Probability Theory and Related Fields manuscript No.
(will be inserted by the editor)

Patrick Cattiaux[1 2] · Laurent Mesnager[2]

Hypoelliptic non homogeneous diffusions

the date of receipt and acceptance should be inserted later

Abstract. Let $L_t = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(t,x)\partial_i\partial_j + \sum_{i=1}^{d} b_i(t,x)\partial_i = \frac{1}{2}\sum_{j=1}^{m} X_j^2(t,x) + X_0(t,x)$ be a time dependent second order operator, written in usual or Hörmander form. We study the regularity of the law of the associated non homogeneous (time dependent) diffusion process, under Hörmander's like conditions. Coefficients are only Hölder continuous in time. The main tool is Malliavin calculus. Our results extend and correct previous ones ([17] and related works, [15]). Related topics like filtering theory, killed or reflected processes, parabolic hypoellipticity are also discussed.

1. Introduction and Summary

For $j = 0, \ldots, m$, consider measurable flows $t \mapsto X_j(t,\bullet)$ of $\mathcal{C}^\infty$ vector fields on $\mathbb{R}^d$, such that for any multi-index α, $\partial_x^\alpha X_j(t,\bullet)$ are bounded on any time-space compact subset $[s,T]\times K$.
One can thus find a pathwise unique solution of the stochastic differential equation

$$\text{(1.1.a)}\qquad \begin{cases} dx_t = \sum_{j=1}^{m} X_j\left(t,x_t\right) dw_t^j \;+\; X_0\left(t,x_t\right) dt, \quad t>s \\ x_s = x \end{cases}$$

written in Stratonovich form, where $(w^1,\ldots,w^m)$ is a standard Brownian motion, or equivalently of the Ito stochastic differential equation

$$\text{(1.1.b)}\qquad \begin{cases} dx_t = \sum_{j=1}^{m} X_j\left(t,x_t\right) \delta w_t^j \;+\; \bar{X}_0\left(t,x_t\right) dt, \quad t>s \\ x_s = x \end{cases}$$

where $\bar{X}_0 \;=\; X_0 \;+\; \frac{1}{2}\sum_{j=1}^{m} \frac{\partial X_j}{\partial x}\, X_j$.
We will assume that the solution is non explosive, for instance assuming when necessary that $\partial_x^\alpha X_j(t,x)$ are bounded on $[s,T]\times\mathbb{R}^d$. If we denote by $x_t^s(x)$ the solution of (1.1) , then

$$\begin{aligned} T_{st} : \mathscr{C}_b &\longmapsto \mathscr{C}_b \\ f &\longmapsto \mathbb{E}\left(f(x_t^s(x))\right) \end{aligned}$$

Ecole Polytechnique, CMAP, F-91128 Palaiseau Cedex
Modal'X, Université de Paris X, UFR SEGMI, 200 avenue de la République, F-92001 Nanterre Cedex
e-mail: `cattiaux@cmapx.polytechnique.fr`
e-mail: `mesnager@modalx.u-paris10.fr`

is a non-homogeneous (strong) Markov semigroup with (local) generator

$$L_t \;=\; \frac{1}{2}\sum_{j=1}^{m} X_j^2(t,\bullet) \;+\; X_0(t,\bullet)$$

written in Hörmander form. If we denote by $\mu_{st}(x,dy)$ the law of $x_t^s(x)$, then $\mu_{st}(x,dy)$ is a solution of the weak forward equation $\left(\frac{\partial}{\partial t}+L_t\right)^{\star}\mu_{st} \;=\; 0$, $\mu_{ss} \;=\; \delta_x$. We are first interested in the existence of (smooth) densities for the laws, i.e.

$$\mu_{st}(x,dy) \;=\; p_{st}(x,y)\,dy \tag{1.2}$$

where p_{st} is a (smooth) function on $\mathbb{R}^d$ (or a smooth kernel on $\mathbb{R}^d\otimes\mathbb{R}^d$). Since the homogeneous case (when L_t does not depend on t) is now very well known, we shall really focus on the non-homogeneous case. However, we will extensively use the time-space process (which is homogeneous), namely the solution of

$$\begin{cases} dy_t = \sum_{j=1}^{m} X_j\left(u_t,y_t\right)dw_t^j \;+\; X_0\left(u_t,y_t\right), \\ du_t = dt, \\ (u_0,y_0) = (u,x) \end{cases} \tag{1.3}$$

which is given by a flow $\psi_t(w,u,x)$ such that $x\mapsto\phi_t(w,u,x) = \Pi_x\psi_t(w,u,x)$ is , for each u, a $\mathcal{C}^\infty$ diffeomorphism of $\mathbb{R}^d$. Here of course Π_x denotes the projection operator from $\mathbb{R}\otimes\mathbb{R}^d$ onto $\mathbb{R}^d$. Note that the standard proof, lying on Kolmogorov continuity theorem, allows to choose such a version for each u, but non necessarily for all u simultaneously. In particular if one wants to use the pathwise composition of flows formula, one has to make an additional regularity assumption on the X_j. For instance one may assume that $\partial_x^\alpha X_j(\bullet,\bullet)$ are β-Hölder continuous, for some $\beta>0$. In this case one can build continuous versions of $\partial_x^\alpha\phi_t(w,\bullet,\bullet)$.
Of course the key point is that

$$x_t^s(x) = \phi_{t-s}(w,s,x) \quad \text{in Law} \quad t>s. \tag{1.4}$$

Without loss of generality, we may and will assume that $s=0$, and write $\phi_t(w,x)$ instead of $\phi_t(w,0,x)$. This will not introduce any confusion and will simplify the notations. However some statements in sections 4 and 5 and just below, are written with a starting time s.

In order to prove (1.2) one can call upon celebrated Hörmander's sum of squares theorem ([18]) i.e.

(1.5) If the X_j's are $\mathcal{C}_b^\infty$ on $[s,T]\times\mathbb{R}^d$ and the Lie algebra generated by X_j, $1\leq j\leq m$ and $\frac{\partial}{\partial t}+X_0$ is full at each $(t,y)\in]s,T]\times\mathbb{R}^d$, then (1.2) holds for all $t\in]s,T]$, all $x\in\mathbb{R}^d$ and $y\mapsto p_{st}(x,y)$ is smooth (i.e. $\mathcal{C}^\infty$).

Actually Hörmander's result is stronger since it deals with parabolic-hypoellipticity. Note that the statement of (1.5) includes the time component i.e. for each fixed (t,y),

$$\dim\text{ span Lie}\left(\frac{\partial}{\partial t}+X_0, X_j, 1\leq j\leq m,\right)(t,y) = d+1. \tag{1.6}$$

Extending Kohn's proof of Hörmander's result, Chaleyat-Maurel and Michel [14] have shown a weak form of parabolic-hypoellipticity, for coefficients which are only measurable in time. This yields

(1.7) If all $\partial_x^\alpha X_j, 0 \le j \le m$ are measurable and bounded on $[s,T] \times \mathbb{R}^d$ and

$$\text{dim span Lie } (X_j, 1 \le j \le m)(t,y) = d$$

for all $(t,y) \in [s,T] \times \mathbb{R}^d$, then the same conclusion as in (1.5) holds.

Note that the time component (as vector fields) disappears, but that Lie brackets with X_0 are no more allowed (this is sometimes called the restricted Hörmander's hypothesis). The role of $\dfrac{\partial}{\partial t}$ will be enlightened by a very simple example in a moment.
Notice that when the vector fields are real analytic, the Hörmander's condition is almost necessary (see [16]).

Malliavin stochastic calculus of variations initiated in the 80th's ([26][25]) , proposed a new approach based on a variational method for s.d.e. (1.3) which leads to the following principle

(1.8) Let $C_t(x) = \displaystyle\sum_{j=1}^m \int_0^t \phi_s^{\star -1} X_j(x) >< \phi_s^{\star -1} X_j(x)\, ds$ be (part of) the Malliavin covariance matrix. If $C_t^{-1}(x)$ belongs to all the L^p's ($p \in [1, +\infty[$), then $\mu_{0t}(x,dy) = p_t(x,y)dy$ with

$$y \mapsto p_t(x,y) \in \mathcal{C}^\infty\left(\mathbb{R}^d\right).$$

In (1.8), we are using Bismut notations

$$\begin{cases} \phi_s^{\star -1} X_j(x) = \left(\dfrac{\partial \phi_s}{\partial x}(w,x)\right)^{-1} X_j\left(s, \phi_s(w,x)\right) \\ Y >< Z \quad = Y^t Z, \end{cases}$$

so that C_t is a nonnegative symmetric matrix. In order to study C_t the standard way is to use Ito-Kunita-Bismut formula

(1.9) for all $\xi \in \mathbb{R}^d$, $j = 1, \ldots, m$

$$\begin{aligned} \langle \xi, \phi_t^{\star -1} X_j \rangle(x) = \langle \xi, X_j(0,x)\rangle + \sum_{k=1}^m \int_0^t \langle \xi, \phi_s^{\star -1}\left[X_k, X_j\right]\rangle(x) \delta w_s^k \\ + \int_0^t \left\langle \xi, \phi_s^{\star -1}\left(\left[\frac{\partial}{\partial t} + X_0, X_j\right] + \frac{1}{2}\sum_{k=1}^m [X_k, [X_k, X_j]]\right)\right\rangle(x) ds \end{aligned}$$

where $[Y,Z]$ is the Lie bracket between vector fields Y and Z. It follows from a now well known procedure (see e.g. [32] or [5]) that

(1.10) When the X_j's, $j = 0, \ldots, m$ are $\mathcal{C}_b^\infty$ on $[0,T] \times \mathbb{R}^d$, a sufficient condition for $C_t^{-1}(x)$ to be in all the L^p's for all $t \in]0,T]$ is that

$$\text{dim span Lie}\left(X_j, 1 \le j \le m, \frac{\partial}{\partial t} + X_0\right)(0,x) = d+1$$

(for another approach of this result also see Bally [4])

The main difference with (1.5) is that the assumption on the Lie brackets is only made at the starting point $(0,x)$. This could be considered as a slight improvement, but Malliavin calculus has now proved its strength in many areas where analytic methods are very difficult or impossible to use (see e.g. the recent book of Malliavin himself [27]). Let us also mention that one can recover exactly Hörmander's result using the Markov property. This is partly explained by the following elementary remark, which has already been made by numerous authors

Remark 1.11 *Assume that for $s_0 > s$, $\mu_{ss_0}(x,dy) = p_{ss_0}(x,y)dy$. Then for $t > s_0$,*

$$\begin{aligned}\mathbb{E}\left(f(x_t^s(x))\right) &= \mathbb{E}\left(f(\phi_{t-s}(w,s,x))\right)\\ &= \mathbb{E}\left[\mathbb{E}\left(f(\phi_{t-s_0}(w',s_0,\phi_{s_0-s}(w,s,x)))\right)\right] \ \textit{(where (w,w') is the generic pair on $\Omega\otimes\Omega$)}\\ &= \int p_{ss_0}(x,y)\,\mathbb{E}\left(f(\phi_{t-s_0}(w',s_0,y))\right)\,dy\\ &= \int f(z)\,\mathbb{E}\left(p_{ss_0}(x,\phi^{-1}_{t-s_0}(w',s_0,z))\ \left|\text{Jacobian }\phi^{-1}_{t-s_0}(w',s_0,z)\right|\right)\,dz\end{aligned}$$

provided all manipulations are allowed. In particular, if p_{ss_0} is smooth with a good behaviour at infinity, p_{st} is also smooth thanks to the regularity of the flow ϕ^{-1} and $\mu_{st}(x,dy) = p_{st}(x,y)dy$ with

$$p_{st}(x,y) = \mathbb{E}\left(p_{ss_0}\left(x,\phi^{-1}_{t-s_0}(w,s_0,y)\right)\left|\text{Jacobian }\phi^{-1}_{t-s_0}(w,s_0,y)\right|\right).$$

This allows to improve (1.7) assuming the restricted Hörmander's hypothesis only near the time origin (but everywhere in space).
At this point we shall make an additional important remark about the role of the vector field $\frac{\partial}{\partial t}$. In the homogeneous case an equivalent formulation for (1.6) is

$$\text{dim span}\left(X_j, 1\leq j\leq m, \text{Lie brackets of order} \geq 2 \text{ of the } X_j'\,s, 0\leq j\leq m\right)(t,y) = d.$$

In the non homogeneous case, cancellations may occur because of $\frac{\partial}{\partial t}$. Here is an example

(1.12) <u>example</u> [36]. Take $d=2$, $m=1$ and consider the process

$$x_t(x) = \left(x_1 + w_t, x_2 + t\,(x_1+w_t)^2\right)$$

which law, supported by the curve $y_2 = x_2 + ty_1^2$, is singular. The associated vector fields are

$$X_0(t,x) = x_1^2\frac{\partial}{\partial x_2}, \quad X_1(t,x) = \frac{\partial}{\partial x_1} + 2tx_1\frac{\partial}{\partial x_2}$$

So X_1 and $[X_1,[X_1,X_0]] = 2\dfrac{\partial}{\partial x_2}$ span $\mathbb{R}^2$ at each (t,x).

Of course (1.6) is not satisfied because $\left[\dfrac{\partial}{\partial t} + X_0, X_1\right] = 0$ at each (t,x).

This fact is not related to the unboundedness of the coefficients. One can easily modify this example. For instance taking

$$(y_2 - x_2 = f(t,y_1) = \exp\left(-\tfrac{1}{2}\left(y_1^2 + (1+t)^2\right)\right)$$

and looking at the starting point $y=0$, the situation is unchanged. In particular, the statement of Theorem 1.1.3 in [17] is wrong.
Let us now turn to the case when the vector fields are no more regular in the time component. From Malliavin calculus point of view, the only difficulty is that (1.9) (which involves $\left[\frac{\partial}{\partial t}, X_j\right]$) is no more valid or cannot be iterated. Nevertheless three cases are almost immediate

(1.13.1) vector fields are measurable in time, and L_t is uniformly elliptic in a neighborhood of $(0,x)$,i.e. $\sum_{j=1}^{m}\langle\xi, X_j(t,x)\rangle^2 \geq c\,|\xi|^2$ for some $c>0$ and all $(t,y)\in[0,s_0]\times B(x,R)$;

(1.13.2) $\partial_x^\alpha X_j$ belong to $\mathcal{C}_b^1(\mathbb{R}\times\mathbb{R}^d)$ for all α and dim span Lie $(X_j, 1\leq j\leq m)(0,x)=d$;

(1.13.3) $X_j(t,x)=f_j(t)W_j(x)$ where $0<c\leq|f_j|\leq\dfrac{1}{c}$ and usual Hörmander's hypothesis holds for the W_j's.

In these three cases (1.2) holds (with $s=0$). (1.13.1) is really immediate while (1.13.2) and (1.13.3) are mainly contained in Taniguchi [36]. As we already said the tentative of Florchinger [17] to relax ellipticity or regularity is incorrect. It seems that the derivations of pages 211-212 are incorrect. Some related works by some authors which use the same discretization method also contain wrong results.

Another approach was proposed by Chen and Zhou [15]. Since one cannot directly use the stochastic calculus for $X_j(s,\phi_s(w,x))$ once $s\mapsto X_j(s,y)$ is not $\mathcal{C}^1$, one can freeze the time, writing

(1.14)

$$\begin{aligned}\langle\xi,\phi_t^{\star-1}X_j\rangle(x) &= \langle\xi, X_j(t,x)\rangle + \sum_{k=1}^{m}\int_0^t\langle\xi,\phi_s^{\star-1}\,[X_k(s,\bullet),X_j(t,\bullet)]\rangle(x)\,\delta w_s^k \\ &\quad + \int_0^t\left\langle\xi,\phi_s^{\star-1}\left([X_0(s,\bullet),X_j(t,\bullet)]+\frac{1}{2}\sum_{k=1}^{m}[X_k(s,\bullet),[X_k(s,\bullet),X_j(t,\bullet)]]\right)\right\rangle(x)ds.\end{aligned}$$

The authors use Kusuoka-Stroock method [21] for obtaining some expansion of C_t in terms of repeated integrals. They then state (see [15] Thm 1.1)

(1.15) if the $\partial_x^\alpha X_j'$ s are uniformly β-Hölder continuous in time, for some $\beta>\frac{1}{2}$, and dim span Lie $(X_j,\,1\leq j\leq m)=d$, then (1.2) holds.

Condition $\beta>\frac{1}{2}$ is used in the derivation of lemma 4.3 of [15]. But it is too restrictive in many applications (for example in filtering) where dependence in time is given by some extra stochastic process, like another diffusion process, for which we have to consider $\beta<\frac{1}{2}$.

Actually the proof of (1.15) given in [15] is still somewhat obscure for us because it lies on controls on repeated integrals without checking possible cancellations. Let us explain this point on example (1.12).

(1.16) <u>example (continuation)</u> Of course vector fields being smooth in time, one can on one hand apply (1.9). For $\xi=(0,1)$ one gets

$$\langle\xi,\phi_t^{\star-1}X_1\rangle(x)=\langle\xi,X_1(0,x)\rangle=0$$

On the other hand if we use (1.14) we have to calculate

$$\begin{aligned}[X_1(s,\bullet),X_1(t,\bullet)] &= 2(t-s)\,\frac{\partial}{\partial x_2}\\ [X_0(s,\bullet),X_1(t,\bullet)] &= -2x_1\,\frac{\partial}{\partial x_2}\\ [X_1(s,\bullet),[X_1(s,\bullet),X_1(t,\bullet)]] &= 0\\ \left(\frac{\partial\phi_s}{\partial x}(w,x)\right)^{-1} &= \begin{pmatrix}1 & 0\\ -2s(w_s+x_1) & 1\end{pmatrix}\end{aligned}$$

Thus

$$\langle \xi, \phi_t^{\star -1} X_1(t,\bullet)\rangle(x) = \langle \xi, X_1(t,x)\rangle + \int_0^t 2\,\xi_2(t-s)\delta w_s - \int_0^t 2\,\xi_2(x_1+w_s)ds$$
$$= 2tx_1 + 2\int_0^t (t-s)\delta w_s - 2\int_0^t (x_1+w_s)ds.$$

Of course this is equal to 0, as the terminal value at time t of $\langle \xi, \phi_\theta^{\star -1} X_1\rangle(x) = 2(t-\theta)(x_1+w_\theta)$. But this is due to cancellations at time t of terms which do not play the same role: initial condition, bounded variation term, stochastic integral. In other words repeated integrals with deterministic (but non constant) time dependent integrands may cancel at time t.

Remark and Apologies: Let us add to the previous remark that our first tentative, based on the same idea (freezing the time) in connection with Norris lemma, contains a similar mistake. All our apologies for those who had in hands the first version of the paper. We warmly thank an anonymous referee for pointing out this mistake.

In order to bypass this difficulty, one can use a time discretization procedure and use time regularity of the process. This is what the authors do in [15], but in a quite intricate way.
In section 2 we shall give an alternate proof of (1.15). It is essentially based on the same ideas, but in the spirit of Norris lemma instead of Kusuoka-Stroock. We hope it will help to clarify Chen and Zhou paper. Condition $\beta > \frac{1}{2}$ will clearly appear as a limitation due to the use of time regularity of the process. To overcome this limitation other ideas are necessary. These ideas are introduced in section 3. In addition to all that was used before (time discretization, time freezing, stochastic calculus), the new points are: first replace L^2 norms by L^∞ norms (thanks to the regularity of the coefficients), next use the continuity of the *argsup* of some one dimensional semi-martingales. The strategy is quite intricate, so a brief account of it is given at the beginning of section 3. In section 4 we collect some useful estimates on the density. We underline that these estimates are certainly not sharp. But they are sufficient to derive a parabolic-hypoelliptic theorem which is an improvement of Chaleyat-Maurel and Michel. In section 5 various applications are quickly discussed: filtering theory, killed and reflected processes, infinite degeneracy. In all these cases we shall only mention new difficulties arising from non homogeneity. Precise statements are mainly left to the reader.

Once again we warmly thank an anonymous referee for a careful reading of a first version of this work, and a second referee for the same careful and courageous reading of this version. We also wish to thank D. Nualart who mentioned to us the paper of Chen and Zhou [15].

2. Estimates on Malliavin covariance matrix : first approach

For a $\mathbb{R}^l$ valued function $t \mapsto h(t)$ and $\beta > 0$, we define the β-Hölder norm of h as

$$\|h\|_\beta = \sup_{0\le s,t\le T} (t-s)^{-\beta}\,|h(t)-h(s)| \tag{2.1}$$

Throughout this section we will assume that : for all $j = 0, ..., m$, all multi-index α, there exists $\beta(\alpha) > 0$ such that

$$\sup_{y\in\mathbb{R}^d} \left\|\partial_y^\alpha X_j(\bullet,y)\right\|_{\beta(\alpha)} < +\infty \tag{2.2}$$

In particular $\sup_{0\le t\le T} \left\|\partial_y^\alpha X_j(t,\bullet)\right\|_\infty < \infty$, i.e. the X_j's belong to $\mathcal{C}_b^\infty$ uniformly in time on $[0,T]$. In addition we assume that the restricted Hörmander's hypothesis holds i.e.

$$\dim \operatorname{span} \operatorname{Lie}(X_j, 1 \le j \le m)(0,x) = d \tag{2.3}$$

or equivalently, there exist $N(x) = N \in \mathbb{N}$, $c_N(x) = c_N > 0$ such that for all $\xi \in S^{d-1}$

$$\sum_{k=0}^{N} \sum_{Z\in\Sigma_k} \langle \xi, Z\rangle^2 (0,x) \ge c_N \tag{2.3$'$}$$

where $\Sigma_0 = \{X_j,\, 1 \le j \le m\}$, $\Sigma_{k+1} = \{[X_j, Z]\,,\, 1 \le j \le m,\, Z \in \Sigma_k\}$. N being given by (2.3′) we can define some constants

$$K = \sup_{0\le s\le T} \max_{0\le|\alpha|\le N+2} \max_{0\le j\le m} \left\|\partial_y^\alpha X_j(s,\bullet)\right\|_\infty \tag{2.4.a}$$

(2.4.b) β a common Hölder coefficient for all α s.t. $|\alpha| \le N+2$

$$K_\beta = \sup_{y\in\mathbb{R}^d} \max_{0\le|\alpha|\le N+2} \max_{0\le j\le m} \left\|\partial_y^\alpha X_j(\bullet,y)\right\|_\beta \tag{2.4.c}$$

Using time-space continuity, one can then find $(t_0, R) \in \mathbb{R}^+\times\mathbb{R}^+$ s.t. for all $s \in [0,t_0]$, all $y \in B(x,R)$, all $\xi \in S^{d-1}$

$$\sum_{k=0}^{N} \sum_{Z\in\Sigma_k} \langle \xi, Z\rangle^2 (s,y) \ge \frac{c_N}{2} \tag{2.5}$$

We shall prove the following version of classical estimates (see [31])

Theorem 2.6 *Let (2.2), (2.3) and (2.4) hold, for some $\beta > \frac{1}{2}$. Then for all $p \in [1,+\infty[$, there exists c_p which only depends on p, t, x, N, c_N, K, β and K_β such that for all $\varepsilon > 0$*

$$\sup_{\xi\in S^{d-1}} \mathbb{P}\left(\int_0^t \sum_{j=1}^m \left\langle \xi, \phi_s^{\star-1} X_j\right\rangle^2 (x)\, ds \le \varepsilon\right) \le c_p\, \varepsilon^p.$$

Note that for $\beta = 1$, this result essentially follows from the estimates of [31, p. 119], thanks to (1.9). This is what can be done for obtaining (1.13.2).

Proof: The proof will follow the standard scheme of [31], but requires additional work. For $k = 0$, $\ldots$, N, let $m(k)$ be a positive constant we will choose later. Then for a given $\xi \in S^{d-1}$ we introduce the sets

$$E_k = \left\{\sum_{Z\in\Sigma_k} \int_0^t \left\langle \xi, \phi_s^{\star-1} Z\right\rangle^2 (x)\, ds \le \varepsilon^{m(k)}\right\}.$$

Choosing $m(0) = 1$, what we need is a bound for $\mathbb{P}(E_0)$. But

$$\mathbb{P}(E_0) \leq \sum_{k=0}^{N-1} \mathbb{P}(E_k \cap E_{k+1}^c) + \mathbb{P}(E) \text{ with } E = \bigcap_{k=0}^{N} E_k.$$

Step 1 : **Upper bound for** $\mathbb{P}(E)$.

It can be obtained as in [31, p. 121] which yields $\mathbb{P}(E) \leq c_p \varepsilon^p$, where c_p depends only on t, x, p, K, N, c_N, R, t_0 (defined in (2.5)) and $\min_{k=0,\dots,N} m(k)$.

Step 2 : **Upper bound for** $\mathbb{P}(E_k \cap E_{k+1}^c)$.

We will more generally get a bound for

$$\mathbb{P}(B) = \mathbb{P}\left(\int_0^t \left\langle \xi, \phi_s^{\star -1} Z \right\rangle^2 (x) ds \leq \varepsilon^q, \sum_{j=1}^{m} \int_0^t \left\langle \xi, \phi_s^{\star -1} [X_j, Z] \right\rangle^2 (x) ds \geq \varepsilon \right) \tag{2.7}$$

for a vector field Z satisfying (2.2) with bounds (2.4). The ad-hoc value of q will determine the ratio $m(k)/m(k+1)$. As usual it is useful to bound the process introducing for $\theta > 0$

$$\tau = \inf \left\{ z \geq 0, \quad \left| \phi_s^{\star -1}(\omega, x) \right| \geq \varepsilon^{-\theta} \right\} \wedge t \tag{2.8}$$

which satisfies $\mathbb{P}(\tau < t) \leq c_p \varepsilon^p$ where c_p only depends on p, t, x, K and θ. So we only have to bound $\mathbb{P}(B, \tau = t)$. To overcome the difficulty explained in (1.16) we introduce a time discretization

$$s_i = \frac{it}{n}, \quad i = 0, \dots, n, \quad u_i = \frac{s_i + s_{i+1}}{2} \tag{2.9}$$

and define

$$B_i = \left\{ \int_{s_i}^{s_{i+1}} \left\langle \xi, \phi_s^{\star -1} Z \right\rangle^2 (x) ds \leq \varepsilon^q, \sum_{j=1}^{m} \int_{s_i}^{s_{i+1}} \left\langle \xi, \phi_s^{\star -1} [X_j, Z] \right\rangle^2 (x) ds \geq \frac{\varepsilon}{n} \right\} \tag{2.10}$$

so that

$$\mathbb{P}(B, \tau = t) \leq \sum_{i=0}^{n-1} \mathbb{P}(B_i, \tau = t).$$

Again

$$\mathbb{P}(B_i, \tau = t) \leq \mathbb{P}(B_i^1, \tau = t) + \mathbb{P}(B_i^2, \tau = t)$$

where

$$
\begin{aligned}
B_i^1 &= B_i \cap \left\{ \sum_{j=1}^m \int_{s_i}^{u_i} \left\langle \xi, \phi_s^{\star -1} \left[X_j, Z\right] \right\rangle^2 (x) ds \geq \frac{\varepsilon}{2n} \right\} \\
B_i^2 &= B_i \cap \left\{ \sum_{j=1}^m \int_{u_i}^{s_{i+1}} \left\langle \xi, \phi_s^{\star -1} \left[X_j, Z\right] \right\rangle^2 (x) ds \geq \frac{\varepsilon}{2n} \right\}
\end{aligned} \tag{2.11}
$$

We shall get an estimate for $\mathbb{P}(B_i^1, \tau = t)$. The same estimate holds for $\mathbb{P}(B_i^2, \tau = t)$ just changing s_i into s_{i+1} in the derivation below.
It should be surprising to the reader that we separate the quadratic variation part in two terms. The reason will become apparent when studying C_i.
To get this estimate we use the Ito formula as in (1.14), freezing the time for $Z(\bullet, y)$. Hence for $s \in [s_i, s_{i+1}]$, $\mathbb{P}$ a.s.

$$
\begin{aligned}
\left\langle \xi, \phi_s^{\star -1} Z \right\rangle^2 (x) = \left\langle \xi, \phi_{s_i}^{\star -1} Z(s, \bullet) \right\rangle^2 (x) + \sum_{j=1}^m \int_{s_i}^s \left\langle \xi, \phi_u^{\star -1} \left[X_j(u, \bullet), Z(s, \bullet)\right] \right\rangle^2 (x) du \\
+ A_s^i + \sum_{j=1}^m M_s^{ij}
\end{aligned} \tag{2.12}
$$

with

$$
\begin{aligned}
M_s^{ij} &= 2 \int_{s_i}^s \left\langle \xi, \phi_u^{\star -1} Z(s, \bullet) \right\rangle (x) \left\langle \xi, \phi_u^{\star -1} \left[X_j(u, \bullet), Z(s, \bullet)\right] \right\rangle (x) \delta w_u^j \\
A_s^i &= 2 \int_{s_i}^s \left\langle \xi, \phi_u^{\star -1} Z(s, \bullet) \right\rangle (x) \left\langle \xi, \phi_u^{\star -1} \left[X_0(u, \bullet), Z(s, \bullet)\right] \right\rangle (x) du \\
&\quad + \int_{s_i}^s \left\langle \xi, \phi_u^{\star -1} Z(s, \bullet) \right\rangle (x) \sum_{j=1}^m \left\langle \xi, \phi_u^{\star -1} \left[X_j(u, \bullet), \left[X_j(u, \bullet), Z(s, \bullet)\right]\right] \right\rangle (x) du
\end{aligned}
$$

Now using Kolmogorov continuity criterion, it is not hard to see that both hand sides in (2.12) have a continuous version, so that (2.12) holds for $\mathbb{P}$ a.s. for all $s \in [s_i, s_{i+1}]$. So we can integrate (2.12) with respect to ds (up to a fixed negligible set). Applying Fubini and a stochastic Fubini theorem (see e.g. [19, p. 116], replacing boundedness by integrability of any order), we get

$$
\int_{s_i}^{s_{i+1}} \left\langle \xi, \phi_s^{\star -1} Z \right\rangle^2 (x) \, ds \geq C_i + A_i + \sum_{j=1}^m M_{ij} \tag{2.13}
$$

with

$$A_i = \int_{s_i}^{s_{i+1}} A_s^i \, ds$$

$$C_i = \int_{s_i}^{s_{i+1}} \int_s^{s_{i+1}} \sum_{j=1}^m \left\langle \xi, \phi_s^{\star -1} \left[X_j(s,\bullet), Z(u,\bullet)\right]\right\rangle^2 (x) \, du \, ds$$

$$M_{ij} = \int_{s_i}^{s_{i+1}} \alpha_s^{ij} \, \delta w_s$$

$$\alpha_s^{ij} = \int_s^{s_{i+1}} \left\langle \xi, \phi_s^{\star -1} Z(u,\bullet)\right\rangle (x) \left\langle \xi, \phi_s^{\star -1} \left[X_j(s,\bullet), Z(u,\bullet)\right]\right\rangle (x) \, du$$

But on $\{\tau = t\}$, for u and s less than t, the following holds

$$\left|\left\langle \xi, \phi_s^{\star -1} Z\right\rangle (x) - \left\langle \xi, \phi_s^{\star -1} Z(u,\bullet)\right\rangle (x)\right| \le \varepsilon^{-\theta} K_\beta \left|s-u\right|^\beta \tag{2.14}$$

and

$$\left|\left\langle \xi, \phi_s^{\star -1} \left[X_j, Z\right]\right\rangle (x) - \left\langle \xi, \phi_s^{\star -1} \left[X_j(s,\bullet), Z(u,\bullet)\right]\right\rangle (x)\right| \le 2\varepsilon^{-\theta} K K_\beta \left|s-u\right|^\beta \tag{2.14'}$$

Using (2.14) and (2.14′), we shall successively get a lower bound of C_i, and upper bounds for $|A_i|$ and $|M_{i,j}|$, up to subsets of small Probability.

Lower bound for C_i

According to (2.14′) and $(a+b)^2 \ge a^2 - 2|ab|$, one has

$$\begin{aligned} C_i \ge & \int_{s_i}^{s_{i+1}} (s_{i+1} - s) \sum_{j=1}^m \left\langle \xi, \phi_s^{\star -1} \left[X_j, Z\right]\right\rangle^2 (x) \, ds \\ & -4\varepsilon^{-\theta} K K_\beta \int_{s_i}^{s_{i+1}} \frac{(s_{i+1} - s)^{1+\beta}}{1+\beta} \sum_{j=1}^m \left|\left\langle \xi, \phi_s^{\star -1} \left[X_j, Z\right]\right\rangle (x)\right| \, ds. \end{aligned}$$

In order to get a lower bound for the first term, we have to bound $s_{i+1} - s_i$ from below in the integral. This explains why we have to divide the time interval in two subintervals.
Hence, on $B_i^1 \cap \{\tau = t\}$ one has,

$$\begin{aligned} C_i & \ge (s_{i+1} - u_i) \frac{\varepsilon}{2n} - 8m\varepsilon^{-2\theta} K^3 K_\beta \frac{(s_{i+1} - s_i)^{2+\beta}}{(1+\beta)(2+\beta)} \\ & \ge \frac{\varepsilon \, t}{4n^2} - 4mK^3 K_\beta \varepsilon^{-2\theta} \, \frac{t^{2+\beta}}{n^{2+\beta}} \end{aligned}$$

We will choose $n = [\varepsilon^{-\mu}]$ (integer part of $\varepsilon^{-\mu}$). Hence, if $\beta\mu - 2\theta > 1$, one can find ε_0 which only depends on m, K, K_β, μ, θ and t such that for $\varepsilon < \varepsilon_0$, (2.15) $$C_i \geq \frac{1}{8}\,\varepsilon^{1+2\mu} \text{ on } B_i^1 \cap \{\tau = t\}\,.$$

($\beta\mu - 2\theta > 1$ ensures that $-2\theta + (2+\beta)\mu > 1 + 2\mu$, so that the principal part of C_i is $\dfrac{\varepsilon}{n^2} \sim \varepsilon^{1+2\mu}$)

Upper bound for $|A_i|$

On $\{\tau = t\}$,

$$\begin{aligned}\left|A_s^i\right| &\leq c\,\varepsilon^{-\theta}\int_{s_i}^{s}\left|\left\langle \xi, \phi_u^{\star -1} Z(s,\bullet)\right\rangle(x)\right|\,du \\ &\leq c\,\varepsilon^{-\theta}\int_{s_i}^{s}\left|\left\langle \xi, \phi_u^{\star -1} Z\right\rangle(x)\right|\,du + c\varepsilon^{-2\theta}K_\beta\,\frac{(s-s_i)^{1+\beta}}{1+\beta}\end{aligned}$$

where c only depends on K and m according to (2.14). It follows that on $\{\tau = t\}$

$$|A_i| \leq c\varepsilon^{-\theta}\int_{s_i}^{s_{i+1}}(s_{i+1}-s)|\left\langle \xi, \phi_s^{\star -1} Z\right\rangle(x)|ds + c\varepsilon^{-2\theta}K_\beta\,\frac{(s_{i+1}-s_i)^{2+\beta}}{(1+\beta)(2+\beta)}$$

and on $B_i^1 \cap \{\tau = t\}$, by Cauchy-Schwarz inequality,

$$|A_i| \leq c\varepsilon^{-\theta}\varepsilon^{\frac{q}{2}}\,\frac{t^{\frac{3}{2}}}{n^{\frac{3}{2}}} + c\varepsilon^{-2\theta}K_\beta\,\frac{t^{2+\beta}}{n^{2+\beta}}.$$

Thus if $\beta\mu - 2\theta > 1$ and $-\theta + \dfrac{q}{2} + \dfrac{3}{2}\mu > 1 + 2\mu$, on can find ε_0' which only depends on m, K, K_β, t, β, μ, θ and q such that on $B_i^1 \cap \{\tau = t\}$

(2.16) $$|A_i| \leq \frac{1}{32}\varepsilon^{1+2\mu} \text{ for } \varepsilon < \varepsilon_0'.$$

Upper bound for $|M_{i,j}|$

As usual, we shall use the classical martingale inequality

(2.17) $$\mathbb{P}(\int_a^b \alpha_s^2\,ds \leq \eta^{2k},\, \left|\int_a^b \alpha_s\,\delta w_s\right| \geq \eta^l) \leq 2\exp\left(-\frac{1}{2}\eta^{2(l-k)}\right).$$

As before on $\{\tau = t\}$, $\left|\alpha_s^{ij}\right| \leq c\varepsilon^{-\theta}\displaystyle\int_s^{s_{i+1}}\left|\left\langle \xi, \phi_s^{\star -1} Z(u,\bullet)\right\rangle(x)\right|du$ where c only depends on K. According to (2.14), on $\{\tau = t\}$,

$$\left|\alpha_s^{ij}\right|^2 \leq 2c^2\varepsilon^{-2\theta}(s_{i+1}-s)^2\left\langle \xi, \phi_s^{\star -1} Z\right\rangle^2(x) + 2c^2\varepsilon^{-4\theta}K_\beta^2(s_{i+1}-s)^{2+2\beta}.$$

Thus on $\{\tau = t\} \cap B_i^1$,

$$\int_{s_i}^{s_{i+1}} \left|\alpha_s^{ij}\right|^2 ds \le 2c^2 \varepsilon^{-2\theta} \varepsilon^q \frac{t^2}{n^2} + 2c^2 \varepsilon^{-4\theta} K_\beta^2 \frac{t^{2+2\beta}}{n^{3+2\beta}}.$$

Hence, if $q + 2\mu - 2\theta > 2 + 4\mu + 2r$ and $-4\theta + (3 + 2\beta)\mu > 2 + 4\mu + 2r$, one can find ε_0'' which only depends on K, K_β, t, β, μ, θ and r such that for $\varepsilon \le \varepsilon_0''$, on $\{\tau = t\} \cap B_i^1$,

$$\int_{s_i}^{s_{i+1}} \left|\alpha_s^{ij}\right|^2 ds \le \frac{\varepsilon^{2+4\mu+2r}}{1024m^2}.$$

Applying (2.17), we get

$$\mathbb{P}(\tau = t, B_i^1, |M_{ij}| \ge \frac{1}{32m}\varepsilon^{1+2\mu}) \le 2\exp\left(-\frac{1}{2}\varepsilon^{-2r}\right). \tag{2.18}$$

To conclude, assume that all conditions (2.15), (2.16) and (2.18) are satisfied, and recall (2.13). Then for ε small enough ($\varepsilon \le \min(\varepsilon_0, \varepsilon_0', \varepsilon_0'')$) the set

$$\{\tau = t\} \cap B_i^1 \cap \bigcap_{j=1}^m \left\{|M_{ij}| < \frac{1}{32m}\varepsilon^{1+2\mu}\right\}$$

will be empty provided $q > 1 + 2\mu$. Hence, in this case,

$$\mathbb{P}(\tau = t, B_i^1) \le 2m \exp\left(-\frac{1}{2}\varepsilon^{-2r}\right)$$

and if $r > 0$,

$$\mathbb{P}(B, \tau = t) \le 4m\, \varepsilon^{-\mu} \exp\left(-\frac{1}{2}\varepsilon^{-2r}\right) \le c_p \varepsilon^p$$

for ε small enough (universal). Thanks to (2.8), we finally get $\mathbb{P}(B) \le c_p\varepsilon^p$ hence a bound for $\mathbb{P}(E_{k+1}^c \cap E_k)$ (recall that $q = \dfrac{m(k)}{m(k+1)}$). This achieves step 2 and the proof of Theorem 2.6, provided we can find θ, q, μ, r which satisfies all conditions cited above i.e.

i) $r > 0$, $\theta > 0$, $\mu > 0$, $q > 0$.
ii) $\beta\mu - 2\theta > 1$.
iii) $-4\theta + (3 + 2\beta)\mu > 2 + 4\mu + 2r$.
iv) $q > 2\mu + 1$, $q + 2\mu - 2\theta > 2 + 4\mu + 2r$, $-2\theta + q + 3\mu > 2 + 4\mu$.

Fix $r > 0$, $\theta > 0$. Then choose $\mu > \dfrac{1 + 2\theta}{\beta}$ and $q > \max(2\mu + 1, 2\mu + 2\theta + 2r + 2, \mu + 2 + 2\theta)$ so that i), ii) and iv) are satisfied. The delicate point is iii) which yields

$$(2\beta - 1)\mu > 2 + 4\theta + 2r \tag{2.19}$$

and can be satisfied only if $2\beta - 1 > 0$. □

Remark 2.20 *As the preceding proof clearly shows, condition $\beta > \frac{1}{2}$ appears in the control of the martingale term. More precisely, Ito formula leads to bound $\left\langle \xi, \phi_u^{\star -1} Z(s, \bullet) \right\rangle (x)$, while the initial control is on $\left\langle \xi, \phi_s^{\star -1} Z(s, \bullet) \right\rangle (x)$. The error $(s-u)^\beta$ is not important for the bounded variation term (it only imposes $\beta\mu - 2\theta > 1$), and becomes important for the martingale term (imposing condition iii) above). Since the martingale term appears to be the leading one, one should try to get a better control on it. In a sense, it is what we are going to do in the next section, but the strategy will require more refined tools.*

Remark 2.21 *One can relax the (uniform) boundedness conditions on the coefficients and their modulus of continuity; i.e. K and K_β may have polynomial growth in R, where R is the radius of the ball in $\mathbb{R}^d$. Indeed, this will only change the constants in front of θ, and will not change the final $\beta > \frac{1}{2}$. But in order to show regularity of the density in section 4 we shall use a localization procedure, which enables to only look at the uniformly bounded case.*

One can also see that ε_0 has some polynomial growth in $\frac{1}{t}$ when t goes to 0. This will be used in section 4.

3. Estimates on Malliavin covariance matrix : second approach

To help the reader we first explain the flavor of the strategy on an example

example 3.1: Consider the 2-dimensional process

$$\begin{cases} x_t = x + w_t \\ y_t = y + \displaystyle\int_0^t h(u, w_u)\,\delta B_u \end{cases}$$

where (w, B) is a Brownian motion. Suppose that $h(0,0) = 0$ but $\dfrac{\partial h}{\partial x}(0,0) \neq 0$. Then $X_1(0,0)$ and $[X_1, X_2](0,0)$ span $\mathbb{R}^2$ ($X_1 = \binom{1}{0}$, $X_2 = \binom{0}{h(0,x)}$, $X_0 = \binom{0}{0}$), so that Hörmander's condition is satisfied. Take $\xi = (0,1)$. It is easy to check

$$\left\langle \xi, \phi_t^{\star -1} X_1 \right\rangle (0,0) = -\int_0^t \frac{\partial h}{\partial x}(u, w_u)\,\delta B_u, \quad \left\langle \xi, \phi_t^{\star -1} X_2 \right\rangle (0,0) = h(t, w_t).$$

Hence it is enough to control $\mathbb{P}\left(\int_0^t h^2(s, w_s) ds \leq \varepsilon \right)$ at least for small t. If h satisfies the β-Hölder condition (2.2), we may assume $\left| \dfrac{\partial h}{\partial x}(\lambda, 0) \right| \geq \dfrac{1}{2} \left| \dfrac{\partial h}{\partial x}(0,0) \right| > 0$ for λ small enough. We may first freeze the time component in the vector field and thus apply Norris Lemma [31, lemma 2.3.2] to the process

$$h(\lambda, w_s) = h(\lambda, 0) + \int_0^s \frac{\partial h}{\partial x}(\lambda, w_u)\,\delta w_u + \frac{1}{2}\int_0^s \frac{\partial^2 h}{\partial x^2}(\lambda, w_u)\,du$$

and get that

$$\mathbb{P}\left(\int_0^t h^2(\lambda, w_s)\,ds \leq \varepsilon\right) \leq c_p \varepsilon^p$$

where c_p does not depend on λ (step 3 in the general case where we shall also need a time discretization). We claim that a regularity argument, or a modification of Norris Lemma, shows that actually (step 1 in the general case where we shall replace the L^2 norm by the L^∞ norm since the Brownian paths are Hölder continuous)

$$\text{for all } \lambda \in [0,t], \quad \mathbb{P}\left(\sup_{0\leq s\leq t} h^2(\lambda, w_s) \leq \varepsilon\right) \leq c_p \varepsilon^p.$$

Here again c_p does not depend on λ.
The second claim is the following : thanks to the regularity of h in λ, we can show

$$\mathbb{P}\left(\forall \lambda \in [0,t], \sup_{0\leq s\leq t} h^2(\lambda, w_s) \geq \varepsilon\right) \geq 1 - c_p(\beta)\varepsilon^p.$$

Now observe that because of $\left|\dfrac{\partial h}{\partial x}(\lambda, x)\right| \geq \alpha > 0$ for $\lambda \in [0,t]$, $x \in \mathbb{R}$ (which is a stronger assumption than the one we did, but can be used if we modify h and introduce stopping times), the process $h(\lambda, w_\bullet)$ is up to a Girsanov transform and a time change, just a Brownian motion. For Brownian motion the uniqueness of the supremum is well known [20, Remark 8.16, p.102]. Hence , we shall assume that for almost all w, for all λ, the supremum is attained at some single time s_λ, i.e. $s_\lambda = \operatorname{argsup}_{0\leq s\leq t} h^2(\lambda, w_s)$. Of course s_λ is random, but for a fixed w, $\lambda \mapsto s_\lambda(w)$ will then be continuous from $[0,t]$ into itself, hence admits at least one fixed point $s^\star$. On the above set of big Probability, we will then have $h^2(s^\star, w_{s^\star}) \geq \varepsilon$, thus $\sup_{0\leq s\leq t} h^2(s, w_s) \geq \varepsilon$. Using Hölder continuity of $s \mapsto h(s, w_s)$, this yields (step 4 in the general case)

$$\mathbb{P}\left(\int_0^t h^2(s, w_s)\,ds \leq \varepsilon^{1+\frac{1}{\beta}}\right) \leq c'_p(\beta)\,\varepsilon^p$$

for some β i.e. the desired result up to an obvious change of notation. Unfortunately our assumption

$$\mathbb{P}\text{-a.s.}, \quad \forall \lambda \in [0,t], \text{ the supremum is attained at a single point } s_\lambda$$

has very few chances to be true, because of the exchange of $\mathbb{P}$-a.s. and for all λ. Here again one has to use continuity to modify this assumption, without modifying the conclusion.

¿From now on, we assume that hypotheses (2.2), (2.3)and (2.4) are satisfied and we use the notations of section 2. Our aim is to prove the following

Proposition 3.2 *There exists $q > 0$ such that for all Z satisfying (2.2) with bounds (2.4)*

$$\mathbb{P}(B) = \mathbb{P}\left(\int_0^t \left\langle \xi, \phi_s^{\star-1} Z\right\rangle^2(x)ds \leq \varepsilon^q, \sum_{j=1}^m \int_0^t \left\langle \xi, \phi_s^{\star-1}\left[X_j, Z\right]\right\rangle^2(x)ds \geq \varepsilon\right) \leq c_p \varepsilon^p$$

where c_p only depends on β, K, K_β, p, x, m and t.

This statement is exactly (2.7) in the previous section, but we do no more assume $\beta > \frac{1}{2}$ from now. In the proof, we shall use similar arguments as those exposed in the example above, and first to get rid of the assumption on β, we shall deeply use β-Hölder continuity. To this end, we state a straightforward application of Kolmogorov continuity criterion

Lemma 3.3 *Let Y be a flow of vector fields satisfying (2.2) with the bounds (2.4). Then for all $\beta' < \beta \wedge \frac{1}{2}$, there exists a version of*

$$\left\langle \xi, \phi_s^{\star -1} Y \right\rangle (x) \;=\; \left\langle \xi, \left(\frac{\partial \phi_s}{\partial x}(w,x) \right)^{-1} Y\left(\phi_s(w,x)\right) \right\rangle$$

which is β'-Hölder continuous on $[0,T]$. Furthermore if we set

$$K_{\beta'}(w,Y) \;=\; \left\| \left\langle \xi, \phi_\bullet^{\star -1} Y \right\rangle (x) \right\|_{\beta'}$$

then

$$\mathbb{E}\left(K_{\beta'}^p(w,Y) \right) \le c_p < +\infty$$

where c_p only depends on p, x, β', K and K_β.

The proof is left to the reader. Since the proof of Proposition 3.2 is quite long, we split it into several steps.

Step 1 : **From L^2 estimates to L^∞ estimates**.
Thanks to Lemma 3.3, we first replace integral terms by supremums. In all what follows, we fix once for all β' given by Lemma 3.3. As a first consequence of β-Hölder continuity, we get

Lemma 3.4 *Let Y be as in Lemma 3.3 and $\varepsilon < \frac{1}{4} \wedge t$. Then if $a \ge 2\left(1 + \frac{1}{\beta'}\right) + 1$, for all $p \in [1, +\infty[$,*

$$\mathbb{P}\left(\int_0^t \left\langle \xi, \phi_s^{\star -1} Y \right\rangle^2 (x) ds < \varepsilon^a, \sup_{0 \le s \le t} \left| \left\langle \xi, \phi_s^{\star -1} Y \right\rangle (x) \right| \ge \varepsilon \right) \le c_p \varepsilon^p$$

where c_p only depends on t, p, x, β', K and K_β (we shall call it a "good" constant).

Proof: Let $z_s = \left\langle \xi, \phi_s^{\star -1} Y \right\rangle (x)$. According to 3.3, $|z_s(w) - z_u(w)| \le K_{\beta'}(w,Y)|s-u|^{\beta'}$ and

$$\mathbb{P}\left(K_{\beta'}(w,Y) \ge \frac{1}{2\varepsilon} \right) \le c_p \varepsilon^p$$

where c_p is a "good" constant. So we may assume that $w \in \left\{ K_{\beta'}(w,Y) < \frac{1}{2\varepsilon} \right\}$ from now on. For such an w and $|s^\star - s| \le \varepsilon^\eta$, one has $|z_s(w)| \ge |z_{s^\star}(w)| - \frac{1}{2}\varepsilon^{\beta'\eta - 1}$. Moreover assume that $\sup_{0 \le s \le t} |z_s(w)| \ge \varepsilon$. Then one has for $\varepsilon < t$ and $\eta > 1$

$$\int_0^t z_s^2(w)\,ds \;\geq\; \int_{\text{interval of size } \varepsilon^\eta} \left(|z_{s^\star}(w)| - \frac{1}{2}\varepsilon^{\beta'\eta-1}\right)^2 \,ds$$

where $|z_{s^\star}(w)| = \sup_{0\leq s\leq t} |z_s(w)|$. So choosing $\eta = \frac{2}{\beta'}$, we get $\int_0^t z_s^2(w)ds \geq \frac{1}{4}\varepsilon^{2+\frac{2}{\beta'}}$. Since $a \geq 2\left(1+\frac{1}{\beta'}\right)+1$ and $\varepsilon \leq \frac{1}{4}$, it holds $\varepsilon^a < \frac{1}{4}\varepsilon^{2+\frac{2}{\beta'}}$ and accordingly the set

$$\left\{\int_0^t \left\langle \xi, \phi_s^{\star-1} Y\right\rangle^2(x)ds < \varepsilon^a, \sup_{0\leq s\leq t} \left|\left\langle \xi, \phi_s^{\star-1} Y\right\rangle(x)\right| \geq \varepsilon\right\} \cap \left\{K_{\beta'}(w,Y) < \frac{1}{2\varepsilon}\right\}$$

is empty. □

Thanks to Lemma 3.4, if we put $q = ra$ with $r > 0$ and $a = 2\left(1+\frac{1}{\beta'}\right)+1$, the proof of Proposition 3.2 reduces to show that for some r,

$$P(C) = \mathbb{P}\left(\sup_{0\leq s\leq t} \left|\left\langle \xi, \phi_s^{\star-1} Z\right\rangle(x)\right| \leq \varepsilon^r, \sum_{j=1}^m \int_0^t \left\langle \xi, \phi_s^{\star-1}\left[X_j, Z\right]\right\rangle^2(x)ds \geq \varepsilon\right) \leq c_p\varepsilon^p \tag{3.5}$$

for some "good" constant c_p.

Step 2 : **Time discretization**.
We introduce a dicretization of the time interval $[0, t]$

$$s_i = \frac{it}{n} \text{ with } n = \left[\varepsilon^{-\frac{4}{\beta'}}\right]. \tag{3.6}$$

Introduce the sets

$$C_i = \left\{\int_0^t \sum_{j=1}^m \left\langle \xi, \phi_s^{\star-1}\left[X_j, Z\right]\right\rangle^2(x)ds \geq \varepsilon \text{ and } \exists s \in [s_i, s_{i+1}], \sum_{j=1}^m \left\langle \xi, \phi_s^{\star-1}\left[X_j, Z\right]\right\rangle^2(x) \;<\; \varepsilon^{\frac{3}{2}}\right\}$$

We then show

Lemma 3.7 *If $\varepsilon^{\frac{1}{2}} < \frac{1}{2t}$ then*

$$\mathbb{P}\left(\bigcap_{i=1}^{n-1} C_i\right) \leq c_p\varepsilon^p$$

for some "good" c_p.

Proof: If $\int_0^t$ is greater than ε, then $\sup_{0\leq s\leq t}$ is greater than $\frac{\varepsilon}{t}$. So as before, up to a set of probability less than $c_p\varepsilon^p$, we may assume that $K_{\beta'}(w,Y)\leq\varepsilon^{-\frac{1}{2}}$ for all Y of the form $[X_j,Z]$. Thus there exists at least one interval I of length $\varepsilon^{\frac{2}{\beta'}}$ such that for $s\in I$,

$$\sum_{j=1}^m \left\langle \xi,\phi_s^{\star-1}\left[X_j,Z\right]\right\rangle^2(x)\geq\frac{\varepsilon}{t}-\varepsilon^2\varepsilon^{-\frac{1}{2}}\geq\frac{\varepsilon}{2t}>\varepsilon^{\frac{3}{2}}.$$

Finally to conclude, it suffices that I contains at least one interval $[s_i,s_{i+1}]$ which ensures that the set $\bigcap_{i=0}^{n-1}C_i\cap\left\{K_{\beta'}(w,Y)\leq\varepsilon^{-\frac{1}{2}}\right\}$ is empty. □

Introduce the family of sets

$$D_i=\left\{\sup_{s_i\leq s\leq s_{i+1}}\left|\left\langle\xi,\phi_s^{\star-1}Z\right\rangle(x)\right|\leq\varepsilon^r,\sum_{j=1}^m\left\langle\xi,\phi_s^{\star-1}\left[X_j,Z\right]\right\rangle^2(x)\geq\varepsilon\text{ for all }s\in[s_i,s_{i+1}]\right\}$$

and the set $D=\bigcup_{i=0}^{n-1}D_i$. As in section 2, we can introduce the stopping time τ defined by (2.8) and remark that $\mathbb{P}(\tau<t)\leq c_p\varepsilon^p$. Hence we are reduced to estimate $\mathbb{P}(C\cap\{\tau=t\})$. According to Lemma 3.7 (just changing ε into $\varepsilon^{\frac{3}{2}}$ and replacing r by $\frac{3}{2}r$), one has $\mathbb{P}(C\cap D^c)\leq c_p\varepsilon^p$. So we only have to estimate $\mathbb{P}(C\cap D\cap\{\tau=t\})$. Note now that

$$\mathbb{P}(C\cap D\cap\{\tau=t\})\leq\mathbb{P}(D\cap\{\tau=t\})\leq\sum_{i=0}^{n-1}\mathbb{P}(D_i\cap\{\tau=t\})\leq\left[\varepsilon^{-\frac{4}{\beta'}}\right]\max_i\mathbb{P}(D_i\cap\{\tau=t\}).$$

Since p is arbitrary, we are reduced to show that

$$\mathbb{P}(D_i\cap\{\tau=t\})\leq c_p\varepsilon^p \tag{3.8}$$

for some "good" constant independent of i.

Step 3 : **Freezing the time**.
On $\{\tau=t\}$, one has

$$\sup_{(s,\lambda)\in[s_i,s_{i+1}]^2}\left|\left\langle\xi,\phi_s^{\star-1}\left[X_j,Z\right]\right\rangle(x)-\left\langle\xi,\phi_s^{\star-1}\left[X_j(s,\bullet),Z(\lambda,\bullet)\right]\right\rangle(x)\right|\leq 2KK_\beta\varepsilon^{-\theta}\varepsilon^{\frac{4\beta}{\beta'}} \tag{3.9}$$

Set

$$E_i=\left\{\inf_{(s,\lambda)\in[s_i,s_{i+1}]^2}\sum_{j=1}^m\left\langle\xi,\phi_s^{\star-1}\left[X_j(s,\bullet),Z(\lambda,\bullet)\right]\right\rangle^2(x)\geq\frac{\varepsilon}{2}\right\}.$$

Thanks to (3.9), one can find ε_0 which only depends on m, K, K_β such that, provided $\theta < \dfrac{1}{2}$ in order to ensure that $-\theta + \frac{4\beta}{\beta'} > \frac{3}{2}$, for all $\varepsilon < \varepsilon_0$

$$D_i \cap \{\tau = t\} \subset D_i \cap \{\tau = t\} \cap E_i = F_i \cap \{\tau = t\}$$

where

$$F_i = \left\{ \sup_{s_i \leq s \leq s_{i+1}} \left| \left\langle \xi, \phi_s^{\star -1} Z \right\rangle (x) \right| \leq \varepsilon^r, \inf_{(s,\lambda)\in[s_i,s_{i+1}]^2} \sum_{j=1}^m \left\langle \xi, \phi_s^{\star -1} \left[X_j(s,\bullet), Z(\lambda,\bullet) \right] \right\rangle^2 (x) \geq \frac{\varepsilon}{2} \right\}.$$

So we are reduced to show

$$\mathbb{P}\left(F_i, \tau = t\right) \leq c_p \varepsilon^p. \tag{3.10}$$

Without loss of generality, we shall only deal with the case $i = 0$. This will avoid notational intricacies, since the only important point in our future arguments will be the length $s_{i+1} - s_i$.We now use Ito formula which yields

$$\left\langle \xi, \phi_s^{\star -1} Z(\lambda,\bullet) \right\rangle (x) =: y_s(\lambda) = y_0(\lambda) + \int_0^s \beta_u(\lambda) du + \sum_{j=1}^m \int_0^s \alpha_u^j(\lambda) \delta w_u^j \tag{3.11}$$

with

$$\alpha_u^j(\lambda) = \left\langle \xi, \phi_u^{\star -1} \left[X_j(u,\bullet), Z(\lambda,\bullet) \right] \right\rangle (x)$$
$$\beta_u(\lambda) = \left\langle \xi, \phi_u^{\star -1} \left\{ \left[X_0(u,\bullet), Z(\lambda,\bullet) \right] + \frac{1}{2} \sum_{j=1}^m \left[X_j(u,\bullet), \left[X_j(u,\bullet), Z(\lambda,\bullet) \right] \right] \right\} \right\rangle (x).$$

Note that there exists a constant C depending only on m, K and $|\xi|$ such that, on $\{\tau = t\}$,

$$|\alpha_u^j(\lambda)| + |\beta_u(\lambda)| \leq C\varepsilon^{-\theta}$$

for all j, u and λ.
As we explained in example 3.1, we want to study the supremum of $y_s(\lambda)$, but be sure that this supremum is attained at a single point. So we will have to restrict ourselves to a random time interval rather than working on $[0, s_1]$ and then evaluate the supremum. The next step will consist in showing the existence of the argsup and is deeply related to the existence of the argsup for Brownian motion.

Step 4 : **Existence of the argsup**.
Since we will use a Girsanov transform, we have to work on the full Probability space Ω and not only on F_0. To this end, we have first to modify $y_\bullet(\lambda)$. Taking if necessary an extension of Ω, we may find another Brownian motion w^{m+1} which is independent of $(w^1, ..., w^m)$ (when we work on $[s_i, s_{i+1}]$, we have to consider $w_s^j - w_{s_i}^j$ rather than w_s^j and choose w^{m+1} which is independent of $y_{s_i}(\lambda)$ too). Next we define

$$\bar{y}_s(\lambda) = y_0(\lambda) + \int_0^s \bar{\beta}_u(\lambda) du + \sum_{j=1}^{m+1} \int_0^s \bar{\alpha}_u^j(\lambda) \delta w_u^j \tag{3.12}$$

where $\bar{\beta}$, $\bar{\alpha}^j$ are defined as follows. We choose two smooth functions $M_{\theta,\varepsilon}$ and χ_ε defined on $\mathbb{R}$ such that

$$M_{\theta,\varepsilon}(z) = z \text{ if } |z| < C\varepsilon^{-\theta}, \quad \pm 2C\varepsilon^{-\theta} \text{ if } |z| > 3C\varepsilon^{-\theta} \text{ and } |M_{\theta,\varepsilon}| \text{ is bounded by } 2C\varepsilon^{-\theta}$$

so that we may also assume that its derivative is bounded by 1, and

$$\chi_\varepsilon(z) = 1 \text{ if } |z| \geq \frac{\varepsilon}{2}, \quad 0 \text{ if } |z| \leq \frac{\varepsilon}{4}, \quad \chi_\varepsilon \text{ is even and non decreasing on } \mathbb{R}^+.$$

These functions may be viewed as smooth versions of the cut-off by $\pm 2C\varepsilon^{-\theta}$ and the indicator $\mathbb{1}_{|z|\geq\frac{\varepsilon}{2}}$. We then set

$$\begin{cases} \bar{\beta}_u(\lambda) = M_{\theta,\varepsilon}\left(\beta_u(\lambda)\right) \quad (\text{roughly } \beta_u(\lambda) \wedge C\varepsilon^{-\theta} \vee -C\varepsilon^{-\theta}) \\ \bar{\alpha}_u^j(\lambda) = M_{\theta,\varepsilon}\left(\alpha_u^j(\lambda)\right) \chi_\varepsilon\left(\sum_{k=1}^m \left(\alpha_u^k(\lambda)\right)^2\right) \quad \text{for } j = 1, ..., m \\ \bar{\alpha}_u^{m+1}(\lambda) = \sqrt{\frac{\varepsilon}{2}}\left(1 - \chi_\varepsilon\left(\sum_{k=1}^m \left(\alpha_u^k(\lambda)\right)^2\right)\right) \end{cases} \tag{3.13}$$

Notice that

$$\text{if } \omega \in \{\tau = t\}, \quad \text{for } s \leq s_1, \quad \bar{y}_s(\lambda, \omega) = y_s(\lambda, \omega) \tag{3.14}$$

i.e. $\bar{y}$ is actually a modification of y. The second and main fact is

$$\sum_{j=1}^{m+1} \left(\bar{\alpha}_u^j(\lambda)\right)^2 = \bar{\alpha}_u^2(\lambda) \geq \frac{\varepsilon}{8} \text{ for all } u. \tag{3.15}$$

Indeed if $\sum_{j=1}^m \left(\bar{\alpha}_u^j(\lambda)\right)^2$ is greater than $\frac{\varepsilon}{2}$ or less that $\frac{\varepsilon}{4}$ then $\bar{\alpha}_u^2(\lambda) \geq \frac{\varepsilon}{2}$. Otherwise

$$\bar{\alpha}_u^2(\lambda) \geq \sum_{j=1}^m \left(\bar{\alpha}_u^j(\lambda)\right)^2 \left(\chi_\varepsilon^2 + (1 - \chi_\varepsilon)^2\right) \geq \frac{\varepsilon}{4} \times \frac{1}{2}.$$

Thanks to (3.15), we can introduce a new Probability measure $\tilde{\mathbb{P}}$ given by the Girsanov density

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}|_{\mathcal{F}_t} = \exp\left(-\sum_{j=1}^{m+1} \int_0^t \frac{\bar{\alpha}_u^j(\lambda)\bar{\beta}_u(\lambda)}{\bar{\alpha}_u^2(\lambda)} \delta w_u^j - \frac{1}{2}\int_0^t \frac{\bar{\beta}_u^2(\lambda)}{\bar{\alpha}_u^2(\lambda)} du\right)$$

which is equivalent to $\mathbb{P}$ on each $\mathcal{F}_t$. Furthermore, $\tilde{\mathbb{P}}$-almost surely,

$$\bar{y}_s(\lambda) = y_0(\lambda) + \sum_{j=1}^{m+1} \int_0^s \bar{\alpha}_u^j(\lambda) \delta \tilde{w}_u^j$$

where $(\tilde{w}^1, ..., \tilde{w}^{m+1})$ is a $\tilde{\mathbb{P}}$-Brownian motion. Now if we define

$$A_s(\lambda) = \int_0^s \bar{\alpha}_u^2(\lambda) du \text{ and } V_s(\lambda) = \inf \{u \geq 0, \quad A_u(\lambda) \geq s\}$$

there exists a $\tilde{\mathbb{P}}$-Brownian motion B such that for all s

$$\bar{y}_s(\lambda) = y_0(\lambda) + B_{A_s(\lambda)} \text{ i.e. } \bar{y}_{V_s(\lambda)}(\lambda) = y_0(\lambda) + B_s \quad \tilde{\mathbb{P}}\text{-a.s.}$$

But as we recall in example 3.1, it is well known that the supremum of the absolute value of a linear Brownian motion on a deterministic time interval is a.s. attained at a single time. Take $T_1 = \frac{\varepsilon}{8} s_1$. Since

$$\text{(3.16)} \qquad \frac{\varepsilon}{8} s \leq A_s(\lambda) \leq 4C^2 \varepsilon^{-2\theta}(m+1)s$$

then

$$\frac{C^2 \varepsilon^{2\theta}}{4(m+1)} s \leq V_s(\lambda) \leq \frac{8}{\varepsilon} s.$$

Hence

$$\text{(3.17)} \qquad \frac{C^2 t}{32(m+1)} \varepsilon^{1+2\theta+\frac{4}{\beta'}} \leq V_{T_1}(\lambda) \leq t\varepsilon^{\frac{4}{\beta'}} = s_1.$$

¿From now on, to simplify the notation, we set $S_1(\lambda) = V_{T_1}(\lambda)$. We get from what precedes

(3.18) $\forall \lambda \in [0, s_1], \tilde{\mathbb{P}}$-a.s.,there exists an unique (random) time $u_1(\omega, \lambda) = \underset{0 \leq u \leq S_1(\lambda)}{\text{argsup}} |\bar{y}_u(\lambda)|$.

Since $\mathbb{P}$ and $\tilde{\mathbb{P}}$ are equivalent on $\mathcal{F}_{s_1}$ and $S_1(\lambda) \leq s_1$, (3.18) holds $\mathbb{P}$-a.s. and we can thus define

$$\text{(3.19)} \qquad u_1(\omega, \lambda) = \underset{0 \leq u \leq S_1(\lambda)}{\text{argsup}} |\bar{y}_u(\lambda)| \quad \mathbb{P}\text{-a.s. for all } \lambda \in [0, s_1].$$

Since we only have the existence of the argsup for the modified process $\bar{y}(\lambda)$, we need to control the supremum of this process.

Step 5 : **Estimates for the supremum**.

Define $t_1 = \dfrac{C^2 t}{32(m+1)}\varepsilon^{1+2\theta+\frac{4}{\beta'}}$. Thanks to (3.17), $\sup_{0\le u\le t_1} |\bar{y}_u(\lambda)| \le \sup_{0\le u\le S_1(\lambda)} |\bar{y}_u(\lambda)|$. We will estimate $\sup_{0\le u\le t_1} |\bar{y}_u(\lambda)|$ similarly as in the classical Norris lemma. Ito's formula yields

$$\text{(3.20)}\quad \bar{y}^2_{t_1}(\lambda) = \bar{y}^2_0(\lambda) + 2\int_0^{t_1} \bar{y}_u(\lambda)\bar{\beta}_u(\lambda)du + 2\sum_{j=1}^{m+1}\int_0^{t_1} \bar{y}_u(\lambda)\bar{\alpha}^j_u(\lambda)\delta w^j_u + \int_0^{t_1} \bar{\alpha}^2_u(\lambda)du.$$

We then show

Lemma 3.21 *If $r' > \frac{3}{2} + 3\theta + \frac{2}{\beta'}$, there exists $\varepsilon_0(m,t)$ s.t. if $\varepsilon < \varepsilon_0$*

$$\mathbb{P}\left(\sup_{0\le u\le t_1} |\bar{y}_u(\lambda)| \le \varepsilon^{r'}\right) \le ce^{-c'\varepsilon^{-2\theta}}$$

where c and c' depend on m and t.

Proof: on $\left\{\sup_{0\le u\le t_1} |\bar{y}_u(\lambda)| \le \varepsilon^{r'}\right\}$, one has

$$\text{(i)}\qquad \left|2\int_0^{t_1} \bar{y}_u(\lambda)\bar{\beta}_u(\lambda)du\right| \le 4\varepsilon^{-\theta}\varepsilon^{r'}t_1 = \frac{C^2 t}{8(m+1)}\varepsilon^{1+\frac{4}{\beta'}+r'+\theta}$$

$$\text{(ii)}\qquad \int_0^{t_1} \left(\bar{y}_u(\lambda)\bar{\alpha}^j_u(\lambda)\right)^2 du \le 4\varepsilon^{-2\theta}\varepsilon^{r'}t_1 = \frac{C^2 t}{8(m+1)}\varepsilon^{1+2r'+\frac{4}{\beta'}}$$

so that

$$\mathbb{P}\left(\sup_{0\le u\le t_1} |\bar{y}_u(\lambda)| \le \varepsilon^{r'}, \left|2\sum_{j=1}^{m+1}\int_0^{t_1} \bar{y}_u(\lambda)\bar{\alpha}^j_u\delta w^j_u\right| \ge \varepsilon^{r'+\frac{2}{\beta'}+\frac{1}{2}-\theta}\right)$$

is less than $ce^{-c'\varepsilon^{-2\theta}}$ thanks to the martingale inequality we recall in (2.17).

$$\text{(iii)}\qquad \int_0^{t_1} \bar{\alpha}^2_u(\lambda)du \ge \frac{\varepsilon}{8}t_1 = \frac{C^2 t}{256(m+1)}\varepsilon^{2+2\theta+\frac{4}{\beta'}}$$

so if we choose r' such that

$$2r' > 2 + 2\theta + \frac{4}{\beta'},\quad 1 + \frac{4}{\beta'} + r' + \theta > 2 + 2\theta + \frac{4}{\beta'},\quad \frac{1}{2} + r' + \frac{2}{\beta'} - \theta > 2 + 2\theta + \frac{4}{\beta'}$$

i.e.

$$r' > \frac{3}{2} + 3\theta + \frac{2}{\beta'},$$

then there is an ε_0 depending on m and t s.t. for all $\varepsilon < \varepsilon_0$, $\sup_{0\le u\le t_1} |\bar{y}_u(\lambda)| \le \varepsilon^{r'}$ implies that the martingale term in (3.20) is bigger than $\varepsilon^{r'+\frac{2}{\beta'}+\frac{1}{2}-\theta}$ and we may apply (ii). □

We will now get an estimate for all λ simultaneously i.e. we prove

Lemma 3.22 *Let r' be as in lemma 3.21. Then there exists ε_0' s.t. for $\varepsilon < \varepsilon_0'$*

$$\mathbb{P}\left(\exists\lambda \in [0, s_1],\ \sup_{0\le u\le t_1} |\bar{y}_u(\lambda)| \le \varepsilon^{r'}\right) \le c_p\varepsilon^p$$

for some "good" c_p.

Proof: First of all remark that we can again apply Kolmogorov continuity criterion in order to build a β'-Hölder continuous version of $(s, \lambda) \mapsto \bar{y}_s(\lambda)$ thanks to the mollifying (smooth) functions $M_{\theta,\varepsilon}$ and χ_ε. Denote by $\Gamma_{\beta',\varepsilon}(\omega)$ the β'-Hölder norm $\|\bar{y}_\bullet(\bullet)\|_{\beta'}$. The only point is that moments of $\Gamma_{\beta',\varepsilon}(\omega)$ depend on ε. Indeed if $\varepsilon < 1$ then the first derivative of $M_{\theta,\varepsilon}$ is bounded but the first derivative of χ_ε is bounded by $\dfrac{4}{\varepsilon}$. Standard applications of BDG inequalities and easy calculations then yield

$$\mathbb{E}\left(\Gamma^p_{\beta',\varepsilon}(\omega)\right) \le c_p\varepsilon^{-p} \tag{3.23}$$

for some "good" c_p independent of ε. Divide $[0, s_1]$ into $K = \left[s_1\varepsilon^{-2\frac{r'}{\beta'}}\right] + 1$ intervals of length at most $\varepsilon^{\frac{2r'}{\beta'}}$. Denote by $(\lambda_i)_{i=0,\dots,K}$ the extremities of these intervals. Then

$$\mathbb{P}\left(\exists\lambda \in [0, s_1],\ \sup_{0\le u\le t_1} |\bar{y}_u(\lambda)| \le \varepsilon^{r'}\right) \le \sum_{i=0}^{K-1} \mathbb{P}\left(\exists\lambda \in [\lambda_i, \lambda_{i+1}],\ \sup_{0\le u\le t_1} |\bar{y}_u(\lambda)| \le \varepsilon^{r'}\right).$$

But if $|\lambda - \lambda'| \le \varepsilon^{2\frac{r'}{\beta'}}$ then $\sup_{0\le u\le t_1} |\bar{y}_u(\lambda) - \bar{y}_u(\lambda')| \le \Gamma_{\beta',\varepsilon}\varepsilon^{2r'}$ will be less than $\varepsilon^{r'}$ on the set $\left\{\Gamma_{\beta',\varepsilon} \le \varepsilon^{-r'}\right\}$. In particular

$$\left\{\exists\lambda \in [\lambda_i, \lambda_{i+1}],\ \sup_{0\le u\le t_1} |\bar{y}_u(\lambda)| \le \varepsilon^{r'},\ \Gamma_{\beta',\varepsilon} \le \varepsilon^{-r'}\right\} \subset \left\{\sup_{0\le u\le t_1} |\bar{y}_u(\lambda_i)| \le 2\varepsilon^{r'}\right\} = G_i$$

using

$$|\bar{y}_u(\lambda_i)| \le |\bar{y}_u(\lambda) - \bar{y}_u(\lambda_i)| + |\bar{y}_u(\lambda_i)| \le \Gamma_{\beta',\varepsilon}\varepsilon^{2r'} + \varepsilon^{r'}.$$

Moreover one has

$$\left\{\exists\lambda \in [\lambda_i, \lambda_{i+1}],\ \sup_{0\le u\le t_1} |\bar{y}_u(\lambda)| \le \varepsilon^{r'},\ \Gamma_{\beta',\varepsilon} > \varepsilon^{-r'}\right\} \subset \left\{\Gamma_{\beta',\varepsilon} > \varepsilon^{-r'}\right\} = H.$$

But $\mathbb{P}(G_i) \le c\exp\left(-c'\varepsilon^{-2\theta}\right)$ by lemma 3.21, and $\mathbb{P}(H) \le c_\gamma\varepsilon^{(r'-1)\gamma}$ thanks to (3.23) and Markov inequality. Accordingly

$$\mathbb{P}\left(\exists\lambda \in [0, s_1], \sup_{0\le u\le t_1} |\bar{y}_u(\lambda)| \le \varepsilon^{r'}\right) \le s_1\varepsilon^{-2\frac{r'}{\beta'}}ce^{-c\varepsilon^{-2\theta}} + c_\gamma s_1\varepsilon^{-2\frac{r'}{\beta'}}\varepsilon^{(r'-1)\gamma}. \tag{3.24}$$

Recall that $s_1 = \varepsilon^{\frac{4}{\beta'}}$ and $r' > \dfrac{3}{2} + \dfrac{2}{\beta'}$. Hence choosing $\dfrac{\gamma - r'}{\beta'} = p$, the right hand side of (3.24) will be less than $c_p\varepsilon^p$ for $\varepsilon < \varepsilon_0'$ universal. □

Step 6 : **Conclusion**.
Let us return to $y_\bullet(\lambda)$. We may of course choose a continuous version of $(s,\lambda) \mapsto y_s(\lambda)$, and accordingly of $(s,\lambda) \mapsto (A_s(\lambda), V_s(\lambda))$. So (3.14) may be rewritten

$$\text{if } \omega \in F_0, \text{ for all } \lambda \in [0,s_1], \text{ all } s \le s_1, \; \bar{y}_s(\lambda,\omega) = y_s(\lambda,\omega).$$

Lemma 3.22 (recall that $\sup_{0\le u\le t_1} |\bar{y}_u(\lambda)| \le \sup_{0\le u\le S_1(\lambda)} |\bar{y}_u(\lambda)|$) yields

$$\mathbb{P}(F_0) \le \mathbb{P}(\tilde{F}_0) + c_p\varepsilon^p$$

where $\tilde{F}_0 = F_0 \cap \left\{ \forall\lambda \in [0,s_1], \sup_{0\le u\le S_1(\lambda)} |y_u(\lambda)| > \varepsilon^{r'} \right\}$.
We also know that for each fixed λ, the supremum is a.s. attained at a single point. But here again we have to be careful, because the negligible set depends on λ. So first define $u_1(\omega,\lambda) = \text{argsup}_{0\le u\le S_1(\lambda)} \bar{y}_u(\lambda)$ for all $\lambda \in \mathbb{Q}\cap[0,s_1]$ which is well defined $\mathbb{P}$-a.s. since $\mathbb{Q}$ is countable. We claim that for $\mathbb{P}$-almost all ω, $\lambda \mapsto u_1(\omega,\lambda)$ is continuous from $[0,s_1]\cap\mathbb{Q}$ into $[0,s_1]$.
This claim is almost standard (continuity of the argsup), but let us prove it. If λ_n goes to λ, $S_1(\lambda_n)$ goes to $S_1(\lambda)$ and $u_1(\omega,\lambda_n)$ is bounded. If $\bar{u}(\omega,\lambda)$ is a cluster point, we have (using index n, instead of subsequence)

$$\forall\eta>0, \, \forall s \le S_1(\lambda)-\eta, \, \exists N \text{ s.t if } n \ge N, \, S_1(\lambda_n) \ge S_1(\lambda)-\eta \text{ and } y_s(\lambda_n) \le y_{u_1(\omega,\lambda_n)}(\lambda_n).$$

Taking limits in both hand sides we get $\forall s \le S_1(\lambda)-\eta$, $y_s(\lambda) \le y_{\bar{u}(\omega,\lambda)}(\lambda)$, and by letting η go to 0, we get $y_s(\lambda) \le y_{\bar{u}(\omega,\lambda)}(\lambda)$, $\forall s \le S_1(\lambda)$ which implies $\bar{u}(\omega,\lambda) = u_1(\omega,\lambda)$ by uniqueness.

So we can extend $(\lambda \mapsto u_1(\omega,\lambda), \lambda \in \mathbb{Q})$ into a continuous function $\bar{u}_1(\omega,\lambda)$, $\lambda \in [0,s_1]$. Since $\bar{u}_1$ is continuous from $[0,s_1]$ into itself, it admits at least one fixed point $\bar{s} = \bar{u}_1(\omega,\bar{s}) \in [0,s_1]$.

For $\mathbb{P}$-almost all $\omega \in \tilde{F}_0$, $y_{u_1(\omega,\lambda)}(\lambda) > \varepsilon^{r'}$ for all $\lambda \in \mathbb{Q}$. By continuity $y_{\bar{u}_1(\omega,\lambda)}(\lambda) \ge \varepsilon^{r'}$ for all $\lambda \in [0,s_1]$, hence $y_{\bar{u}_1(\omega,\bar{s})}(\bar{s}) = y_{\bar{s}}(\bar{s}) \ge \varepsilon^{r'}$. But $y_{\bar{s}}(\bar{s}) = \left\langle \xi, \phi_{\bar{s}}^{\star-1} Z(\bar{s},\bullet)\right\rangle(x) = \left\langle \xi, \phi_{\bar{s}}^{\star-1} Z\right\rangle(x)$. If $r' < r$, it follows that $\tilde{F}_0$ is empty.
Hence $\mathbb{P}(F_0) \le c_p\varepsilon^p$, as well as all $\mathbb{P}(F_i) \le c_p\varepsilon^p$, i.e. (3.10) is satisfied and the proof of proposition 3.2 is complete.
As a consequence we get exactly as in section 2

Theorem 3.25 *Let (2.2), (2.3) and (2.4) holds for some $\beta > 0$. Then for all $p \in [1,+\infty[$, there exists c_p and ε_0 which only depends on p, t, x, N, c_N, K, β and K_β such that for all $\varepsilon \le \varepsilon_0$*

$$\sup_{\xi\in S^{d-1}} \mathbb{P}\left(\int_0^t \sum_{j=1}^m \left\langle \xi, \phi_s^{\star-1} X_j \right\rangle^2 (x) ds \le \varepsilon \right) \le c_p\varepsilon^p.$$

Remark 3.26 *The statements of Remark 2.21 are still available, with a different polynomial growth.*

4. Existence of regular densities, estimates, parabolic-hypoellipticity

Theorem 3.25 (or Theorem 2.6) implies that Malliavin covariance matrix $C_t(x)$ (see (1.8)) is $\mathbb{P}$ almost surely invertible ad that $C_t^{-1}(x)$ belongs to all the L^p. Furthermore one can estimate these moments, namely

Proposition 4.1 *If (2.2), (2.3) and (2.4) hold for some $\beta > 0$, then one can find constants c_p, K_p, l_p depending on p, x, c_N, β, K and K_β, such that*

$$\mathbb{E}\left[\left(C_t^{-1}(x)\right)^p\right] \leq c_p\, t^{-K_p(Nl_p+1)}.$$

The proof is not very different as in the homogeneous case (see [32]). Indeed, although one can't use the scaling properties since coefficients are time dependent, we can nevertheless use remark 3.26 on the growth in time of ε_0. Notice that in the proof of Theorem 3.25, the size of $q = \dfrac{m(k)}{m(k+1)}$ is rather big (it contains some part in $(\frac{1}{\beta})^2$) when β is small. The constants c_p, K_p, l_p will then really be big, and certainly not sharp. Also notice that the methods developed in both previous sections cannot be extended to the case of (uniformly) continuous flows of vector fields, rather than β-Hölder flows. Indeed for each time discretization we need a polynomial control on the cardinality of the discretization (i.e. $\varepsilon^{-\frac{1}{\beta'}}$ or $\varepsilon^{-\frac{4}{\beta'}}$ at some points). However the very strong global assumption (2.2) can be weakened into a local one, thanks to the following argument:

Lemma 4.2 *If the hypotheses of Proposition 4.1 are fulfilled, then for all stopping times τ satisfying $\dfrac{1}{\tau} \in \displaystyle\bigcap_{1\leq p<\infty} L^p$, $C_\tau(x)$ is $\mathbb{P}$ almost surely invertible and $C_\tau^{-1}(x) \in \displaystyle\bigcap_{1\leq p<\infty} L^p$.*

This is an easy consequence of the estimates 4.1 (see e.g. [7] or [9] for an extensive use). In particular if (2.2) is replaced by a local assumption, one can use for τ the exit time of some ball centered at x. All these remarks, combined with the usual methodology of Malliavin calculus, yield the following Theorem

Theorem 4.3 *Let $(x_t)_{t\geq s}$ be the solution of the stochastic differential equation (1.1). Assume that*

(i) $t \mapsto X_j(t, \bullet)$ is a measurable flow of $\mathcal{C}^\infty$ vector fields such that $\partial_x^\alpha X_j$, $0 \leq j \leq m$, are bounded on $[s,T] \times \mathbb{R}^d$ for any multiindex α of length at least 1;

(ii) there exists $s_0 > s$ and $R > 0$ such that for all $j = 0, \dots, m$ and all multiindex α, there exists $\beta(\alpha) > 0$ s.t. $\partial_x^\alpha X_j(\bullet, y)$ is $\beta(\alpha)$-Hölder continuous for all y with $|y - x| \leq R$ and $\displaystyle\sup_{|y-x|\leq R} \|\partial_x^\alpha X_j(\bullet, y)\|_{\beta(\alpha)} < +\infty$ on the time interval $[s, s_0]$;

(iii) dim span Lie $(X_j,\, 1 \leq j \leq m)\,(s,x) = d$.

Then for all $t \in]s,T]$, the law $\mu_{st}(x,dy)$ of x_t has a density $p_{st}(x,y)$ and $y \mapsto p_{st}(x,y) \in \mathcal{C}_b^\infty(\mathbb{R}^d)$. Furthermore there exists a constant K which only depends on the assumptions $i) - iii)$ such that for all multiindex α

$$\left\|\partial_y^\alpha p_{st}(x,\bullet)\right\|_\infty \leq c(|\alpha|)(1 \wedge (t-s))^{-K(|\alpha|+\frac{d}{2})}\, exp\left(c(|\alpha|, x)(t-s)\right). \tag{4.4}$$

where $c(\bullet)$ stands for constants depending only on their arguments.

Actually one can expect more, i.e. that $p_{st}(x, \bullet)$ belongs to the Schwartz space of rapidly decreasing functions. Recall that Malliavin integration by parts formula yields the existence of a random variable $\theta_t(\alpha)$ with moments of any order such that

$$\mathbb{E}\left[\partial^\alpha f(x_t)\right] = \mathbb{E}\left[f(x_t)\theta_t(\alpha)\right]$$

In particular if f is compactly supported by Γ and $x \notin \Gamma$, one can introduce the hitting time of Γ in the right hand side. As in [11] this furnishes the more precise estimate

$$(4.5)\quad |\partial_y^\alpha p_{st}(x,y)| \leq c_1(1\wedge(t-s))^{-K(|\alpha|+\frac{d}{2})}\,|x-y|^{-\frac{1}{2}}\,exp\,c_2(t-s)\,exp\,-c_3\Big(\frac{|x-y|^2}{1\wedge(t-s)}\Big)$$

where c_1, c_2, c_3 depend on the assumptions $i)-iii)$, c_1 and c_2 also depend on α.

As we already said these estimates are not sharp, and when $\beta > \frac{1}{2}$, the proof of Theorem 2.6 furnishes better estimates than the ones one get thanks to Theorem 3.25. As in the homogeneous case one can use the diffeomorphism property of the flow in order to study the regularity of the kernels $p_{st}(\bullet,\bullet)$. One can also use the Markov property of the time-space process (we called $\psi_t(w,s,x)$) to localize regularity in an open domain (see e.g. [11] section 1). In the latter case however, Hörmander's hypothesis has to be satisfied for all $u\in[s,t]$, since it requires the use of stopping times. Up to this restriction, results like Proposition 1.12 in [11] are still true in the non homogeneous case.

We shall now study parabolic-hypoellipticity as discussed in [11], [14] or [22], that is the regularity of solutions of

$$(4.6)\qquad \frac{\partial u}{\partial t} + L_t\,u + cu = f\ ;\ u(T,x)=g(x)\ ;\quad t\in[S,T].$$

A natural candidate for solving (4.6) is

$$(4.7)\quad u(t,x) = \mathbb{E}\left[g(x_{tT})\,exp\int_t^T c(s,x_{ts})ds\right] - \mathbb{E}\left[\int_t^T f(s,x_{ts})\,exp\left(\int_t^s c(v,x_{tv})dv\right)ds\right]$$

where x_{ts} is the solution of (1.1) which starts at x at time t.
One can easily show that if all coefficients are $\mathcal{C}^\infty$ in space, uniformly in time (as well as f, g, c) then u is actually a smooth (in space) solution of (4.6) in the sense of Schwartz distributions. It easily follows that $\frac{\partial u}{\partial t}$ is a bounded function. When everything is continuous in time, so is u and then $\frac{\partial u}{\partial t}$. Conversely by applying Ito formula, any classical solution of (4.6) is given by (4.7).

Definition 4.8 *Let U be an open subset of $\mathbb{R}^d$. The operator $\frac{\partial}{\partial t} + L_t + c$ is said to be parabolic-hypoelliptic in $]S,T[\times U$ if for any $u\in\mathcal{D}'\,(]S,T[\times U)$,*

$$(\frac{\partial}{\partial t} + L_t + c)u \in L^\infty((]S',T'[\,,\,\mathcal{C}^\infty(V))$$

implies $u\in L^\infty((]S',T'[\,,\,\mathcal{C}^\infty(V))$ for any $]S',T'[\subseteq]S,T[$ and any open subset V of U.

Recall that in the homogeneous case, Hörmander's theorem implies hypoellipticity of $\frac{\partial}{\partial t}+L_t+c$. This result was derived, using Malliavin calculus, by Kusuoka and Stroock ([21], Thm 8.6 and Thm 8.13). Another, slightly less complete, but simpler approach is proposed in [11] section 2. In the non homogeneous case, Thm 1.1 of Chaleyat-Maurel and Michel is some kind of parabolic-hypoellipticity in $]S,T[\times\mathbb{R}^d$ (actually it is a global statement, but it can be rephrased in local terms).
Namely they show that, if Hörmander's hypothesis is everywhere satisfied (with some uniform bounds on $[S,T]\times\Gamma$, Γ compact), then any solution of (4.6) which can be written $d\mu_t\times dt$ for a locally bounded flow of Radon measures μ_t, belongs to $L^\infty\big(([S,T]\,,\,\mathcal{C}^\infty(\mathbb{R}^d)\big)$.
Remark than in order to prove parabolic-hypoellipticity, it is enough to prove

(4.9) for all $h \in \mathcal{C}_0^\infty(]S,T[\times U)$, all multiindex α and all multiindex η of length $|\eta| = 0$ or 1, $|\langle \partial_t^\eta \, \partial_x^\alpha \, h \, , \, u \rangle| \le C \, \|h\|_\infty$, where C only depends on the data and the support of h.

But if u solves (4.6) in $\mathcal{D}'$, then

$$\langle \partial_t \, \partial_x^\alpha \, h \, , \, u \rangle = \langle L_t^\star \, \partial_x^\alpha \, h + c \, \partial_x^\alpha \, h \, , \, u \rangle \, - \, \langle \partial_x^\alpha \, h \, , \, f \rangle$$

hence it is enough to show (4.9) with $\eta = 0$, i.e. integrate by parts in the space direction. So we can state

Theorem 4.10 *Assume that 4.3 i) is fulfilled, 4.3 ii) holds on any compact subset of $\bar{U}$ on the full time interval $[S,T]$, and 4.3 iii) holds at each $x \in \bar{U}$. Then $\dfrac{\partial}{\partial t} + L_t + c$ is parabolic-hypoelliptic in $]S,T[\times U$.*

The proof can be done exactly as in [11] p.433-446, using smoothness of the densities $p_{st}^c(x, \bullet)$ of the law of the process exponentially killed at rate c, which follows from 4.1, 4.2 and usual tools of Malliavin calculus. We do not know how to adapt Kusuoka-Stroock method which is based on small time precise estimates.

In regard with Chaleyat-Maurel and Michel result, Theorem 4.10 holds for any distribution solution instead of measure solution, but we have to assume β-Hölder continuity in time, while these authors only need L^∞. However it is quite possible that their proof can be improved, while, as we said, our approach really needs Hölder continuity.

In view of the intricacies of sections 2 and 3, we hope it is possible to get a simpler proof of existence and regularity for p_{st}. But Malliavin calculus has shown to be a very efficient tool in studying various problems, so that Theorem 2.6 and Theorem 3.25 have their own interest. In the next section we shall briefly indicate some possible applications.

5. Further applications

5.1. Filtering

One of the object of [14] or [17] was to use Malliavin's calculus for time dependent systems in order to get smoothness of conditional laws as the ones which appear in filtering theory. Our results give the correct hypothesis for section 2 of [17]. However the best way to study conditional laws is the use of the partial Malliavin calculus introduced by Bismut and Michel [8]. The same can be done for time dependent coefficients, the only difference being that the drift X_0 is forbidden in the statement of partial Hörmander's hypotheses.

5.2. Infinitely degenerate diffusions

The terminology is due to Bell and Mohammed [6] (also see [30]). The non homogeneous case is studied in [33], but with wrong hypotheses and proofs. With the material of the present paper, [33] can be corrected.

5.3. Killed and reflected processes

The study of reflected diffusions in a domain M with non characteristic boundary ∂M or of the killed process when it exits such a domain was done, in the homogeneous case in [12] [13][10][9].

One of the main geometric tools is to perform a change of coordinates in a neighborhood of boundary points such that the transversal coordinate becomes a drifted (reflected) Brownian motion.

In the non homogeneous case, one can mimic what is done in the above papers, except this change of coordinates. Indeed it leads to use Ito formula with a time dependent function. If the coefficients are $\mathcal{C}^1$ in time, there is no problem. For less time-regular coefficients the situation is more delicate. We will not enter into details here.

5.4. Miscellaneous

We should mention various others applications like systems driven by an infinite number of B.M. as in [28] (this would correct [34]), manifold values s.d.e and so on.
In the numerous applications of Malliavin calculus, small time behaviour of heat kernels takes an important place (see [22], [2] [3] [23] [24]). We do not know what can be done for non homogeneous heat kernels in the same spirit.
Another open question is the study of the set of strict positivity of the density (see [1] or [29] for the homogeneous case).

References

1. N. Aida, S. Kusuoka, and D. Stroock. On the support of Wiener functionals. In K.D. Elworthy and N.Ikeda, editors, *Asymptotic problems in Probability theory*, number 284 in Pitman research notes in Mathematics, Sanda & Kyoto, 1990. Taniguchi international Symposium.
2. G. Ben Arous. Développement asymptotique du noyau de la chaleur hypoelliptique hors du cut-locus. *Ann. Sci. École Norm. Sup.*, 21:307–331, 1988.
3. G. Ben Arous. Développement asymptotique du noyau de la chaleur hypoelliptique sur la diagonale. *Ann. Institut Fourier*, 39:73–99, 1989.
4. V. Bally. On the connection between the Malliavin covariance matrix and Hörmander's condition. *J. Funct. Anal.*, 96:219–255, 1991.
5. D. Bell. *The Malliavin Calculus*, volume 34 of *Pitman Monographs and Surveys in Pure and Applied Math.* Longman and Wiley, 1987.
6. D. Bell and S. Mohammed. An extension of Hörmander's theorem for infinitely degenerate second order operators. *Duke Math. J.*, 78:453–476, 1995.
7. J.M. Bismut. Last exit decompositions and regularity at the boundary of transition probabilities. *Z. Wahrsch. verw. Gebiete*, 69:65–98, 1985.
8. J.M. Bismut and D. Michel. Diffusions conditionnelles I. *J. Funct. Anal.*, 44:174–211, 1981.
9. P. Cattiaux. Hypoellipticité et hypoellipticité partielle pour les diffusions avec une condition frontière. *Ann. Institut Henri Poincaré*, 22:67–112, 1986.
10. P. Cattiaux. Régularité au bord pour les densités et les densités conditionnelles d'une diffusion réfléchie hypoelliptique. *Stochastics*, 20:309–340, 1987.
11. P. Cattiaux. Calcul stochastique et opérateurs dégénérés du second ordre I. Résolvantes et théorème de Hörmander. *Bull. Sciences. Math.*, 114:421–462, 1990.
12. P. Cattiaux. Calcul stochastique et opérateurs dégénérés du second ordre II. Problème de Dirichlet. *Bull. Sciences. Math.*, 115:81–122, 1991.
13. P. Cattiaux. Stochastic calculus and degenerate boundary value problems. *Ann. Institut Fourier*, 42:541–624, 1992.
14. M. Chaleyat-Maurel and D. Michel. Hypoellipticy theorems and conditional laws. *Z. Wahrsch. verw. Gebiete*, 65:573–597, 1984.
15. M. Chen and X. Zhou. Applications of Malliavin calculus to stochastic differential equations with time-dependent coefficients. *Acta Appl. Math. Sinica*, 7(3):193–216, 1991.
16. M. Derridj. Un problème aux limites pour une classe d'opérateurs du second ordre hypoelliptiques. *Ann. Institut Fourier*, 21:99–148, 1971.
17. P. Florchinger. Malliavin calculus with time-dependent coefficients and application to non linear filtering. *Prob. Theor. and Rel. Fields*, 86:203–223, 1990.
18. L. Hörmander. Hypoelliptic second order differential equations. *Acta Math.*, 117:147–171, 1967.
19. N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes.* North-Holland, second edition, 1989.
20. I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus.* Springer-Verlag, second edition, 1997.
21. S. Kusuoka and D.W Stroock. Applications of Malliavin calculus, part II. *J. Fac. Sc. Univ. Tokyo*, 32:1–76, 1985.
22. S. Kusuoka and D.W Stroock. Applications of Malliavin calculus, part III. *J. Fac. Sc. Univ. Tokyo*, 34:391–442, 1987.
23. R. Léandre. Intégration dans la fibre associée à une diffusion dégénérée. *Prob. Theor. and Rel. Fields*, 76:341–35, 1987.

24. R. Léandre. Développement asymptotique de la densité d'une diffusion dégénérée. *Forum Mathematicum*, 4:45–75, 1992.
25. P. Malliavin. C^k-hypoellipticity with degeneracy. In A. Friedmann and M. Pinsky, editors, *Stochastic Analysis*, pages 199–214. Acad. Press, New York, 1978.
26. P. Malliavin. Stochastic calculus of variations and hypoelliptic operators. In *Conference on Stochastic Differential Equations, Kyoto 1976*, pages 195–263. Kinokuniya, Tokyo and Wiley, New York, 1978.
27. P. Malliavin. *Stochastic Analysis*, volume 313 of *Series of comprehensive studies in Mathematics*. Springer, 1997.
28. M.D. Nguyen, D. Nualart and M. Sanz. Application of the Malliavin calculus to a class of stochastic differential equations. *Prob. Theor. and Rel. Fields*, 84:549–571, 1990.
29. A. Millet and M. Sanz. A simple proof of the support theorem for diffusion processes. *Séminaire de Probabilités XXVIII*, Lecture Notes in Math. 1583:36–48, 1994.
30. Y. Morimoto. Hypoellipticity for infinitely degenerate elliptic operators. *Osaka J. math*, 24:13–35, 1987.
31. J.R. Norris. Simplified Malliavin calculus. *Séminaire de Probabilités XX*, Lecture Notes in Math. 1204:101–130, 1986.
32. D. Nualart. *The Malliavin Calculus and Related Topics*. Probability and its Application. Springer-Verlag, 1995.
33. J. Schiltz. Le théorème de Hörmander pour des opérateurs du second ordre infiniment dégénérés avec des coefficients dépendant du temps. *Stochastics*, 59:259–281, 1996.
34. J. Schiltz. Malliavin's calculus with time depending coefficients applied to a class of stochastic differential equations. *Stochastic Analysis and Appl.*, 17, 1999.
35. D.W. Stroock. Some applications of stochastic calculus and partial differential equations. In *Ecole d'été de Probabilités de Saint-Flour*, volume 976, pages 267–280. Lecture Notes in Math., 1983.
36. S. Taniguchi. Applications of Malliavin calculus to time-dependent system of heat equations. *Osaka J. Math*, 89:457–485, 1991.