

**SPDES IN DIVERGENCE FORM WITH VMO
COEFFICIENTS AND FILTERING THEORY OF
PARTIALLY OBSERVABLE DIFFUSION PROCESSES
WITH LIPSCHITZ COEFFICIENTS**

N.V. KRYLOV

ABSTRACT. We present several results on the smoothness in L_p sense of filtering densities under the Lipschitz continuity assumption on the coefficients of a partially observable diffusion processes. We obtain them by rewriting in divergence form filtering equation which are usually considered in terms of formally adjoint to operators in nondivergence form.

1. INTRODUCTION

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

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

$$du_t = (L_t u_t + D_i f_t^i + f_t^0) dt + (\Lambda_t^k u_t + g_t^k) dw_t^k, \quad (1.1)$$

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

$$L_t \psi(x) = D_j (a_t^{ij}(x) D_i \psi(x) + a_t^j(x) \psi(x)) + b_t^i(x) D_i \psi(x) + c_t(x) \psi(x),$$

$$\Lambda_t^k \psi(x) = \sigma_t^{ik}(x) D_i \psi(x) + \nu_t^k(x) \psi(x),$$

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

One can rewrite (1.1) in the nondivergence form assuming that the coefficients a_t^{ij} and a_t^j are differentiable in x and then one could apply

2000 *Mathematics Subject Classification.* 60H15, 35R60.

Key words and phrases. Stochastic partial differential equations, divergence equations, filtering equations.

This work was partially supported by NSF grant DMS-0653121.

the results from [5]. It turns out that the differentiability of a_t^{ij} and a_t^j is not needed for the corresponding counterparts of the results in [5] to be true and showing this and generalizing the corresponding results of [3] is one of the main purposes of Section 2 of the present article. We assume, roughly speaking, that $a_t^{ij}(x)$ are measurable in t and of class VMO with respect to x .

One of the main motivations for developing the theory of SPDEs comes from filtering theory of partially observable diffusion processes. This problem is stated as follows. Let $d \geq 1$, $d_1 > d$ be integers .

Consider a d_1 -dimensional two component process $z_t = (x_t, y_t)$ with x_t being d -dimensional and y_t ($d_1 - d$)-dimensional. We assume that z_t is a diffusion process defined as a solution of the system

$$\begin{aligned} dx_t &= b(t, z_t)dt + \theta(t, z_t)dw_t, \\ dy_t &= B(t, z_t)dt + \Theta(t, y_t)dw_t \end{aligned} \tag{1.2}$$

with some initial data.

The coefficients of (1.2) are assumed to be vector- or matrix-valued functions of appropriate dimensions defined on $[0, T] \times \mathbb{R}^{d_1}$. Actually $\Theta(t, y)$ is assumed to be independent of x , so that it is a function on $[0, T] \times \mathbb{R}^{d_1-d}$ rather than $[0, T] \times \mathbb{R}^{d_1}$ but as always we may think of $\Theta(t, y)$ as a function of (t, z) as well.

The component x_t is treated as unobservable and y_t as the only observations available. The problem is to find a way to compute the density $\pi_t(x)$ of the conditional distribution of x_t given $y_s, s \leq t$. Finding an equation satisfied by π_t (filtering equation) is considered to be a solution of the (filtering) problem. The filtering equations turn out to be particular cases of SPDEs.

In 1964 in [14] the filtering equations were proposed in a somewhat nonrigorous way and most likely some terms in these equations appeared from stochastic integrals written in the Stratonovich form and the others appeared from the Itô integrals. Perhaps, the author of [14] realized this too and published an attempt to rescue some results of [14] in 1967 in [15]. This attempt turned successful for simplified models without the so-called cross terms.

Meanwhile, in 1966 in [20] the correct filtering equations in full generality, yet assuming some regularity of the filtering density, were presented. This is the reason we propose to call the filtering equations in the case of partially observable diffusion processes *Shiryayev's equations* and their particular case without cross terms *Kushner's equations*.

In case $d = 1$ the result of [20] is presented in [17] on the basis of the famous Fujisaki-Kallianpur-Kunita theorem (see [2]) about the

filtering equations in a very general setting. Some authors even call the filtering equation for diffusion processes the Fujisaki-Kallianpur-Kunita equation.

By adding to the Fujisaki-Kallianpur-Kunita theorem some simple facts from the theory of SPDEs, the a priori regularity assumption was removed in [9] and under the Lipschitz and uniform nondegeneracy assumption the L_2 -version of Theorem 3.2 was proved. The basic result of [9] is that $\pi_t \in W_2^1$. It is also proved that if the coefficients are smoother, $\pi_t(x)$ is smoother too. The nondegeneracy assumption was later removed (see [19]) on the account of assuming that $\theta\theta^*$ is three times continuously differentiable in x . It is again proved that $\pi_t \in W_2^1$ and π_t is even smoother if the coefficients are smoother.

In [5] the results of [9] were improved, $\theta\theta^*$ is assumed to be twice continuously differentiable in x and it is shown that $\pi_t \in W_p^2$ with any $p \geq 2$.

The above mentioned results of [9], [19], and [5] use filtering theory in combination with the theory of SPDEs, the latter being stimulated by certain needs of filtering theory. It turns out that the theory of SPDEs alone can be used to obtain the above mentioned regularity results about π_t without knowing anything from filtering theory itself. It also can be used to solve other problems from filtering theory.

The first “direct” (only using the theory of SPDEs) proof of regularity of π_t is given in [11] in the case that system (1.2) defines a nondegenerate diffusion process and $\theta\theta^*$ is twice continuously differentiable in x . It is proved that $\pi_t \in W_p^2$ with any $p \geq 2$ as in [5]. Advantages of having arbitrary p are seen from results like our Theorem 3.3. Of course, on the way of investigating π_t in [11] the filtering equations are derived “directly” in an absolutely different manner than before (on the basis of an idea from [10]).

In Section 3 of this article we relax the smoothness assumption in [11] to the assumption that the coefficients of (1.2) are merely Lipschitz continuous, the assumption which is almost always supposed to hold when one deals with systems like (1.2). We find that $\pi_t \in W_p^1$. Thus, under the weakest smoothness assumptions we obtain the best (in the author’s opinion) regularity result on π_t . In particular, we prove that if the initial data is sufficiently regular, then the filtering density is almost Lipschitz continuous in x and $1/2$ Hölder continuous in t . However, we still assume z_t to be nondegenerate. Our approach is heavily based on analytic results. There is also a probabilistic approach developed in [13] and based on explicit formulas for solutions introduced in [16]

and later developed in [10] and [12] (also see references therein). This approach cannot give as sharp results as ours in our situation.

It seems to the author that under the same assumptions of Lipschitz continuity, by following an idea from [4] one can solve another problem from filtering theory, the so-called innovation problem, and obtain the equality

$$\sigma\{y_s, s \leq t\} = \sigma\{\check{w}_s, s \leq t\},$$

where \check{w}_t is the innovation Wiener process of the problem (its definition is reminded in Section 3). Recall that for degenerate diffusion processes the positive solution of the innovation problem is obtained in [18] again on the basis of the theory of SPDEs under the assumption that the coefficients are more regular.

By the way, in our situation, if the coefficients are more regular, the filtering equation can be rewritten in a nondivergence form and then additional smoothness of the filtering density, existence of which is already established in this article, is obtained on the basis of regularity results from [5].

Although for the proof of the above mentioned results concerning the filtering equations it suffices to use article [3] about SPDEs in divergence form with continuous coefficients, we prefer to give more general results borrowed from [7] in Section 2. In Section 3 we present some results about the filtering equations from [8].

We finish this section by introducing some notation. Let $K, \delta > 0$ be fixed finite constants, $p \in [2, \infty)$. Denote $L_p = L_p(\mathbb{R}^d)$, $C_0^\infty = C_0^\infty(\mathbb{R}^d)$. Introduce

$$D_i = \frac{\partial}{\partial x^i}, \quad i = 1, \dots, d.$$

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

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

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

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

Recall that τ is a stopping time and introduce

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

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

- (i) $u_0 \in L_p(\Omega, \mathcal{F}_0, L_p)$;
- (ii) $u \in \mathbb{W}_p^1(\tau)$;
- (iii) There exist $f^i \in \mathbb{L}_p(\tau)$, $i = 0, \dots, d$, and $g = (g^1, g^2, \dots) \in \mathbb{L}_p(\tau)$ such that for any $\varphi \in C_0^\infty$ with probability 1 for all $t \in [0, \infty)$ we have

$$\begin{aligned} (u_{t \wedge \tau}, \varphi) &= (u_0, \varphi) + \sum_{k=1}^{\infty} \int_0^t I_{s \leq \tau}(g_s^k, \varphi) dw_s^k \\ &+ \int_0^t I_{s \leq \tau}((f_s^0, \varphi) - (f_s^i, D_i \varphi)) ds, \end{aligned} \quad (1.3)$$

where by (f, φ) we mean the action of a generalized function f on φ , in particular, if f is a locally summable,

$$(f, \varphi) = \int_{\mathbb{R}^d} f(x) \varphi(x) dx.$$

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

The reader can find in [5] a discussion of (ii) and (iii), in particular, the fact that the series in (1.3) converges uniformly in probability on every finite subinterval of $[0, \tau]$. In case that property (iii) holds, we write

$$du_t = (D_i f_t^i + f_t^0) dt + g_t^k dw_t^k \quad (1.4)$$

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

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

2. SPDES IN DIVERGENCE FORM WITH VMO COEFFICIENTS

We are considering (1.1) under the following assumptions.

Assumption 2.1. (i) The coefficients a_t^{ij} , a_t^i , b_t^i , σ_t^{ik} , c_t , and ν_t^k are measurable with respect to $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$, where $\mathcal{B}(\mathbb{R}^d)$ is the Borel σ -field on \mathbb{R}^d .

- (ii) For all values of indices and arguments

$$|a_t^i| + |b_t^i| + |c_t| + |\nu|_{\ell_2} \leq K, \quad c_t \leq 0.$$

- (iii) For all values of the arguments and $\xi \in \mathbb{R}^d$

$$a_t^{ij} \xi^i \xi^j \leq \delta^{-1} |\xi|^2, \quad (a_t^{ij} - \alpha_t^{ij}) \xi^i \xi^j \geq \delta |\xi|^2, \quad (2.1)$$

where $\alpha_t^{ij} = (1/2)(\sigma^{i \cdot}, \sigma^{j \cdot})_{\ell_2}$.

It is worth emphasizing that we do not require the matrix (a^{ij}) to be symmetric. Assumption 2.1 (i) guarantees that equation (1.1) makes perfect sense if $u \in \mathcal{W}_p^1(\tau)$.

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

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

where $|B|$ is the volume of B . If $\rho \geq 0$, set $B_\rho = \{x : |x| < \rho\}$ and for locally integrable $h_t(x)$ and continuous \mathbb{R}^d -valued function $x_r, r \geq 0$, introduce

$$\text{osc}_\rho(h, x) = \sup_{s \geq 0} \frac{1}{\rho^2} \int_s^{s+\rho^2} (|h_r - h_{r(B+x_r)}|)_{(B+x_r)} dr,$$

where $B = B_\rho$. Also for $y \in \mathbb{R}^d$ set

$$\text{Osc}_\rho(h, y) = \sup_{|x|_C \leq \rho} \sup_{r \leq \rho} \text{osc}_r(h, y + x),$$

where $|x|_C$ is the sup norm of $|x|$. Observe that $\text{osc}_\varepsilon h = 0$ if $h_t(x)$ is independent of x .

Denote by β_0 one third of the constant $\beta_0(d, p, \delta) > 0$ from Lemma 5.1 of [7].

Assumption 2.2. There exist a constant $\varepsilon \in (0, 1]$ such that for any $y \in \mathbb{R}_+^d$ (and ω) we have

$$\text{Osc}_\varepsilon(a^{ij}, y) \leq \beta_0, \quad \forall i, j. \quad (2.2)$$

Furthermore,

$$(a_t^{jk}(x) - \alpha_t^{jk}(y)) \xi^j \xi^k \geq \delta |\xi|^2$$

for all t, ξ , and x satisfying $|x - y| \leq \varepsilon$.

Let $\beta_1 = \beta_1(d, p, \delta, \varepsilon) > 0$ be the constant from Lemma 5.2 of [7].

Assumption 2.3. There exists a constant $\varepsilon_1 > 0$ such that for any $t \geq 0$ we have

$$|\sigma_t^{i \cdot}(x) - \sigma_t^{i \cdot}(y)|_{\ell_2} \leq \beta_1,$$

whenever $x, y \in \mathbb{R}_+^d, |x - y| \leq \varepsilon_1, i = 1, \dots, d$.

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

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

where $\mathbb{R}_+ = (0, \infty)$ and $H_p^{-1} = (1 - \Delta)^{-1/2} L_p$. For such functions there is a (unique) modification denoted again u such that u_t is a continuous

L_p -valued function on $[0, \infty)$ so that u_0 is well defined. Any such u_t is called an extension of u_0 .

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

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

over all extensions u_t of elements u_0 .

Theorem 2.1. *Let $f^j, g \in \mathbb{L}_p(\tau)$ and let $u_0 \in L_p(\Omega, \mathcal{F}_0, W_p^{1-2/p})$. Then*

(i) *Equation (1.1) for $t \leq T \wedge \tau$ has a unique solution $u \in \mathcal{W}_p^1(T \wedge \tau)$ with initial data u_0 for any constant $T \in (0, \infty)$.*

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

Assertion (ii) of Theorem 2.1 follows from assertion (i) and Theorem 2.4.

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

Theorem 2.2. *Assume that for each $n = 1, 2, \dots$ we are given functions $a_{nt}^{ij}, a_{nt}^i, b_{nt}^i, c_{nt}, \sigma_{nt}^{ik}, \nu_{nt}^k, f_{nt}^j, g_{nt}^k$, and u_{n0} having the same meaning and satisfying the same assumptions with the same $\delta, K, \varepsilon, \varepsilon_1, \beta_0$, and β_1 as the original ones. Assume that for $i, j = 1, \dots, d$ and almost all (ω, t, x) we have*

$$\begin{aligned} (a_{nt}^{ij}, a_{nt}^i, b_{nt}^i, c_{nt}) &\rightarrow (a_t^{ij}, a_t^i, b_t^i, c_t), \\ |\sigma_{nt}^i - \sigma_t^i|_{\ell_2} + |\nu_{nt} - \nu_t|_{\ell_2} &\rightarrow 0, \end{aligned}$$

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

$$\sum_{j=0}^d (\|f_n^j - f^j\|_{\mathbb{L}_p(\tau)} + \|g_n - g\|_{\mathbb{L}_p(\tau)} + \|u_{n0} - u_0\|_{L_p(\Omega, \mathcal{F}_0, W_p^{1-2/p})}) \rightarrow 0$$

as $n \rightarrow \infty$. Let u_n be the unique solutions of equations (1.1) for $t \leq \tau$ constructed from $a_{nt}^{ij}, a_{nt}^i, b_{nt}^i, c_{nt}, \sigma_{nt}^{ik}, \nu_{nt}^k, f_{nt}^j$, and g_{nt}^k and having initial values u_{n0} .

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

$$E \sup_{t \leq \tau \wedge T} \|u_{nt} - u_t\|_{L_p}^p \rightarrow 0.$$

In many situation the following maximum principle based on the results of [6] is useful.

Theorem 2.3. *Suppose that, for $q \in [2, p]$, Assumptions 2.2 and 2.3 are satisfied with $\beta_0 \leq \beta_0(d, q, \delta)$ and $\beta_1 \leq \beta_1(d, q, \delta, \varepsilon)$. Also suppose that $u_0 \in L_p(\Omega, \mathcal{F}_0, W_q^{1-2/q})$, $q \in [2, p]$, $u_0 \geq 0$, $f^i = 0$, $i = 1, \dots, d$,*

$f^0 \geq 0$, $g = 0$. Then for the solution u almost surely we have $u_t \geq 0$ for all finite $t \leq \tau$.

Part of the proofs of the above results is based on the following Itô's formula.

Theorem 2.4. *Let $u \in \mathcal{W}_p^1(\tau)$, $f^j \in \mathbb{L}_p(\tau)$, $g = (g^k) \in \mathbb{L}_p(\tau)$ and assume that (1.4) holds for $t \leq \tau$ in the sense of generalized functions. Then there is a set $\Omega' \subset \Omega$ of full probability such that*

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

$$\begin{aligned} \int_{\mathbb{R}^d} |u_{t \wedge \tau}|^p dx &= \int_{\mathbb{R}^d} |u_0|^p dx + p \int_0^{t \wedge \tau} \int_{\mathbb{R}^d} |u_s|^{p-2} u_s g_s^k dx dw_s^k \\ &+ \int_0^{t \wedge \tau} \left(\int_{\mathbb{R}^d} [p|u_t|^{p-2} u_t f_t^0 - p(p-1)|u_t|^{p-2} f_t^i D_i u_t \right. \\ &\quad \left. + (1/2)p(p-1)|u_t|^{p-2} |g_t|_{\ell_2}^2] dx \right) dt. \end{aligned} \quad (2.3)$$

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

$$\begin{aligned} E \sup_{t \leq \tau \wedge T} \|u_t\|_{L_p}^p &\leq 2E\|u_0\|_{L_p}^p + NT^{p-1} \|f^0\|_{\mathbb{L}_p(\tau)}^p \\ &+ NT^{(p-2)/2} \left(\sum_{i=1}^d \|f^i\|_{\mathbb{L}_p(\tau)}^p + \|g\|_{\mathbb{L}_p(\tau)}^p + \|Du\|_{\mathbb{L}_p(\tau)}^p \right), \end{aligned} \quad (2.4)$$

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

We have a direct proof of this result. However, (2.3) can also be obtained by extending some arguments from [1].

3. FILTERING EQUATIONS

Fix a constant $T \in (0, \infty)$ and for simplicity assume that w_t in (1.2) is finite dimensional. First we state and discuss our assumptions.

Assumption 3.1. The functions b , θ , B , and Θ are Borel measurable and bounded functions of their arguments. Each of them satisfies the Lipschitz condition in z with the constant K .

Assumption 3.2. The process z_t is uniformly nondegenerate: for any $\lambda, z \in \mathbb{R}^{d_1}$ and $t \in [0, T]$ we have

$$\tilde{a}_t^{ij}(z) \lambda^i \lambda^j \geq \delta |\lambda|^2,$$

where $2\tilde{a}_t(z) = 2(\tilde{a}_t^{ij}(z)) = \theta(t, z)\theta^*(t, z) + \Theta(t, y)\Theta^*(t, y)$.

Traditionally, Assumption 3.2 is split into two following assumptions the combination of which is equivalent to Assumption 3.2 and in which some useful objects are introduced. These assumptions were also used in the past to reduce system (1.2) to the so-called triangular form by replacing w_t with a different Brownian motion.

Assumption 3.3. The symmetric matrix $\Theta\Theta^*$ is invertible and

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

is a bounded function of (t, y) .

Assumption 3.4. For any $\xi \in \mathbb{R}^d$, $z = (x, y) \in \mathbb{R}^{d_1}$, and $t > 0$, we have

$$|Q(t, y)\theta^*(t, z)\xi|^2 \geq \delta|\xi|^2,$$

where Q is the orthogonal projector on $\text{Ker } \Theta$. In other words,

$$(\theta(I - \Theta^*\Psi^2\Theta)\theta^*\xi, \xi) \geq \delta|\xi|^2. \quad (3.1)$$

Assumption 3.5. The random vectors x_0 and y_0 are independent of the process w_t . The conditional distribution of x_0 given y_0 has a density, which we denote by $\pi_0(x) = \pi_0(\omega, x)$. We have $\pi_0 \in L_p(\Omega, W_p^{1-2/p})$.

Next we introduce few more notation. Let

$$\Psi_t = \Psi(t, y_t), \quad \Theta_t = \Theta(t, y_t), \quad a_t(x) = \frac{1}{2}\theta\theta^*(t, x, y_t), \quad b_t(x) = b(t, x, y_t),$$

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

In the remainder of the article we use the notation

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

only for $i = 1, \dots, d$ and set

$$L_t(x) = a_t^{ij}(x)D_iD_j + b_t^i(x)D_i, \quad (3.2)$$

$$\begin{aligned} L_t^*(x)u_t(x) &= D_iD_j(a_t^{ij}(x)u_t(x)) - D_i(b_t^i(x)u_t(x)) \\ &= D_j(a_t^{ij}(x)D_iu_t(x) - b_t^j(x)u_t(x) + u_t(x)D_ia_t^{ij}(x)), \end{aligned} \quad (3.3)$$

$$\Lambda_t^k(x)u_t(x) = \beta_t^k(x)u_t(x) + \sigma_t^{ik}(x)D_iu_t(x), \quad (3.4)$$

$$\begin{aligned} \Lambda_t^{k*}(x)u_t(x) &= \beta_t^k(x)u_t(x) - D_i(\sigma_t^{ik}(x)u_t(x)) \\ &= -\sigma_t^{ik}(x)D_iu_t(x) + (\beta_t^k(x) - D_i\sigma_t^{ik}(x))u_t(x), \end{aligned} \quad (3.5)$$

where $t \in [0, T]$, $x \in \mathbb{R}^d$, $k = 1, \dots, d_1 - d$, and as above we use the summation convention. Observe that Lipschitz continuous functions have bounded generalized derivatives and by

$$D_ia_t^{ij}, \quad D_i\sigma_t^{ik}$$

we mean these derivatives. Obviously, the operator L defined by (3.2) is uniformly elliptic with constant of ellipticity δ .

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

Let us consider the following initial value problem

$$d\bar{\pi}_t(x) = L_t^*(x)\bar{\pi}_t(x) dt + \Lambda_t^{k*}(x)\bar{\pi}_t(x)\Psi_t^{kr} dy_t^r, \quad (3.6)$$

$$\bar{\pi}_0(x) = \pi_0(x),$$

where $t \in [0, T]$, $x \in \mathbb{R}^d$, and $\bar{\pi}_t(x) = \bar{\pi}_t(\omega, x)$. Equation (3.6) is called the Duncan-Mortensen-Zakai or just the Zakai equation.

We understand this equation and the initial condition in the following sense. We are looking for a function $\bar{\pi} = \bar{\pi}_t(x) = \bar{\pi}_t(\omega, x)$, $\omega \in \Omega$, $t \in [0, T]$, $x \in \mathbb{R}^d$, such that

- (i) For each (ω, t) , $\bar{\pi}_t(\omega, x)$ is a generalized function on \mathbb{R}^d ,
- (ii) We have $\bar{\pi} \in L_p(\Omega \times [0, T], \mathcal{P}, W_p^1)$,
- (iii) For each $\varphi \in C_0^\infty(\mathbb{R}^d)$ with probability one for all $t \in [0, T]$ it holds that

$$\begin{aligned} (\bar{\pi}_t, \varphi) &= (\pi_0, \varphi) - \int_0^t (a_t^{ij} D_i \bar{\pi}_t - b_t^j \bar{\pi}_t + \bar{\pi}_t D_i a_t^{ij}, D_j \varphi) dt \\ &- \int_0^t (\sigma_t^{ik} D_i \bar{\pi}_t + (D_i \sigma_t^{ik} - \beta_t^k) \bar{\pi}_t, \varphi) \Psi_t^{kr} (B^r(t, z_t) dt + \Theta^{rs}(t, y_t) dw_t^s). \end{aligned} \quad (3.7)$$

Observe that all expressions in (3.7) are well defined due to the fact that the coefficients of $\bar{\pi}$ and of $D_i \bar{\pi}$ are bounded and appropriately measurable and $\bar{\pi}, D_i \bar{\pi} \in L_p(\Omega \times [0, T], \mathcal{P}, L_p)$.

Hence, equation (3.6) has the same form as (1.1) and the existence and uniqueness part of Lemma 3.1 below follow from Theorem 2.1. The second assertion of the lemma follows from Theorem 2.3.

Lemma 3.1. *There exists a unique solution $\bar{\pi}$ of (3.6) with initial condition π_0 in the sense explained above. In addition, $\bar{\pi}_t \geq 0$ for all $t \in [0, T]$ (a.s.).*

Here is a basic result of filtering theory for partially observable diffusion processes. Its relation to the previously known ones is discussed above.

Theorem 3.2. *Let $\bar{\pi}$ be the function from Lemma 3.1. Then*

$$0 < \int_{\mathbb{R}^d} \bar{\pi}_t(x) dx = (\bar{\pi}_t, 1) < \infty \quad (3.8)$$

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

$$E[f(x_t)|\mathcal{F}_t^y] = \frac{(\bar{\pi}_t, f)}{(\bar{\pi}_t, 1)} \quad (a.s.). \quad (3.9)$$

Equation (3.9) shows (by definition) that

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

is a conditional density of distribution of x_t given $y_s, s \leq t$. Since, generally, $(\bar{\pi}_t, 1) \neq 1$, one calls $\bar{\pi}_t$ an unnormalized conditional density of distribution of x_t given $y_s, s \leq t$.

We derive Theorem 3.2 from Theorem 2.2 and the result of [11] where more regularity on the coefficients is assumed.

The following is a direct corollary of embedding theorems from [5].

Theorem 3.3. *Let π_0 be a nonrandom function and $\pi_0 \in W_p^{1-2/p}$ for all $p \geq 2$, which happens for instance, if π_0 is a Lipschitz continuous function with compact support. Then for any $\varepsilon \in (0, 1/2)$ almost surely $\bar{\pi}_t(x)$ is $1/2 - \varepsilon$ Hölder continuous in t with a constant independent of x , $\bar{\pi}_t(x)$ is $1 - \varepsilon$ Hölder continuous in x with a constant independent of t , and the above mentioned (random) constants have all moments.*

REFERENCES

- [1] Z. Brzezniak, J. M. A. M. van Neerven, M. C. Veraar, and L. Weis, *Ito's formula in UMD Banach spaces and regularity of solutions of the Zakai equation*, J. Differential Equations, Vol. 245 (2008), No. 1, 30-58.
- [2] M. Fujisaki, G. Kallianpur, and H. Kunita, *Stochastic differential equations for the non linear filtering problem*, Osaka J. Math., Vol. 9 (1972), 19-40.
- [3] Kyeong-Hun Kim, *On L_p -theory of stochastic partial differential equations of divergence form in C^1 domains*, Probab. Theory Related Fields, Vol. 130 (2004), No. 4, 473-492.
- [4] N.V. Krylov, *On the equivalence of σ -algebras in the filtering problem of diffusion processes*, Teor. Veroyatnost. i Primenen, Vol. 24 (1979), No. 4, 771-780 in Russian; English translation: Theor. Probability Appl., Vol. 24 (1980), No. 4, 772-781.
- [5] 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.
- [6] 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.
- [7] N.V. Krylov, *On divergence form SPDEs with VMO coefficients*, SIAM J. Math. Anal. Vol. 40 (2009), No. 6, 2262-2285.

- [8] N.V. Krylov, *Filtering equations for partially observable diffusion processes with Lipschitz continuous coefficients*, to appear in The Oxford Handbook of Non-linear Filtering, Oxford University Press.
- [9] N.V. Krylov and B.L. Rozovskii, *On conditional distributions of diffusion processes*, Izvestiya Akademii Nauk SSSR, seriya matematicheskaya, Vol. 42 (1978), No. 2, 356-378 in Russian; English translation in Math. USSR Izvestija, Vol. 12 (1978), No. 2, 336-356.
- [10] N.V. Krylov and B.L. Rozovskii, *On the first integrals and Liouville equations for diffusion processes*, pp. 117-125 in "Stochastic Differential Systems, Proc. 3rd IFIP-WG 7/1 Working Conf., Visegrád, Hungary, Sept. 15-20, 1980", Lecture Notes in Contr. Inform. Sci., Vol. 36, 1981.
- [11] N.V. Krylov and A. Zatezalo, *A direct approach to deriving filtering equations for diffusion processes*, Applied Math. Optimiz., Vol. 42 (2000), no. 3, 315-332.
- [12] H. Kunita, *First order stochastic partial differential equations*, Stochastic analysis (Katata/Kyoto, 1982), 249-269, North-Holland Math. Library, Vol. 32, North-Holland, Amsterdam, 1984.
- [13] H. Kunita, "Stochastic flows and stochastic differential equations", Reprint of the 1990 original, Cambridge Studies in Advanced Mathematics, Vol. 24, Cambridge University Press, Cambridge, 1997.
- [14] H. J. Kushner, *On the differential equations satisfied by conditional probability densities of Markov processes, with applications*, J. Soc. Indust. Appl. Math. Ser. A Control, Vol. 2 (1964), 106-119.
- [15] H. J. Kushner, *Dynamical equations for optimal nonlinear filtering*, J. Differential Equations, Vol. 3 (1967), 179-190.
- [16] E. Pardoux, *Stochastic partial differential equations and filtering of diffusion processes*, Stochastics, Vol. 3 (1979), No. 2, 127-167.
- [17] 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.
- [18] O.G. Purtukhia, *The innovation problem for degenerate diffusion processes (growing coefficients)*, Uspekhi Mat. Nauk, Vol. 39 (1984), No. 4 (238), 177-178 in Russian; English translation: Russ. Math. Surv., Vol. 39 (1984), No. 4, 137-138.
- [19] B.L. Rozovskii, "Stochastic evolution systems", Kluwer, Dordrecht, 1990.
- [20] A. N. Shiryaev, *On stochastic equations in the theory of conditional Markov process*, Teor. Veroyatnost. i Primenen., Vol. 11 (1966), 200-206 in Russian; English translation: Theor. Probability Appl. 11 (1966), 179-184.

127 VINCENT HALL, UNIVERSITY OF MINNESOTA, MINNEAPOLIS, MN, 55455,
USA

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