t\|^2_l \, dt.
\]

Proof. Obviously, it suffices to concentrate on \( n \leq l \). Take a \( \zeta \in C_0^\infty(B_{R_0}) \) such that \( \zeta = 1 \) in a neighborhood of \( B_r \). Then by Theorem 3.1 for any \( l \)
and multi-index $\alpha$ satisfying (5.1) and $n \in \mathbb{R}$, with probability one, there exists a (random, finite) constant $N$ independent of $u$ such that

$$
\sup_{(t,x) \in [S,T] \times B_r} |D^\alpha u_t(x)|^2 \leq N \int_S^T \left[ \|f_t\|_l^2 + \|u_t\|^2 \right] dt. \tag{5.3}
$$

Owing to Theorem 2.7

$$
\|u_t\|^2 \leq N\|u_0\|^2 + N \int_0^T \|f_s\|^2 ds
$$

and this proves the lemma.

**Lemma 5.2.** In the setting of Lemma 5.1 suppose that $S = 0$ and $u_0$ is infinitely differentiable in $B_{R_0}$ for every $\omega$. Then for any $\zeta \in C_0^\infty (B_{R_0})$ such that $\zeta = 1$ in a neighborhood of $\bar{B}_r$, any $l$ and multi-index $\alpha$ satisfying (5.1), with probability one, there exists a (random, finite) constant $N$ independent of $u$ such that

$$
\sup_{(t,x) \in [0,T] \times B_r} |D^\alpha u_t(x)|^2 \leq N\|\zeta u_0\|^2_{l+1} + N \int_0^T \|f_t\|^2 dt + N\|u_0\|^2.
$$

This lemma is derived from Theorem 3.1 and Theorem 2.7 in the same way as Lemma 5.1.

**Theorem 5.3.** Assume that $S = 0$ and take $s_0 \in (0,T)$ and $0 < r < r_1 < r_2 < R_0$. Suppose that $u_0 \in H_2^0 (B_r)$. Finally, let $f_1(x) = 0$ for $|x| \geq r$.

Then for any $l$ and multi-index $\alpha$ satisfying (5.1), and $n \in \mathbb{R}$, with probability one, there exists a (random, finite) constant $N$ independent of $u$ such that

$$
\sup_{(t,x) \in \Gamma} |D^\alpha u_t(x)|^2 \leq N J, \tag{5.4}
$$

where

$$
\Gamma = ([0,T] \times B_{r_2}) \setminus ([0,s_0] \times B_{r_1}), \quad J = \int_0^T \|f_t\|^2 dt + N\|u_0\|^2.
$$

Proof. We take $r_2 < r_3 < R_0$, $\zeta \in C_0^\infty (B_{r_3})$, and $\eta \in C_0^\infty (B_{r_2})$ such that $\zeta = 1$ in a neighborhood of $B_{r_2}$ and $\eta = 1$ in a neighborhood of $B_{r_1}$.

As in the proof of Theorem 3.2 by covering $B_{r_2} \setminus B_{r_1}$ with appropriate balls and applying Theorem 3.1 we see that

$$
\sup_{(t,x) \in [0,T] \times (B_{r_2} \setminus B_{r_1})} |D^\alpha u_t(x)|^2 \leq N \int_0^T \|f_t\|^2 + \|\zeta u_t\|^2 dt,
$$

where the (random) constant $N$ is independent of $u, f$. All such constants will be denoted by $N$. Furthermore, by the same theorem

$$
\sup_{(t,x) \in [s_0,T] \times B_{r_2}} |D^\alpha u_t(x)|^2 \leq N \int_0^T \|f_t\|^2 + \|\zeta u_t\|^2 dt.
$$
Obviously, we may assume that \( n < l \) and then we note that to prove (5.4), it suffices to show that with probability one, there exists a (random, finite) constant \( N \) independent of \( u \) such that
\[
\int_0^T \| \zeta u_t \|_n^2 \, dt \leq NJ. \tag{5.5}
\]
To prove (5.5) we need an auxiliary function. Take a \( \xi \in C_0^\infty(B_{2R_0}) \) such that \( \xi = 1 \) on \( B_{R_0} \) and define \( v_t \) as a normal \( H_n^2(T) \)-solution of the equation
\[
dv_t = (L v_t + c_t v_t + f_t) \, dt + (L \xi \sigma^k v_t + \xi v_t^k) \, dw_t^k
\]
with initial condition \( u_0 \). By Theorem 2.7
\[
\sup_{t \leq T} \| v_t \|_n^2 \leq NJ. \tag{5.6}
\]
By Theorem 3.1
\[
\sup_{[0, T]} \sup_{|x| = r_1} |v_s(x)|^2 \leq N \int_0^T \left[ \| f_t \|_n^2 + \| v_t \|_n^2 \right] \, dt,
\]
which after being combined with (5.6) shows that
\[
\sup_{[0, T]} \sup_{|x| = r_1} |v_s(x)|^2 \leq NJ
\]
and
\[
\sup_{[0, t]} \sup_{|x| = r_1} |u_s(x)|^2 \leq 2 \sup_{[0, t]} \sup_{|x| = r_1} |u_s(x) - v_s(x)|^2 + NJ.
\]
By applying Theorem 3.1 to \( u_t - v_t \) we conclude that for \( t \in [0, T] \)
\[
\sup_{[0, t]} \sup_{B_{r_1}} |u_s - v_s|^2 \leq N \int_0^t \| \eta(u_s - v_s) \|_n^2 \, ds. \tag{5.7}
\]
Hence,
\[
\sup_{[0, t]} \sup_{|x| = r_1} |u_s(x)|^2 \leq N \int_0^t \| \eta(u_s - v_s) \|_n^2 \, ds + NJ \leq N \int_0^t \| \eta u_s \|_n^2 \, ds + NJ.
\]
Next, set
\[
P_t = \| u_t \zeta \|_n^2, \quad Q_t = \| u_t \eta \|_n^2
\]
and observe that
\[
P_t \leq 2Q_t + 2\| u_t (\zeta - \eta) \|_n^2,
\]
where, as in the proof of Theorem 3.2 \( u_t (\zeta - \eta) \) is infinitely differentiable with respect to \( x \) for \( (t, x) \in [0, T] \times \mathbb{R}^d \) and each derivative is a bounded continuous function in \( [0, T] \times \mathbb{R}^d \). In particular,
\[
\| u_t (\zeta - \eta) \|_n \leq N \sup_{[0, t] \times \mathbb{R}^d} |u_s (\zeta - \eta)(x)|
\]
\[
\leq N \sup_{[0, t]} \sup_{B_{r_1}} |u_s(x)| \leq N \sup_{[0, t]} \sup_{|x| = r_1} |u_s(x)|,
\]
where the last inequality is taken from Theorem 2.9. Hence,

\[ P_t \leq 2Q_t + N \int_0^t \| \eta u_s \|_n^2 \, ds + NJ \leq 2Q_t + N \int_0^t P_s \, ds + NJ. \quad (5.8) \]

Then

\[ Q_t \leq 2 \| (u_t - v_t) \eta \|_n^2 + 2 \| v_t \|_n^2 \leq N \sup_{B_{r_2}} | u_t - v_t |^2 + NJ \]

and similarly to (5.7)

\[ \sup_{B_{r_2}} | u_t - v_t |^2 \leq N \int_0^t \| \zeta (u_s - v_s) \|_n^2 \, ds \leq N \int_0^t P_s \, ds + NJ. \]

By coming back to (5.8) and using Gronwall’s inequality we conclude \( P_t \leq NJ \) and this proves (5.5) and the theorem.

In the following theorem we prove the existence of a kernel for our SPDE. It is worth drawing the reader’s attention to the fact that no continuity with respect to \((t, y, x)\) is claimed in (i) and (iv) and no continuity with respect to \(y\) is claimed in (iii).

**Theorem 5.4.** Assume that \( S = 0 \) and take \( r \in (0, R_0) \) and \( s_0 \in (0, T) \). Then there exists a nonnegative function \( p_t(x, y) = p_t(\omega, y, x) \) defined for \((\omega, t, y, x) \in \Omega \times (0, T] \times B_r \times \mathbb{R}^d\) such that

(i) it is infinitely differentiable with respect to \(y\) in \(B_r\), each of its \(y\)-derivative of any order is a bounded function of \((t, y, x) \in [s_0, T] \times B_r \times \mathbb{R}^d\), and the functions

\[ \int_{B_r} | D_y^\alpha p_t(y, x) |^2 \, dy \]

are bounded on \([s_0, T] \times B_r\) for any \( \omega \in \Omega \) and multi-index \(\alpha\),

(ii) if \( u_0 \in H^1_0(B_r) \) and \( f \equiv g^k \equiv 0, k = 1, \ldots, d_1 \), then with probability one \( u_t(x) \) coincides (as a generalized function with respect to \(x\)) with

\[ (u_0, p_t(\cdot, x)) = \int_{B_r} u_0(y)p_t(y, x) \, dy \]

for all \((t, x) \in (0, T] \times \mathbb{R}^d\).

(iii) for any \( y \in B_r \) with probability one for any \( t \in (0, T] \) and multi-index \(\alpha\) the function \( D_y^\alpha p_t(y, x) \) is infinitely differentiable with respect to \(x\) and each of its \(x\)-derivative of any order is a bounded and continuous function of \((t, x) \in [s_0, T] \times \mathbb{R}^d\),

(iv) the function \( p_t(y, x) \) is infinitely differentiable with respect to \((x, y) \in B_r \times B_r\) and each its derivative of any order is a bounded function of \((t, y, x) \in [s_0, T] \times B_r^2\) for any \( \omega \in \Omega \).
Proof. Let \( Z_+ = \{ \zeta_1, \zeta_2, \ldots \} \) be a countable subset of \( C_0^\infty(B_r) \) consisting of nonnegative functions such that \( Z_+ \) is dense in the sup-norm in the subset of \( C_0^\infty(B_r) \) consisting of nonnegative functions. Also set \( Z = Z_+ - Z_+ \).

For \( \zeta \in \mathcal{H}_2^{-l}(B_r) \) denote by \( u_t[\zeta](x) \) the normal solution of (1.4) with initial condition \( \zeta \) and with \( f \equiv g^k \equiv 0, k = 1, \ldots, d_1 \). By Theorems 5.3 and 2.9 for any \( l \geq 0 \) and \( t \in (0,T) \) with probability one there exists a constant \( N_t(l) \) such that for any \( x \in \mathbb{R}^d \) and \( \zeta \in \mathcal{H}_2^{-l}(B_r) \) we have
\[
|u_t[\zeta](x)| \leq N_t(l)
\]
and if \( \zeta \in Z_+ \) then by the maximum principle
\[
u_t[\zeta](x) \geq 0.
\]

Furthermore, owing to uniqueness for any collection of rational numbers \( p_1, p_2, \ldots \) only finite number of which is different from zero with probability one
\[
\left| \sum_i p_i u_t[\zeta_i](x) \right| = \left| u_t \left( \sum_i p_i \zeta_i \right)(x) \right| \leq N_t(l) \left\| \sum_i p_i \zeta_i \right\|_{-l}.
\]

Obviously, one can choose \( N_t(l) \) so that it is a monotone function of \( t \). Owing to this and the fact that other expressions entering (5.13) and (5.12) are continuous in \( (t, x) \) (a.s.) and the fact that the set of rational numbers is countable, there is a set \( \Omega' \) of full probability such that (5.13) is satisfied for any \( l \), any collection of rational numbers \( p_1, p_2, \ldots \), only finite number of which is different from zero, any \( \omega \in \Omega' \), and any \( (t, x) \in (0,T) \times \mathbb{R}^d \), and (5.12) is satisfied for any \( \omega \in \Omega' \) and any \( (t, x) \in (0,T) \times \mathbb{R}^d \) and \( \zeta \in Z \). By setting \( u_t[\zeta](x) = 0 \) outside \( \Omega' \), we may assume that \( \Omega' = \Omega \).

By a theorem of Hahn (see, for instance, Section II.5 of [6]), for any \( x \in \mathbb{R}^d, t \in (0,T), l \geq 0 \), and \( \omega \) there exists a linear functional \( Q_t[:](x) \) on \( H_2^{-l} \) such that
\[
Q_t[\zeta_i](x) = u_t[\zeta_i](x), \quad \forall i, \quad \|Q_t[:](x)\| \leq N_t(l).
\]
The general form of linear functionals \( Q \) on \( H_2^{-l} \) is well known and for smooth elements \( f \in H_2^{-l} \) it is given by
\[
Q(f) = \int_{\mathbb{R}^d} f(x)p(x) \, dx,
\]
where \( p \in H_2^l \) and \( \|p\|_l = \|Q\| \).

Hence, for any \( x \in \mathbb{R}^d, t \in (0,T), l \geq 0 \), and \( \omega \) there exists a function \( p_t(\cdot, x) \in H_2^l \) such that
\[
\|p_t(\cdot, x)\|_l \leq N_t(l), \quad \int_{\mathbb{R}^d} \zeta_i(y)p_t(y, x) \, dy = u_t[\zeta_i](x) \quad \forall i.
\]

In principle \( p_t(y, x) \) is not unique, it might be changed for \( y \notin B_r \). In addition it may depend on \( n \). To choose a better representative, without
loosing generality we take \( l = 2m \), where \( m = 1, 2, \ldots \), so that
\[
\|u\|_l = \|(1 - \Delta)^m u\|_0
\]
and then set \( p_t(y, x) = 0 \) for \( y \not\in B_r \). Then the new \( p_t(\cdot, x) \) for which we keep the same notation will satisfy the second relation in (5.14), hence will be independent of \( l \). Also for thus modified \( p_t(y, x) \) the functions (5.9) are bounded on \([s_0, T] \times B_r \) for any \( s_0 \in (0, T] \), \( \omega \in \Omega \), and multi-index \( \alpha \) since this was true for the initial \( p_t(y, x) \) even with \( \mathbb{R}^d \) in place of \( B_r \) in (5.9).

Since this holds for any multi-index, applying embedding theorems we obtain assertion (i). The second relation in (5.14) along with (5.13) and the choice of \( \{\zeta_i\} \) implies that \( p_t(y, x) \geq 0 \) for any \( x \in \mathbb{R}^d \), \( y \in B_r \), \( t \in (0, T] \), and \( \omega \).

Estimate (5.11), equations in (5.14), and the fact that \( \mathcal{Z} \) is dense in \( H^0(B_r) \) implies assertion (ii). The expressions in (5.10) are (s.s.) infinitely differentiable with respect to \( x \) and each of the derivatives is bounded and continuous in \([s_0, T] \times \mathbb{R}^d \) by Theorem 3.2. This proves assertion (iii) if we set \( u_0 = D^\alpha \delta_y \), where \( \delta_y \) is the delta-function concentrated at \( y \in B_r \).

To prove assertion (iv) observe that, as above, owing to Theorem 5.3, we can modify, if necessary, the functions \( u_0(\zeta)(x) \) for \( \zeta \in \mathcal{Z} \) in such a way that for any \( \omega \) and \( t \in (0, T] \) they will be infinitely differentiable with respect to \( x \) in \( B_r \) and for any multi-index \( \alpha \) and \( l > 0 \) satisfy
\[
\sup_{B_r} |D^\alpha u_l(\zeta)(x)| \leq N_l(|\alpha|, l)\|\zeta\|_{-l}, \tag{5.15}
\]
where \( N_l(|\alpha|, l) \) are independent of \( \zeta \). In particular, for \( x, x', x'' \in B_r \) and \( \zeta, \zeta', \zeta'' \in \mathcal{Z} \)
\[
|u_0(\zeta)(x') - u_0(\zeta)(x')| \leq N_l(1, l)\|\zeta\|_{-l}|x' - x''|,
\]
\[
|u_0(\zeta')(x) - u_0(\zeta'')(x)| \leq N_l(0, l)\|\zeta' - \zeta''\|_{-l}. \tag{5.16}
\]

Since \( \mathcal{Z} \) is dense in \( H^{-1}_0(B_r) \) and (5.14) holds, estimates (5.16) hold for all \( l \), \( \zeta, \zeta', \zeta'' \in H^{-1}_0(B_r) \), which after taking an appropriate \( l \) and \( \zeta, \zeta', \zeta'' \) as \( \delta \)-functions show that \( p_t(x, y) \) is a continuous function of \( (x, y) \) on \( B_r^2 \).

Estimate (5.15) also implies that the generalized derivatives of \((\zeta, p_t(\cdot, x))\) of order \( |\alpha| \) are bounded on \( B_r \) by the right-hand side of (5.15) for any \( \zeta \in H^{-1}_0(B_r) \). Since this holds for any \( \alpha \) and \((\zeta, p_t(\cdot, x))\) is continuous with respect to \( x \in B_r \) (owing to (5.16)), the generalized derivatives are, actually, usual ones, which are bounded and continuous.

By taking \( \zeta = D^\alpha y \) in (5.15), we find that the usual (and hence generalized) functions \( D^\alpha_y [D^\alpha_y p_t(y, x)] \) are bounded for \((x, y) \in B_r^2 \). It follows that \( p_t(y, x) \) admits a modification with respect to \((y, x)\) which is infinitely differentiable. However, the modification coincides with \( p_t(y, x) \) on \( B_r^2 \) since \( p_t(y, x) \) is continuous there.

The asserted boundedness of the derivatives of \( p_t(y, x) \) is easily derived from the above argument. The theorem is proved.
In the following theorem we assert the regularity of \( p_t(y, x) \) not only for \( y \in B_r \) but for all \( y \in \mathbb{R}^d \) albeit for \( x \in B_r \), the latter being of course inevitable. But the result is proved under a somewhat restrictive assumption. This assumption arose because of our inability to control \( \|u_t\|_n \) through \( \|u_0\|_n \times \) a (random) constant independent of \( u_0 \).

**Theorem 5.5.** Suppose that \( S > 0 \), take an \( R \in [0, \infty) \), \( r \in (0, R_0) \), \( s_0 \in (S, T) \), and assume that \( \sigma_t^k = 0 \) and \( \nu_t^k = 0 \) outside \( B_R \) for any \( k = 1, ..., d_1 \), \( t \), and \( \omega \). Then there exists a nonnegative function \( p_t(x, y) = p_t(\omega, y, x) \) defined for 

\[
(\omega, t, y, x) \in \Omega \times (S, T) \times \mathbb{R}^d \times B_r
\]

such that

(i) it is infinitely differentiable with respect to \((x, y) \in \mathbb{R}^d \times B_r\), each its derivative of any order is a bounded function of \((t, y, x) \in [s_0, T] \times \mathbb{R}^d \times B_r \) and the functions \( \|p_t(\cdot, x)\|_l \) are bounded on \([s_0, T] \times B_r \) for any \( \omega \in \Omega \) and \( l \geq 0 \),

(ii) if \( f \equiv g^k \equiv 0 \), \( k = 1, ..., d_1 \), then

\[
u_t(x) = \int_{\mathbb{R}^d} p_t(y, x)u_0(y) \, dy
\]

for all \((t, x) \in (S, T) \times B_r \).

Proof. It suffices to repeat the proof of Theorem 5.4 using the same notation as there, replacing \( Z \) with a countable subset of \( C_0^\infty(\mathbb{R}^d) \), and using Lemma 5.1 to conclude that for any and \( l \geq 0 \) and \( t \in (S, T) \) with probability one there exists a constant \( N_t(l) \) such that, for any \( \zeta \in \mathcal{H}_2^{0, l} \) and \( x \in B_r \), estimate (5.11) holds. After that the proof goes almost exactly the same way as that of Theorem 5.4. The theorem is proved.

**References**

[1] J.-M. Bismut and D. Michel, *Diffusions conditionnelles. I. Hypoellipticité partielle* (French), J. Funct. Anal., Vol. 44 (1981), No. 2, 174–211.

[2] Ju. N. Blagovesčenskii and M. I. Freidlin, *Some properties of diffusion processes depending on a parameter*, Dokl. Akad. Nauk SSSR, Vol. 138 (1961), 508–511, Russian, English translation in Sov. Math. Dokl., Vol. 2 (1961), 633–636.

[3] M. Chaleyat-Maurel, *Malliavin calculus application to the study of nonlinear filtering*, pp. 195–230 in The Oxford handbook of nonlinear filtering, Oxford Univ. Press, Oxford, 2011.

[4] M. Chaleyat-Maurel and M. Michel, *Hypoellipticity theorems and conditional laws*, Z. Wahrsch. Verw. Gebiete, Vol. 65 (1984), No. 4, 573–597.

[5] P. Cattiaux and L. Mesnager, *Hypoelliptic non-homogeneous diffusions*, Probab. Theory Related Fields. Vol. 123 (2002), No. 4, 453–483.

[6] N. Dunford and J.T. Schwartz, “Linear operators, Part I, General theory”, Interscience Publishers, New York, 1958.

[7] N.V. Krylov, *An analytic approach to SPDEs*, pp. 185-242 in Stochastic Partial Differential Equations: Six Perspectives, Mathematical Surveys and Monographs, Vol. 64, AMS, Providence, RI, 1999.
[8] N.V. Krylov, *Maximum principle for SPDEs and its applications*, pp. 311-338 in “Stochastic Differential Equations: Theory and Applications, A Volume in Honor of Professor Boris L. Rozovskii”, P.H. Baxendale, S.V. Lototsky eds., Interdisciplinary Mathematical Sciences, Vol. 2, World Scientific, 2007.

[9] N.V. Krylov, *Hörmander’s theorem for stochastic partial differential equations*, preprint.

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

[11] N.V. Krylov and B.L. Rozovsky, *Stochastic evolution equations*, “Itogy nauki i tekhniki”, Vol. 14, VINITI, Moscow, 1979, 71-146 in Russian; English translation in J. Soviet Math., Vol. 16 (1981), No. 4, 1233-1277.

[12] N.V. Krylov and B.L. Rozovsky, *On the characteristics of degenerate second order parabolic Itô equations*, Trudy seminara imeni Petrovskogo, Vol. 8 (1982), 153–168 in Russian, English transl. in J. Soviet Math., Vol. 32 (1986), No. 4, 336–348.

[13] H. Kunita, *Stochastic partial differential equations connected with nonlinear filtering*, Nonlinear filtering and stochastic control (Cortona, 1981), Lecture Notes in Math., Vol. 972 (1982), 100–169.

[14] H. Kunita, *Densities of a measure-valued process governed by a stochastic partial differential equation*, Systems Control Lett., Vol. 1 (1981/82), No. 2, 100–104.

[15] H. Kunita, “Stochastic flows and stochastic differential equations”, Cambridge Studies in Advanced Mathematics, Vol. 24, Cambridge University Press, Cambridge, 1990.

[16] S. Kusuoka and D. Stroock, *The partial Malliavin calculus and its application to nonlinear filtering*, Stochastics, Vol. 12 (1984), No. 2, 83–142.

[17] B. L. Rozovskii, “Stochastic evolution systems”, Kluwer, Dordrecht, 1990.

[18] A.D. Ventcel’, *On equations of the theory of conditional Markov processes*, Teoriya Veroyatnostei i ee Primeneniya, Vol. 10 (1965), No. 2, 390–393 in Russian; English transl. Theory Probability and Appl., Vol. 10 (1965), 357–361.

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

127 Vincent Hall, University of Minnesota, Minneapolis, MN, 55455