Densities of one dimensional backward sde's

Fabio Antonelli * **Arturo Kohatsu-Higa** †

Università di L'Aquila - Italy *Universitat Pompeu Fabra - Spain*

May 27, 2004

Abstract

This work is devoted to the study of the existence and smoothness of the marginal densities of the solution of one dimensional backward stochastic differential equations. Under monotonicity conditions of a function of the coefficients, we obtain that the smoothness properties of the forward process influencing the backward equation, transfer to the densities of the solution. Once established these conditions, we apply the result to study the tail behavior of the solution process.

Keywords: Backward SDE's, Malliavin Calculus.

2000 Mathematics Subject Classification. Primary: 60H10

1 Introduction

The main focus of this work is to find sufficient conditions to ensure that the first component of the solution (Y, Z) of a backward stochastic differential equation (BSDE) of the following type

$$Y_t = \Gamma + \int_t^T h(s, Y_s, Z_s)ds - \int_t^T Z_s dW_s, \tag{1}$$

with random Γ and h, has a smooth marginal density at each time $t \in (0, T)$.

This is a well known problem in the classical theory of forward SDE's, but it has not been fully addressed yet in the case of backward ones. To our knowledge, the only partial study of the weak differentiability of the solution process Y is to be found in [PP2] and partially in [MZ].

As the main tool in this field is Malliavin calculus, we will deal only with backward equations in a Brownian environment, since more general formulations would lack the support of such theory. One exception is the theory of Malliavin calculus for Lévy processes as developed by Picard or Bichteler, Gravereaux and Jacod. One may as well study the qualitative properties of the fundamental solution of the associated quasilinear parabolic equation if one studies the properties of the Malliavin covariance matrix associated with (Γ, Y_t, Z_t).

When Γ and h are specified as deterministic functions of the solution process of an independent forward equation, some of the properties of the latter pass on to Y. This type of structure is very

*The first author was partially supported by the MIUR COFIN grant 2000

†The second author was partially supported by grants BFM2000-0807 and BFM 2000-0598.

The authors wish to thank the referee for his/her comments.

common for backward equations, particularly in applications in mathematical finance like pricing of options of various kinds, based on some underlying risky asset.

The main difficulty in the study of the existence and regularity of the density of Y lies in the possibility of cancelling effects between Γ and h that could eventually generate mass points in the law of Y (see example 3.3). In particular, this problem appears when analyzing the Malliavin variance associated with Y. Therefore we have to assume some kind of monotonicity condition of a function joining the two terms. This is our main condition in all our results (see e.g. conditions (7) and (8)).

This condition is the parallel of the uniform elliptic condition for diffusions. We also obtain a higher order condition, probably an equivalent of a Hörmander type condition for backward stochastic differential equations. In order to avoid the uniformity in our condition we have also localized them in space obtaining a somewhat weaker version of uniform ellipticity. This leads to a result of existence and regularity of the density of Y. To obtain these results it is important to have estimates on the support (for existence) and lower bound estimates (for regularity) for the process that drives Γ. Finally, under the conditions that guarantee the smoothness of the marginal densities, we find the tail behavior of the process Y.

When interpreted from the analytical point of view, our work is related to the regularity properties of the fundamental solution of a, possibly degenerate, nonlinear parabolic PDE, usually not obtainable using standard tools. This will be addressed in future publications.

2 Weak Differentiability

As we mentioned before, the main results in this field are found in [PP2]. Here we are going to summarize as well as adapting them to our needs.

Consider the interval $[0,T]$ and a complete probability space $(\Omega,\mathcal{F},P)$ on which a standard one dimensional Brownian motion W is defined; $\{\mathcal{F}_t\}_{t\in[0,T]}$ denotes the filtration generated by W, augmented with the $P-$null sets and made right continuous. Since all the results in the paper rely heavily on Malliavin calculus, we want to introduce very briefly some of its terminology.

We denote by $\mathbb{C}_b^\infty(\mathbb{R}^n)$ the set of $\mathbb{C}^\infty$ bounded functions f from $\mathbb{R}^n$ to $\mathbb{R}$, with bounded derivatives of all orders. If S is the class of real random variables F that can be represented as $f(W_{t_1},\ldots,W_{t_n})$ for some $n\in\mathrm{N}$, $t_1,\ldots,t_n\in[0,T]$ and $f\in\mathbb{C}_b^\infty(\mathbb{R}^n)$, we can complete this space under the Sobolev norm $\|\cdot\|_{1,p}$ given by

$$\| F\|_{1,p}^p = E(|F|^p) + E\left((\int_0^T |D_sF|^2 ds)^{\frac{p}{2}} \right),$$

where D is defined as $D_sF = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(W_{t_1},\ldots,W_{t_n})1_{[0,t_i]}(s)$, obtaining a Banach space, usually indicated with $\mathrm{D}^{1,p}$. Analogously, we can construct the space $\mathrm{D}^{k,p}$ by completing S under the

Sobolev norm

$$\| F \|_{k,p}^p = E(|F|^p) + \sum_{i=1}^{k} E\Big[\Big(\int_0^T \ldots \int_0^T |D^i_{s_i \ldots s_1} F|^2 ds_1 \ldots ds_i \Big)^{\frac{p}{2}} \Big],$$

where $D^i_{s_i \ldots s_1} F = D_{s_i} \ldots D_{s_1} F$. Finally, we denote $\mathbf{D}^\infty = \bigcap_{p \geq 1} \bigcap_{k \geq 1} \mathbf{D}^{k,p}$.

We indicate the adjoint of the closable unbounded operator

$$D : \mathbf{D}^{1,2} \subseteq L^2(\Omega) \longrightarrow L^2([0,T] \times \Omega)$$

by δ_0^T and it is called the Skorohod integral. The domain of δ_0^T is the set of all processes u in $L^2([0,T] \times \Omega)$ such that

$$\left| E \left(\int_0^T D_t F u_t dt \right) \right| \leq C \| F \|_2 \qquad \forall\, F \in S,$$

for some constant C possibly depending on u.

If $u \in Dom(\delta_0^T)$, then $\delta_0^T(u)$ is the square integrable random variable determined by the duality relation

$$E(\delta_0^T(u) F) = E\Big(\int_0^T D_t F u_t dt \Big) \qquad \forall\, F \in \mathbf{D}^{1,2}.$$

We remark that the above construction can be carried through for any fixed time interval $[s,S]$, in the space $L^2([s,S] \times \Omega)$.

Finally, for a one dimensional random variable F we denote its Malliavin variance by

$$\gamma_F = \int_0^T (D_s F)^2 \, ds.$$

The Malliavin variance plays a key role when one wants to determine the existence and the smoothness of the densities of the solutions of stochastic differential equations. Following [N1] (Proposition 2.1.1, page 78), for any random variable $F \in \mathbf{D}^{1,p}_{loc}$ for some $p > 1$, if γ_F is a. s. non-zero, then the law of F is absolutely continuous with respect to Lebesgue measure and the density of any one dimensional $F \in \mathbf{D}^{1,2}$ with $\gamma_F^{-1} DF \in Dom(\delta_0^T)$ is given by $f(x) = E\left(1\{F > x\} \delta_0^T(\gamma_F^{-1} DF)\right)$. Besides we will use the fact that if $F \in \mathbf{D}^\infty$ and $|\gamma_F^{-1}| \in \bigcap_{p>1} L^p$ then F has an infinitely differentiable density (see [N1], Corollary 2.1.2).

When considering smoothness of densities, one generalizes the above concepts using the integration by parts formula repeatedly. That is, (see [N2]) we say that the random variable $F \in \mathbf{D}^\infty$ is non-degenerate if $\gamma_F^{-1} \in \cap_{p>1} L^p$ and in such a case we have that for any $f \in \mathbb{C}_p^\infty$ and $F, G \in \mathbf{D}^\infty$, the following integration by parts formula holds

$$E(f^{(m)}(F) G) = E(f(F) A_m(F,G)) \qquad \text{for} \quad m \geq 1, \ \text{where}$$
$$A_m(F,G) = A(F, A_{m-1}(F,G)), \quad A_1(F,G) = A(F,G) = \delta_0^T(G \gamma_F^{-1} D_\cdot F).$$

Moreover (see [N2], page 41), for any $p > 1$ there exist indices $p_1, p_2, p_3, \alpha_1, \alpha_2$, depending on m and p and a constant $C = C(m, p, p_1, p_2, p_3)$ such that

$$\| A_m(F,G) \|_p \leq C \, \| \gamma_F^{-1} \|_{p_1}^{\alpha_1} \| F \|_{m+1,p_2}^{\alpha_2} \| G \|_{m,p_3}, \quad \text{with } \frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = \frac{1}{p},$$

where by $\| \cdot \|_{d,b}$ we mean the Sobolev norm.

In particular, one can state the following slight generalization of Corollary 3.2.1 in [N2] combined with Lemma 2.3.1 in [N1].

Proposition 2.1 *Let $F \in \mathrm{D}^{\infty}$. Then for any $\rho > 0$, there exists $p \equiv p(\rho)$ such that if $\gamma_F^{-1} \in L^p$ and consequently F has a density which is differentiable up to order ρ. Furthermore for any $p > 1$ there exist $\beta \equiv \beta(p)$ and $\epsilon_0 > 0$ such that if*

$$P(\gamma_F \leq \epsilon) \leq \epsilon^{\beta}, \text{ for all } \epsilon \leq \epsilon_0$$

then $\gamma_F^{-1} \in L^p$.

The existence and smoothness of densities of the solution of (1) is necessarily influenced by Γ and h. Since in most of the theory and applications of BSDE's, those are deterministic functions of an extra process X, solution of a forward equation, we decide to adopt this formulation. Namely, let us consider for $s \geq t$ and $x \in \mathbb{R}$

$$X_s^{t,x} = x + \int_t^s b(u, X_u^{t,x}) du + \int_t^s \sigma(u, X_u^{t,x}) dW_u \tag{2}$$

$$Y_s^{t,x} = g(X_T^{t,x}) + \int_s^T h(u, X_u^{t,x}, Y_u^{t,x}, Z_u^{t,x}) du - \int_s^T Z_u^{t,x} dW_u, \quad \text{where} \tag{3}$$

(i) $b, \sigma : [0,T] \times \mathbb{R} \longrightarrow \mathbb{R}$ are continuously differentiable in space for any fixed t, with $|b_x(t,x)| \leq k_b, |\sigma_x(t,x)| \leq k_\sigma$ for some $k_b, k_\sigma > 0$ for all x. Besides $b(t,0)$, $\sigma(t,0)$ are bounded functions of t.

(ii) $g : \mathbb{R} \longrightarrow \mathbb{R}$ is uniformly Lipschitz with constant $k_g > 0$;

(iii) $h : [0,T] \times \mathbb{R}^3 \longrightarrow \mathbb{R}$ is such that for any $x_1, x_2, y_1, y_2, z_1, z_2 \in \mathbb{R}$,

$$|h(s, x_1, y_1, z_1) - h(s, x_2, y_2, z_2)| \leq k_x |x_1 - x_2| + k_y |y_1 - y_2| + k_z |z_1 - z_2|$$

for some $k_x, k_y, k_z > 0$, uniformly in s. Besides $\int_0^T |h(s,0,0,0)|^2 ds < \infty$.

In the following, if $x = x_0$ is the initial point of (2) at time 0, we will write $\Theta_t = \Theta_t^{0,x_0}$ for $\Theta = X, Y, Z$. From well known results ([N1], Section 2.2, page 99), we have that

Theorem 2.1 : *Under (i), there exists a unique continuous solution of (2), such that*

$$E(\sup_{0 \leq t \leq T} |X_t|^p) \leq C,$$

for any $p \geq 2$, where C is a constant depending only on $k_b \vee k_\sigma, T, p$ and x. Moreover $X_t \in \mathbf{D}^{1,\infty}$ and the following representation holds

$$D_r X_t = \zeta_t \zeta_r^{-1} \sigma(r, X_r), \quad \text{for} \quad r \leq t, \quad D_r X_t = 0 \quad \text{for} \quad r > t,$$

where $\zeta_s^{(t,x)}$ is the derivative of the flow $X_s^{t,x}$, $s \geq t$ (with the usual notation $\zeta = \zeta^{0,x_0}$).

The integrability and weak differentiability properties of the forward equation actually transfer to the backward one. The proof of the following theorems is the same as in [PP2], hence we refer the reader to their work, but we point out that considering coefficients with no regularity in time does not affect their proofs.

Theorem 2.2 : *Under the hypotheses (i)-(iii), there exists a pair of progressively measurable, $\mathbb{R}^2$-valued process (Y, Z) solution of (3), such that for any $p \geq 1$*

$$E(\sup_{0 \leq t \leq T} |Y_t|^p)^{\frac{1}{p}} < \infty, \qquad E\left[\left(\int_0^T |Z_s|^2 ds\right)^{\frac{p}{2}}\right] < \infty.$$

Moreover if we assume that g and h are differentiable in space in their respective arguments, then $Y, Z \in L^2([0,T]; \mathbf{D}^{1,p})$ for any $p \geq 2$ and a version of $\{D_r Y_t, D_r Z_t, r \in [0,T]\}$ is given by: $D_r Y_t = D_r Z_t = 0$ for any $r > t$ and for any fixed r, $\{D_r Y_t, D_r Z_t, r \leq t \leq T\}$ is the unique solution of the BSDE

$$D_r Y_t = g'(X_T) D_r X_T + \int_t^T H(s, D_r X_s, D_r Y_s, D_r Z_s) ds - \int_t^T D_r Z_s dW_s, \tag{4}$$

where $H(\omega, s, u, y, z) = h_x(s,X_s,Y_s,Z_s)u + h_y(s,X_s,Y_s,Z_s)y + h_z(s, X_s, Y_s, Z_s)z$. Finally a version of Z_s is given by $D_s Y_s = \lim_{u \uparrow s} D_u Y_s$ and $\sup_{0 \leq r \leq T} E\Big(\sup_{r \leq t \leq T} |D_r Y_t|^p\Big) < \infty$.

Due to the linearity of equation (4) (see [PP2]), we can give an explicit representation of $D_r Y_t$. From now on, we denote $h_x(s) = h_x(s, X_s, Y_s, Z_s)$ and similarly the other derivatives. We introduce a probability $\tilde{P}$ equivalent to P, given by

$$\frac{d\tilde{P}}{dP} = \mathcal{E}_T := \exp\{\int_0^T h_z(s) dW_s - \frac{1}{2}\int_0^T h_z^2(s) ds\}, \tag{5}$$

which is well defined because of the boundedness of the derivative h_z. The Radon Nikodym derivative, that defines $\tilde{P}$, is therefore strictly positive $P-$a.s., giving the equivalence between the two measures. Hence all the a.s. statements in $\tilde{P}$ transfer to similar statements w.r.t. P and viceversa. Under $\tilde{P}$, $\tilde{W}_t = W_t - \displaystyle\int_0^t h_z(s) ds$ is a Brownian motion and equation (4) becomes

$$D_r Y_t = g'(X_T) D_r X_T + \int_t^T [h_x(s) D_r X_s + h_y(s) D_r Y_s] ds + \int_t^T D_r Z_s d\tilde{W}_s,$$

so taking the conditional expectation under $\tilde{P}$ with respect to $\mathcal{F}_t$, we get for $r \leq t$

$$D_r Y_t = E_{\tilde{P}}(g'(X_T) D_r X_T + \int_t^T [h_x(s) D_r X_s + h_y(s) D_r Y_s] ds | \mathcal{F}_t),$$

that has the explicit solution (which can be checked replacing in the above equation)

$$D_rY_t = E_{\tilde{P}}(\mathrm{e}^{\int_t^T h_y(u)du} g'(X_T)D_rX_T + \int_t^T \mathrm{e}^{\int_t^s h_y(u)du} h_x(s)D_rX_s ds|\mathcal{F}_t).$$

Switching back to P and substituting the expression of D_rX_t, we obtain

$$\begin{aligned} D_rY_t &= E_P(\mathcal{E}_T\mathcal{E}_t^{-1}\mathrm{e}^{\int_t^T h_y(u)du} g'(X_T)D_rX_T + \int_t^T \mathcal{E}_s\mathcal{E}_t^{-1}\mathrm{e}^{\int_t^s h_y(u)du} h_x(s)D_rX_s ds|\mathcal{F}_t) \\ &= E(\mathrm{e}^{\int_t^T [h_y(u)-\frac{1}{2}h_z^2(u)]du+\int_t^T h_z(u)dW_u} g'(X_T)\zeta_T \\ &+ \int_t^T \mathrm{e}^{\int_t^s [h_y(u)-\frac{1}{2}h_z^2(u)]du+\int_t^s h_z(u)dW_u} h_x(s)\zeta_s ds|\mathcal{F}_t)\zeta_r^{-1}\sigma(r,X_r) = \xi_t\zeta_r^{-1}\sigma(r,X_r), \end{aligned}$$

where $\xi_s^{t,x} = \dfrac{dY_s^{t,x}}{dx}$ denotes the derivative of the flow associated with (3). This derivative can be written explicitly as

$$\xi_t = E\Big(\psi_T\psi_t^{-1}g'(X_T)\zeta_T + \int_t^T \psi_s\psi_t^{-1}h_x(s)\zeta_s ds|\mathcal{F}_t\Big),$$

where ψ is the solution of the equation

$$\psi_t = 1 + \int_0^t h_y(s,X_s,Y_s,Z_s)\psi_s ds + \int_0^t h_z(s,X_s,Y_s,Z_s)\psi_s dW_s.$$

Further, if we assume that the coefficients b, σ and h are smooth in the spatial variables, then we can associate with Y the following quasilinear parabolic partial differential equation

$$\begin{cases} [\,u_t + \frac{1}{2}\sigma^2 u_{xx} + bu_x + h(\cdot,\cdot,u,u_x\sigma)\,](t,x) = 0 \\ u(T,x) = g(x), \end{cases} \tag{6}$$

then it is known (see [PP2]) that $Y_s^{t,x} = u(s,X_s^{t,x})$ and $Z_s^{t,x} = (u_x\sigma)(s,X_s^{t,x})$. We remark that one can easily prove that the function $u(s,\cdot)$ is smooth with uniformly bounded derivative.

Indeed differentiating $Y^{t,x}$ we obtain

$$u_x(t,x) = E(\psi_T^{t,x}\zeta_T^{t,x}g'(X_T^{t,x}) + \int_t^T \psi_s^{t,x}\zeta_s^{t,x}h_x(r,X_s^{t,x},Y_s^{t,x},Z_s^{t,x})ds).$$

Therefore u_x results bounded from the boundedness of the derivatives g' and h_x, since the derivatives of the flows, $\psi^{t,x}$ and $\zeta^{t,x}$ have bounded moments.

3 Existence of Densities

Once established the weak differentiability of the process Y, we are interested in understanding the sufficient conditions to ensure that Y has marginal laws which are absolutely continuous with respect to the Lebesgue measure.

To do so, we need to evaluate the Malliavin variance of Y

$$\begin{aligned} \gamma_{Y_t} &= \xi_t^2\int_0^t [\zeta_r^{-1}\sigma(r,X_r)]^2 dr \\ &= E\Big(\psi_T g'(X_T)\zeta_T + \int_t^T \psi_s h_x(s)\zeta_s ds|\mathcal{F}_t\Big)^2(\psi_t^{-1})^2\int_0^t [\zeta_r^{-1}\sigma(r,X_r)]^2 dr. \end{aligned}$$

This is clearly composed by two parts: $\int_0^t [\zeta_r^{-1}\sigma(r, X_r)]^2 dr$ that comes from the forward equation and ξ_t, coming directly from the backward one, so the problem is reduced to analyzing the non-degeneracy of the two components.

We start with the forward part, related with the existence of densities for forward equations with time dependent coefficients. This is a long standing problem; a first answer was given by Taniguchi (see [T]), assuming a restricted Hörmander condition (that in the one-dimensional case reduces to ellipticity) and smoothness of the coefficients b and σ in time, later some attempts to remove the regularity in time or the restricted Hörmander condition were made. Florchinger ([F]) tries to do both, but the proofs are wrong as it can be seen when comparing them with a counterexample given in Taniguchi's paper, while Cattiaux and Mesnager's results ([CM]) apply for coefficients Hölder in time, but still require the restricted Hörmander condition to be satisfied. Instead Chen and Zhou ([CZ]) are able to use an unrestricted Hörmander condition, but they are forced to assume that the coefficients near zero have an order of Hölder continuity which depends on the order of the Hörmander condition. Here we give a characterization using Hölder properties that seem to be more appropriate, under the unrestricted Hörmander condition.

To proceed, we need to introduce the following additional hypothesis on the coefficients of (2)-(3)

(A) σ, b, g and h have continuous derivatives in space as many times as needed with derivatives bounded independently of t. Besides σ and b are continuous at the initial point $(0, x_0)$.

To carry out the extension of Hörmander's condition we have to define extensions of derivatives to any real order between 0 and 1.

Let $\alpha \in \mathbb{R}_+$, for $\alpha = 0$ we define

$$A_0(t,x) = \sigma(t,x), \qquad B_0(t,x) = A_0(t,x), \qquad D_0(t,x) = A_0(t,x) - A_0(0,x) \quad x \in \mathbb{R}.$$

By induction on $n \in \mathbb{N} \cup \{0\}$, we take A_n, B_n and D_n as defined and we define further

$$\begin{aligned} B_{n+1}(t, s_1, \ldots, s_{n+1}, x) &= \frac{\partial B_n}{\partial x}(t, s_1, \ldots, s_n, x) b(s_{n+1}, x) \qquad x \in \mathbb{R}, \\ \frac{\partial^\beta A_n}{\partial t^\beta}(0, x_0) &= \lim_{t \to 0} \frac{D_n(t, x_0)}{t^{n+\beta}}, \qquad \text{for} \quad \beta \in (0,1], \end{aligned}$$

where the limits are to be understood in the extended real line. That is, they may take the values $\pm\infty$. Next for $\alpha \in (n, n+1]$, we set

$$A_\alpha(0, x_0) = \frac{\partial^{\alpha-n} A_n}{\partial t^{\alpha-n}}(0, x_0) + \frac{1_{\{\alpha=n+1\}}}{(n+1)!} B_{n+1}(0, \ldots 0, x_0).$$

Finally we define

$$\begin{aligned} D_{n+1}(t, x_0) &= D_n(t, x_0) - \frac{\partial}{\partial t} A_n(0, x_0) t^n \\ &+ \int_0^t \int_0^{s_{n+1}} \cdots \int_0^{s_2} [B_{n+1}(t, s_1, \ldots, s_{n+1}, x_0) - B_{n+1}(0, \ldots, 0, x_0)] ds_1 \ldots ds_{n+1}. \end{aligned}$$

Our main hypothesis is

(H0) There exists $\alpha \geq 0$ such that $A_\beta(0, x_0) = 0$ for all $0 \leq \beta < \alpha$ and $A_\alpha(0, x_0) \neq 0$.

In particular this assumption implies that all terms above, in particular A_β, are well defined in the extended real line for $\beta \leq \alpha$. Also note that $A_\alpha(t, x_0)$ has not been defined for $t \neq 0$.

Theorem 3.1 : *Let us assume hypotheses (i)-(iii), (A) and (H0) are satisfied and let us set* $K = k_b + k_y + k_\sigma k_z$. *For fixed* $t \in (0, T)$ *and a measurable set* $A \subseteq \mathbb{R}$ *with* $P(X_T \in A | X_t) > 0$, *we denote*

$$\begin{aligned} \underline{g} &= \min_{\mathbb{R}} g'(x), \quad \bar{g} = \max_{\mathbb{R}} g'(x), \\ \underline{g}^A &= \min_{x \in A} g'(x), \quad \bar{g}^A = \max_{x \in A} g'(x), \\ \underline{h}(t) &= \min_{[t,T]\times\mathbb{R}^3} h_x(s, x, y, z), \quad \bar{h}(t) = \max_{[t,T]\times\mathbb{R}^3} h_x(s, x, y, z). \end{aligned}$$

If one of the following holds for the same fixed t

$$(7) \quad \underline{g}\mathrm{e}^{-\mathrm{sgn}\,(\underline{g})KT} + \underline{h}(t)\int_t^T \mathrm{e}^{-\mathrm{sgn}\,(\underline{h}(s))Ks} ds \geq 0, \quad \underline{g}^A \mathrm{e}^{-\mathrm{sgn}\,(\underline{g}^A)KT} + \underline{h}(t)\int_t^T \mathrm{e}^{-\mathrm{sgn}\,(\underline{h}(s))Ks} ds > 0$$

$$(8) \quad \bar{g}\mathrm{e}^{\mathrm{sgn}\,(\bar{g})KT} + \bar{h}(t)\int_t^T \mathrm{e}^{\mathrm{sgn}\,(\bar{h}(s))Ks} ds \leq 0, \quad \bar{g}^A \mathrm{e}^{\mathrm{sgn}\,(\bar{g}^A)KT} + \bar{h}(t)\int_t^T \mathrm{e}^{\mathrm{sgn}\,(\bar{h}(s))Ks} ds < 0$$

then the random variable Y_t *has a probability distribution that is absolutely continuous with respect to the Lebesgue measure.*

Proof: As previously stated in Section 2, we need to prove that $Y_t \in \mathbb{D}^{1,p}_{\mathrm{loc}}$ for $p > 1$ and that γ_{Y_t} is invertible $a.s.$

The first point comes directly from Theorem 2.2, where it was shown that $Y_t \in \mathbb{D}^{1,p}$ for any $p > 2$. The second point needs an argument for the diffusion part and another for the backward component.

First we prove that $E\Big(\psi_T g'(X_T)\zeta_T + \int_t^T \psi_s h_x(s)\zeta_s ds | \mathcal{F}_t\Big) \neq 0$. By Itô's formula, it is easy to show that $\psi_t\zeta_t$ is the stochastic exponential

$$\begin{aligned} \psi_t\zeta_t &= \exp\Big\{\int_0^t [b_x(s, X_s) + h_y(s) + \sigma_x(s, X_s)h_z(s)]ds\Big\} \times \\ &\quad \exp\Big\{\int_0^t (\sigma_x(s, X_s) + h_z(s))dW_s - \frac{1}{2}\int_0^t (\sigma_x(s, X_s) + h_z(s))^2 ds\Big\} = H_t M_t. \end{aligned}$$

The process H verifies $\mathrm{e}^{-Kt} \leq H_t \leq \mathrm{e}^{Kt}$, where $K = k_b + k_y + k_\sigma k_z$, because all the coefficients are Lipschitz, while the process M is a stochastic exponential $P-$martingale, since it verifies the Novikov condition thanks to the conditions (i)-(iii).

If we define a probability measure Q equivalent to P by setting $\frac{dQ}{dP} = M_T$, Q a.e. we have

$$\begin{aligned} E\Big(\psi_T g'(X_T)\zeta_T + \int_t^T \psi_s h_x(s)\zeta_s ds | \mathcal{F}_t\Big) &= E\Big(g'(X_T)H_T M_T + \int_t^T h_x(s)H_s M_s ds | \mathcal{F}_t\Big) \\ &= E_Q\big(g'(X_T)H_T + \int_t^T h_x(s)H_s ds | \mathcal{F}_t\Big). \end{aligned}$$

On the other hand

$$\begin{aligned} & \underline{g}\mathrm{e}^{-\operatorname{sgn}(\underline{g})KT} + \underline{h}(t)\int_t^T \mathrm{e}^{-\operatorname{sgn}(\underline{h}(s))Ks}ds \le \underline{g}H_T + \underline{h}(t)\int_t^T H_s ds \\ \le\ & g'(X_T)H_T + \int_t^T h_x(s)H_s ds \le \bar{g}H_T + \bar{h}(t)\int_t^T H_s ds \\ \le\ & \bar{g}\mathrm{e}^{\operatorname{sgn}(\bar{g})KT} + \bar{h}(t)\int_t^T \mathrm{e}^{\operatorname{sgn}(\bar{h}(s))Ks}ds \end{aligned}$$

and either (7) or (8) gives $\xi_t \geq 0$ or $\xi_t \leq 0$ a.s. (under either P or Q)

Assume (7) is verified, by the previous argument we have

$$\begin{aligned} \xi_t\psi_t &= E_Q\Big((g'(X_T)H_T + \int_t^T h_x(s)H_s ds)(\mathbf{1}_{\{X_T\in A\}} + \mathbf{1}_{\{X_T\in A^c\}})|\mathcal{F}_t\Big) \\ &\geq E_Q\Big((g'(X_T)H_T + \int_t^T h_x(s)H_s ds)\mathbf{1}_{\{X_T\in A\}}|\mathcal{F}_t\Big) \\ &\geq E_Q\Big((\underline{g}^A H_T + \underline{h}(t)\int_t^T H_s ds)\mathbf{1}_{\{X_T\in A\}}|\mathcal{F}_t\Big) \\ &\geq (\underline{g}^A\mathrm{e}^{-\operatorname{sgn}(\underline{g}^A)KT} + \underline{h}(t)\int_t^T \mathrm{e}^{-\operatorname{sgn}(\underline{h}(s))Ks}ds)Q(X_T\in A|\mathcal{F}_t) > 0 \end{aligned}$$

because of the equivalence between P and Q. To finish we need to prove that

$$\int_0^t \left[\zeta_r^{-1}\sigma(r,X_r)\right]^2 dr > 0.$$

This follows from the result in the appendix. □

Remark 3.2 : *The above result holds for $t < T$. When $t = T$, $Y_T = g(X_T)$ needs a study of its own. One straightforward case is given when X_T has a density and g is an invertible function.*

The above conditions (7) and (8) are required to avoid the degenerate case when g' and h_x cancel each other. This may happen easily so that the law of Y_t has a point mass, even if X_s has a smooth density for all s.

Example 3.3 :

Let us consider $X_t = W_t$, a standard Brownian motion, and

$$Y_t = E(W_1 + \int_t^1 f(s)W_s ds|\mathcal{F}_t) \quad \textit{for} \quad f(s) = \begin{cases} 0 & s \leq \frac{1}{2} \\ 4(-2s+1) & s > \frac{1}{2}. \end{cases}$$

In this context, for $t \leq 1/2$, we have $g' \equiv \underline{g} = \bar{g} = 1$, $h(s,x,y,z) = f(s)x$, $\bar{h} \equiv 0$, $K = 0$ $\underline{h} \equiv -4$, thus condition (7) (or (8)) is not verified. In fact, by the martingale property of W, for all $t \leq \frac{1}{2}$

$$Y_t = W_t(1 + \int_t^1 f(s)ds) = W_t(1-1) = 0$$

and it does not have a density.

Remark 3.4 : *Incidentally the previous theorem provides information about the sign of the derivative of the solution of the associated PDE (6). Indeed, under (7) with the strict inequality verified on some subset A, we have that* $\xi_s^{t,x} > 0$ *a.s. for any* $s \geq t$ *and* $D_r Y_s^{t,x} = \xi_s^{t,x}(\zeta_r^{t,x})^{-1}\sigma(r, X_r^{t,x})$.

On the other hand, a version of $Z_s^{t,x}$ *is given by* $\lim_{v \uparrow s} D_v Y_s^{t,x}$ *and if u results differentiable, also* $Z_s^{t,x} = u_x(s, X_s^{t,x})\sigma(s, X_s^{t,x})$.

Therefore, applying all our equalities to the flow $X^{t,x}$ *and specializing them at the initial time t, we obtain*

$$\xi_t^{t,x}(\zeta^{t,x})_t^{-1}\sigma(t,x) = Z_t^{t,x} = u_x(t,x)\sigma(t,x)$$

but $(\zeta^{t,x})_t^{-1} = 1$ *and the previous inequality implies* $u_x(t,x) > 0$ *a.e.* x *for* $\sigma(t,x) \neq 0$.

We now give a result that guarantees the condition $P(X_T \in A|\mathcal{F}_t) > 0$. If the support of this conditional law is $\mathbb{R}$, then the positivity is implied simply by $\lambda(A) > 0$ (λ denotes the Lebesgue measure). Characterization of the support of probability laws is a well known problem. Usually, one expects that the uniform ellipticity condition is sufficient to obtain that the support of the probability law is the whole space.

Various versions of the so called support theorems are available. For the sake of completeness and consistency we briefly present one here, adapted to our current situation. In the proof we use a method that appears in [MS]. Their proof exploits two convergence results in $\alpha-$Hölder norm ($\alpha \in [0, \frac{1}{2})$) for approximations of the diffusion and its skeleton, obtained by a linear interpolation of the Brownian motion. We were able to extend it to the time dependent case, as long as the coefficients are uniformly Hölder of some order $\beta_1 > \frac{1}{2}$ for the diffusion and $\beta_2 > 0$ for the drift as functions of t for any choice of x (that is to say β_1, β_2 do not depend on t and x). Another related result appears in [GP]. This problem can also be treated using analytical tools (see [D], chapter 3).

Proposition 3.5 : *Let the hypotheses of theorem 3.1 hold. Besides we assume that* σ *is uniformly Hölder of order* $\beta_1 > \frac{1}{2}$*, while* b, σ_x *are uniformly Hölder of order* $\beta_2 > 0$ *as functions of* t *for all* $x \in \mathbb{R}$ *and that there exists a positive constant c such that* $|\sigma(t,x)| \geq c$ *for all* $(t,x) \in [0,T] \times \mathbb{R}$. *Then* $\lambda(A) > 0$ *implies that* $P(X_T \in A|\mathcal{F}_t) > 0$ *a.s. for all* $t \in (0,T)$.

Proof: By the Markov and the flow properties, we have that

$$P(X_T \in A|\mathcal{F}_t) = P(X_T \in A|X_t) = P(X_T^{t,X_t} \in A|X_t) = P(X_T^{t,y} \in A)|_{y=X_t}.$$

Therefore we need to show that conditions (A1) and (A2) of proposition 2.1 in [MS] are verified for the sequences $S(W^n)_t - X_t$ and $X_t^n - S(\theta)_t$, where

$$S(\theta)_t = x + \int_0^t \sigma(s, S(\theta)_s)\dot{\theta}_s ds + \int_0^t [b(s, S(\theta)_s) - \frac{1}{2}\sigma\sigma_x(s, S(\theta)_s)]ds \tag{9}$$

with $\theta \in \mathcal{H}^2 = \{\theta \in L^2(0,T] : \dot{\theta} \in L^2(0,T]\}$ and

$$\begin{aligned}
W^n_t &= W_{\frac{k-1}{2^n}\vee 0} + 2^n(t - \frac{k}{2^n})[W_{\frac{k}{2^n}} - W_{\frac{k-1}{2^n}\vee 0}] \\
& \qquad n \in \mathbb{N}, \quad 0 \leq k \leq 2^n, \quad \frac{k}{2^n} \leq t \leq \frac{k+1}{2^n} \\
X^n_t &= x + \int_0^t b(s, X^n_s)ds + \int_0^t \sigma(s, X^n_s)dW_s - \int_0^t \sigma(s, X^n_s)\dot{W}^n_s ds + \int_0^t \sigma(s, X^n_s)\dot{\theta}_s ds \\
S(W^n)_t &= x + \int_0^t \sigma(s, S(W^n)_s)\dot{W}^n_s ds + \int_0^t [b - \frac{1}{2}\sigma\sigma_x(s, S(W^n)_s)]ds.
\end{aligned}$$

Actually condition (A1) is straightforward verified, because of the Lipschitz property of the coefficients, while the Hölder property in time becomes important in order to verify the second condition. In particular, as the above approximations are of Wong-Zakai type, the diffusion coefficient has to be Hölder of order higher than 1/2. The calculations involved are lengthy and we omit them here for briefness, but they go exactly along the lines of the result in [MS]. Besides it is to be noted that a careful computation allows to consider Lipschitz coefficients, relaxing the boundedness hypothesis used in the proof of Millet and Sanz. This argument proves the characterization of the support of the law as the topological closure of the set $\{S(h)_T;\ h \in \mathcal{H}^2\}$. Now we prove that this set is $\mathbb{R}$.

We remark that up to this stage, it is not necessary to employ the elliptic hypothesis on σ, which we use in the following part.

From now on, using Girsanov's theorem we may assume, without loss of generality, that the drift of X is given by $b = \frac{1}{2}\sigma\sigma_x$. Given the characterization of the support, it is enough to prove that for any $y,\ z \in \mathbb{R}$, there exists a function $\theta \in \mathcal{H}^2$ such that $S(\theta)^{(t,y)}_t = y$, $S(\theta)^{(t,y)}_T = z$.

For this, let us notice that, for a smooth path $\varphi : [t,T] \to \mathbb{R}$ so that $\varphi_t = y$ and $\varphi_T = z$, we may choose $\dot{\bar{\theta}}_s = \dfrac{\varphi'_s}{\sigma(s, \varphi_s)}$, which implies $S(\bar{\theta})^{(t,y)} \equiv \varphi$. From here the result follows. □

We may generalize Theorem 3.1 applying Itô's formula to obtain a second order condition, when h is independent of z.

Theorem 3.6 : *Let us assume that h does not depend on z, that the hypotheses (i)-(iii), (A),(H0) are verified and set $K = k_b + k_y + k_\sigma k_z$. For fixed $t \in (0,T)$ and a measurable set $A \subseteq \mathbb{R}$ with $P(X_T \in A|X_t) > 0$, a.s., we denote by*

$$\begin{aligned}
\tilde{g}(x) &= g'(x) + (T-t)h_x(T, x, g(x)), \\
\tilde{h}(s,x,y) &= -\{h_{xt} + bh_{xx} - hh_{xy} + \frac{1}{2}(\sigma^2 h_{xxx} + 2z\sigma h_{xxy} + z^2 h_{xyy}) + (b_x + h_y)h_x\}(s,x,y), \\
\underline{g}_1 &= \min_{\mathbb{R}} \tilde{g}(x), \quad \bar{g}_1 = \max_{\mathbb{R}} \tilde{g}(x) \\
\underline{g}^A_1 &= \min_{x\in A} \tilde{g}(x), \quad \bar{g}^A_1 = \max_{x\in A} \tilde{g}(x) \\
\underline{h}_1(t) &= \min_{[t,T]\times\mathbb{R}^3} \tilde{h}(s,x,y), \quad \bar{h}_1(t) = \max_{[t,T]\times\mathbb{R}^3} \tilde{h}(s,x,y).
\end{aligned}$$

If one of the following is satisfied

$$(10)\quad \underline{g}_1 \mathrm{e}^{-\mathrm{sgn}\,(\underline{g}_1)KT} + \underline{h}_1(t)\int_t^T \mathrm{e}^{-\mathrm{sgn}\,(\underline{h}_1(s))Ks} ds \geq 0, \quad \underline{g}_1^A \mathrm{e}^{-\mathrm{sgn}\,(\underline{g}_1^A)KT} + \underline{h}_1(t)\int_t^T \mathrm{e}^{-\mathrm{sgn}\,(\underline{h}_1(s))Ks} ds > 0$$

$$(11)\quad \bar{g}_1 \mathrm{e}^{\mathrm{sgn}\,(\bar{g}_1)KT} + \bar{h}_1(t)\int_t^T \mathrm{e}^{\mathrm{sgn}\,(\bar{h}_1(s))Ks} ds \leq 0, \quad \bar{g}_1^A \mathrm{e}^{\mathrm{sgn}\,(\bar{g}_1^A)KT} + \bar{h}_1(t)\int_t^T \mathrm{e}^{\mathrm{sgn}\,(\bar{h}_1(s))Ks} ds < 0$$

then the random variable Y_t has a probability distribution that is absolutely continuous with respect to the Lebesgue measure.

Proof: With the same notation as before, it is enough to prove that

$$E\Big(M_T H_T g'(X_T) + \int_t^T M_s H_s h_x(s, X_s, Y_s) ds | \mathcal{F}_t\Big) \neq 0.$$

Changing the probability measure and applying Itô's formula, we obtain

$$\begin{aligned}
& E_Q\Big(H_T[g'(X_T) + h_x(T)(T-t)] + \int_t^T [H_s h_x(s) - H_T h_x(T)] ds | \mathcal{F}_t\Big) \\
= \; & E_Q\Big(H_T[g'(X_T) + h_x(T)(T-t)] - \int_t^T \int_s^T H_r[h_{xt} + b h_{xx} - h h_{xy}](r, X_r, Y_r) dr ds \\
- \; & \int_t^T \int_s^T H_r[\frac{\sigma^2}{2} h_{xxx} + Z_. \sigma h_{xxy} + \frac{1}{2} Z_.^2 h_{xyy} + (b_x + h_y) h_x](r, X_r, Y_r) dr ds \\
- \; & \int_t^T \int_s^T H_r[\sigma h_{xx} + Z_. h_{xy}](r, X_r, Y_r) dW_r ds | \mathcal{F}_t\Big)
\end{aligned}$$

By virtue of the conditional expectation, the martingale part gives no contribution.

To conclude, we can follow the same steps as before applied to the functions $\tilde{g}$ and $-\tilde{h}$, defined in the statement. □

Remark 3.7 : *In general, this method is not useful to obtain higher order conditions as they would involve the differentiability of the process Z. One might think to circumvent this difficulty by exploiting the relation $Z_s = \sigma(s, X_s) u_x(s, X_s)$ and the possible regularity of u, but this would generate a condition depending on the solution of the quasilinear parabolic partial differential equation (6).*

4 Smoothness of Densities

In this section we study the smoothness of the marginal densities of Y. Under the assumptions of the previous theorems we know γ_{Y_t} is invertible a.s. To ensure the smoothness of densities, we need to find what conditions make $\gamma_{Y_t}^{-1} \in \bigcap_{p>1} L^p$, so that we can apply theorem 2.3.3 of [N1]. Throughout the section we will assume that the coefficients b and σ are time independent and that they are infinitely differentiable in space with bounded derivatives of all orders uniformly in time.

We start with a preliminary result on the solution of equations.

Lemma 4.1 : *For adapted processes* $\Gamma, F^1, F^2, K \in \mathbf{D}^{1,\infty}$*, such that* F^1, F^2 *are bounded and*

$$E(|\Gamma_T|^p) < \infty, \quad E(\sup_{0\le t\le T}|K_t|^p) < \infty \quad \sup_{0\le s\le T} E(|D_s\Gamma_T|^p) < \infty,$$
$$\sup_{0\le s\le T} E(\sup_{s\le t\le T}|D_sK_t|^p) < \infty \quad \sup_{0\le s\le T} E(\sup_{s\le t\le T}|D_sF^i_t|^p) < \infty, \quad \textit{for } i=1,2$$

for any $p \ge 2$*, let us consider the backward equation*

$$V_t = \Gamma_T + \int_t^T (F^1_sV_s + F^2_sU_s + K_s)ds - \int_t^T U_s dW_s. \tag{12}$$

Then $V_t, U_t \in \mathbf{D}^{1,\infty}$ *for any* t*, and the derivatives* (D_sV_t, D_sU_t) *verify the equation for* $r \le t$*,*

$$D_rV_t = D_r\Gamma_T + \int_t^T (F^1_sD_rV_s + F^2_sD_rU_s + D_rK_s + D_rF^1_sV_s + D_rF^2_sU_s)ds - \int_t^T D_rU_sdW_s$$

and $\sup_{0\le r\le T} E\sup_{r\le t\le T}|D_rV_t|^p < \infty$.

Proof: The proof of this lemma goes exactly along the same lines as the proof of theorem 2.2 and we refer the reader to [PP2]. We just describe briefly the idea. Considering the Picard iterations associated to equation (12)

$$\begin{aligned} V^0_t &= E(\Gamma_T|\mathcal{F}_t) \qquad U^0_t = -D_tV^0_t \\ V^{n+1}_t &= E(\Gamma_T + \int_t^T (F^1_sV^n_s + F^2_sU^n_s + K_s)ds|\mathcal{F}_t) \quad U^{n+1}_t = -D_tV^{n+1}_t, \end{aligned}$$

by the hypotheses on the processes F^1, F^2, K, Γ, one can show by induction on n, that each iterate is in $\mathbf{D}^{1,\infty}$ and that the estimates are independent of n. From the convergence of V^n, U^n in $L^p([0,T]\times\Omega)$ and proposition 1.5.5 of [N1] we obtain that also the solution of (12) belongs to $\mathbf{D}^{1,\infty}$. □

Lemma 4.2 : *Let* b, σ, h, g *be infinitely differentiable in space with bounded derivatives of all orders greater than one, then* $Y_t \in \mathbf{D}^\infty$ *for each* t.

Proof: To verify that $Y_t \in \mathbf{D}^{k,p}$ for all t and any choice of k, p, we proceed by induction on k. Theorem 2.2 gives the statement for $k=1$ and any p, so we assume that $Y_t \in \mathbf{D}^{k,p}$ for any p and prove the same holds for $k+1$. Taken a sequence of times $0 \le r_1 \le r_2 \le \cdots \le r_k \le t$, the $k-$th Malliavin derivative of Y_t verifies the following linear equation

$$\begin{aligned} D^k_{r_1\ldots r_k}Y_t = G^k(X_T) &+ \int_t^T [h_y(s)D^k_{r_1\ldots r_k}Y_s + h_z(s)D^k_{r_1\ldots r_k}Z_s + H^k(X_s,Y_s,Z_s)]ds \\ &- \int_t^T D^k_{r_1\ldots r_k}Z_sdW_s, \quad \text{where} \\ G^k(X_T) &= \sum_{j=1}^k g^{(j)}(X_T)\sum D^{k(I_1)}_{r(I_1)}X_T\ldots D^{k(I_j)}_{r(I_j)}X_T \\ H^k(X_s,Y_s,Z_s) &= h_x(s)D^k_{r_1\ldots r_k}X_s + \sum_{j=2}^k \sum_{j_1+j_2+j_3=j} \frac{\partial^j h(s)}{\partial x^{j_1}\partial y^{j_2}\partial z^{j_3}} \\ &\sum D^{k(I_1)}_{r(I_1)}X_s\cdots D^{k(I_{j_1})}_{r(I_{j_1})}X_sD^{k(J_1)}_{r(J_1)}Y_s\cdots D^{k(J_{j_2})}_{r(J_{j_2})}Y_sD^{k(L_1)}_{r(L_1)}Z_s\cdots D^{k(L_{j_3})}_{r(L_{j_3})}Z_s \end{aligned}$$

and the sums are taken over all possible partitions $I_1,\ldots,\ I_j,\ I_1,\ldots,I_{j_1},\ J_1,\ldots,J_{j_2},\ L_1,\ldots,L_{j_3}$ of $\{1,\ldots,k\}$. By hypothesis, $g^{(j)},h_y(s),\ h_z(s)$ and $\frac{\partial^j h(s)}{\partial x^{j_1}\partial y^{j_2}\partial z^{j_3}}$ are bounded for any choice of the indices, consequently $G^k(X_T)$ and $H^k(X_s,Y_s,Z_s)$ satisfy the hypotheses of Lemma 4.1, thus $D^k_{r_1\ldots r_k}Y_t, D^k_{r_1\ldots r_k}Z_t$ belong to $\mathbb{D}^{1,\infty}$. □

Remark 4.3 : *Actually $D.Y_t$ can be defined even if g and h are not regular in x, as long as $\gamma^{-1}_{X_T}\in L^p$ for all $p>1$.*

For simplicity, consider $Y_t = E(\delta_a(X_T)|\mathcal{F}_t),\ t<T$, where δ_a stands for the Dirac delta function. If X_T is non-degenerate conditioned to $\mathcal{F}_t$ for any $t<T$, then one can use the integration by parts to write

$$E(\delta_a(X_T)|\mathcal{F}_t) = E(1_{\{X_T\ge a\}}\delta^T_t((D_rX_T)\gamma^{-1}_{X_T})|\mathcal{F}_t).$$

In this sense Y_t is uniquely defined. To prove that $Y_t\in\mathbb{D}^\infty$, we choose a sequence of smooth bounded functions g_k converging weakly to δ_a. Since the conditional density of X_T given $\mathcal{F}_t$ is smooth and bounded, then $Y^k_t = E(g_k(X_T)|\mathcal{F}_t)$ converges to Y_t in $L^p(\Omega)$ for any $t<T$. Besides, by the integration by parts formula, we have for any $r\le t$

$$D_rY^k_t = E(g'_k(X_T)D_rX_T|\mathcal{F}_t) = E(g_k(X_T)\delta^T_t((D_rX_T)\gamma^{-1}_{X_T}D.X_T)|\mathcal{F}_t). \tag{13}$$

Integrating by parts once more, by dominated convergence theorem, the right side of equality (13) converges, which, by Lemma 1.2.3 in [N1], gives D_rY_t.

The above remark becomes important when considering binary options ($g(x) = 1_{\{x\ge K\}}$) or when studying the properties of fundamental solutions of the associated quasilinear parabolic pde's ($g(x)=\delta_a(x)$).

Using this remark one may extend the next two theorems in certain cases when g and h are not regular in x.

Now we are going to describe two ways to obtain the regularity of the density of Y_t. In the first, we require global conditions similar to those in Theorem 3.1. In the second, in order to localize our conditions, we add the hypothesis of uniform ellipticity for the diffusion, since we need to apply the lower bounds estimates as in [A] or [KSIII] for the densities of X.

Theorem 4.4 : *Let $b,\sigma:\mathbb{R}\to\mathbb{R}$ and h,g be infinitely differentiable in space with bounded derivatives of all orders greater than one and the Hörmander hypothesis be satisfied, that is either $\sigma(x_0)\ne 0$ or there exists $n\in\mathbb{N}$ such that $\sigma^{(n)}b(x_0)\ne 0$.*

Let $t\in(0,T)$ be fixed and assume the same notation as in theorem 3.1. If there exists a constant $\mu>0$ such that one of the following holds

$$\underline{g}\mathrm{e}^{-\mathrm{sgn}\,(\underline{g})KT} + \underline{h}(t)\int_t^T \mathrm{e}^{-\mathrm{sgn}\,(\underline{h}(s))Ks}ds > \mu, \tag{14}$$

$$\bar{g}\mathrm{e}^{\mathrm{sgn}\,(\bar{g})KT} + \bar{h}(t)\int_t^T \mathrm{e}^{\mathrm{sgn}\,(\bar{h}(s))Ks}ds < -\mu, \tag{15}$$

then Y_t has an infinitely differentiable density for $t \in (0,T)$.

Proof: It is enough to prove that $\gamma_{Y_t}^{-1} \in \bigcap_{p>1} L^p$.

$$E((\gamma_{Y_t}^{-1})^p) = E\Big(\frac{1}{[\xi_t^2 \int_0^t (\zeta_r^{-1}\sigma(X_r))^2 dr]^p}\Big).$$

By Hölder inequality we have

$$E((\gamma_{Y_t}^{-1})^p) \leq \Big[E\Big(\frac{1}{\xi_t^{4p}}\Big)\Big]^{1/2}\Big[E\Big(\frac{1}{(\int_0^t (\zeta_r^{-1}\sigma(X_r))^2 dr)^{2p}}\Big)\Big]^{1/2}$$

the second factor is bounded by virtue of the Hörmander hypothesis, hence we only have to show the first is bounded. We assume now that (14) is satisfied, as the other case follows similarly. So we have

$$\xi_t = E\Big(\psi_T g'(X_T)\zeta_T + \int_t^T \psi_s h_x(s)\zeta_s ds|\mathcal{F}_t\Big)\psi_t^{-1} = E_Q\Big(H_T g'(X_T) + \int_t^T H_s h_x(s) ds|\mathcal{F}_t\Big)\psi_t^{-1} > \mu\psi_t^{-1}.$$

Here Q is the measure determined by the change of measure in (5). Therefore we obtain immediately $E\Big(\frac{1}{\xi_t^{4p}}\Big) \leq \frac{C}{\mu^{4p}}$, which is finite because the properties of the flows. □

As before it is not difficult to obtain second order conditions of smoothness. Instead, we show how to localize condition (14).

Theorem 4.5 : *Let b, σ be bounded and b, σ, h, g be infinitely differentiable in space with bounded derivatives of all orders greater or equal than one and let $|\sigma(\cdot)| \geq c > 0$ (i.e., σ is uniformly elliptic). If for fixed $t \in (0,T)$ the first inequality in either (7) or (8) holds and there exist $\mu > 0$ and a set $A \subseteq \mathbb{R}$ with $\lambda(A) > 0$ such that correspondingly one of the following is verified*

$$\underline{g}^A \mathrm{e}^{-\mathrm{sgn}\,(\underline{g}^A)KT} + \underline{h}(t)\int_t^T \mathrm{e}^{-\mathrm{sgn}\,(\underline{h}(s))Ks} ds > \mu \tag{16}$$

$$\bar{g}^A \mathrm{e}^{\mathrm{sgn}\,(\bar{g}^A)KT} + \bar{h}(t)\int_t^T \mathrm{e}^{-\mathrm{sgn}\,(\bar{h}(s))Ks} ds < -\mu, \tag{17}$$

then Y_t has a smooth density.

Proof: The proof is done in two steps. For the first we consider the smoothness of the density of $Y_t^{(s,x)}$ for s close to t with t fixed. For this, consider the Malliavin variance of $Y_t^{(s,x)}$,

$$\gamma_{Y_t^{(s,x)}} = (\xi_t^{(s,x)})^2 \int_s^t [(\zeta_r^{(s,x)})^{-1}\sigma(r, X_r^{(s,x)})]^2 dr.$$

We have that for $s < t$,

$$\Big(\int_s^t [(\zeta_r^{(s,x)})^{-1}\sigma(r, X_r^{(s,x)})]^2 dr\Big)^{-1} \in \bigcap_{p>1} L^p,$$

therefore is enough to prove that $(\xi_t^{(s,x)})^{-1} \in \bigcap_{p>1} L^p$.

We assume hypotheses (7) and (17) are satisfied and we show that for $s \leq t$, $Y_t^{(s,x)}$ has a smooth density of order $\nu > 0$, if $t - s$ is small enough. From theorem 3.1, we get (with the obvious modifications) that $H_T^{(s,x)} g'(X_T^{(s,x)}) + \int_t^T H_r^{(s,x)} h_x^{(s,x)}(r) dr$ is (Q or P) a.s. nonnegative, then

$$\begin{aligned}
\xi_t^{(s,x)} &= E_Q\Big(H_T^{(s,x)} g'(X_T^{(s,x)}) + \int_t^T H_r^{(s,x)} h_x^{(s,x)}(r) dr | \mathcal{F}_t\Big)(\psi_t^{(s,x)})^{-1} \\
&\geq \Big(\underline{g}^A \mathrm{e}^{-\mathrm{sgn}\,(\underline{g}^A) KT} + \underline{h}(t) \int_t^T \mathrm{e}^{-\mathrm{sgn}\,(\underline{h}(r)) Kr} dr\Big) Q(X_T^{(s,x)} \in A | \mathcal{F}_t)(\psi_t^{(s,x)})^{-1} \\
&> \mu Q(X_T^{(s,x)} \in A | \mathcal{F}_t)(\psi_t^{(s,x)})^{-1}.
\end{aligned}$$

Since $\psi_t^{(s,x)} \in \bigcap_{p>1} L^p$, according to Proposition 2.1 it only remains to prove that for any $\rho > 0$, $Q(X_T^{(s,x)} \in A|\mathcal{F}_t)^{-1} \in L^p$ for some $p \equiv p(\rho)$, or equivalently that $P(Q(X_T^{(s,x)} \in A|\mathcal{F}_t) < \epsilon) \leq \epsilon^\beta$ for $\epsilon \leq \epsilon_0$ and $\beta = \beta(p)$.

To obtain this result we use the lower bounds estimates for densities of diffusions as proven by [KSIII], Theorem 4.13. By virtue of the uniform ellipticity of σ (see Lemma 4.6), there exists a positive non-random constant C such that

$$Q(X_T^{(s,x)} \in A|\mathcal{F}_t) \geq \frac{C}{(T-t)^{1/2}} \int_A \exp(-\frac{|X_t^{(s,x)} - y|^2}{C(T-t)}) dy.$$

Therefore we can conclude

$$\begin{aligned}
Q(X_T^{(s,x)} \in A|\mathcal{F}_t) &\geq \frac{C}{\sqrt{T-t}} \int_A \exp(-\frac{(X_t^{(s,x)} - y)^2}{C(T-t)}) dy \\
&\geq \frac{C}{\sqrt{T-t}} \exp(-\frac{(X^{(s,x)})_t^2}{C(T-t)}) \int_A \mathrm{e}^{-\frac{y^2}{C(T-t)}} dy = C_A(T-t) \exp(-\frac{(X_t^{(s,x)})^2}{C(T-t)}).
\end{aligned}$$

Let $\epsilon > 0$ be so that $\frac{\epsilon}{C_A(T-t)} < 1$ (note that $T - t$ is fixed), using a similar Gaussian type upper bound for the density, one obtains

$$\begin{aligned}
P(Q(X_T^{(s,x)} \in A|\mathcal{F}_t) < \epsilon) &\leq P(C_A(T-t) \exp(-\frac{(X_t^{(s,x)})^2}{C(T-t)}) < \epsilon) \\
&= P\Big(|X_t^{(s,x)}| > \sqrt{-C(T-t) \ln \frac{\epsilon}{C_A(T-t)}}\Big) \\
&\leq C_1 \int_{\{|x| > \sqrt{-C(T-t)\ln \frac{\epsilon}{C_A(T-t)}}\}} \frac{1}{\sqrt{t-s}} \exp(-C_2 \frac{(y-x_0)^2}{t-s}) dy \\
&\leq C_3 \left(\frac{\epsilon}{C_A(T-t)}\right)^{\frac{T-t}{C_3(t-s)}},
\end{aligned}$$

for some proper constants C_1, C_2, C_3 and $\epsilon \leq \epsilon_0$, independent of $s < t < T$. Therefore, by Proposition 2.1, $Y_t^{(s,x)}$ has a degree of smoothness depending on $t - s$, increasing to infinity as $t - s$ gets closer to 0.

In the following, we denote by $p_Y^{(s,z)}(t, y)$ the density of Y_t at point y when X starts at z at time s. Note that this density exists due to Theorem 3.1. To finish the proof we use the Markov property for X.

$$\begin{aligned} p_Y^{0,x_0}(t,y) &= E(\delta_y(Y_t^{0,x_0})) = E(E(\delta_y(u(t,X_t^{0,x_0}))|\mathcal{F}_s)) \\ &= E(E(\delta_y(u(t,X_t^{s,X_s}))|\mathcal{F}_s)) = \int E(\delta_y(u(t,X_t^{s,z})))p_X^{(0,x_0)}(s,z)dz \\ &= \int E(\delta_y(Y_t^{s,z}))p_X^{(0,x_0)}(s,z)dz = \int p_Y^{s,z}(t,y)p_X^{0,x_0}(s,z)dz. \end{aligned}$$

Therefore in order to prove the differentiability of order k of $p_Y^{0,x_0}(t,y)$ is enough to prove the differentiability of order k of $p_Y^{s,z}(t,y)$ for some $s \le t$ that may depend on k. Using the previous result that establishes that if s is close to t then the density of $Y_t^{s,z}$ is smooth to a high degree, we end the proof. □

Lemma 4.6 *Under the same conditions of Theorem 4.5*

$$Q(X_T^{(s,x)} \in A|\mathcal{F}_t) \ge \frac{C}{(T-t)^{1/2}} \int_A \exp(-\frac{|X_t^{(s,x)}-y|^2}{C(T-t)})dy.$$

for a non-random constant C.

Proof: To simplify the notation, we assume without loss of generality that $s=0$. Recall that

$$\frac{dQ}{dP} = M_T = \exp\{\int_0^T (\sigma_x(X_s)+h_z(s))dW_s - \frac{1}{2}\int_0^T (\sigma_x(X_s)+h_z(s))^2 ds\}.$$

We consider a regular approximation of the identity ψ_λ such that $\psi_\lambda(x) = x$ if $|x| \le \lambda - 1$ and $|\psi_\lambda(x)| \le \lambda$ for all $x \in \mathbb{R}$ and we denote

$$M_T^\lambda = \exp\left\{\psi_\lambda\left(\int_t^T (\sigma_x(s,X_s)+h_z(s))dW_s\right) - \frac{1}{2}\int_t^T (\sigma_x(s,X_s)+h_z(s))^2 ds\right\}.$$

Therefore we can write

$$\begin{aligned} Q(X_T \in A|\mathcal{F}_t) &= E(\mathbf{1}_{\{X_T\in A\}}M_T M_t^{-1}|\mathcal{F}_t) = E(\mathbf{1}_{\{X_T\in A\}}M_T^\lambda|\mathcal{F}_t) + E(\mathbf{1}_{\{X_T\in A\}}[M_T M_t^{-1} - M_T^\lambda]|\mathcal{F}_t) \\ &= E(\mathbf{1}_{\{X_T\in A\}}M_T^\lambda|\mathcal{F}_t) + R_\lambda. \end{aligned}$$

By the boundedness of σ_x, h_z and ψ_λ, we have $\mathrm{e}^{-\lambda-K(T-t)} \le M_T^\lambda \le \mathrm{e}^{\lambda+K(T-t)}$, thus

$$Q(X_T \in A|\mathcal{F}_t) \ge \mathrm{e}^{-\lambda-K(T-t)}P(X_T \in A|\mathcal{F}_t) + R_\lambda$$

and applying Theorem 4.13 in [KSIII], the following estimate holds

$$Q(X_T \in A|\mathcal{F}_t) \ge \frac{C(T)e^{-\lambda}}{(T-t)^{1/2}} \int_A \exp(-\frac{|X_t-y|^2}{C(T-t)})dy + R_\lambda.$$

Now we estimate the residue, assuming $\lambda > 1$

$$\begin{aligned} |R_\lambda| &\le E(\mathbf{1}_{\{X_T\in A\}}M_T^\lambda|M_T M_t^{-1}(M_T^\lambda)^{-1} - 1|\,|\mathcal{F}_t) \\ &\le \mathrm{e}^{\lambda+K(T-t)}E(\mathbf{1}_{\{X_T\in A\}}\left[\mathrm{e}^\lambda \mathrm{e}^{\int_t^T(\sigma_x(X_s)+h_z(s))dW_s} + 1\right]\mathbf{1}_{\{|\int_t^T(\sigma_x(X_s)+h_z(s))dW_s|>\lambda-1\}}|\mathcal{F}_t) \\ &\le C\mathrm{e}^\lambda(\mathrm{e}^\lambda+1)E(\mathbf{1}_{\{X_T\in A\}}\left[\mathrm{e}^\lambda \mathrm{e}^{\int_t^T(\sigma_x(X_s)+h_z(s))dW_s} \vee 1\right]\mathbf{1}_{\{|\int_t^T(\sigma_x(X_s)+h_z(s))dW_s|>\lambda-1\}}|\mathcal{F}_t) \\ &\le \frac{C(T)\mathrm{e}^{3\lambda}}{(T-t)^{1/8}}\Big(\int_A \exp(-\frac{|X_t-y|^2}{C(T-t)})dy\Big)^{\frac{1}{4}} \exp\{-C\frac{(\lambda-1)^2}{T-t}\}, \end{aligned}$$

where we used the Hölder inequality and the flow property. The proof finishes by choosing λ appropriately large. $\square$

5 Tail Behavior

Here we give some upper bounds for the tail behavior of the density of Y. As in the previous section we assume that all coefficients are smooth with bounded derivatives and that b and σ do not depend on time. The tails behave as lognormal tails if the coefficients grow linearly and as Gaussian tails if the coefficients are bounded. The time dependence does not seem to be optimal except in the uniformly elliptic case.

Theorem 5.1 : *Assume the same conditions as in Theorem 4.4 or 4.5. Then we have that the density of Y_t satisfies for large $|y|$*

$$p_Y^{0,x_0}(t,y) \leq A_0 \frac{(1+|x_0|^p)}{t^{(n+\frac{1}{2})\vee(4n)}} \exp\Big\{ - \frac{[\ln(1+|y-u(t,x_0)|)+A_1]^2}{A_2 t}\Big\}$$

in the case of Theorem 4.4. and

$$p_Y^{0,x_0}(t,y) \leq A_0 \frac{(1+|x_0|^p)}{\sqrt{t}} \exp\Big\{ - \frac{|y-u(t,x_0)|^2}{A_2 t}\Big\}$$

in the case of Theorem 4.5. Here $p>1$ and A_0, A_1, A_2 are positive constants, while $n=0$ in the case that $\sigma(x_0)\neq 0$ and $n\geq 1$ is the index such that $\sigma^{(n)}b(x_0)\neq 0$ in the case of Theorem 4.4.

Proof: Let us fix $t\in(0,T)$. For one dimensional random variables we know that

$$p_Y^{0,x_0}(t,y) = E[\mathbf{1}_{\{Y_t>y\}}\delta(\gamma_{Y_t}^{-1}D.Y_t)] \leq P(Y_t>y)^{\frac{1}{2}}\Big[E([\delta(\gamma_{Y_t}^{-1}D.Y_t)]^2)\Big]^{\frac{1}{2}}. \tag{18}$$

First we study the last factor. One can dominate this term using (1.48) in [N1]

$$E([\delta(\gamma_{Y_t}^{-1}D.Y_t)]^2) \leq E\Big(\int_0^t (\gamma_{Y_t}^{-1}D_sY_t)^2 ds\Big) + E\Big(\int_0^t\int_0^t (D_u(\gamma_{Y_t}^{-1}D_sY_t))^2 ds du\Big). \tag{19}$$

For the first term the following estimate holds for any $p>1$

$$\begin{aligned} & E\Big(\int_0^t (\gamma_{Y_t}^{-1}D_sY_t)^2 ds\Big) = E\Big(\gamma_{Y_t}^{-2}\int_0^t (D_sY_t)^2 ds\Big) = E(\gamma_{Y_t}^{-1}) \\ = \;& E\Big([\xi_t^2\zeta_t^{-2}\gamma_{X_t}]^{-1}\Big) \leq \Big[E\Big([\xi_t^{-1}\zeta_t]^{2q}\Big)\Big]^{\frac{1}{q}}\Big[E\Big([\gamma_{X_t}^{-1}]^p\Big)\Big]^{\frac{1}{p}} \leq C_0\frac{(1+|x_0|^p)}{t^{2n+1}}, \end{aligned}$$

where in the last inequality we used either theorem 4.4 or 4.5 and the upper bound estimates for forward SDE's, as in Corollary 3.25 in [KSII], with n coming from the order of the Hörmander condition verified by X (hence $n=0$ in the case of theorem 4.5). Similarly one can prove that $\|\gamma_{Y_t}^{-1}\|_p \leq C_0\dfrac{(1+|x_0|^p)}{t^{2n+1}}$, with n chosen as before.

Thus it remains to estimate the second term in (19), which by the chain rule is

$$E\Big(\int_0^t\int_0^t (D_u(\gamma_{Y_t}^{-1}D_sY_t))^2 ds du\Big) = E\Big(\int_0^t\int_0^t [\gamma_{Y_t}^{-1}D_uD_sY_t - \gamma_{Y_t}^{-2}\int_0^t 2D_uD_vY_tdvD_sY_t]^2 ds du\Big).$$

To bound efficiently these quantities, one needs to obtain that for $p \geq 1$

$$E\Big(\int_0^t (D_sY_t)^2 ds\Big)^p \leq Ct^p, \quad E\Big(\int_0^t \int_0^t (D_uD_sY_t)^2 duds\Big)^p \leq Ct^{2p}$$

and the last immediately implies, by Jensen's inequality,

$$E\Big(\int_0^t \Big(\int_0^t D_uD_sY_t ds\Big)^2 du\Big)^p \leq Ct^{3p}.$$

The first two inequalities follow directly from Lemma 4.2. Applying the three bounds we obtain

$$E\Big(\int_0^t \int_0^t (D_u(\gamma_{Y_t}^{-1} D_sY_t))^2 dsdu\Big) \quad \leq \quad C\Big(\|\gamma_{Y_t}^{-1}\|_4^2 t^2 + \|\gamma_{Y_t}^{-1}\|_{12}^4 t^4\Big) \leq C(1+|x_0|^p)^4 t^{-8n}.$$

We still need to estimate the first factor in (18). We recall that $Y_s^{t,x} = u(t, X_s^{t,x})$ where u is a $\mathcal{C}^{1,2}$ solution of equation (6). Using the flow properties and the explicit formula for the solution of linear backward sde's (like in section 2) one obtains that u_x is bounded uniformly in t.

Now we first consider the situation of Theorem 4.4. Without loss of generality, we may take $y > |u(t, x_0)|$, then

$$\begin{aligned} P(|Y_t| > y) \quad &\leq \quad P\Big(|u(t, X_t^{0,x_0}) - u(t, x_0)| + |u(t, x_0)| > y\Big) \leq P\Big(c|X_t^{0,x_0} - x_0| > y - |u(t, x_0)|\Big) \\ &\leq \quad P\Big(\ln(1 + |X_t^{0,x_0} - x_0|^2) > \ln(1 + c^{-2}(y - |u(t, x_0)|)^2)\Big) \end{aligned}$$

Using Itô's formula we decompose $\ln(1 + |X_t^{0,x_0} - x_0|^2) = A_t + M_t$, where the total variation part verifies $|A_t| \leq C_1 t$ for some constant C_1 and the martingale part

$$M_t = \int_0^t \frac{2(X_s^{0,x_0} - x_0)\sigma(X_s^{0,x_0})}{1 + |X_s^{0,x_0} - x_0|^2} dW_s$$

verifies $< M >_t \leq C_2 t$, for some constant C_2. By the martingale exponential inequality (see (A.5) in [N1]), we finally obtain

$$\begin{aligned} P(|Y_t| > y) \quad &\leq \quad P\Big(A_t + M_t > \ln(1 + c^{-2}(y - |u(t, x_0)|)^2)\Big) \\ &= \quad P\Big(< M >_t \leq C_2 t, M_t > \ln(1 + c^{-2}(y - |u(t, x_0)|)^2) - C_1 t\Big) \\ &\leq \quad 2\exp\Big\{ - \frac{[-C_1 t + \ln(1 + c^{-2}(y - |u(t, x_0)|)^2)]^2}{2C_2 t}\Big\}, \end{aligned}$$

for some positive constants. Joining the two estimates, the result follows.

In the case of Theorem 4.5, the situation is similar but one does not need to apply the logarithmic transformation. □

References

[A] Azencott, Robert (1984) "Densite des diffusions en temps petit: developpements asymptotiques. I". (French) [Density of diffusions in small time: asymptotic expansions. I] *Seminar on probability, XVIII*, 402–498, Lecture Notes in Math., 1059, Springer, Berlin.

[BH] N. Bouleau and F. Hirsch (1991) "Dirichlet Forms and Analysis on Wiener Space". de Gruyter Studies in Math. 14. em Walter de Gruyter

[CM] P. Cattiaux and L. Mesnager (2002) "Hypoelliptic non homogeneous diffusions". *Probab. Theory Related Fields* **123**, no.4, 453–483.

[CZ] M. Chen and X. Zhou (1991) "Applications of Malliavin Calculus to Stochastic Differential Equations with Time Dependent Coefficients". *Acta Math. Appl. Sinica* **7**, no. 3, 193–216.

[D] E. B. Davies (1989) "Heat kernels and spectral theory". *Cambridge Tracts in Mathematics*, 92, Cambridge University Press, Cambridge.

[EPQ] N. El Karoui, S. Peng, M.C. Quenez (1997) " Backward stochastic differential equations in finance". *Math. Finance* 7, no. 1, 1–71.

[F] P. Florchinger (1990) "Malliavin calculus with time dependent coefficients and application to nonlinear filtering". *Probab. Theory Related Fields* **86**, no. 2, 203–223.

[GP] I. Gyöngy and T. Pröhle (1990) "On the approximation of stochastic differential equation and on a Stroock-Varadhan's Support theorem " *Computers Math. Applic.***19**, no. 1, 65–70.

[KSII] S. Kusuoka and D. Stroock (1985) "Applications of the Malliavin calculus II". *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **32**, 1–76.

[KSIII] S. Kusuoka and D. Stroock (1987) "Applications of the Malliavin calculus III". *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **34**, no. 2, 391–442.

[MZ] J. Ma, J. Zhang (2002) "Representation Theorems for Backward Stochastic Differential Equations". *Annals of Applied Probability*, **12**, no. 4, 1390–1418

[MS] A. Millet, M.Sanz-Solé (1994) "A simple proof of the support theorem for diffusion processes". *Séminaire de Probabilités, XXVIII*, Lecture Notes in Math. vol. 1583, 36–48 Springer, Berlin.

[N1] D. Nualart (1996) "The Malliavin Calculus and Related Topics". *Probability and its Applications* Springer Verlag.

[N2] D. Nualart (1998) "Analysis on Wiener space and anticipating stochastic calculus". in *Ecole d'été de Probabilités de Saint-Flour XXV (1995).* Lecture Notes in Mathematics vol. 1690 Springer Verlag.

[PP1] E. Pardoux, S. Peng (1990) “Adapted solutions of Backward SDE’s”. *System and Control Letters*, **14** , 55-61.

[PP2] E. Pardoux, S. Peng (1993) “BSDE’s and Quasilinear Parabolic PDE’s”. Stochastic partial differential equations and their applications (Charlotte, NC, 1991), 200–217 *Lecture Notes in Control and Inform. Sci.*

[T] S. Taniguchi (1985) “Applications of Malliavin Calculus to time-dependent systems of Heat equations”. *Osaka J. Math*, **22** , 307-320.

6 Appendix

Here we prove the result on existence of densities for non-homogeneous diffusions. This is related with previous results in [CZ] and [CM] and it is of interest in itself.

Theorem 6.1 *Let us assume that, uniformly in t, $\sigma(t,\cdot), b(t,\cdot)$ are in $C_b^\infty(\mathbb{R})$ and besides that they are continuous at $(0,x_0)$. If condition (H0) is satisfied, then X_t has a law that is absolutely continuous with respect to the Lebesgue measure.*

Proof: We show our result by contradiction, hence we assume that $\int_0^t \zeta_s^{-2}\sigma(s,X_s)^2 ds = 0$, which, by the continuity of the processes and of the coefficient, implies that $\sigma(s,X_s)=0$ a.s. for all $s \le t$.

If hypothesis (H0) is verified for $\alpha = 0$, that is $\sigma(0,x_0) = A_0(0,x_0) \neq 0$, we immediately reach a contradiction. Next we proceed by induction on α.

If $\alpha \in (0,1]$ and $A_\beta(0,x_0)=0$ for all $0 \le \beta < \alpha$, by using Itô’s formula we may write for all $s \le t$

$$\begin{aligned}\int_0^s B_1(s,r,X_r)dr + D_0(s,x) &= \int_0^s \sigma_x(s,X_r)b(r,X_r)dr + \sigma(s,x_0) - \sigma(0,x_0)\\ &= \sigma(s,X_s) - \sigma(s,x_0) + \sigma(s,x_0) - \sigma(0,x_0) = \sigma(s,X_s) = 0.\end{aligned}$$

The term of biggest order in $\int_0^s B_1(s,r,X_r)dr + D_0(s,x_0)$ is exactly A_α. Indeed dividing the latter by t^α and taking limits as $t \to 0$ we obtain

$$A_\alpha(0,x_0) = 1_{\{\alpha=1\}}B_1(0,0,x_0) + \lim_{t\to 0}\frac{\sigma(t,x_0)-\sigma(0,x_0)}{t^\alpha} = 1_{\{\alpha=1\}}B_1(0,0,x_0) + \lim_{t\to 0}\frac{D_0(t,x_0)}{t^\alpha} = 0,$$

which contradicts our hypothesis $A_\alpha(0,x_0) \neq 0$.

Let $\alpha > 1$ and $A_\beta(0,x_0)=0$ for all $\beta < \alpha$. In particular, for $j \le [\alpha]$ ($[\cdot]$ = integer part), we have

$$\frac{\partial A_{j-1}}{\partial t}(0,x_0) + \frac{1}{j!}B_j(0,\ldots,0,x_0) = 0.$$

Therefore, using Itô's formula we obtain

$$\begin{aligned}
&\int_0^s \int_0^{s_j} ... \int_0^{s_2} B_j(s,s_1,...,s_j,X_{s_j}) ds_1...ds_j + D_{j-1}(s,x_0) \\
=& \int_0^s \int_0^{s_j} ... \int_0^{s_2} \Big[B_j(s,s_1,...,s_j,X_{s_j}) - B_j(0,...,0,x_0) \Big] ds_1...ds_j + D_{j-1}(s,x_0) - \frac{\partial A_{j-1}}{\partial t}(0,x_0) s^j \\
=& \int_0^s \int_0^{s_j} ... \int_0^{s_2} \Big[B_j(s,s_1,...,s_j,X_{s_j}) - B_j(s,s_1,...,s_j,x_0) \Big] ds_1 \dots ds_j + D_{j-1}(s,x_0) \\
-& \frac{\partial A_{j-1}}{\partial t}(0,x_0) s^j + \int_0^s \int_0^{s_j} ... \int_0^{s_2} \Big[B_j(s,s_1,...,s_j,x) - B_j(0,...,0,x_0) \Big] ds_1...ds_j \\
=& \int_0^s \int_0^{s_{j+1}} ... \int_0^{s_2} B_{j+1}(s,s_1,...,s_{j+1},X_{s_{j+1}}) ds_1...ds_{j+1} + D_j(s,x_0).
\end{aligned}$$

This proves by iteration (under the inductive hypothesis) that for all $j \leq [\alpha]$ and $s \leq t$

$$\int_0^s \int_0^{s_j} ... \int_0^{s_2} B_j(s,s_1,...,s_j,X_{s_j}) ds_1...ds_j + D_{j-1}(s,x_0) = 0. \tag{20}$$

To finish the proof we have to consider two cases. If $[\alpha] \neq \alpha$, then writing (20) for $[\alpha]$, dividing both sides by t^α and taking limits for $t \to 0$ we find that

$$0 = \lim_{t \to 0} \frac{D_{[\alpha]}(t,x_0)}{t^\alpha},$$

because the integrand $B_{[\alpha]}(s,s_1,\dots,s_j,X_{s_{[\alpha]}})$ is a.s. bounded, giving the contradiction with (H0).

If $[\alpha] = \alpha$, first note that $B_{[\alpha]}$ is continuous at $(0,...,0,x_0)$. Then using the same procedure we instead obtain

$$0 = \lim_{t \to 0} \frac{D_{\alpha-1}(t,x_0)}{t^\alpha} + \frac{1}{\alpha!} B_\alpha(0,\dots,0,x_0).$$

that is different from 0 by hypothesis. This ends the proof. □

Università di L'Aquila
Dipartimento di Matematica
Via Vetoio, loc. Coppito
67100 L'Aquila - ITALY
antonf@univaq.it

Universitat Pompeu Fabra
Departament d'Economia
Ramon Trias Fargas 25-27
Barcelona 08005 - SPAIN
kohatsu@econ.upf.es