Density estimates for solutions to one dimensional Backward SDE's

Omar Aboura, Solesne Bourguin

Density estimates for solutions to one dimensional Backward SDE's

Omar Aboura[*] and Solesne Bourguin[†‡]

Abstract: In this paper, we derive sufficient conditions for each component of the solution to a general backward stochastic differential equation to have a density for which upper and lower Gaussian estimates can be obtained.

Keywords: Backward stochastic differential equations, Malliavin calculus, density estimates.

2010 AMS Classification Numbers: 60H10, 60H07.

1 Introduction

In [NV09], I. Nourdin and F.G. Viens have introduced sufficient conditions to prove the existence of a density for a Malliavin differentiable random variable and to provide upper and lower Gaussian estimates for this density.

This result has led to several research papers, such as those by D. Nualart and L. Quer-Sardanyons ([NQ09a], [NQ09b]), in which these authors applied Nourdin and Viens result to solutions of quasi-linear stochastic partial differential equations and to a class of stochastic equations with additive noise.

In this paper, we use Nourdin and Viens's approach to prove that, under proper conditions on the coefficients, each component of the solution (X_t, Y_t, Z_t) to a backward stochastic differential equation

$$
\begin{cases}
X_t = x_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s & (1.1) \\[2mm]
Y_t = \xi + \int_t^T f(X_s, Y_s, Z_s)\,ds - \int_t^T Z_s dW_s & (1.2)
\end{cases}
$$

has a density for which upper and lower Gaussian bounds can be derived. This implies to study the relation between the coefficients of the diffusion equation (1.1) and the coefficients of the backward SDE (1.2).

Our paper is organized as follows: in a first part, we study the component Y_t of the solution and in a second and last part, we focus on the component Z_t (for which, up to our knowledge, no density existence results exist). We will not develop a specific study of the first component X_t of the solution to the BSDE due to the fact that the question of the existence of a density for the solution to an SDE of the type (1.1) and the properties of this density has been intensively

---

[*]SAMM, EA 4543, Université Paris 1 Panthéon Sorbonne, 90 Rue de Tolbiac, 75634 Paris Cedex France. Email: `omar.aboura@malix.univ-paris1.fr`

[†]SAMM, EA 4543, Université Paris 1 Panthéon Sorbonne, 90 Rue de Tolbiac, 75634 Paris Cedex France.

[‡]Faculté des Sciences, de la Technologie et de la Communication; UR en Mathématiques. 6, rue Richard Coudenhove-Kalergi, L-1359 Luxembourg. Email: `solesne.bourguin@uni.lu`

studied and we refer the reader to [Nua06] for an extensive survey of the existing litterature and results on this topic.

Equations of the type (1.2) were introduced in [PP90] and are closely related with viscosity solution to PDEs. These equations have been intensively studied and have many applications in control theory and financial methods among others.

The existence of the density for the random variable Y_t at a fixed time $t \in (0, T)$, as well as upper bounds for its tail behavior, have been proven by F. Antonelli and A. Kohatsu-Higa [AK05], using the Bouleau-Hirsch Theorem. We retrieve Antonelli and Kohatsu-Higa's existence result for the density of Y_t, and we also derive Gaussian estimates for it. In order to provide (additionally to the existence result itself) estimates for the density of Y_t, we need to strengthen the hypotheses of Antonelli and Kohatsu-Higa.

We also address the question of the existence of a density for the random variable Z_t as well as the possibility of deriving Gaussian estimates for it. This question has not been solved in [AK05]. We need the same hypotheses as in the case of Y_t, as well as additional ones, since Z_t can be expressed as a function of the Malliavin derivative of Y_t.

In order to be self contained, we first give an overview of some elements of Malliavin calculus in Section 2, and the corresponding notations. Section 3 is dedicated to the component Y_t of the solution to the BSDE. Section 4 deals with the question of the existence of a density for Z_t and is organized in two subsections, dealing respectively with the question of the existence of a density and conditions for this density to be bounded by Gaussian upper and lower estimates.

2 Framework, main tools and notations

2.1 Elements of Malliavin calculus

Consider the real separable Hilbert space $L^2([0, T])$ and let $(W(\varphi), \varphi \in L^2([0, T]))$ be an isonormal Gaussian process on a probability space $(\Omega, \mathfrak{A}, P)$, that is, a centered Gaussian family of random variables such that $\mathbf{E}(W(\varphi)W(\psi)) = \langle \varphi, \psi \rangle_{L^2([0,T])}$. For any integer $n \geq 1$, denote by I_n the multiple stochastic integral with respect to W (see [Nua06] for an extensive survey on Malliavin calculus). The map I_n is actually an isometry between the Hilbert space $L^2([0, T]^n)$ equipped with the scaled norm $\frac{1}{\sqrt{n!}} \| \cdot \|_{L^2([0,T]^n)}$ and the Wiener chaos of order n, which is defined as the closed linear span of the random variables $H_n(W(\varphi))$ where $\varphi \in L^2([0, T]), \|\varphi\|_{L^2([0,T])} = 1$ and H_n is the Hermite polynomial of degree $n \geq 1$, that is defined by

$$H_n(x) = \frac{(-1)^n}{n!} \exp\left(\frac{x^2}{2}\right) \frac{d^n}{dx^n} \left(\exp\left(-\frac{x^2}{2}\right)\right), \quad x \in \mathbb{R}.$$

The isometry of multiple integrals can be written as follows: for positive integers m, n,

$$\begin{aligned}
\mathbf{E}\left(I_n(f)I_m(g)\right) &= n!\langle f, g \rangle_{L^2([0,T]^n)} \quad \text{if } m = n, \\
\mathbf{E}\left(I_n(f)I_m(g)\right) &= 0 \quad \text{if } m \neq n.
\end{aligned}$$

It also holds that

$$I_n(f) = I_n\big(\tilde{f}\big)$$

2

where $\tilde{f}$ denotes the symmetrization of f defined by

$$\tilde{f}(x_1, \ldots, x_n) = \frac{1}{n!} \sum_{\sigma \in \mathcal{S}_n} f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}).$$

We recall that any square integrable random variable F which is measurable with respect to the σ-algebra generated by W can be expanded into an orthogonal sum of multiple stochastic integrals

$$F = \sum_{n \geq 0} I_n(f_n) \tag{2.1}$$

where $f_n \in L^2([0,T]^n)$ are (uniquely determined) symmetric functions and $I_0(f_0) = \mathbf{E}[F]$.

Let L be the Ornstein-Uhlenbeck operator defined by $LF = -\sum_{n \geq 0} n I_n(f_n)$ if F is given by (2.1) and satisfies $\sum_{n \geq 1} n^2 n! \|f_n\|^2 < \infty$. For $p > 1$ and $\alpha \in \mathbb{R}$ we introduce the Sobolev-Watanabe space $\mathbb{D}^{\alpha,p}$ as the closure of the set of random variables of the form (2.2) (see (1.28) in [Nua06]) with respect to the norm defined by

$$\|F\|_{\alpha,p} = \|(I - L)^{\frac{\alpha}{2}} F\|_{L^p(\Omega)},$$

where I represents the identity. We denote by D the Malliavin derivative operator that acts on smooth random variables of the form

$$F = g(W(\varphi_1), \ldots, W(\varphi_n)), \tag{2.2}$$

where g is a smooth function with compact support and $\varphi_i \in L^2([0,T])$, as follows:

$$DF = \sum_{i=1}^{n} \frac{\partial g}{\partial x_i}(W(\varphi_1), \ldots, W(\varphi_n))\varphi_i.$$

The operator D is continuous from $\mathbb{D}^{\alpha,p}$ into $\mathbb{D}^{\alpha-1,p}\left(L^2([0,T])\right)$. The adjoint of D is denoted by δ and is called the divergence (or Skorohod) integral. It is a continuous operator from $\mathbb{D}^{\alpha,p}\left(L^2([0,T])\right)$ into $\mathbb{D}^{\alpha+1,p}$. More generally, we can introduce iterated weak derivatives of order k. If F is a smooth random variables and k is a positive integer, we set

$$D_{t_1,\ldots,t_k}^k F = D_{t_1} D_{t_2} \ldots D_{t_k} F.$$

We have the following duality relationship between D and δ for $F \in \mathbb{D}^{1,2}$ and $u \in dom\ \delta$

$$\mathbf{E}(F\delta(u)) = \mathbf{E}\langle DF, u \rangle_{L^2([0,T])}.$$

For adapted integrands, the divergence integral coincides with the classical Itô integral. We will use the notation

$$\delta(u) = \int_0^T u_s dW_s.$$

Note that the following integration by parts relation between D and δ holds

$$D_t(\delta(u)) = u_t + \int_0^T D_t u_s dW_s,$$

where $u \in \mathbb{D}^{1,2}(L^2([0,T]))$ is such that $\delta(u) \in \mathbb{D}^{1,2}$.

3

2.2 Density existence and Gaussian estimates

A classical density existence result is the celebrated Bouleau-Hirsch theorem (see [Nua06] for an extensive survey on this result). This result provides conditions in term of Malliavin derivatives for a random variable to have a density.

Theorem 2.1 (Bouleau–Hirsch). *Let F be a random variable of the space $\mathbb{D}^{1,2}$ and suppose that $\|DF\|_{L^2([0,T])} > 0$ a.s. Then the law of F is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$.*

In [NV09], Corollary 3.5, Nourdin and Viens have given the following sufficient condition for a weakly differentiable random variable to have a density with lower and upper Gaussian estimates.

Proposition 2.2. *Let F be in $\mathbb{D}^{1,2}$ and let the function g be defined for all $x \in \mathbb{R}$ by*

$$g(x) = \mathbf{E}\left(\langle DF, -DL^{-1}F\rangle_{L^2([0,T])}\Big| F - \mathbf{E}(F) = x\right). \tag{2.3}$$

If there exist positive constants $\gamma_{\min}, \gamma_{\max}$ such that, for all $x \in \mathbb{R}$, almost surely

$$0 < \gamma_{\min}^2 \leqslant g(x) \leqslant \gamma_{\max}^2$$

then F has a density ρ satisfying, for almost all $z \in \mathbb{R}$

$$\frac{\mathbf{E}|F - \mathbf{E}(F)|}{2\,\gamma_{\max}^2} \exp\left(-\frac{(z - \mathbf{E}(F))^2}{2\gamma_{\min}^2}\right) \leqslant \rho(z) \leqslant \frac{\mathbf{E}|F - \mathbf{E}(F)|}{2\,\gamma_{\min}^2} \exp\left(-\frac{(z - \mathbf{E}(F))^2}{2\gamma_{\max}^2}\right).$$

Furthermore, Nourdin and Viens have also provided the following useful result, which gives some rather explicit description of $g(x)$. Recall that $W = \big(W(\phi), \phi \in L^2([0,T])\big)$.

Proposition 2.3. *Let F be in $\mathbb{D}^{1,2}$ and write $DF = \Phi_F(W)$ with a measurable function $\Phi_F : \mathbb{R}^{L^2([0,T])} \to L^2([0,T])$. Then, if $g(x)$ is defined by (2.3), we have*

$$g(x) = \int_0^\infty e^{-u}\, \mathbf{E}\left(\mathbf{E}'\big(\langle\Phi_F(W), \widetilde{\Phi_F^u}(W)\rangle_{L^2([0,T])}\big)|F - \mathbf{E}(F) = x\right) du,$$

where $\widetilde{\Phi_F^u}(W) = \Phi_F(e^{-u}W + \sqrt{1 - e^{-2u}}W')$, W' stands for an independent copy of W, and is such that W and W' are defined on the product probability space $(\Omega \times \Omega', \mathcal{F} \otimes \mathcal{F}', \mathbb{P} \times \mathbb{P}')$ and $\mathbf{E}'$ denotes the mathematical expectation with respect to $\mathbb{P}'$.

2.3 Notations

We denote by $\mathcal{C}_b^n(\mathbb{R}^p)$ the space of n–times differentiable functions on $\mathbb{R}^p$ with bounded partial derivatives up to order n.

Let f be a three times differentiable function of three variables x, y and z. We will use the following notations : $\frac{\partial f}{\partial x} = f_x$, $\frac{\partial f}{\partial y} = f_y$, $\frac{\partial f}{\partial z} = f_z$ $\frac{\partial^2 f}{\partial x^2} = f_{xx}$, $\frac{\partial^2 f}{\partial y^2}(x,y) = f_{yy}$, $\frac{\partial^2 f}{\partial x \partial y} = f_{xy}$, $\frac{\partial^2 f}{\partial y \partial x} = f_{yx}$.

We will also use the following notation for the Lie bracket : $[h,g] = hg' - gh'$, where $h, g : \mathbb{R} \to \mathbb{R}$.

In the whole paper, c and C will denote constants that may vary from line to line.

3 Density of Y_t : existence and Gaussian estimates

The following backward stochastic differential equation was introduced in Pardoux and Peng [PP90] (see also [PP92]) and density properties of its solutions were investigated in [AK05]:

$$
\begin{cases}
X_t = x_0 + \displaystyle\int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s & (3.1) \\[2mm]
Y_t = \xi + \displaystyle\int_t^T f(X_s, Y_s, Z_s)\, ds - \int_t^T Z_s dW_s & (3.2)
\end{cases}
$$

In this section, we give conditions for the random variable Y_t to have a density which can be bounded from above and below by Gaussian ones.

3.1 Hypotheses

We consider b, σ and f to be appropriately smooth functions to ensure the existence and uniqueness of solutions to equations (3.1) and (3.2). We also impose additional conditions needed to state our main result:

$$
\begin{cases}
\mathbf{H1:}\ \xi \in L^2(\Omega, \mathcal{F}_T) \cap \mathbb{D}^{1,2} \quad \text{and} \quad \forall\, \theta \leq T,\ 0 < c \leq D_\theta \xi \leq C \quad a.s. \\
\mathbf{H2:}\ f \in \mathcal{C}_b^1(\mathbb{R}^3) \quad \text{and} \quad 0 \leq f_x \leq C. \\
\mathbf{H3:}\ b \in \mathcal{C}_b^1(\mathbb{R}), \quad \sigma \in \mathcal{C}_b^2(\mathbb{R}), \quad 0 \leq \sigma \leq C \quad \text{and} \quad |[b,\sigma]| \leq M\sigma.
\end{cases}
$$

where $[b, \sigma]$ denotes the Lie bracket between b and σ.

Remark 3.1. Note that it is natural to have a condition on the Lie bracket between b and σ as this quantity is the one that appearing in the classical density and smoothness results (e.g. Hörmander's brackets condition).

Remark 3.2. The hypotheses on the positivity of σ and f_x are made in order to make the proofs as readable as possible. In fact, one only needs σ and f_x to have the same sign to draw to same conclusions.

3.2 Main result

Under the above assumptions, Y_t has a density for which the following Gaussian estimates can be derived.

Theorem 3.3. *Under the above hypotheses, for $t \in (0, T)$ the random variable Y_t defined in (3.2) has a density ρ_{Y_t}. Furthermore, there exist strictly positive constants c and C such that, for almost all $y \in \mathbb{R}$ and all $t \in [0, T]$, ρ_{Y_t} satisfies the following:*

$$
\frac{\mathbf{E}|Y_t - \mathbf{E}(Y_t)|}{2ct}\exp\left(-\frac{(y - \mathbf{E}(Y_t))^2}{2Ct}\right) \leq \rho_{Y_t}(y) \leq \frac{\mathbf{E}|Y_t - \mathbf{E}(Y_t)|}{2Ct}\exp\left(-\frac{(y - \mathbf{E}(Y_t))^2}{2ct}\right).
$$

The rest of this section is devoted to the proof of Theorem 3.3 which is divided in three steps. In the first step, we will prove that the Malliavin derivative of X_t is bounded and non-negative. This property will be necessary in Step 2 (as DX_t appears in the Malliavin derivative of Y_t).

In the second step, we will derive upper and lower bounds for the Malliavin derivative of Y_t. Indeed, our purpose is to use the Nourdin–Viens formula (Proposition 2.2) in which one needs to bound a function of the Malliavin derivative of random variable for which density results are investigated. We will make use of the properties of DX_t prooved in Step 1.

In the third step, we will use the bounds obtained on DY_t and the Nourdin–Viens formula to conclude the proof.

Proof of Theorem 3.3:

Step 1: Boundedness and positivity of DX_t

Consider equation (3.1). Using a Lamperti transform (see [Lam64] or [KS91] pp. 294–295 exercise 2.20), we compute the Malliavin derivative of X_t. The Lamperti transform of X_t, hereafter denoted by U_t, is given by

$$U_t = g(x_0) + \int_0^t \beta \circ g^{-1}(U_s)ds + W_t,$$

where

$$g(x) = \int_0^x \frac{du}{\sigma(u)} \quad \text{and} \quad \beta(x) = \frac{b}{\sigma}(x) - \frac{\sigma'(x)}{2}.$$

Computing the Malliavin derivative of U_t yields, for $\theta \in [0, t]$,

$$D_\theta U_t = 1 + \int_\theta^t (\beta \circ g^{-1})'(U_s)D_\theta U_s ds = \exp\left[\int_\theta^t (\beta \circ g^{-1})'(U_s)ds\right]. \tag{3.3}$$

Deriving the identity $g \circ g^{-1}(x) = x$ on $g(supp(X_t))$ yields $(g^{-1})'(x) = \sigma \circ g^{-1}(x)$. Using this fact we get $(\beta \circ g^{-1})'(x) = \beta' \circ g^{-1}(x)(g^{-1})'(x) = (\beta'\sigma) \circ g^{-1}(x)$. In addition, it is easy to check that on $g(supp(X_t))$,

$$(\beta'\sigma)(x) = \frac{[\sigma, b](x)}{\sigma(x)} - \frac{(\sigma\sigma'')(x)}{2}. \tag{3.4}$$

Gathering those results and using hypothesis (**H3**) of Subsection 3.1 immediately yields on $g(supp(X_t))$

$$-C \leq (\beta \circ G^{-1})' \leq C,$$

where C is a positive constant. Using (3.3), we deduce, $\mathbb{P}$-a.s,

$$0 < c \leq D_\theta U_t \leq C. \tag{3.5}$$

Furthermore, as $X_t = g^{-1}(U_t)$, it holds that, for $0 < \theta < t \leq T$,

$$D_\theta X_t = (g^{-1})'(U_t)D_\theta U_t = \sigma \circ g^{-1}(U_t)D_\theta U_t. \tag{3.6}$$

Combining (3.5) and (3.6) with the fact that σ is bounded and non-negative yields, $\mathbb{P}$-a.s,

$$0 \leq D_\theta X_t \leq C. \qquad (3.7)$$

Step 2: Computation of bounds on DY_t

We at first represent $D_\theta Y_t$ by means of an equivalent probability; this is similar to [AK05] and the proof is included for the sake of completeness. It is well known (see for example Theorem 2.2 in [AK05]) that, for every $t \in (0, T]$, $Y_t \in \mathbb{D}^{1,2}$ and $Z \in L^2\left(0, T; \mathbb{D}^{1,2}\right)$. Furthermore, since $\theta < t$, we have

$$
D_\theta Y_t = D_\theta \xi - \int_t^T D_\theta Z_s dW_s
$$
$$
+ \int_t^T \left[f_x \left(X_s, Y_s, Z_s \right) D_\theta X_s + f_y \left(X_s, Y_s, Z_s \right) D_\theta Y_s + f_z \left(X_s, Y_s, Z_s \right) D_\theta Z_s \right] ds. \qquad (3.8)
$$

The product $e^{\int_0^t f_y(X_s,Y_s,Z_s)ds} D_\theta Y_t$ yields a more suitable representation of $D_\theta Y_t$; indeed, for $t \in (0, T]$, and $0 \leq \theta < t$

$$
d\left[e^{\int_0^t f_y(X_s,Y_s,Z_s)ds} D_\theta Y_t \right] = \left[D_\theta Y_t e^{\int_0^t f_y(X_s,Y_s,Z_s)ds} f_y\left(X_t, Y_t, Z_t \right) \right.
$$
$$
- e^{\int_0^t f_y(X_s,Y_s,Z_s)ds} \left(f_x\left(X_t, Y_t, Z_t \right) D_\theta X_t + f_y\left(X_t, Y_t, Z_t \right) D_\theta Y_t \right.
$$
$$
\left. + f_z\left(X_t, Y_t, Z_t \right) D_\theta Z_t \right] dt + e^{\int_0^t f_y(X_s,Y_s,Z_s)ds} D_\theta Z_t dW_t.
$$

Integrating from t to T yields, for $\theta < t$,

$$
e^{\int_0^T f_y(X_s,Y_s,Z_s)ds} D_\theta Y_T - e^{\int_0^t f_y(X_s,Y_s,Z_s)ds} D_\theta Y_t = -\int_t^T e^{\int_0^s f_y(X_r,Y_r,Z_r)dr} \left[f_x\left(X_s, Y_s, Z_s \right) D_\theta X_s \right.
$$
$$
\left. + f_z\left(X_s, Y_s, Z_s \right) D_\theta Z_s \right] ds + \int_t^T e^{\int_0^s f_y(X_r,Y_r,Z_r)dr} D_\theta Z_s dW_s.
$$

Note that $D_\theta Y_T = D_\theta \xi$; therefore, for $t \in (0, T]$,

$$
D_\theta Y_t = e^{\int_t^T f_y(X_s,Y_s,Z_s)ds} D_\theta \xi + \int_t^T e^{\int_t^s f_y(X_r,Y_r,Z_r)dr} \left[f_x\left(X_s, Y_s, Z_s \right) D_\theta X_s \right.
$$
$$
\left. + f_z\left(X_s, Y_s, Z_s \right) D_\theta Z_s \right] ds - \int_t^T e^{\int_t^s f_y(X_r,Y_r,Z_r)dr} D_\theta Z_s dW_s.
$$

Let $\widetilde{W}_t = W_t - \int_0^t f_z\left(X_s, Y_s, Z_s \right) ds$. Because $f_z \in \mathcal{C}_b^0(\mathbb{R})$, Novikov's condition is verified and $\widetilde{W}$ is a Brownian motion under some equivalent probability $\widetilde{\mathbb{P}}$. Girsanov's theorem yields

$$
D_\theta Y_t = e^{\int_t^T f_y(X_s,Y_s,Z_s)ds} D_\theta \xi + \int_t^T e^{\int_t^s f_y(X_r,Y_r,Z_r)dr} f_x\left(X_s, Y_s, Z_s \right) D_\theta X_s ds
$$
$$
- \int_t^T e^{\int_t^s f_y(X_r,Y_r,Z_r)dr} D_\theta Z_s d\widetilde{W}_s.
$$

Conditionning by $\mathcal{F}_t$ under $\widetilde{\mathbb{P}}$ and taking into account the fact that Y_t and $D_\theta Y_t$ are adapted with respect to $\mathcal{F}_t$ while $\int_t^s f_y(X_r, Y_r, Z_r)\,dr$ and $D_\theta Z_s$ are $\mathcal{F}_s$-adapted for $\theta < t \leq s \leq T$, we obtain

$$
\begin{aligned}
D_\theta Y_t =& \widetilde{\mathbf{E}}\left(e^{\int_t^T f_y(X_s, Y_s, Z_s)ds} D_\theta \xi \,\Big|\, \mathcal{F}_t \right) \\
&+ \widetilde{\mathbf{E}}\left(\int_t^T e^{\int_t^s f_y(X_r, Y_r, Z_r)dr} f_x(X_s, Y_s, Z_s) D_\theta X_s ds \,\Big|\, \mathcal{F}_t \right).
\end{aligned}
\tag{3.9}
$$

Using hypotheses (**H1**) and (**H2**), the first summand in (3.9) can be bounded by two positive constants c and C in the following manner:

$$
0 < c \leq \widetilde{\mathbf{E}}\left(e^{\int_t^T f_y(X_s, Y_s, Z_s)ds} D_\theta \xi \,\Big|\, \mathcal{F}_t \right) \leq C.
\tag{3.10}
$$

Using the results on DX_t proved in Step 1 along with hypothesis (**H1**), we deduce that the second summand in (3.9) is bounded and non-negative: there is a positive constant C such that

$$
0 \leq \widetilde{\mathbf{E}}\left(\int_t^T e^{\int_t^s f_y(X_r, Y_r, Z_r)dr} f_x(X_s, Y_s, Z_s) D_\theta X_s ds \,\Big|\, \mathcal{F}_t \right) \leq C.
\tag{3.11}
$$

Combining the bounds (3.10) and (3.11), we immediately deduce that there exist two positive constants c and C such that

$$
0 < c \leq D_\theta Y_t \leq C.
\tag{3.12}
$$

Step 3: Conclusion of the proof by the Nourdin–Viens formula

Write $D_\bullet Y_t = \Phi_{Y_t}^\bullet(W)$ with a measurable function $\Phi_{Y_t}^\bullet : \mathbb{R}^{L^2([0,T])} \longrightarrow L^2([0,T])$. Then the bounds obtained in (3.12) yield, for $\theta < t$,

$$
0 < c \leq \Phi_{Y_t}^\theta(W) \leq C.
$$

Define $\widetilde{\Phi_{Y_t}^{\bullet,u}}(W) = \Phi_{Y_t}^\bullet(e^{-u}W + \sqrt{1 - e^{-2u}}W')$ for $u \in [0, +\infty[$, where W' stands for an independent copy of W and is such that W and W' are defined on the product probability space $(\Omega \times \Omega', \mathcal{F} \otimes \mathcal{F}', \mathbb{P} \times \mathbb{P}')$. It is clear that, for $\theta < t$, we have for any $u \in [0, \infty)$,

$$
0 < c \leq \widetilde{\Phi_{Y_t}^{\theta,u}}(W) \leq C.
$$

Combining the two previous bounds yields, for $\theta < t$, $u \in [0, \infty)$,

$$
0 < c^2 \leq \Phi_{Y_t}^\theta(W)\widetilde{\Phi_{Y_t}^{\theta,u}}(W) \leq C^2.
\tag{3.13}
$$

Using the notation from Propositions 2.2 and 2.3, we define

$$
g(y) = \int_0^\infty e^{-u} \mathbf{E}\left(\mathbf{E}'\left(\int_0^t \Phi_{Y_t}^\theta(W)\widetilde{\Phi_{Y_t}^{\theta,u}}(W)d\theta \right) \Big| Y_t - \mathbf{E}(Y_t) = y \right) du.
$$

The bounds obtained in (3.13) immediately yield

$$
0 < ct \leq g(y) \leq Ct,
$$

with strictly positive constants c and C. Thus, Propositions 2.2 and 2.3 conclude the proof of Theorem 3.3. $\qquad\square$

4 Density of Z_t : existence and Gaussian estimates

We consider equations (3.1) and (3.2) with a function $f^\star$ that only has a linear dependency on Z, i.e.

$$\begin{cases} X_t = x_0 + \displaystyle\int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s & (4.1) \\[4mm] Y_t = \xi + \displaystyle\int_t^T f^\star\left(X_s, Y_s, Z_s\right)ds - \int_t^T Z_s dW_s & (4.2) \end{cases}$$

where $f^\star(x, y, z) = f(x, y) + \alpha z$, $\alpha \in \mathbb{R}$.

Because of the dependency of f on Z, the Malliavin derivative DZ will depend on $D^2 Z$, which is not suitable for analyzing it within our framework. One can circumvent the above mentioned issue by using the Girsanov theorem to dispose of the impeding terms (similarly as done in the proof of Theorem 3.3). In order to clarify the proofs and to improve readability, we will consider that this step has already been performed in all of our proofs. This procedure leaves us with an equation of the type

$$\begin{cases} X_t = x_0 + \displaystyle\int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s & (4.3) \\[4mm] Y_t = \xi + \displaystyle\int_t^T f\left(X_s, Y_s\right)ds - \int_t^T Z_s dW_s, & (4.4) \end{cases}$$

which is the one that will be referred to in the proofs of the upcoming results.

In the following subsections, we will prove our two main results concerning Z_t. We will begin by giving sufficient conditions for Z_t to have a density. Up to our knowledge, this is the first result on density existence for the component Z of the solution to equation (4.2). We will then study in what framework and under what conditions this density can be bounded by Gaussian estimates.

4.1 Existence of a density for Z_t

We list in the next section the full set of hypotheses we need in this section.

4.1.1 Hypotheses

We consider b, σ and $f^\star$ to be appropriately smooth functions to ensure the existence and uniqueness of solutions to equations (4.1) and (4.2). We also impose additional conditions needed to prove Theorem 4.3.

$$\begin{cases} \textbf{H4 :}\ \xi \in L^2(\Omega, \mathcal{F}_T) \cap \mathbb{D}^{2,2},\ \forall \theta \leq T,\ D_\theta \xi \geq 0\ a.s\ \text{and}\ \forall\, \theta < t \leq T,\ D^2_{\theta,t}\xi > 0\ a.s. \\ \textbf{H5 :}\ f \in \mathcal{C}^2(\mathbb{R}^2)\ \text{and}\ f_x, f_y, f_{xy}, f_{xx}, f_{yy} \geq 0\ a.s. \\ \textbf{H6 :}\ b \in \mathcal{C}^2(\mathbb{R}),\ \sigma \in \mathcal{C}^3(\mathbb{R}),\ \sigma, \sigma', -\sigma'', -\sigma''' \geq 0\ a.s\ \text{and}\ [\sigma, [\sigma, b]] \geq 0\ a.s. \end{cases}$$

where $[b, \sigma]$ denotes the Lie bracket between b and σ.

Remark 4.1. Note that it is natural to have a condition on the iterated Lie bracket $[\sigma, [\sigma, b]]$ between b and σ as second order Malliavin derivatives appear in the expression of Z_t.

Remark 4.2. The hypotheses on the signs of σ and f_x are made in order to make the proofs as readable as possible. It is possible to have more complex hypotheses for the signs of the products of derivatives of σ and derivatives of f.

The following theorem states that under the above hypotheses, Z_t has a density on $\mathbb{R}$

Theorem 4.3. *Under the above hypotheses, for $t \in (0, T)$ the law of the random variable Z_t is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$.*

Before proving Theorem 4.3, we will first give a technical Lemma and a Proposition which will play a key role in the upcoming proof of this Theorem. First recall a lemma used to calculate the Malliavin derivative of a product of random variables in $\mathbb{D}^{1,2}$ (for example, see [Nua06], p.36, exercice 1.2.12).

Lemma 4.4. *(i) Let $s, t \in [0, T]$ and $F \in \mathbb{D}^{1,2}$; then we have $\mathbf{E}(F|\mathcal{F}_t) \in \mathbb{D}^{1,2}$ and*

$$D_s \mathbf{E}(F|\mathcal{F}_t) = \mathbf{E}(D_s F|\mathcal{F}_t) \, 1_{s \leq t}.$$

(ii) If $F, G \in \mathbb{D}^{1,2}$ are such that F and $\|DF\|_{L^2([0,T])}$ are bounded, then $FG \in \mathbb{D}^{1,2}$ and

$$D(FG) = FDG + GDF.$$

The rest of this section is devoted to the proof of Theorem 4.3 which is divided in three steps. In the first step, we will prove that under the conditions of subsection 4.1.1 the second-order Malliavin derivatives of X and Y are non-negative. This will be of importance in Step 2 (as these second-order derivatives appear in the expression of DZ_t.

In the second step, we will show that DZ_t is positive almost surely. This will ensure that $\|DZ_t\|_{L^2([0,T])} > 0$ a.s.

In the third and last step, we will use the Bouleau–Hirsch Theorem to conclude the proof (see Theorem 2.1).

Proof of Theorem 4.3:

Step 1: Non-negativity of $D^2 X$ ***and*** $D^2 Y$

We start by proving that for $0 < \theta < t < s \leq T$, $\mathbb{P}$–a.s $D^2_{\theta,t} X_s$ is non-negative.

Applying the Malliavin derivative to (3.6) and using the second point in Lemma 4.4, we deduce for $\theta, t \leq s \leq T$, since $U_s = g(X_s)$,

$$
\begin{aligned}
D^2_{\theta,t} X_s &= (\sigma \circ g^{-1})'(U_s) D_\theta U_s D_t U_s + (\sigma \circ g^{-1})(U_s) D^2_{\theta,t} U_s \\
&= (\sigma'\sigma)(X_s) D_\theta U_s D_t U_s + \sigma(X_s) D^2_{\theta,t} U_s.
\end{aligned}
\tag{4.5}
$$

Hypothesis (**H6**) ensures that the term $(\sigma'\sigma)(X_s) D_\theta U_s D_t U_s$ is non-negative. It remains to prove that the second summand in (4.5) is also non-negative. As σ is non-negative, we focus on proving

that $D^2_{\theta,t}U_s$ is too. Applying once again the Malliavin derivative operator to (3.3) and using the second point in Lemma 4.4, we deduce for $\theta < t \leq s$,

$$
\begin{aligned}
D^2_{\theta,t}U_s &= \int_t^s (\beta \circ g^{-1})''(U_r)D_tU_rD_\theta U_r dr + \int_t^s (\beta \circ g^{-1})'(U_r)D^2_{\theta,t}U_r dr \\
&= \int_t^s e^{\int_r^s (\beta \circ g^{-1})'(U_v)dv}(\beta \circ g^{-1})''(U_r)D_tU_rD_\theta U_r dr \\
&= \int_t^s (\beta \circ g^{-1})''(U_r)D_rU_sD_tU_rD_\theta U_r dr.
\end{aligned}
$$

Further calculations yield the following expression

$$
\begin{aligned}
\left(\beta \circ g^{-1}\right)''(x) &= \left(\sigma\left(\frac{[\sigma,b]'}{\sigma} - \frac{[\sigma,b]\,\sigma'}{\sigma^2}\right) - \frac{1}{2}\left(\sigma''\sigma\right)'\sigma\right) \circ g^{-1}(x) \\
&= \left(\frac{[\sigma,[\sigma,b]]}{\sigma} - \frac{1}{2}\left(\sigma''\sigma\right)'\sigma\right) \circ g^{-1}(x).
\end{aligned}
$$

Again, Hypothesis (**H6**) ensures that the term $\left(\beta \circ g^{-1}\right)''(x)$ is non-negative and thus that $D^2_{\theta,t}U_s$ is too. We have finally obtained that $D^2_{\theta,t}U_s \geq 0$ a.s.

We will now focus on $D^2_{\theta,t}Y_s$ and prove that for $0 < \theta < t < s \leq T$, $\mathbb{P}$–a.s it is also non-negative. Applying once more the Malliavin derivative operator to $D_\theta Y_s$ in (3.8) and using the second point in Lemma 4.4, since f does not depend on Z we obtain, for $0 \leq \theta < t \leq s \leq T$,

$$
\begin{aligned}
D^2_{\theta,t}Y_s =\; &D^2_{\theta,t}\xi - \int_s^T D^2_{\theta,t}Z_r dW_r \\
&+ \int_s^T \Big\{ f_{xx}\left(X_r,Y_r\right)D_\theta X_r D_t X_r + f_x\left(X_r,Y_r\right)D^2_{\theta,t}X_r \\
&\quad + f_{yx}\left(X_r,Y_r\right)\left(D_\theta Y_r D_t X_r + D_\theta X_r D_t Y_r\right) \\
&\quad + f_{yy}\left(X_r,Y_r\right)D_\theta Y_r D_t Y_r + f_y\left(X_r,Y_r\right)D^2_{\theta,t}Y_r \Big\} dr.
\end{aligned}
$$

Since $D^2_{\theta,t}Y_s$ solves a linear equation and is $\mathcal{F}_s$-measurable, we have that, for $0 \leq \theta < t \leq s \leq T$,

$$
\begin{aligned}
D^2_{\theta,t}Y_s =\; &\mathbf{E}\left(e^{\int_s^T f_y(X_r,Y_r)dr}D^2_{\theta,t}\xi\middle|\mathcal{F}_s\right) \\
&+ \mathbf{E}\bigg(\int_s^T e^{\int_s^r f_y(X_u,Y_u)du}\Big\{ f_{xx}\left(X_r,Y_r\right)D_\theta X_r D_t X_r + f_x\left(X_r,Y_r\right)D^2_{\theta,t}X_r \\
&\quad + f_{yx}\left(X_r,Y_r\right)\left(D_\theta Y_r D_t X_r + D_\theta X_r D_t Y_r\right) + f_{yy}\left(X_r,Y_r\right)D_\theta Y_r D_t Y_r \Big\} dr\,\middle|\mathcal{F}_s\bigg). \quad (4.6)
\end{aligned}
$$

Using hypotheses (**H4**) and (**H5**) along with the fact that for $0 < \theta < t < s \leq T$, $\mathbb{P}$–a.s $D^2_{\theta,t}X_s$ is non-negative, we obtain that for $0 < \theta < t < s \leq T$, $\mathbb{P}$–a.s $D^2_{\theta,t}Y_s$ is non-negative.

Step 2: Positivity of DZ_t

Using a representation of Z, we compute its Malliavin derivative and show that under the hypotheses of Subsection 4.1.1, it is almost surely positive. We begin by giving a representation of

Z. We do not use the one from [PP92] in terms of gradient, that is $Z_t = \sigma\left(X_t\right)\left(\nabla X_t\right)^{-1}\nabla Y_t$, but rather use the fact that Z_t can be represented by use of the Clark-Ocone formula. Indeed, by the uniqueness of the solution (Y, Z), Z_t can be written as

$$Z_t = \mathbf{E}\left(D_t\xi + D_t\int_0^T f\left(X_s, Y_s\right)ds \Big| \mathcal{F}_t\right) \in \mathbb{D}^{1,2}. \tag{4.7}$$

Using this fact, we get for $t \in [0, T]$

$$Z_t = \mathbf{E}\left(D_t\xi + \int_t^T \left\{f_x\left(X_s, Y_s\right)D_tX_s + f_y\left(X_s, Y_s\right)D_tY_s\right\}ds | \mathcal{F}_t\right).$$

Let $\theta \le t$. We use both points of Lemma 4.4 in order to calculate the first order Malliavin derivative of Z_t. This leads, for $\theta \le t$:

$$D_\theta Z_t = \mathbf{E}\Big(D_{\theta,t}^2\xi + \int_t^T \Big\{ f_{xx}\left(X_s, Y_s\right)D_\theta X_s D_t X_s + f_{yx}\left(X_s, Y_s\right)\left(D_\theta Y_s D_t X_s + D_\theta X_s D_t Y_s\right)$$

$$+ f_{yy}\left(X_s, Y_s\right)D_\theta Y_s D_t Y_s + f_x\left(X_s, Y_s\right)D_{\theta,t}^2 X_s + f_y\left(X_s, Y_s\right)D_{\theta,t}^2 Y_s \Big\}ds\Big|\mathcal{F}_t\Big). \tag{4.8}$$

Using Hypotheses (**H4**) and (**H5**) along with the results obtained in Step 1, we obtain that for $0 < \theta < t \le T$, $\mathbb{P} - a.s$, $D_\theta Z_t > 0$.

Step 3: Conclusion of the proof by the Bouleau–Hirsch Theorem

For all $t \le T$, we have

$$\|DZ_t\|_{L^2([0,T])}^2 = \int_0^T (D_\theta Z_t)^2 d\theta.$$

Using the fact that for $0 < \theta < t \le T$, $\mathbb{P} - a.s$, $D_\theta Z_t > 0$ proved in Step 2, we deduce that $\|DZ_t\|_{L^2([0,T])} > 0$ a.s. Applying the Bouleau–Hirsch Theorem (see Theorem 2.1) concludes the proof. □

Remark 4.5. Theorem 4.3 has been proven under a set of hypotheses (those of Subsection 4.1.1) based on the fact that σ is positive. The case where σ is negative was included neither in the proof nor in the hypotheses for the sake of clarity and readability of the paper. However, as already mentioned in Remark 4.2, this case can be addressed (without any further difficulties) by using the following transformations: $\sigma \to \tilde{\sigma} := -\sigma$ and $W \to \tilde{W} := -W$. After performing those tranformations, it suffices to consider $(\tilde{X}, \tilde{Y}, \tilde{Z}) = (X, Y, -Z)$ to be the solution of

$$\begin{cases} \tilde{X}_t = x_0 + \int_0^t b(\tilde{X}_s)ds + \int_0^t \tilde{\sigma}(\tilde{X}_s)d\tilde{W}_s \\ \tilde{Y}_t = \xi + \int_t^T f\left(\tilde{X}_r, \tilde{Y}_r\right)dr - \int_t^T \tilde{Z}_r d\tilde{W}_r \end{cases}$$

This brings the problem back to the set of hypotheses of Subsection 4.1.1 and it can be dealt with using the techniques presented in the last section.

4.2 Gaussian bounds for the density of Z_t

In this section, we study a particular case of equations (4.3) and (4.4) and show that under proper conditions, the density of Z_t can be bounded from above and below by Gaussian estimates. The backward equation we study is the following:

$$Y_t = \phi(W_T) + \int_t^T f(Y_s)ds - \int_t^T Z_s dW_s, \tag{4.9}$$

4.2.1 Hypotheses

We consider f to be an appropriately smooth function to ensure the existence and uniqueness of solutions to equation (4.9). We also impose additional conditions needed to prove Theorem 4.6.

$$\begin{cases} \textbf{H7 :} \ \phi \in \mathcal{C}_b^2(\mathbb{R}) \text{ and } \phi'' \geq c > 0. \\ \textbf{H8 :} \ f \in \mathcal{C}_b^2(\mathbb{R}) \text{ and } f', f'' \geq 0. \end{cases}$$

The following theorem states that under the above hypotheses, Z_t has a density that can be bounded from above and below by Gaussian estimates.

Theorem 4.6. *Under the above hypotheses, for $t \in (0, T)$ the random variable Z_t defined in (4.9) has a density ρ_{Z_t}. Furthermore, there exist strictly positive constants c and C such that, for almost all $y \in \mathbb{R}$, ρ_{Z_t} satisfies the following:*

$$\frac{\mathbf{E}|Z_t - \mathbf{E}(Z_t)|}{2ct}\exp\left(-\frac{(z - \mathbf{E}(Z_t))^2}{2Ct}\right) \leq \rho_{Z_t}(z) \leq \frac{\mathbf{E}|Z_t - \mathbf{E}(Z_t)|}{2Ct}\exp\left(-\frac{(z - \mathbf{E}(Z_t))^2}{2ct}\right).$$

Proof: We will proceed in two steps, the first one being dedicated to proving that the Malliavin derivative of Z_t is bounded and bigger than a positive constant. The second step will be to use the Nourdin–Viens formula to conclude the proof.

Step 1: Boundedness and positivity of DZ_t

Y_t being defined as in equation (4.9), equation (3.9) becomes

$$D_\theta Y_t = \mathbf{E}\left(e^{\int_t^T f'(Y_s)ds}\phi'(W_T)\Big|\mathcal{F}_t\right),$$

and equation (4.6) becomes

$$D_{\theta,t}^2 Y_s = \mathbf{E}\left(e^{\int_s^T f'(Y_r)dr}\phi''(W_T)|\mathcal{F}_s\right) + \mathbf{E}\left(\int_s^T e^{\int_s^r f'(Y_u)du}f''(Y_r)D_\theta Y_r D_t Y_r dr|\mathcal{F}_s\right).$$

Using hypotheses (**H7**) and (**H8**), we obtain that $0 \leq |D_\theta Y_t| \leq C$ and $0 < c \leq D_{\theta,t}^2 Y_s \leq C$. We finally compute $D_\theta Z_t$ from equation (4.8) and we get

$$D_\theta Z_t = \mathbf{E}\left(\phi''(W_T) + \int_t^T \left\{f''(Y_s)D_\theta Y_s D_t Y_s + f'(Y_s)D_{\theta,t}^2 Y_s\right\}ds\Big|\mathcal{F}_t\right).$$

Using Hypotheses (**H7**) and (**H8**) again, we finally get

$$0 < c \leq D_\theta Z_t \leq C. \tag{4.10}$$

Step 2: Conclusion of the proof by the Nourdin–Viens formula

Write $D_\bullet Z_t = \Phi^\bullet_{Z_t}(W)$ with a measurable function $\Phi^\bullet_{Z_t} : \mathbb{R}^{L^2([0,T])} \longrightarrow L^2([0,T])$. Then (4.10) yields, for $\theta < t$, $0 < c \leq \Phi^\theta_{Z_t}(W) \leq C$. As previously done, define $\widetilde{\Phi^{\bullet,u}_{Z_t}}(W) = \Phi^\bullet_{Z_t}(e^{-u}W + \sqrt{1 - e^{-2u}}W')$ for $u \in [0, +\infty[$. Using (4.10), it is clear that, for $\theta < t$, we have for $u \in [0, +\infty)$, $0 < c \leq \widetilde{\Phi^{\theta,u}_{Z_t}}(W) \leq C$. Combining the bounds on $\Phi^\theta_{Z_t}$ and $\widetilde{\Phi^{\theta,u}_{Z_t}}$ yields, for $\theta < t$ and $u \in [0, +\infty)$,

$$0 < c^2 \leq \Phi^\theta_{Z_t}(W)\widetilde{\Phi^{\theta,u}_{Z_t}}(W) \leq C^2. \tag{4.11}$$

Finally, let

$$g(z) = \int_0^\infty e^{-u}\mathbf{E}\left(\mathbf{E}'\big(\langle\Phi^\bullet_{Z_t}(W), \widetilde{\Phi^{\bullet,u}_{Z_t}}(W)\rangle_{L^2([0,T])}\big)\Big|Z_t - \mathbf{E}(Z_t) = z\right)du$$

$$= \int_0^\infty e^{-u}\mathbf{E}\left(\mathbf{E}'\left(\int_0^t \Phi^\theta_{Z_t}(W)\widetilde{\Phi^{\theta,u}_{Z_t}}(W)d\theta\right)\Big|Z_t - \mathbf{E}(Z_t) = z\right)du.$$

The bounds obtained in (4.11) immediatly yield $0 < ct \leq g(z) \leq Ct$. Thus, Proposition 2.2 concludes the proof of Theorem 4.6. □

Remark 4.7. It is also possible to derive Gaussian density estimates for more complex equations than the one dealt with in this section. Hypotheses have to be changed in each case, making it difficult to state a general result with reasonable hypotheses covering most cases.

References

[AK05] F. Antonelli and A. Kohatsu-Higa (2005): Densities of One-Dimensional Backward SDEs. *Potential Analysis,* **22(3)**, 263–287.

[KS91] I. Karatzas and S. Shreve (1991): *Brownian motion and stochastic calculus. Second Edition.* Springer-Verlag, Berlin.

[Lam64] J. Lamperti (1964): A simple construction of certain diffusion processes. *J. Faculty Science Univ. Tokyo,* **32**, 1–76.

[NV09] I. Nourdin and F.G. Viens (2009): Density formula and concentration inequalities with Malliavin calculus. *Electronic Journal of Probability,* **14**, 2287–2309.

[Nua06] D. Nualart (2006): *The Malliavin calculus and related topics. Second Edition.* Springer-Verlag, Berlin.

[NQ09a] D. Nualart and L. Quer-Sardanyons (2009): Gaussian density estimates for solutions to quasi-linear stochastic partial differential equations. *Stochastic Process. Appl.,* **119** , no. 11, 3914–3938.

[NQ09b] D. Nualart and L. Quer-Sardanyons (2009): Optimal Gaussian density estimates for a class of stochastic equations with additive noise. *Infinite Dimensional Analysis, Quantum Probability and Related Topics,* to appear.

[PP90] E. Pardoux and S. Peng (1990): Adapted solution of a backward stochastic differential equation. *Systems Control Lett.,* **14**, no. 1, 55–61.

[PP92] E. Pardoux and S. Peng (1992): Backward stochastic differential equation and quasilinear parabolic partial differential equations. *Stochastic partial equations and their applications. Lect. Notes control Inf. Sci.,* **176**, 200–217.