November 97

Mannheimer Manuskripte 229/97
Fakultät für Mathematik und Informatik

Explicit Strong Solutions of SPDE's with Applications to Non-Linear Filtering

F. E. Benth, T. Deck, J. Potthoff und G. Våge

Explicit Strong Solutions of SPDE's with Applications to Non-Linear Filtering

F. E. Benth, Th. Deck[1], J. Potthoff[1], G. Våge[2]

Abstract. A differential calculus for random fields is developed and combined with the S–transform to obtain an explicit strong solution of the Cauchy problem

$$\begin{aligned} du(t,x) &= (Lu+cu)(t,x)\,dt + \sum_{i=1}^{m} h_i u(t,x)\,dY_t^i, \\ u(0,x) &= u_0(x), \quad x \in \mathbb{R}^d. \end{aligned}$$

Here L is a linear second order elliptic operator, h_i and c are real functions, and $Y_t^i = \int_0^t \psi^i(s)\,ds + W_t^i$, where W_t is a Brownian motion. An application of the solution to non-linear filtering and mathematical finance is also considered.

1 Introduction

Let W_t^1 and W_t^2 be independent Brownian motions on the probability space $(\Omega, \mathcal{F}, P)$, and suppose the one–dimensional diffusion

$$dX_t = b(X_t)\,dt + \sigma(X_t)\,dW_t^1$$

is observed by

$$dY_t = h(X_t)\,dt + dW_t^2.$$

The real-valued functions b and σ are assumed to be Lipschitz continuous and of linear growth, and h is assumed to be bounded and measurable. The non-linear filtering problem is to find the conditional expectation $E[f(X_t)|\mathcal{F}_t]$, where $\mathcal{F}_t$ is the σ-algebra generated by $\{Y_s;\ 0 \le s \le t\}$, and f is some bounded measurable function. From general probability theory, we know that

$$E[f(X_t)|\mathcal{F}_t](\omega) = \int_{\mathbb{R}} f(x) P[X_t \in dx|\mathcal{F}_t](\omega) = \int_{\mathbb{R}} f(x) p(t,x,\omega)\,dx,$$

if the conditional probability $P[X_t \in B|\mathcal{F}_t]$ is absolutely continuous with respect to Lebesgue measure. From the Fujisaki-Kallianpur-Kunita equation (see e.g. [1]), we obtain the following stochastic partial differential equation for the conditional density $p(t,x,\omega)$:

$$dp(t,x,\omega) = A^* p(t,x,\omega)\,dt + p(t,x,\omega)\left[h(x) - \int_{\mathbb{R}} h(y)p(t,y,\omega)\,dy\right] d\nu_t(\omega). \qquad (1.1)$$

[1]Partially supported by the DFG.
[2]Supported by the DFG and the Research Council of Norway.

A^* is the adjoint of the infinitesimal generator of X_t and ν_t is the so-called *innovation process* defined by $d\nu_t = dY_t - E[X_t|\mathcal{F}_t]\,dt$. Since (1.1) is difficult to solve, M. Zakai (see [2]) introduced the so-called *unnormalized conditional density* in the following manner: Define the probability measure $\overline{P}$ by $d\overline{P} = \Lambda_T\,dP$, where $T > 0$ is fixed, and

$$\Lambda_t = \exp\left(\int_0^t h(X_s)\,dY_s - \frac{1}{2}\int_0^t h(X_s)^2\,ds\right).$$

Under $\overline{P}$, $(Y_t)_{0\le t\le T}$ becomes a standard Brownian motion. Let $E_{\overline{P}}$ denote the expectation with respect to $\overline{P}$. The unnormalized conditional density $u(t,x,\omega)$ is now defined as

$$u(t,x,\omega) = E_{\overline{P}}[\Lambda_t|\mathcal{F}_t](\omega)\,p(t,x,\omega).$$

This relation readily implies

$$p(t,x,\omega) = \frac{u(t,x,\omega)}{\int_{\mathbb{R}} u(t,y,\omega)\,dy},$$

which explains the name of u. It also shows that the original problem of finding $E[f(X_t)|\mathcal{F}_t]$ can be reduced to finding the density u. But this latter problem is less difficult to treat, because u must satisfy the following *linear* (in contrast to (1.1)) stochastic partial differential equation (see [2]),

$$\begin{aligned} du(t,x,\omega) &= A^*u(t,x,\omega)\,dt + h(x)u(t,x,\omega)\,dY_t(\omega) \\ u(0,x,\omega) &= u_0(x). \end{aligned} \tag{1.2}$$

Here u_0 is the density function of the initial condition X_0. (1.2) is known as the *Zakai equation.* We will construct an explicit strong solution for this type of Cauchy problem. By a *strong solution* of (1.2) on $[0,T]\times\mathbb{R}$, we mean a random field $u : [0,T]\times\mathbb{R}\times\Omega \to \mathbb{R}$ which has the following properties:

(S1) There is an $N \in \mathcal{F}$ with $P(N) = 0$ so that $u(\cdot,\cdot,\omega) \in C^{0,2}([0,T]\times\mathbb{R})$ for all ω in the complement N^c of N.

(S2) For all $x \in \mathbb{R}$ the process $(t,\omega) \mapsto u(t,x,\omega)$ is jointly measurable, $u(t,x,\cdot)$ is $\mathcal{F}_t$-adapted and Itô-integrable over $[0,T]$.

(S3) For all $(t,x) \in [0,T]\times\mathbb{R}$ the following equation is satisfied:

$$u(t,x) = u_0(x) + \int_0^t A^*u(s,x)\,ds + \int_0^t h(x)u(s,x)\,dY_s \quad \text{a.s.}$$

The concept of strong solutions is extended to higher space dimensions and time-dependent A^* and h in the obvious way.

Existence and uniqueness results for (1.2) have been found by many authors. See e.g. the works by I. Gyöngy and N. V. Krylov [3, 4] and the references therein. Explicit formulas for the solution of equations closely related to (1.2) have been derived by E. Pardoux [5, 6], H. Kunita [7] and F. E. Benth [8].

The present work extends those results of Benth which deal with deterministic initial conditions. Our main improvement is that we obtain strong solutions. (The "generator definition" of A used by Benth in [8] essentially avoids differentiability considerations.) Pardoux derives a formula for the solution similar to ours, but he considers the backward equation related to (1.2), and his solution concept is weaker than ours. He uses Sobolev spaces with the associated notion of weak derivatives. In principle, it should be possible to prove that his formula provides a strong solution by applying the Sobolev embedding theorem. But this would require to increase the smoothness of h and of the coefficients in A^* considerably (depending on the dimension of the space variable x). On page 308 in [7], Kunita gives a general formula for the solution in the strong sense. This formula also applies for equations of type (1.2) which contain stochastic drifts in addition. Our formula can be derived from the one in [7], if one sets the drift terms equal to zero, transforms the backward integrals into forward integrals (with respect to the Brownian motion $\tilde{B}_s := B_{t-s} - B_t$, $s \in [0,t]$), and rewrites the resulting formula from Stratonovich to Itô form. We noticed this non-obvious relation after having finished the main body of the present paper.

The main rationale for the present paper consists in the following: The most important point is that our result applies for a class of unbounded, degenerate diffusion and drift coefficients. (The example discussed in Section 5 is of this type.) It is also stronger than the corresponding result by Pardoux (but we should mention that Pardoux can handle related filtering problems which are beyond our method), and it covers Kunita's result under weaker conditions (e.g., we require C^2 instead of C^4 differentiability for coefficients σ^{ij}.) Moreover, our proof is much more elementary than the method of inverse stochastic flows used by Kunita. Our proof is based essentially on Kolmogorov's continuity theorem for random fields.

Beside these mathematical reasons we believe that the formula for the solution can be of some practical, numerical interest, in particular in applications of non-linear filtering. (The statement of the conditions and the results in Pardoux [5, 6] and in Kunita [7] may be not so well suited for this purposes.) We tried to state our main theorem (Theorem 2.7) in such a way that it is accessible without too much effort for those who are interested in concrete applications.

The paper is organized as follows. After preparations and the statement of the main result in Section 2, we develop a differential calculus (with respect to space–time parameters) for random fields in Section 3. As far as we know this has not been done before. For instance, it can not be found in the standard references [7, 9], although the main techniques are provided there. This calculus may also be useful in other related contexts. In Section 4 we prove the main result. The final Section 5 is devoted to an application in mathematical finance. We point out how one can use our solution numerically to obtain optimal estimates of the parameters arising in the log-normal model for stocks. As a by-product we obtain an existence and uniqueness result, and a Feynman-Kac formula for the (deterministic) Cauchy problem

$$\frac{\partial u}{\partial t} = \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 u}{\partial x^2} + \mu x \frac{\partial u}{\partial x} - q(t,x)u, \qquad u(0,\cdot) = u_0.$$

2 Preparations and statement of the main result

We will consider the following spaces of real valued functions:

$C_0^\infty(\mathbb{R}_+)$: Infinitely differentiable functions on $\mathbb{R}_+$ with compact support.

$C_b^{n,\beta}(\mathbb{R}^d)$: n times continuously differentiable functions on $\mathbb{R}^d$ with all derivatives bounded, and highest derivative Hölder continuous of order $0 < \beta \leq 1$.

$C^{n,m}(\mathbb{R}_+ \times \mathbb{R}^d)$: Continuously differentiable functions f (n times in $t \in \mathbb{R}_+$, m times in $x \in \mathbb{R}^d$). An index b denotes that f and all derivatives are bounded.

$C^{n+\beta}(\mathbb{R}_+ \times \mathbb{R}^d)$: The space of functions on $\mathbb{R}_+ \times \mathbb{R}^d$ for which there exists a constant K such that

$$|D_x^\gamma f(t,x) - D_x^\gamma f(t',x')| \leq K(|t-t'|^{\beta/2} + |x-x'|^\beta),$$

for all $(t,x), (t',x') \in \mathbb{R}_+ \times \mathbb{R}^d$, $\gamma \in \mathbb{N}_0^d$ with $0 \leq |\gamma| \leq k$, and $0 < \beta \leq 1$.

Function spaces with $\mathbb{R}_+ \times \mathbb{R}^d$ substituted by $D_T := [0,T] \times \mathbb{R}^d$ are defined analogously.

Let vw be the usual scalar product for vectors $v, w \in \mathbb{R}^d$, and $\|\xi\|_2^2 := \int_0^\infty \xi^2(t)dt$ for $\xi \in L^2(\mathbb{R}_+)^m$. Let $(W_t)_{t\geq 0}$ be an m-dimensional standard Brownian motion on a probability space $(\Omega_1, \mathcal{F}_1, P_1)$, define $(\mathcal{F}_1)_t := \sigma\{W_s, 0 \leq s \leq t\}$, and

$$\mathcal{E}^\xi := \exp\left(\int_0^\infty \xi(s)\, dW_s - \frac{1}{2}\|\xi\|_2^2\right).$$

We suppose that $\mathcal{F}_1$ is the σ-algebra generated by the Brownian motion. Recall that under this assumption the algebra of functions generated by $\{\mathcal{E}^\xi; \xi \in C_0^\infty(\mathbb{R}_+)^m\}$ is dense in $L^2(\Omega_1, \mathcal{F}_1, P_1)$. For $f \in L^2(P_1)$ we define the $\mathcal{S}$-transform of f, $\mathcal{S}f : C_0^\infty(\mathbb{R}_+)^m \to \mathbb{R}$, as

$$\mathcal{S}f(\xi) = E\left[f\mathcal{E}^\xi\right].$$

Remark. Usually the $\mathcal{S}$-transform is defined (and applied) in the context of white noise analysis. In [10] this transform is discussed (and its usefulness demonstrated) in the general setting of probability theory. In the present paper – except for the $\mathcal{S}$-transform – we do not refer to white noise techniques at all. The few elementary facts about the $\mathcal{S}$-transform which we need are collected in the following two examples, and can be found in [10].

Example 2.1 *Fix $t \geq 0$ and $h \in L^2(\mathbb{R}_+)^m$. Let $\mathcal{E}_t^h := \mathcal{E}^{h1_{[0,t]}}$ denote the exponential martingale. Then the $\mathcal{S}$-transform of $\mathcal{E}_t^h$ reads*

$$\mathcal{S}\mathcal{E}_t^h(\xi) = \exp\int_0^t h(s)\xi(s)\, ds. \tag{2.1}$$

Example 2.2 *The $\mathcal{S}$-transform of an Itô integral is given by Theorem 3.3 in [10]. If $X_t = (X_t^1, \ldots, X_t^m)$ is Itô integrable over the interval $[0,T]$, then*

$$\mathcal{S}\int_0^T X_t\, dW_t(\xi) = \int_0^T (\mathcal{S}X_t)(\xi)\xi(t)\, dt.$$

Next we consider problem (1.2). Since our result will be independent from the filtering context we write $L+c$ instead of A^*. (A^*, in contrast to A, usually contains a zero order term c, cf. the example in Section 5.) For $(t,x) \in I\!R_+ \times I\!R^d$ let

$$L(t,x) = \frac{1}{2}\sum_{i,j=1}^{d} a^{ij}(t,x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b^i(t,x)\frac{\partial}{\partial x_i}, \tag{2.2}$$

where $a(t,x) = \sigma(t,x)\sigma(t,x)^T$ and σ^{ij}, $b^i : I\!R_+ \times I\!R^d \to I\!R^d$ are continuous, bounded, and globally Lipschitz continuous in x with a Lipschitz constant which does not depend on t. (Different conditions will be used in Theorem 2.7.) Suppose that $c \in C_b(I\!R_+ \times I\!R^d)$, $h = (h_1, \ldots, h_m) \in C_b(I\!R_+ \times I\!R^d)^m$, $u_0 \in C_b(I\!R^d)$, and consider the stochastic partial differential equation

$$u(t,x) = u_0(x) + \int_0^t \{L(s,x) + c(s,x)\}u(s,x)\,ds + \int_0^t h(s,x)u(s,x)\,dY_s, \tag{2.3}$$

where $Y_t^i = \int_0^t \psi_s^i\,ds + W_t^i$, $1 \le i \le d$, and the $(\mathcal{F}_1)_t$-adapted process $\psi_t(\omega_1)$ is such that

$$\mathcal{E}_t^{\psi} := \exp\left(\int_0^t \psi_s\,dW_s - \frac{1}{2}\int_0^t |\psi_s|^2\,ds\right) \tag{2.4}$$

is a martingale. Fix $T > 0$, then Girsanov's theorem implies that $(Y_t)_{0 \le t \le T}$ is an m-dimensional Brownian motion with respect to the measure $\overline{P}_1$ given by

$$d\overline{P}_1 = \mathcal{E}_T^{\psi} dP_1.$$

Let us assume for a moment that u is continuously x-differentiable in the $L^2(\overline{P}_1)$-sense up to the second order, that the $L^2(\overline{P}_1)$-norm $\|D_x^{\gamma}u(t,x)\|$, $0 \le |\gamma| \le 2$, is bounded on every strip $(s,T] \times I\!R^d \subset [0,T] \times I\!R^d$, and that u satisfies (2.3). To determine a representation formula for $u(t,x)$ we proceed as follows. We apply the $\mathcal{S}$-transform with respect to $\overline{P}_1$, i.e., we multiply both sides of (2.3) by the normalized exponential,

$$\bar{\mathcal{E}}^{\xi} := \exp\left(\int_0^{\infty} \xi(s)\,dY_s - \frac{1}{2}\|\xi\|_2^2\right), \; \xi \in C_0^{\infty}(I\!R_+)^m,$$

and compute the expectation with respect to $\overline{P}_1$. Because of $L^2(\overline{P}_1)$-continuity we can interchange the $\mathcal{S}$-transform with the integrals and with the partial derivatives. From Example 2.2 we obtain

$$v(t,x;\xi) = u_0(x) + \int_0^t \{L(s,x) + c(s,x) + h(s,x)\xi(s)\}v(s,x;\xi)\,ds, \tag{2.5}$$

where $v(t,x;\xi) = E_{\overline{P}_1}\left[u(t,x)\bar{\mathcal{E}}^{\xi}\right]$. Since the r.h.s. of (2.5) is t-differentiable we find

$$\frac{\partial v}{\partial t} = (L + c + h\xi)v, \; v\,|_{t=0} = u_0. \tag{2.6}$$

Thus $v(t,x;\xi)$ can be represented by the Feynman-Kac formula (see [11], p. 132)

$$\begin{aligned} v(t,x;\xi) &= E_{P_2}\left[u_0(X_t^{t,x})\exp\int_0^t \left(c(t-s,X_s^{t,x}) + h(t-s,X_s^{t,x})\xi(t-s)\right)ds\right] \\ &= E_{P_2}\left[u_0(X_t^{t,x})\exp\int_0^t \left(c(s,X_{t-s}^{t,x}) + h(s,X_{t-s}^{t,x})\xi(s)\right)ds\right], \end{aligned} \tag{2.7}$$

for $(t,x)\in \mathbb{R}_+\times \mathbb{R}^d$ and $\xi\in C_0^\infty(\mathbb{R}_+)^m$. Here the d-dimensional process $(X_s^{t,x})_{s\geq 0}$ solves the Itô equation

$$dX_s^{t,x} = b(t-s,X_s^{t,x})\,ds + \sigma(t-s,X_s^{t,x})\,dB_s, \quad X_0^{t,x}=x, \tag{2.8}$$

where we have extended the coefficients to negative times by defining $b(-s,x):=b(s,x)$, $\sigma(-s,x):=\sigma(s,x)$ for $(s,x)\in\mathbb{R}_+\times\mathbb{R}^d$, and $(B_s)_{s\geq 0}$ is a d-dimensional Brownian motion defined on an auxiliary probability space $(\Omega_2,\mathcal{F}_2,P_2)$. We are thus led to consider the probability space $\Omega=\Omega_1\times\Omega_2$, $\mathcal{F}=\mathcal{F}_1\otimes\mathcal{F}_2$, and $P=\overline{P}_1\otimes P_2$. In the following the processes $X^{t,x}$ and Y will be extended to processes on the product space $(\Omega,\mathcal{F},P)$, via $X^{t,x}(\omega_1,\omega_2)\equiv X^{t,x}(\omega_2)$, $Y(\omega_1,\omega_2)\equiv Y(\omega_1)$.

In view of Example 2.1 one expects that $v(t,x;\xi)$ is the $\mathcal{S}$-transform of

$$u(t,x,\omega_1) := E_{P_2}\left[u_0(X_t^{t,x})e^{\int_0^t c(s,X_{t-s}^{t,x})\,ds+\int_0^t h(s,X_{t-s}^{t,x})\,dY_s(\omega_1)-\frac{1}{2}\int_0^t|h(s,X_{t-s}^{t,x})|^2\,ds}\right]. \tag{2.9}$$

Proposition 2.3 *Let $(t,x)\in\mathbb{R}_+\times\mathbb{R}^d$, $(X_s^{t,x})_{0\leq s\leq t}$ be a measurable process on $(\Omega_2,\mathcal{F}_2,P_2)$, $c,h_i\in C_b([0,t]\times\mathbb{R}^d)$ for $1\leq i\leq m$, and $u_0\in C_b(\mathbb{R}^d)$. Then $u(t,x)$ defined by (2.9) is in $L^p(\overline{P}_1)$ for all $p\geq 1$, and the $\mathcal{S}$-transform of $u(t,x)$ is given by (2.7).*

Proof: By Schwarz' inequality and Fubini's theorem

$$\begin{aligned} E_{\overline{P}_1}[|u(t,x)|^p] &\leq \|u_0\|_\infty^p e^{pt\|c\|_\infty} E_{P_2}E_{\overline{P}_1}\left[e^{p\int_0^t h(s,X_{t-s}^{t,x})\,dY_s-\frac{p}{2}\int_0^t|h(s,X_{t-s}^{t,x})|^2\,ds}\right] \\ &= \|u_0\|_\infty^p \exp\left(pt(\|c\|_\infty+\frac{1}{2}(p-1)\|h\|_\infty^2)\right) < \infty, \end{aligned}$$

where the equality follows from (2.1). Similarly we obtain for every $\xi\in C_0^\infty(\mathbb{R}_+)^m$ the estimate $E_{\overline{P}_1}[|\bar{\mathcal{E}}^\xi|^p] = \exp\{\frac{1}{2}(p^2-p)\|\xi\|_2^2\} < \infty$ for all $p\geq 1$. These estimates show that one can interchange the expectations $E_{\overline{P}_1}$ and E_{P_2} which arise in the $\mathcal{S}$-transform of $u(t,x)$. Example 2.2 concludes the proof. □

To derive (2.7) we assumed that (2.3) has a solution. In this paper we will assume that (2.6) is solved by (2.7) and we will give a straightforward proof that the random field u defined in (2.9) is the unique strong solution of (2.3): We will verify that the random fields on the r.h.s. of (2.9) are sufficiently smooth and that the standard rules of calculus hold. These rules applied to (2.5) will finally allow us to derive (2.3).

We remark that this method generalizes the direct methods which work for non-stochastic parabolic equations, such as Kolmogorov's backward equation, cf. [13], [14].

To simplify the notation we introduce the difference operators Δ^i for $i = 1, \ldots, d$, for functions defined on $\mathbb{R}_+ \times \mathbb{R}^d$ by

$$\Delta^i_{(t,x,y)} f = f(t, x + ye_i) - f(t, x), \qquad t \in \mathbb{R}_+,\ x \in \mathbb{R}^d,\ y \in \mathbb{R},$$

where e_i denotes the ith unit vector in the Euclidean basis for $\mathbb{R}^d$. In the sequel statements involving coordinate numbers i, j are implicitly understood to hold for all possible values. Most of the random fields $X : \mathbb{R}_+ \times \mathbb{R}^d \times \Omega \to \mathbb{R}$ (or $\mathbb{R}^m$) considered in this paper will satisfy the following condition:

Condition 2.4 *There exists $\delta \in (0, 1]$ such that the following holds: For all R,$T > 0$, and $p > 2$ there exists a constant $C > 0$ such that*

$$\begin{aligned} E\left[|X(t,x)|^p\right] &\leq C & (2.10)\\ E\left[|X(t,x) - X(t',x)|^p\right] &\leq C|t-t'|^{p\delta/2}, \text{ and} & (2.11)\\ E\left[\left|\frac{1}{y}\Delta^i_{(t,x,y)}X - \frac{1}{y'}\Delta^i_{(t',x',y')}X\right|^p\right] &\leq C(|t-t'|^{p\delta/2} + |x-x'|^{p\delta} + |y-y'|^{p\delta}), & (2.12) \end{aligned}$$

for t,$t' \in [0, T]$, x,$x' \in [-R, R]^d$, and y,$y' \in [-R, R] \backslash \{0\}$.

Notice that when X satisfies Condition 2.4, then X also satisfies this condition for all $\delta' \in (0, \delta]$, with an appropriate change of the constant C.

Let $X = \{X(t, x); t \geq 0, x \in D\}$ be a random field with $D \subset \mathbb{R}^d$. We say that X is *continuous* (resp. *continuously differentiable* w.r.t. x), if there exists $N \in \mathcal{F}$ with $P(N) = 0$ such that for all $\omega \in N^c$ the functions $(t, x) \mapsto X(t, x, \omega)$ are continuous (resp. all first order partial x-derivatives are continuous).

The following lemma is basic for the rest of the paper. Its assertions are essentially contained in [7, 9], but detailed proofs are omitted. For the convenience of the reader, we therefore give a proof.

Lemma 2.5 *Let X be a random field on $\mathbb{R}_+ \times \mathbb{R}^d$ which is continuous for every $\omega \in M$, $P(M) = 1$, and which satisfies Condition 2.4. Then there is a subset $M_0 \subset M$ with $P(M_0) = 1$, such that $(t, x, \omega) \mapsto X(t, x, \omega)$ is continuously differentiable in x for all $\omega \in M_0$ and all (t, x). Moreover, the pointwise defined limit (i.e., $\omega \in M_0$ is fixed)*

$$X_{x_i}(t, x, \omega) := \lim_{y \to 0} \frac{X(t, x + ye_i, \omega) - X(t, x, \omega)}{y}$$

also exists as an $L^p(P)$-limit, for any $p \geq 1$.

Proof: For simplicity of notation let $d = 1$. Fix $R > 1$ and let $\xi(t, x, y) := \Delta_{(t,x,y)}X/y$ if $y \neq 0$. In view of (2.12) the sequence $n \mapsto \xi(t, x, 1/n)$ is Cauchy in $L^p([0, R] \times [-R, R] \times \Omega)$. Let $\xi_p(t, x, 0)$ denote the limit as n tends to infinity and let $\xi_p(t, x, y) := \xi(t, x, y)$ when $y \neq 0$. From (2.12) we obtain

$$E[|\xi_p(t, x, y) - \xi_p(t', x', y')|^p] \leq C(|t - t'|^{p\delta/2} + |x - x'|^{p\delta} + |y - y'|^{p\delta})$$

for all $t, t' \in [0, R]$ and all $x, x', y, y' \in [-R, R]$. For sufficiently large p, Kolmogorov's continuity theorem (see [9]) ensures that ξ_p has a version $\tilde{\xi}_p$ which is continuous for $(t, x, y) \in [0, R] \times [-R, R]^2$ and $\omega \in N_{p,R}^c$, where $P(N_{p,R}) = 0$. Moreover, for $y \neq 0$ we have $\xi(t, x, y, \omega) = \tilde{\xi}_p(t, x, y, \omega)$ for all $\omega \in M_{p,R} := M \cap N_{p,R}^c$. Define

$$\tilde{X}(t, x) := X(t, 0) + \int_0^x \tilde{\xi}_p(t, s, 0)\, ds$$

for all $(t, x) \in [0, R] \times [-R, R]^d$. $\tilde{X}$ is clearly continuous and continuously differentiable with $\partial_x \tilde{X}(t, x, \omega) = \tilde{\xi}_p(t, x, 0, \omega)$, for all $\omega \in M_{p,R}$. Since $\xi(t, x, y) = \tilde{\xi}_p(t, x, y)$ when $y \neq 0$ we can estimate $|\tilde{X}(t, x) - X(t, x)|$ by

$$\begin{aligned}
|X(t,0) \quad &+ \quad \int_0^x [\tilde{\xi}_p(t, s, 0) - \xi(t, s, \frac{1}{n})]\, ds + n \int_0^x [X(t, s + \frac{1}{n}) - X(t, s)]\, ds - X(t, x)| \\
&\leq \int_0^x |\tilde{\xi}_p(t, s, 0) - \xi(t, s, \frac{1}{n})|\, ds + |X(t, 0) - \frac{1}{1/n} \int_0^{1/n} X(t, s)\, ds| \\
&\quad + |\frac{1}{1/n} \int_x^{x+1/n} X(t, s)\, ds - X(t, x)|.
\end{aligned}$$

The limit on the r.h.s. for $n \to \infty$ vanishes for every $\omega \in M_{p,R}$ (use uniform continuity of $\tilde{\xi}_p(t, \cdot, \cdot)$ for the first term). Thus X and $\tilde{X}$ coincide on $M_{p,R}$, and therefore also on $M_0 := \bigcap_{p,R \in \mathbb{N}} M_{p,R}$. So $\partial_x X(t, x)$ exists and is continuous for every $\omega \in M_0$ and all (t, x).

Finally consider (2.12) with $t' := t$, $x' := x$, $y' := 1/n$ and $y := 1/m$. It follows that the sequence

$$n \mapsto \frac{X(t, x + 1/n) - X(t, x)}{1/n}$$

is a Cauchy sequence in $L^p(P)$ for any $p > 2$. Since it converges to $\partial_x X(t, x, \omega)$ for every $\omega \in M_0$ it follows that the $L^p(P)$-limit coincides with $\partial_x X(t, x)$ almost surely. The claim for $p \geq 1$ follows now by Hölder's inequality. □

Remark. When X satisfies Condition 2.4, then X has a continuous version $\tilde{X}$. This follows immediately from the estimate (3.2) below and Kolmogorov's continuity theorem. Henceforth we will therefore implicitly assume that a random field X satisfying Condition 2.4 is continuous.

In order to state our main result (Theorem 2.7) we have to assume smoothness properties of the process $X_{t-s}^{t,x}$ which appears in the solution (2.9). It is convenient to define

$$Z_s^{t,x} := \begin{cases} X_{t-s}^{t,x} & \text{if } 0 \leq s \leq t \\ x & \text{if } s > t \end{cases} \tag{2.13}$$

and $\Delta_{(s,t,x,y)}^i Z = Z_s^{t,x+ye_i} - Z_s^{t,x}$ for $(s, t, x, y) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}$. The required smoothness properties are formulated in the conclusion of the next proposition.

Proposition 2.6 *Let $K > 0$, abbreviate $b^i, \sigma^{ij} \in C_b^{0,2}(\mathbb{R}_+ \times \mathbb{R}^d)$ by g and suppose that*

$$|D_x^\gamma g(t,x) - D_x^\gamma g(t',x')| \le K(|t-t'|^{\alpha(|\gamma|)} + |x-x'|^{\alpha(|\gamma|)}),$$

for (t,x), $(t',x') \in \mathbb{R}_+ \times \mathbb{R}^d$, $0 \le |\gamma| \le 2$, where $\alpha(0) = \alpha(1) = 1$ and $\alpha(2) = \alpha \in (0,1]$. Then for any $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$, there is a unique strong solution $(X_s^{t,x})_{s\ge 0}$ of (2.8). For $Z_s^{t,x}$ defined by (2.13) the following conclusion holds:

For any $R,T > 0$, and $p > 2$, there is a constant C such that

$$E\left[|Z_s^{t,x}|^p\right] \le C \tag{2.14}$$

$$E\left[|Z_s^{t,x} - Z_{s'}^{t,x}|^p\right] \le C|s-s'|^{p\delta/2}, \text{ and} \tag{2.15}$$

$$E\left[\left|\frac{1}{y}\Delta^i_{(s,t,x,y)}Z - \frac{1}{y'}\Delta^i_{(s',t',x',y')}Z\right|^p\right] \le C(|s-s'|^{p\delta/2} + |t-t'|^{p\delta/2} + |x-x'|^{p\delta} + |y-y'|^{p\delta}), \tag{2.16}$$

with $\delta = 1$ for $s,s' \in [0,T]$, (t,x), $(t',x') \in [0,R] \times [-R,R]^d$, and $y,y' \in [-R,R]\backslash\{0\}$. Moreover, $Z_s^{t,x}$ is continuously differentiable with respect to x, and the first order partial x-derivatives of $Z_s^{t,x}$ satisfy (2.14-2.16) with $\delta = \alpha$.

Proof: The Lipschitz continuity and boundedness of b^i, σ^{ij} ensure that for any $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$, there is a unique solution $(X_s^{t,x})_{s\ge 0}$ of (2.8). The results in Chapter II of [9] imply that for any $R,T > 0$, and $p > 2$, there is a constant C such that (2.14-2.16) hold with $\delta = 1$.

With the same arguments as in the proof of Lemma 2.5 it follows that there exists N with $P_2(N) = 0$ such that for all $\omega \in N^c$, $\partial_{x_j} Z_s^{t,x}(\omega)$ exists and is continuous for $(s,t,x) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d$. Furthermore, from Chapter II in [9] we obtain a constant C for any $R,T > 0$, and $p > 2$ such that $\tilde{Z}_s^{t,x} := \partial_{x_j} Z_s^{t,x}$ satisfies (2:14-2.16) with $\delta = \alpha$, for $s,s' \in [0,T]$, (t,x), $(t',x') \in [0,R] \times [-R,R]^d$ and $y,y' \in [-R,R]\backslash\{0\}$. □

We can now state our main result. Sufficient conditions which imply the assumptions (A) – (D) of the following theorem are stated and discussed below.

Theorem 2.7 *Let $T > 0$, and $L(t,x)$ be given by (2.2) with $a = \sigma\sigma^T$. Assume that*

(A) *σ^{ij} and b^i are measurable, locally bounded functions on $\mathbb{R}_+ \times \mathbb{R}^d$ such that (2.8) has a weak solution $(X_s^{t,x})_{s\ge 0}$.*

(B) *$Z_s^{t,x}$ given by (2.13) satisfies the conclusion in Proposition 2.6.*

(C) *ψ_t is such that $(\mathcal{E}_t^\psi)_{0\le t\le T}$, defined in (2.4), is a martingale.*

(D) *c, $h_j \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d) \cap C^{2+\beta}(\mathbb{R}_+ \times \mathbb{R}^d)$, and $u_0 \in C_b^{2,\beta}(\mathbb{R}^d)$ with $\beta \in (0,1]$. For any $\xi \in C_0^\infty(\mathbb{R}_+)^m$ a solution v of the Cauchy problem (2.6) is given by (2.7), and v is unique in a subspace $\mathcal{C} \subset C^{1,2}(D_T)$.*

Then u defined in (2.9) is a strong solution of (2.3) on D_T, and $u(t,x) \in L^p(\overline{P}_1)$ for all $p \ge 1$, $(t,x) \in D_T$. If $\tilde{u}$ is another strong solution of (2.3) which is twice continuously x-differentiable in the $L^2(\overline{P}_1)$-sense, and whose S-transform $Su(\xi)$ belongs to $\mathcal{C}$ for every $\xi \in C_0^\infty(\mathbb{R}_+)^m$, then $\tilde{u}(t,x) = u(t,x)$ P_1-a.s., for all (t,x).

Notice that for non-linear filtering, the theorem states roughly the following: *If the signal process $X_s^{t,x}$ is sufficiently smooth and bounded in every L^p, and if the S-transformed Zakai-equation admits a Feynman-Kac solution, then the Zakai-equation is solved by the stochastic Feynman-Kac formula (2.9).*

Remarks. Let us comment on the the conditions (A) – (D) in more detail:

In (A) it is sufficient to impose the local boundedness for the t-variable, for every fixed x, because in the proof of the theorem we only need an argument to interchange the t-integration with the expectation $E_{\overline{P}_1}$. We remark that we do not need to impose uniqueness of the weak solution $X^{t,x}$. This may be useful, for example in the Cox-Ingersoll-Ross model of interest rates [15], where a weak solution can be given explicitly, but whether uniqueness holds is an open question (at least for us).

Assumption (B) holds whenever the conditions in Proposition 2.6 are satisfied. This is the case, for example, if b^i, $\sigma^{ij} \in C_b^{1,3}(I\!R_+ \times I\!R^d)$. But sometimes (e.g. explicit) solutions $X^{t,x}$ are more regular than the coefficients in the differential equation indicate. Because of such cases it is useful to require (A) and (B) separately. It would be desirable to find less restrictive conditions on b and σ which imply (B), in particular conditions which allow for unbounded b and σ. The example in Section 5 indicates that it should be possible to find such conditions.

Assumption (C) holds, for example, if the Novikov–condition is satisfied, i.e. if

$$E_{P_1}[e^{\frac{1}{2}\int_0^T \psi_s^2 ds}] < \infty. \tag{2.17}$$

Related conditions can be found in the literature, cf. [16] and references given there. When (2.17) holds, the conclusion $u(t,x) \in L^p(\overline{P}_1)$ in Theorem 2.7 can be improved: Then one has $u(t,x) \in L^p(P_1)$, for all $p \geq 1$. This follows from Hölder's inequality (for $q = 1 + 1/4$ and q' such that $1/q + 1/q' = 1$) applied to the r.h.s. of

$$E_{P_1}[|u(t,x)|^p] = \int |u(t,x)|^p e^{-\int_0^T \psi_s \, dW_s + \frac{1}{2}\int_0^T |\psi_s|^2 \, ds} \, d\overline{P}_1.$$

Notice that (2.17) holds in particular for bounded processes ψ. In the filtering context $\psi_s = h(s, X_s)$ is such a bounded process.

Assumption (D) will usually be the hardest to verify in applications to non-linear filtering, because the coefficients in A^* are often non-linear, unbounded or degenerate. Conditions which imply that a Feynman-Kac representation (2.7) is valid can be found in [13] and in [16]. (To apply the result in [16] one has to transform the backward equations and integrals into forward equations and integrals. This is straightforward.) (D) holds, for example, if b and σ are Hölder continuous and L is uniformly elliptic, i.e. there are $\lambda_2 \geq \lambda_1 > 0$ such that, for all $(t,x) \in I\!R_+ \times I\!R^d$ and $y \in I\!R^d$:

$$\lambda_2 |y|^2 \geq \sum_{i,j=1}^{d} a^{ij}(t,x) y_i y_j \geq \lambda_1 |y|^2. \tag{2.18}$$

In this case the space $\mathcal{C}$ can be chosen, e.g., as the space of $C^{1,2}$-functions on D_T which are exponentially bounded w.r.t. x, uniformly w.r.t. t. The example in Section 5 illustrates that (2.18) is not a necessary condition. $\mathcal{C}$ in this example will be the space of polynomially bounded functions in $C^{1,2}(D_T)$.

3 A differential calculus for random fields

In this section we prove that differentiable random fields which satisfy Condition 2.4 also satisfy the standard rules of calculus. These rules are potentially useful when one has to verify (by direct calculation) that some given expression, e.g. (2.9), solves a stochastic differential equation. We shall frequently use the elementary inequality

$$|a_1 + \cdots + a_n|^p \leq n^{p-1}(|a_1|^p + \cdots + |a_n|^p),$$

which holds for real a_i and $p, n \in I\!N$. Also we shall use

$$(a+b)^c \leq 2^c(a^c + b^c), \quad a, b, c \geq 0.$$

Lemma 3.1 *Suppose the random field X satisfies Condition 2.4. Then for any $R, T > 0$, and $p > 2$ there is a constant $C_1 > 0$ such that*

$$E\left[\left|\frac{1}{y}\Delta^i_{(t,x,y)}X\right|^p\right] \leq C_1 \tag{3.1}$$

for all $t \in [0,T]$, $x \in [-R,R]^d$ and $y \in [-R,R]\backslash\{0\}$. Moreover, there is a constant $C_2 > 0$ such that for all $t, t' \in [0,T]$ and $x, x' \in [-R,R]^d$:

$$E\left[|X(t,x) - X(t',x')|^p\right] \leq C_2(|t-t'|^{p\delta/2} + |x-x'|^{p\delta}). \tag{3.2}$$

Proof: Let $R, T > 0$, $p > 2$, and $i \in \{1, \ldots, d\}$. Then we obtain from (2.10) and (2.12)

$$\begin{aligned} E\left[\left|\frac{\Delta^i_{(t,x,y)}X}{y}\right|^p\right] &\leq 2^p E\left[\left|\frac{\Delta^i_{(t,x,y)}X}{y} - \frac{\Delta^i_{(0,0,1)}X}{1}\right|^p\right] + 2^{2p}E\left[|X(0,e_i)|^p + |X(0,0)|^p\right] \\ &\leq 2^pC(p, R\vee 1, T)(T^{p\delta/2} + R^{p\delta} + (R+1)^{p\delta}) + 2^{2p+1}C(p,1,1), \end{aligned}$$

where $C(p,R,T)$ denotes the constant in Condition 2.4, for given p, R, T. To prove (3.2), define $y_i \in [-R,R]$ by $x' = x + \sum_{i=1}^d y_ie_i$. Then $|\Delta X|^p := |X(t',x') - X(t,x)|^p$ can be estimated as

$$\begin{aligned} |\Delta X|^p &= |X(t',x') - X(t,x') + X(t, x + \sum_{i=1}^d y_ie_i) - X(t,x)|^p \\ &= |X(t',x') - X(t,x') + \sum_{n=1}^d \{X(t, x + \sum_{i\leq n} y_ie_i) - X(t, x + \sum_{i\leq n-1} y_ie_i)\}|^p \\ &\leq (d+1)^p(|X(t',x') - X(t,x')|^p + \sum_{n=1}^d |X(t, x + \sum_{i\leq n} y_ie_i) - X(t, x + \sum_{i\leq n-1} y_ie_i)|^p). \end{aligned}$$

With (2.10) and (3.1) this yields

$$E[|X(t',x') - X(t,x)|^p] \leq (d+1)^p[C(p,R,T)|t'-t|^{p\delta/2} + \sum_{n=1}^d C(p,2R,T)|y_n|^{p\delta}].$$

Now (3.2) follows from $|y_n| \leq |x-x'|$ and $C_2 := (d+1)^p[C(p,R,T) + dC(p,2R,T)]$. □

Proposition 3.2 (Product rule) *Suppose that the random fields X and Y satisfy Condition 2.4. Then the product XY satisfies Condition 2.4, too. Moreover,*

$$(XY)_{x_i}(t,x) = X_{x_i}(t,x)Y(t,x) + X(t,x)Y_{x_i}(t,x) \qquad a.s. \tag{3.3}$$

for $(t,x) \in I\!R_+ \times I\!R^d$, and the exceptional set does not depend on (t,x).

Proof: Let X and Y satisfy Condition 2.4 with the same δ and C. Fix $R,T>0$ and $p>2$. Then (2.10) follows from

$$E[|X(t,x)Y(t,x)|^p] \le E[|X(t,x)|^{2p}]^{1/2}E[|Y(t,x)|^{2p}]^{1/2} \le C(2p,R,T),$$

and (2.11) follows from

$$\begin{aligned} E[|X(t,x)Y(t,x) - X(t',x)Y(t',x)|^p] &\le 2^pE[|X(t,x)-X(t',x)|^{2p}]^{1/2}E[|Y(t,x)|^{2p}]^{1/2} \\ &\quad +2^pE[|X(t',x)|^{2p}]^{1/2}E[|Y(t,x)-Y(t',x)|^{2p}]^{1/2} \\ &\le 2^{p+1}C(2p,R,T)|t-t'|^{p\delta/2}. \end{aligned}$$

It remains to prove (2.12). Fix $i \in \{1,\dots,d\}$ and write Δ and Δ' for $\Delta^i_{(t,x,y)}$ and $\Delta^i_{(t',x',y')}$, respectively. Then

$$\begin{aligned} &\left|\frac{\Delta(XY)}{y} - \frac{\Delta'(XY)}{y'}\right|^p \\ &\le \left|\frac{\Delta X}{y}Y(t,x+ye_i) + \frac{\Delta Y}{y}X(t,x) - \frac{\Delta' X}{y'}Y(t',x'+y'e_i) - \frac{\Delta' Y}{y'}X(t',x')\right|^p \\ &\le 4^p\left(\left|\left(\frac{\Delta X}{y} - \frac{\Delta' X}{y'}\right)Y(t,x+ye_i)\right|^p + \left|\frac{\Delta' X}{y'}(Y(t,x+ye_i) - Y(t',x'+y'e_i))\right|^p\right. \\ &\quad \left. + \left|\left(\frac{\Delta Y}{y} - \frac{\Delta' Y}{y'}\right)X(t,x)\right|^p + \left|\frac{\Delta' Y}{y'}(X(t,x) - X(t',x'))\right|^p\right), \end{aligned}$$

for $t,t' \in I\!R_+$, $x,x' \in I\!R^d$, and $y,y' \in I\!R\backslash\{0\}$. We denote the terms on the r.h.s. by $I_1^{(i)}$ to $I_4^{(i)}$ and estimate their expectation separately.

From Cauchy's inequality, (2.11) and (2.12) we obtain

$$\begin{aligned} E[I_1^{(i)}] &\le E\left[\left|\frac{\Delta X}{y} - \frac{\Delta' X}{y'}\right|^{2p}\right]^{1/2} E\left[|Y(t,x+ye_i)|^{2p}\right]^{1/2} \\ &\le C(2p,R,T)(|t-t'|^{p\delta/2} + |x-x'|^{p\delta} + |y-y'|^{p\delta}), \end{aligned}$$

for $t,t' \in [0,T]$, $x,x' \in [-R,R]^d$ and $y,y' \in [-R,R]\backslash\{0\}$.

We estimate $E[I_2^{(i)}]$ using Cauchy's inequality, (3.1) and (3.2):

$$\begin{aligned} E[I_2^{(i)}] &\le E\left[\left|\frac{\Delta' X}{y'}\right|^{2p}\right]^{1/2} E[|Y(t,x+ye_i) - Y(t',x'+y'e_i)|^{2p}]^{1/2} \\ &\le \text{Const}\,(|t-t'|^{p\delta/2} + |x-x'|^{p\delta} + |y-y'|^{p\delta}), \end{aligned}$$

for $t, t' \in [0,T]$, $x, x' \in [-R,R]^d$, and $y, y' \in [-R,R]\backslash\{0\}$, where the constant depends on p, R and T.

Because of symmetry $I_3^{(i)}$ and $I_4^{(i)}$ can be estimated similarly. It remains to show (3.3). Since all the derivatives in (3.3) exist, we only have to show the equality. Let $i \in \{1, \ldots, d\}$ be arbitrary, then for almost every $\omega \in \Omega$ and all $(t,x) \in I\!R_+ \times I\!R^d$:

$$
\begin{aligned}
&\left| \frac{\Delta^i_{(t,x,y)} XY}{y} - X_{x_i}(t,x) Y(t,x) - X(t,x) Y_{x_i}(t,x) \right| \\
&\leq \left| \left(\frac{\Delta^i_{(t,x,y)} X}{y} - X_{x_i}(t,x) \right) \cdot Y(t,x) \right| + \left| \frac{\Delta^i_{(t,x,y)} X}{y} \cdot \Delta^i_{(t,x,y)} Y \right| \\
&\quad + \left| X(t,x) \cdot \left(\frac{\Delta^i_{(t,x,y)} Y}{y} - Y_{x_i}(t,x) \right) \right| \to 0 \text{ as } y \to 0. \qquad \square
\end{aligned}
$$

Notice that when X and Y satisfy Condition 2.4 with constants $\delta_X \neq \delta_Y$, we can choose $\delta = \min\{\delta_X, \delta_Y\}$. The foregoing proof shows that also XY satisfies Condition 2.4 with this δ. In the sequel we denote by $\nabla_x f(t,x) = (\partial_{x_1} f(t,x), \ldots, \partial_{x_d} f(t,x))$ the gradient of a function $f(t,x)$ w.r.t. x.

Proposition 3.3 (Chain rule) *Suppose that the random field X satisfies Condition 2.4 and $f \in C^{1+\beta}(I\!R_+ \times I\!R^d)$ with $0 < \beta \leq 1$. Then $(t,x) \mapsto f(t, X(t,x))$ is a random field which satisfies Condition 2.4 with $\delta' = \beta\delta$. Moreover,*

$$
\partial_{x_i} f(t, X(t,x)) = (\nabla_x f)(t, X(t,x)) \cdot X_{x_i}(t,x) \quad a.s. \tag{3.4}
$$

for $(t,x) \in I\!R_+ \times I\!R^d$, and the exceptional set does not depend on (t,x).

Proof: To see that $\tilde{f}(t,x) := f(t, X(t,x))$ satisfies (2.10), let $R, T > 0$ and $p > 2$. Then

$$
\begin{aligned}
E\left[|\tilde{f}(t,x)|^p \right] &\leq 2^p E\left[|f(t, X(t,x)) - f(0,0)|^p + |f(0,0)|^p \right] \\
&\leq 2^p \left(2^p K^p (T^{p\beta/2} + C(p,R,T)^\beta) + |f(0,0)|^p \right),
\end{aligned}
$$

for $t \in [0,T]$ and $x \in [-R,R]$. Using the Hölder continuity of f we obtain

$$
E\left[|\tilde{f}(t,x) - \tilde{f}(t',x)|^p \right] \leq 2^p K^p E\left[|t-t'|^{p\beta/2} + |X(t,x) - X(t',x)|^{p\beta} \right].
$$

Since $X(t,x)$ satisfies (2.11) it follows that $f(t, X(t,x))$ satisfies (2.11) with $\delta' = \beta\delta$.

Fix $i \in \{1, \ldots, d\}$. To show (2.12) we proceed as follows

$$
\begin{aligned}
\left| \frac{\Delta \tilde{f}}{y} - \frac{\Delta' \tilde{f}}{y'} \right|^p &\leq \left| \int_0^1 \nabla_x f(t, X(t,x) + \nu \Delta X)\, d\nu \frac{\Delta X}{y} \right. \\
&\quad \left. - \int_0^1 \nabla_x f(t', X(t',x') + \nu \Delta' X)\, d\nu \frac{\Delta' X}{y'} \right|^p \\
&\leq 2^p \left| \int_0^1 (\nabla_x f)(t, X(t,x) + \nu \Delta X)\, d\nu \right|^p \left| \frac{\Delta X}{y} - \frac{\Delta' X}{y'} \right|^p \\
&\quad + 2^p \left| \int_0^1 [(\nabla_x f)(t, X(t,x) + \nu \Delta X) \right. \\
&\quad \left. - (\nabla_x f)(t', X(t',x') + \nu \Delta' X)]\, d\nu \right|^p \left| \frac{\Delta' X}{y'} \right|^p,
\end{aligned}
$$

for $t, t' \in \mathbb{R}_+$, $x, x' \in \mathbb{R}^d$, and $y, y' \in \mathbb{R}\backslash\{0\}$. Fix $T, R > 0$ and $p > 2$. We now take the expectation on both sides and estimate the two resulting terms on the r.h.s., I_1 and I_2, separately.

From the Hölder continuity of $\nabla_x f$, we obtain that I_1 is bounded,

$$\begin{aligned} I_1 &\leq 2^p E\left[\left|\int_0^1 (\nabla_x f)(t, X(t,x) + \nu\Delta X)\, d\nu\right|^{2p}\right]^{1/2} E\left[\left|\frac{\Delta X}{y} - \frac{\Delta' X}{y'}\right|^{2p}\right]^{1/2} \\ &\leq 2^p d^p (K \vee |(\nabla_x f)(0,0)|)^p E\left[\int_0^1 (1 + t^{\beta/2} + |X(t,x) + \nu\Delta X|^\beta)\, d\nu^{2p}\right]^{1/2} \\ &\quad \cdot C(2p, R, T)^{1/2}(|t - t'|^{p\delta/2} + |x - x'|^{p\delta} + |y - y'|^{p\delta}), \end{aligned}$$

for $t, t' \in [0,T]$, $x, x' \in [-R, R]^d$ and $y, y' \in [-R, R]\backslash\{0\}$. It remains to show that the integral is bounded. This follows from the estimate

$$\begin{aligned} E\left[\left(\int_0^1 |X(t,x) + \nu\Delta X|^\beta\, d\nu\right)^{2p}\right] &\leq E\left[\int_0^1 (|X(t,x)| + \nu|\Delta X|)^{2p\beta}\, d\nu\right] \\ &\leq 2^{2p\beta} E\left[|X(t,x)|^{2p\beta} + |\Delta X|^{2p\beta}\right], \end{aligned}$$

for $t \in [0,T]$, $x \in [-R,R]^d$, and $y \in [-R,R]$. This expression is bounded since X satisfies (2.10). Thus to complete the proof it suffices to show that I_2 has a similar upper bound.

Using the Hölder continuity of $\nabla_x f$, we find that

$$\begin{aligned} I_2 &\leq 2^p d^p K^p E\left[\int_0^1 (|t - t'|^{\beta/2} + \{(1-\nu)|X(t,x) - X(t',x')| \right. \\ &\quad \left. + \nu|X(t, x + ye_i) - X(t', x' + y'e_i)|\}^\beta)\, d\nu^{2p}\right]^{1/2} E\left[\left|\frac{\Delta' X}{y'}\right|^{2p}\right]^{1/2}. \end{aligned} \tag{3.5}$$

(3.1) implies that the second expectation is bounded for $t' \in [0,T]$, $x' \in [-R,R]^d$, and $y' \in [-R,R]\backslash\{0\}$. The first term is also bounded,

$$\begin{aligned} E\left[\left(\int_0^1 \cdots\, d\nu\right)^{2p}\right]^{1/2} &\leq 3^p \Big(|t - t'|^{p\beta} + 2^{p\beta} E[|X(t,x) - X(t',x')|^{2p\beta}]^{1/2} \\ &\quad + 2^{p\beta} E[|X(t, x + ye_i) - X(t', x' + y'e_i)|^{2p\beta}]^{1/2}\Big). \end{aligned} \tag{3.6}$$

Jensen's inequality and (3.2) give

$$\begin{aligned} E\left[|X(t,x) - X(t',x')|^{2p\beta}\right]^{1/2} &\leq E\left[|X(t,x) - X(t',x')|^{2p}\right]^{\beta/2} \\ &\leq 2^{\beta/2} C_2(2p, R, T)^{\beta/2}(|t - t'|^{p\delta\beta/2} + |x - x'|^{p\delta\beta}), \end{aligned} \tag{3.7}$$

for $t, t' \in [0,T]$ and $x, x' \in [-R,R]^d$. A similar estimate holds for the last term in (3.6). Thus combining (3.5), (3.6), and (3.7) we find that $f(t, X(t,x))$ satisfies Condition 2.4.

It remains to show (3.4). Recall that X is continuous. Let $\tilde{f}$ be as above, and fix $i \in \{1,\dots,d\}$. Then, almost surely

$$
\begin{aligned}
&\left|\frac{\Delta^i_{(t,x,y)}\tilde{f}}{y} - (\nabla_x f)(t, X(t,x)) \cdot X_{x_i}(t,x)\right| \\
&\quad \leq \int_0^1 |(\nabla_x f)(t, X(t,x) + \nu \Delta^i_{(t,x,y)} X) - (\nabla_x f)(X(t,x))|\, d\nu \left|\frac{\Delta^i_{(t,x,y)} X}{y}\right| \\
&\qquad + \left|(\nabla_x f)(t, X(t,x)) \cdot \left(\frac{\Delta^i_{(t,x,y)} X}{y} - X_{x_i}(t,x)\right)\right| \to 0 \text{ as } y \to 0.
\end{aligned}
$$

□

Proposition 3.4 (Differentiation under the integral sign) *Let $(B_t)_{t\geq 0}$ be a d-dimensional Brownian motion. If X is Itô integrable and satisfies Condition 2.4, then*

$$
f(t,x) := \int_0^t X(s,x)\, dB_s, \quad g(t,x) := \int_0^t X(s,x)\, ds
$$

satisfy Condition 2.4 with the same δ. Moreover, the continuous versions of f and g satisfy

$$
\partial_{x_i} \int_0^t X(s,x)\, dB_s = \int_0^t X_{x_i}(s,x)\, dB_s \quad \text{and} \quad \partial_{x_i} \int_0^t X(s,x)\, ds = \int_0^t X_{x_i}(s,x)\, ds
$$

almost surely, for $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$, and the exceptional set does not depend on (t,x).

Proof: We only prove the result for f. The proof for g is similar. (2.10) and (2.11) follow immediately from Burkholder's inequality and the assumptions on X. To prove (2.12), suppose $R, T > 0$, $p > 2$. Assume without loss of generality that $0 \leq t \leq t' \leq T$, then

$$
\begin{aligned}
&E\left[\left|\frac{\Delta^i_{(t,x,y)} f}{y} - \frac{\Delta^i_{(t',x',y')} f}{y'}\right|^p\right] \\
&\quad \leq 2^p E\left[\left|\int_0^t \left(\frac{\Delta^i_{(s,x,y)} X}{y} - \frac{\Delta^i_{(s,x',y')} X}{y'}\right) dB_s\right|^p\right] + 2^p E\left[\left|\int_t^{t'} \frac{\Delta^i_{(s,x',y')} X}{y'}\, dB_s\right|^p\right].
\end{aligned}
$$

From Burkholder's inequality and (3.1) this expression is bounded by

$$
\begin{aligned}
&2^p T^{p/2-1} \int_0^T E\left[\left|\frac{\Delta^i_{(s,x,y)} X}{y} - \frac{\Delta^i_{(s,x',y')} X}{y'}\right|^p\right] ds + 2^p |t - t'|^{p/2-1} \int_t^{t'} E\left[\left|\frac{\Delta^i_{(s,x',y')} X}{y'}\right|^p\right] ds \\
&\qquad \leq \tilde{C}(p,R,T)(|t-t'|^{p\delta/2} + |x-x'|^{p\delta} + |y-y'|^{p\delta}),
\end{aligned}
$$

for $t, t' \in [0,T]$, $x, x' \in [-R,R]^d$, and $y, y' \in [-R,R] \backslash \{0\}$, which shows that f satisfies Condition 2.4.

Fix $i \in \{1,\dots,d\}$. Let ξ_i be a continuous version of $(t,x,y) \mapsto \Delta^i_{(t,x,y)} X / y$ and η_i be a continuous version of $(t,x,y) \mapsto \int_0^t \xi_i(s,x,y)\, dB_s$. Thus for a.e. $\omega \in \Omega$ we have

$$
\lim_{y \to 0} \frac{\Delta^i_{(t,x,y)}}{y} \int_0^t X\, dB_s = \eta_i(t,x,0) = \int_0^t \xi_i(s,x,0)\, dB_s = \int_0^t X_{x_i}(s,x)\, dB_s,
$$

for $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$, which concludes the proof. □

Remark. In the following we will always choose continuous versions of stochastic integrals which depend on (t,x).

4 Proof of the main theorem

In this section we study the smoothness of the random field u defined in (2.9) and finally we prove Theorem 2.7. Recall that $(\Omega, \mathcal{F}, P) = (\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, \overline{P}_1 \otimes P_2)$.

We first show that the random field

$$U(t,x) := u_0(X_t^{t,x}) e^{\int_0^t c(s,X_{t-s}^{t,x})\,ds + \int_0^t h(s,X_{t-s}^{t,x})\,dY_s - \frac{1}{2}\int_0^t |h(s,X_{t-s}^{t,x})|^2\,ds} \tag{4.1}$$

and its first order x-derivatives satisfy Condition 2.4 (which implies that U is twice continuously differentiable). Notice that $X_t^{t,x} = Z_0^{t,x}$, by (2.13).

Suppose that $Z_s^{t,x}$ is a continuous random field which satisfies the conclusion in Proposition 2.6 and that $u_0 \in C_b^{2,\beta}(\mathbb{R}^d)$ with $\beta \in (0,1]$. According to the chain rule the random field $(t,x) \mapsto u_0(Z_0^{t,x})$ is continuously x-differentiable with x_i-derivative given by $(\nabla u_0)(Z_0^{t,x}) Z_{x_i}^{t,x}(0)$. Moreover, the product rule and the chain rule imply that $(t,x) \mapsto (\nabla u_0)(Z_0^{t,x}) Z_{x_i}^{t,x}(0)$ satisfies Condition 2.4 with $\delta = \alpha$. Thus, the first factor in (4.1) and its x-derivatives satisfy Condition 2.4. By the product rule it remains to study differentiability of the exponential term in (4.1).

Lemma 4.1 *Let $0 < \beta \le 1$. Suppose that $g \in C^{1+\beta}(\mathbb{R}_+ \times \mathbb{R}^d) \cap L^\infty(\mathbb{R}_+ \times \mathbb{R}^d)$, and that $Z_s^{t,x}$ satisfies (2.14-2.16) with $\delta \in (0,1]$. Let M_s stand for Y_s^j or for s. Then $f(t,x) := \exp \int_0^t g(s, Z_s^{t,x})\,dM_s$ satisfies Condition 2.4 with $\delta' = \delta\beta$. Its continuous x_i-derivative is given by*

$$\partial_{x_i} f(t,x) = f(t,x) \int_0^t (\nabla_x g)(s, Z_s^{t,x}) Z_{x_i}^{t,x}(s)\,dM_s \quad a.s.$$

for $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$, and the exceptional set does not depend on (t,x).

Proof: We only give the proof for the exponential of the stochastic integral

$$X(t,x) := \int_0^t g(s, Z_s^{t,x})\,dY_s^j.$$

The proof for $\exp \int_0^t g(s, Z_s^{t,x})\,ds$ is similar. Note that

$$E_P\left[\left|e^{X(t,x)}\right|^p\right] = E_{P_2} E_{\overline{P}_1}\left[e^{pX(t,x)}\right] = E_{P_2}\left[\exp\left(\frac{p^2}{2}\int_0^t g(s, Z_s^{t,x})^2\,ds\right)\right] \le e^{p^2 T \|g\|_\infty^2/2}, \tag{4.2}$$

for $p \ge 1$, $t \in [0,T]$, and $x \in \mathbb{R}^d$.

Propositions 3.3 and 3.4 imply that $X(t,x)$ satisfies Condition 2.4 with $\delta' = \delta\beta$. The remainder of the proof is similar to that of the chain rule. But since the exponential function fails to be globally Hölder continuous we have to modify the estimates of I_1 and I_2 (defined in the proof of Proposition 3.3), where this assumption was used. Fix $T, R > 0$ and $p > 2$. Then to estimate I_1 it suffices to show that

$$M_1 := E_P\left[\left|\int_0^1 e^{X(t,x) + \nu \Delta^i_{(t,x,y)} X}\,d\nu\right|^{2p}\right]^{1/2} \le C(p,R,T),$$

for $t \in [0,R]$, $x \in \mathbb{R}^d$, $y \in \mathbb{R}$, and to estimate I_2 it suffices to show that

$$M_2 := E_P\left[\left|\int_0^1 e^{X(t,x)+\nu\Delta^i_{(t,x,y)}X} - e^{X(t',x')+\nu\Delta^i_{(t',x',y')}X}\,d\nu\right|^{2p}\right]^{1/2}$$

is bounded from above by

$$C(p,R,T)(|x-x'|^{p\delta\beta} + |y-y'|^{p\delta\beta} + |t-t'|^{p\delta\beta/2}),$$

for $t, t' \in [0,R]$, $x, x' \in [-R,R]^d$, $y, y' \in [-R,R]\backslash\{0\}$.

Since $\int_0^1 e^{a+\nu(b-a)}\,d\nu \leq e^a + e^b$ for all a, $b \in \mathbb{R}$, we find that

$$M_1^2 \leq E_P\left[\left|e^{X(t,x)} + e^{X(t,x+ye_i)}\right|^{2p}\right] \leq 2^{2p}e^{2p^2T\|g\|^2_\infty}.$$

Since $|e^a - e^b| \leq (e^a+e^b)|b-a|$ for a, $b \in \mathbb{R}$ we obtain

$$\begin{aligned}
M_2 \;&\leq\; \int_0^1 E_P\left[\left|e^{X(t,x)+\nu\Delta^i_{(t,x,y)}X} + e^{X(t',x')+\nu\Delta^i_{(t',x',y')}X}\right|^{4p}\right]\,d\nu^{1/4} \\
&\quad\cdot \int_0^1 E_P\left[\left|X(t,x)+\nu\Delta^i_{(t,x,y)}X - X(t',x') + \nu\Delta^i_{(t',x',y')}X\right|^{4p}\right]\,d\nu^{1/4} \\
&\leq\; \text{Const}\,(|x-x'|^{p\delta\beta} + |y-y'|^{p\delta\beta} + |t-t'|^{p\delta\beta/2}),
\end{aligned}$$

where the last inequality follows from (4.2) and Proposition 3.4.

Now fix $i \in \{1,\ldots,d\}$. Proposition 3.4 states that there is a set $N_i \in \mathcal{F}$ with $P(N_i)=0$ such that $X_{x_i}(t,x) = \int_0^t (\nabla_x g)(t, Z^{t,x}_s) Z^{t,x}_{x_i}(s)\,dY^j_s$ for all $\omega \in N_i^c$, and $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$. Since f satisfies Condition 2.4 it is continuously x-differentiable. Therefore, it suffices to show that if $\omega \in N_i^c$, then $\Delta^i_{(t,x,y)}f/y$ converges to $f(t,x)X_{x_i}(t,x)$ as y tends to zero.

Note first that

$$\left|\frac{\Delta^i_{(t,x,y)}f}{y} - X_{x_i}(t,x)f(t,x)\right| \leq f(t,x)\left|\frac{1}{y}(e^{\Delta^i_{(t,x,y)}X} - 1) - X_{x_i}(t,x)\right|,$$

for $y \neq 0$. We will show that the second term converges to zero as y tends to zero. Let

$$F(s) = \begin{cases} \frac{1}{s}(e^s-1)-1 & \text{if } s \in \mathbb{R}\backslash\{0\} \\ 0 & \text{if } s = 0, \end{cases}$$

then

$$\left|\frac{1}{y}(e^{\Delta^i_{(t,x,y)}X} - 1) - X_{x_i}(t,x)\right| \leq \left|f(\Delta^i_{(t,x,y)}X)\frac{\Delta^i_{(t,x,y)}X}{y}\right| + \left|\frac{\Delta^i_{(t,x,y)}X}{y} - X_{x_i}(t,x)\right|. \quad (4.3)$$

If $\omega \in N_i^c$, then $\Delta^i_{(t,x,y)}X \to 0$ and $\Delta^i_{(t,x,y)}X/y \to X_{x_i}(t,x)$ as $y \to 0$ for all t and x. Using that $|F(s)| \leq |s|e^s/2$ for $s \in \mathbb{R}$, it follows that the right hand side of (4.3) tends to zero as y tends to zero. □

Corollary 4.2 *Let c, $h_1, \ldots, h_m \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d) \cap C^{2+\beta}(\mathbb{R}_+ \times \mathbb{R}^d)$, and suppose that $Z_s^{t,x}$ satisfies the conclusion in Proposition 2.6. Then U, given by (4.1), and $\nabla_x U$ satisfy also Condition 2.4. In particular, U is twice continuously differentiable in the pointwise sense and in the L^p-sense, for all $p \geq 1$.*

Proof: The continuous random fields $(t,x) \mapsto \exp \int_0^t h_j(s, Z_s^{t,x})\, dY_s^j$, $\exp \int_0^t c(s, Z_s^{t,x})\, ds$, $\exp\{-\frac{1}{2}\int_0^t h_j(s, Z_s^{t,x})^2\, ds\}$ satisfy Condition 2.4 with $\delta = 1$. Therefore they are continuously x-differentiable. Moreover, from Propositions 3.2, 3.3, and 4.1, their derivatives satisfy Condition 2.4. Lemma 2.5 concludes the proof. $\square$

We can now prove the main result.

Proof of Theorem 2.7: Proposition 2.3 states that $u(t,x) \in L^p(\overline{P}_1)$ for all $p \geq 1$, and

$$v(t,x;\xi) = E_{\overline{P}_1}[u(t,x)\bar{\mathcal{E}}^\xi] = E_{\overline{P}_1} E_{P_2}[U(t,x)\bar{\mathcal{E}}^\xi].$$

Moreover, it is obvious that $u(t,x)$ is $(\mathcal{F}_1)_t$-adapted for every $x \in \mathbb{R}^d$. Assumption (D) states that $v(t,x;\xi)$ is the solution of (2.6) for $(t,x) \in D_T$. Integration of (2.6) gives

$$v(t,x;\xi) - u_0(x) - \int_0^t \{L(s,x) + c(s,x) + h(s,x)\xi(s)\} v(s,x;\xi)\, ds = 0. \tag{4.4}$$

To prove that $u(t,x)$ satisfies (2.3), we will verify that (4.4) can be rewritten as

$$E_{\overline{P}_1}\left[\left(u(t,x) - u_0(x) - \int_0^t \{L(s,x) + c(s,x)\} u(s,x)\, ds - \int_0^t h(s,x) u(s,x)\, dY_s\right) \bar{\mathcal{E}}^\xi\right] = 0.$$

Since the term in front of $\bar{\mathcal{E}}^\xi$ is in $L^2(\overline{P}_1)$, and the algebra generated by $\{\bar{\mathcal{E}}^\xi; \xi \in C_0^\infty(\mathbb{R}_+)^m\}$ is dense in $L^2(\overline{P}_1)$, it follows from this equation that $u(t,x)$ satisfies (2.3) $\overline{P}_1$-a.s., and therefore also P_1-a.s. It thus remains to prove that one can rewrite (4.4) as claimed.

Since (4.2) holds uniformly in t, and $\bar{\mathcal{E}}^\xi$ belongs to $L^p(\overline{P}_1)$ for every $p \geq 1$, the map $(s,\omega) \mapsto g(s) U(s,x) \bar{\mathcal{E}}^\xi(\omega)$ is in $L^p(\lambda^1_{[0,t]} \otimes P)$ for any bounded measurable function g. Thus one can apply Fubini's theorem in the following.

We consider each term under the integral sign in (4.4) separately.

$$\begin{aligned}
\int_0^t c(s,x) E_{\overline{P}_1} E_{P_2}\left[U(s,x)\bar{\mathcal{E}}^\xi\right] ds &= E_{\overline{P}_1}\left[\int_0^t c(s,x) E_{P_2}\left[U(s,x)\right] ds\, \bar{\mathcal{E}}^\xi\right] \\
&= E_{\overline{P}_1}\left[\int_0^t c(s,x) u(s,x)\, ds\, \bar{\mathcal{E}}^\xi\right].
\end{aligned} \tag{4.5}$$

From Example 2.2 we obtain for the next term

$$\begin{aligned}
\int_0^t h(s,x)\xi(s) E_{\overline{P}_1} E_{P_2}\left[U(t,x)\bar{\mathcal{E}}^\xi\right] ds &= \int_0^t E_{\overline{P}_1}\left[h(s,x) E_{P_2}\left[U(s,x)\right] \bar{\mathcal{E}}^\xi\right] \xi(s)\, ds \\
&= E_{\overline{P}_1}\left[\int_0^t h(s,x) u(s,x)\, dY_s\, \bar{\mathcal{E}}^\xi\right].
\end{aligned} \tag{4.6}$$

Finally we consider the derivative terms in (4.4). Let $i \in \{1, \ldots, d\}$, $R, T > 0$ and $p > 2$. Since U satisfies Condition 2.4 with $\delta = 1$ there is $C > 0$ such that

$$\begin{aligned}
E_{\overline{P}_1}\left[\left|\frac{\Delta^i_{(t,x,y)} u}{y} - \frac{\Delta^i_{(t',x',y')} u}{y'}\right|^p\right] &\leq E_P\left[\left|\frac{\Delta^i_{(t,x,y)} U}{y} - \frac{\Delta^i_{(t',x',y')} U}{y'}\right|^p\right] \\
&\leq C(|t-t'|^{p/2} + |x-x'|^p + |y-y'|^p)
\end{aligned}$$

for $t, t' \in [0,T]$, $x, x' \in [-R,R]^d$, and $y, y' \in [-R,R]\backslash\{0\}$. Therefore, $u(t,x,\omega_1)$ is continuously differentiable with respect to x_i for all $(t,x) \in D_T$ and $\omega \in N_i^c$, with $\overline{P}_1(N_i) = 0$. Fix $(t,x) \in D_T$ and let $\varepsilon > 0$ be arbitrary. Since b^i is locally bounded we may choose $y' \neq 0$ such that

$$\begin{aligned}
&\left|\int_0^t b^i(s,x)\partial_{x_i} v(s,x;\xi)\,ds - E_{\overline{P}_1}\left[\int_0^t b^i(s,x)u_{x_i}(s,x)\,ds\,\bar{\mathcal{E}}^\xi\right]\right| \\
&= \left|\int_0^t b^i(s,x)\lim_{y\to 0}\frac{\Delta^i_{(s,x,y)}}{y}E_{\overline{P}_1}E_{P_2}\left[U\bar{\mathcal{E}}^\xi\right]ds - E_{\overline{P}_1}\left[\int_0^t b^i(s,x)\lim_{y\to 0}\frac{\Delta^i_{(s,x,y)}u}{y}\,ds\,\bar{\mathcal{E}}^\xi\right]\right| \\
&\leq \int_0^t |b^i(s,x)|\lim_{y\to 0}E_{\overline{P}_1}E_{P_2}\left[\left|\frac{\Delta^i_{(s,x,y)}U}{y} - \frac{\Delta^i_{(s,x,y')}U}{y'}\right|\bar{\mathcal{E}}^\xi\right]ds \\
&\quad + E_{\overline{P}_1}\left[\int_0^t |b^i(s,x)|\lim_{y\to 0}E_{P_2}\left[\left|\frac{\Delta^i_{(s,x,y')}U}{y'} - \frac{\Delta^i_{(s,x,y)}U}{y}\right|\right]ds\,\bar{\mathcal{E}}^\xi\right] \leq \varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, it follows for all $(t,x) \in D_T$:

$$\int_0^t b^i(s,x)\partial_{x_i} v(s,x;\xi)\,ds = E_{\overline{P}_1}\left[\int_0^t b^i(s,x)u_{x_i}(s,x)\,ds\,\bar{\mathcal{E}}^\xi\right]. \tag{4.7}$$

To see that a corresponding equation holds for the second order derivatives, fix $(t,x) \in D_T$ and note that from above we have

$$\int_0^t a^{ij}(s,x)\partial_{x_i x_j}E_{\overline{P}_1}\left[u(s,x)\,\bar{\mathcal{E}}^\xi\right]ds = \int_0^t a^{ij}(s,x)\partial_{x_i}E_{\overline{P}_1}\left[u_{x_j}(s,x)\,\bar{\mathcal{E}}^\xi\right]ds.$$

We may argue as above once more, to show that when $\omega \in N_{ij}^c$, $\partial_{x_j}u_{x_i}(t,x,\omega)$ exists and is continuous for $(t,x) \in D_T$, where $\overline{P}_1(N_{ij}) = 0$. It follows that

$$\int_0^t a^{ij}(s,x)\partial_{x_i x_j}E_{\overline{P}_1}\left[u(s,x)\bar{\mathcal{E}}^\xi\right]ds = E_{\overline{P}_1}\left[\int_0^t a^{ij}(s,x)u_{x_i x_j}(s,x)\,ds\,\bar{\mathcal{E}}^\xi\right], \tag{4.8}$$

for any $(t,x) \in D_T$.

Finally, let N be the union of all N_i and N_{ij}. Then $P(N) = 0$, and the foregoing considerations show that u is twice continuously x–differentiable for all $(t,x) \in D_T$ and $\omega \in N^c$. Substituting (4.5), (4.6), (4.7), and (4.8) into (4.4) shows that $u(t,x)$ satisfies (2.3).

Consider equation (2.3) with another solution $\tilde{u}$. Because $\tilde{u}$ is twice continuously differentiable in the $L^2(\overline{P}_1)$-sense, we can calculate the $\mathcal{S}$-transform of (2.3) by interchanging the expectations and integrals with the derivatives. It follows that the resulting $\mathcal{S}$-transformed function $\tilde{v} = \mathcal{S}\tilde{u}(\xi)$ satisfies the same deterministic Cauchy problem as the function $v = \mathcal{S}u(\xi)$, for all $\xi \in C_0^\infty(I\!R_+)$. Since $v, \tilde{v} \in \mathcal{C}$ these functions coincide. The injectivity of the $\mathcal{S}$-transform concludes the proof. □

5 An application to mathematical finance

In mathematical finance, assets are frequently modeled as stochastic diffusions. In practice, one is interested in fitting the parameters of such diffusions to given data series from the market. For example, in order to calculate an option price by the well-known Black-Scholes-formula one must know the value of the volatility parameter of the underlying diffusion model. Filtering theory may be one possible way for obtaining "optimal" parameter values. In this section we shall describe how the results derived above can be used to estimate the parameters in the log-normal diffusion model for stocks.

Remark. We have chosen this example for several reasons: From the practical point of view it may be used for actual parameter estimations. Its numerical treatment is comparatively simple because the process $X_s^{t,x}$ is known explicitly. Since one can compare the parameters estimated by filtering with those obtained by the traditional methods one can evaluate whether our new approach is useful. It is not necessary in this example to calculate the conditional density immediately after the observation. (But this is the case in many applications in engineering, and therefore non-linear filtering, in contrast to linear filtering, is seldom used by engineers.) From the mathematical point of view this example is interesting because the coefficients of the diffusion are unbounded and degenerate. As far as we know, there is no other method available which solves this filtering problem.

Geometrical Brownian motion (sometimes called log-normal diffusion) is the most common model for the value S_t of a stock at time t. Based on a data series from the stock market for $t \in [0, T]$, one is interested in *estimating the appreciation rate* μ *and volatility* σ in the model

$$dS_t = \mu S_t \, dt + \sigma S_t \, dB_t.$$

We assume that we have some knowledge about μ and σ at time $t = 0$, so that we can treat μ and σ as stochastic quantities μ_0 and σ_0 (e.g. independent quantities with Gaussian distributions). The estimation of the constants μ and σ may then be formulated as a non-linear filtering problem as follows. Consider the 3-dimensional diffusion

$$\begin{bmatrix} dS_t \\ d\mu_t \\ d\sigma_t \end{bmatrix} = \begin{bmatrix} \mu_t S_t \\ 0 \\ 0 \end{bmatrix} dt + \begin{bmatrix} \sigma_t S_t & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} dW_t^1 \\ dW_t^2 \\ dW_t^3 \end{bmatrix}. \tag{5.1}$$

This gives $\mu_t = \mu_0$ and $\sigma_t = \sigma_0$ for all t, so at first sight we obtained nothing new. But since we improve our knowledge about μ and σ by observing (a function of) $\{S_t, 0 \leq t \leq T\}$, we obtain the conditional distributions of μ_0 and σ_0 based on this observation. (This yields the best estimate of μ_0 and σ_0 with respect to the L^2-norm.) Let us assume that our observations can be described by

$$dY_t = h(S_t) \, dt + d\overline{W}_t. \tag{5.2}$$

We have to suppose that $h : \mathbb{R} \to \mathbb{R}$ is smooth and bounded, so that we can apply our formula. (Other choices for h may be more reasonable, cf. the remarks given below.)

To calculate the conditional densities for μ_0 and σ_0 we first have to compute the unnormalized conditional density u for the filtering problem (5.1) and (5.2), given the joint density $u_0(x, \mu, \sigma)$ of (S_0, μ_0, σ_0). Recall that u has to satisfy the Zakai-equation (1.2) which contains the adjoint A^* of the generator of (5.1). This generator is given by

$$A = \frac{1}{2}\sigma^2 x^2 \frac{\partial^2}{\partial x^2} + \mu x \frac{\partial}{\partial x},$$

and A has the adjoint

$$A^* = \frac{1}{2}\sigma^2 x^2 \frac{\partial^2}{\partial x^2} + x(2\sigma^2 - \mu)\frac{\partial}{\partial x} + \sigma^2 - \mu,$$

which we write as $A^* = L + c$. The operator L is associated to a new diffusion, namely

$$\begin{bmatrix} d\hat{S}_t \\ d\hat{\mu}_t \\ d\hat{\sigma}_t \end{bmatrix} = \begin{bmatrix} (2\hat{\sigma}_t^2 - \hat{\mu}_t)\hat{S}_t \\ 0 \\ 0 \end{bmatrix} dt + \begin{bmatrix} \hat{\sigma}_t \hat{S}_t & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} dW_t^1 \\ dW_t^2 \\ dW_t^3 \end{bmatrix}.$$

Given $(\hat{S}_0, \hat{\mu}_0, \hat{\sigma}_0) = (x, \mu, \sigma)$ we obtain $\hat{\mu}_t = \mu$, $\hat{\sigma}_t = \sigma$ and

$$\hat{S}_t^x = x \exp\left(((2\sigma^2 - \mu) - \frac{1}{2}\sigma^2)t + \sigma W_t^1\right) = x \exp\left((\frac{3}{2}\sigma^2 - \mu)t + \sigma W_t^1\right).$$

Now observe that (1.2), for the present problem, can be considered as a Cauchy problem in the variables (t, x), while (μ, σ) are fixed parameters in (1.2). Therefore the function $c(\mu, \sigma) = \sigma^2 - \mu$ is a constant with respect to this ("reduced") Cauchy problem. It is easily checked that $Z_s^{t,x} := \hat{S}_{t-s}^x$ satisfies condition (B) in Theorem 2.7. Moreover, conditions (A) and (C) are obviously satisfied, and (D) is satisfied in view of the lemma given at the end of this section.

We can thus apply our solution (2.9) to obtain the unnormalized conditional density

$$u(t, x, \mu, \sigma, \omega) = E\left[u_0(\hat{S}_t^x, \mu, \sigma) e^{(\mu - \sigma^2)t - \frac{1}{2}\int_0^t h(\hat{S}_{t-s}^x)^2\, ds + \int_0^t h(\hat{S}_{t-s}^x)\, dY_s(\omega)}\right], \tag{5.3}$$

where the expectation is taken with respect to $\{W_t^1;\, t \geq 0\}$.

Notice that when we know one of the parameters exactly, say μ, then we can do the same calculation with the 2-dimensional diffusion $(dS_t, d\sigma_t)$ which is obtained from (5.1) by erasing the $d\mu_t$-part. The result (5.3) would differ only by the new density $u_0(\hat{S}_t^x, \sigma)$ which is independent of μ.

Remarks on the model. At first sight the observation process Y_t for the stock price might appear somewhat unrealistic. Of course, $(S_t)_{0 \leq t \leq T}$ or equivalently $Y_t = \int_0^t S_u du$ ($t \in [0, T]$), describes the theoretically best possible observation. However, in real observations $\{S_{t_i}, 1 \leq i \leq n\}$ the decimal number for S_{t_i} is not given in full length. This means that the observed value $S(t_i)$ for the stock price can be written as $S(t_i) = S_{t_i} + \delta_i$, where the δ_i can be considered as random errors with zero mean and identical distributions. But then the sum $\sum_{i \leq i(t)} \delta_i$ ($i(t)$ is the largest index so that $t_i \leq t$) behaves approximately as a (small) multiple of a Brownian motion W_t, according to the central limit theorem. Therefore we can think of

$$\tilde{Y}_t := \int_0^t S_u du + \epsilon W_t \tag{5.4}$$

as a continuous model for the quantity

$$\hat{Y}_t = \sum_{i<i(t)} S(s_i)(s_{i+1} - s_i),$$

which can easily be computed from the observed stock prices. (The value of ϵ has to be fixed according to the accuracy of the stock price.) The process $Y_t = \tilde{Y}_t/\epsilon$ satisfies (5.2) with $h(x) = x/\epsilon$. Unfortunately this h is not a bounded function, so we cannot directly apply our formula. Instead of h we choose a smooth bounded function h_n, such that $h_n(x) = x/\epsilon$ for all $x \in [-n, n]$, where n is chosen greater than the maximal observed stock price. (It should be verified numerically that the estimated parameters μ and σ do not depend significantly on such a choice of n. Otherwise, this procedure would not be appropriate for the estimation of parameters.)

Remarks on numerics. In order to actually estimate the parameters above using data series; one has to compute the expectation in (5.3) numerically. This can be done with the methods found in the book by Kloeden and Platen [17]. Here we sketch how σ can be estimated, given data up to time t, and assuming that μ is exactly known: The estimation of σ based on the observation $(Y_t(\omega))_{0\le t\le T}$ reads

$$E[\sigma|\mathcal{F}_T](\omega) = \int_{I\!R^2} \sigma \cdot p(T, x, \sigma, \omega)\, d\sigma\, dx = \int_{I\!R^2} \sigma \cdot \frac{u(T, x, \sigma, \omega)}{\int_{I\!R^2} u(T, y, \rho, \omega)\, d\rho\, dy}\, d\sigma\, dx.$$

The two integrals must be approximated by first restricting to a bounded domain in $I\!R^2$, and then applying standard numerical integration recipes. The major problem is thus to calculate values for $u(t, x, \sigma, \omega)$: For each given point (x, σ) on a grid in $I\!R^2$, simulate a suitable number of realizations of $\hat{S}_t^x$. Techniques for such simulations can be found in [17]. Notice that the increments $dY_t(\omega)$ are given by $S(t_i)(t_{i+1} - t_i)$. Computing the mean of all our realizations of $\hat{S}_t^x$, we end up with a numerical value for $u(T, x, \sigma, \omega)$.

Finally, we prove a lemma which immediately implies that (D) holds for our example. This result may also be useful in related problems on the log-normal diffusion, but we did not find it in the literature.

Lemma 5.1 *Let $\sigma, \mu, T \in I\!R$ with $\sigma, T > 0$. Suppose $u_0 \in C_b^{2,\beta}(I\!R)$, $q \in C_b^{2,\beta}([0,T] \times I\!R)$. Then the Cauchy problem*

$$\frac{\partial u}{\partial t} = \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 u}{\partial x^2} + \mu x \frac{\partial u}{\partial x} - q(t, x)u, \qquad u(0, \cdot) = u_0,$$

has a bounded solution $u \in C^{1,2}([0,T] \times I\!R)$, given by

$$u(t, x) = E[u_0(S_t^x)e^{-\int_0^t q(t-s, S_s^x)\, ds}], \tag{5.5}$$

where $S_t^x = x\exp\{(\mu - \sigma^2/2)t + \sigma B_t\}$, and B_t is a one-dimensional standard Brownian motion. This solution u is unique in the class of functions $f \in C^{1,2}([0,T] \times I\!R)$ for which there exist $M > 0$ and $m \ge 1$ such that

$$\sup_{0\le t\le T} |f(t, x)| \le M(1 + |x|^{2m}).$$

Moreover, $u(t, \cdot) \in C_b^2(I\!R)$ for every $t \in [0, T]$.

Proof: It is obvious that u defined by (5.5) is a well-defined and bounded function on $[0,T] \times \mathbb{R}$. Moreover, as a special case of the results proved in Sections 3 and 4 (choose $h_i = 0$), $u(t,\cdot)$ is twice continuously differentiable for every $t \in [0,T]$. It is straightforward to check (with the calculus derived in Sections 3 and 4) that $x \mapsto \partial_x u(t,x)$ and $x \mapsto \partial_x^2 u(t,x)$ are bounded functions for every $t \in [0,T]$. Because of this the generator A of the Itô diffusion S_t^x acts as a differential operator on u, i.e.

$$Au(t,x) = \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 u}{\partial x^2} + \mu x \frac{\partial u}{\partial x} \tag{5.6}$$

(see [18], Theorem 7.9 and Definition 7.7). Moreover, Theorem 8.6 in [18] shows that u is t-differentiable and satisfies

$$\frac{\partial u}{\partial t} = Au - qu, \quad u(0.\cdot) = u_0. \tag{5.7}$$

(We remark that Theorem 8.6 is stated and proved in [18] only for time-independent functions q and for $u_0 \in C_0^2(\mathbb{R})$. However, a trivial extension of the proof shows that u defined by (5.5) satisfies (5.7) for time-dependent q and for $u_0 \in C_b^{2,\beta}(\mathbb{R})$.) Thus $u \in C^{1,2}([0,T] \times \mathbb{R})$, and (5.6) combined with (5.7) shows that u solves the Cauchy problem, as claimed.

To prove uniqueness, define $v(t,x) := u(T-t,x)$. This function solves the backward Cauchy problem

$$\frac{\partial v}{\partial t} = -Av + qv, \qquad v(T,\cdot) = u_0.$$

Now we obtain uniqueness from Theorem 7.6 in [16]. □

Acknowledgements: One of us (J. P.) is very grateful to M. Hazewinkel for his proposal to use techniques from white noise analysis – such as the $\mathcal{S}$-transform – for problems in non-linear filtering. T. Deck received financial support from the project "Processos Estocásticos" (PRAXIS XXI-FEDER-CITMA) during work with this paper.

Addresses: Fred Espen Benth, Norwegian Computing Centre, Box 114 Blindern, N-0314 Oslo, NORWAY

Thomas Deck, Lehrstuhl für Mathematik V, Universität Mannheim, D-68131 Mannheim, GERMAY.

Jürgen Potthoff, Lehrstuhl für Mathematik V, Universität Mannheim, D-68131 Mannheim, GERMAY.

Gjermund Våge, Department of Mathematics, University of Oslo, Box 1053 Blindern, N-0316 Oslo, NORWAY.

References

[1] G. Kallianpur, *Stochastic Filtering Theory*, Springer–Verlag, 1980.

[2] M. Zakai, *On the Optimal Filtering of Diffusion Processes*, Z. Wahrschein. verw. Geb. **11** (1979), 230–243.

[3] I. Gyöngy and N. V. Krylov, *SPDE's with Unbounded Coefficients and Applications I*, Stochastics **32** (1990), 53–91.

[4] I. Gyöngy and N. V. Krylov, *SPDE's with Unbounded Coefficients and Applications II*, Stochastics **32** (1990), 165-180.

[5] E. Pardoux, *Stochastic Differential Equations and Filtering of Diffusion Processes*, Stochastics **3** (1979), 127–167.

[6] E. Pardoux, *Filtrage non lineaire et equations aux derivees partielle stochastiques associees*, in D. L. Burkholder, E. Pardoux and A. Sznitman: École d'Été de Probabilité de Saint-Flour XIX-1989 **1464** (Springer–Lecture Notes in Mathematics, 1991), 67–163.

[7] H. Kunita, *Stochastic Flows and Stochastic Differnetial Equations*, Cambridge University Press, Cambridge, 1990.

[8] F.E. Benth, *An Explicit Functional Process Solution to a Stochastic Partial Differential Equation with Applications to Nonlinear Filtering*, Stochastics **51**, (1994), 195–216.

[9] H. Kunita, *Stochastic Differential Equations and Stochastic Flows of Diffeomorphisms*, in R. M. Dudley, H. Kunita and F. Ledrappier: École d'Été de Probabilité de Saint-Flour XII-1982 **1079** (Springer–Lecture Notes in Mathematics, 1984) 143–303.

[10] T. Deck, J. Potthoff and G. Vage, *A review of white noise analysis from a probabilistic standpoint*, Acta Applicandae Mathematicae **48** (1997), 91-112.

[11] M.I. Freidlin, *Functional Integration and Partial Differential Equations*, Princeton University Press (1985).

[12] F.E. Benth, *A Generalized Feynman-Kac Formula for the Stochastic Heat Problem with Anticipating Initial Conditions*, in: Stochastic Analysis and Related Topics V: The Silivri Workshop 1994, H. Körezlioglu, B. Øksendal and A. S. Üstünel (eds), Birkhäuser (1996), 121–133.

[13] A. Friedman: *Stochastic differential equations and applications*, Academic Press (1975).

[14] Y.N. Blagovescenskii and M.I. Freidlin, *Certain properties of diffusion processes depending on a parameter*, Soviet Math. Dokl. 2 (1961), 633-636.

[15] J.C. Cox, J. E. Ingersoll and S. A. Ross, *A Theory of the Term Structure of Interest Rates*, Econometrica **53** (1985), 385–407.

[16] I. Karatzas and S.E. Shreve, *Brownian motion and stochastic calculus*, Springer Verlag, 1988.

[17] P.E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer–Verlag (1995).

[18] B. Øksendal, *Stochastic differential equations*, Springer Verlag, 1992 (3rd edition).