Diffusions under a local strong Hörmander condition. Part I: density estimates

Vlad Bally, Lucia Caramellino, Paolo Pigato

Diffusions under a local strong Hörmander condition. Part I: density estimates

Vlad Bally*
Lucia Caramellino†
Paolo Pigato‡

July 19, 2016

Abstract

We study lower and upper bounds for the density of a diffusion process in $\mathbb{R}^n$ in a small (but not asymptotic) time, say δ. We assume that the diffusion coefficients $\sigma_1, \dots, \sigma_d$ may degenerate at the starting time 0 and point x_0 but they satisfy a strong Hörmander condition involving the first order Lie brackets. The density estimates are written in terms of a norm which accounts for the non-isotropic structure of the problem: in a small time δ, the diffusion process propagates with speed $\sqrt{\delta}$ in the direction of the diffusion vector fields σ_j and with speed $\delta = \sqrt{\delta} \times \sqrt{\delta}$ in the direction of $[\sigma_i, \sigma_j]$. In the second part of this paper, such estimates will be used in order to study lower and upper bounds for the probability that the diffusion process remains in a tube around a skeleton path up to a fixed time.

Contents

*Université Paris-Est, LAMA (UMR CNRS, UPEMLV, UPEC), MathRisk INRIA, F-77454 Marne-la-Vallée, France. Email: bally@univ-mlv.fr
†Dipartimento di Matematica, Università di Roma - Tor Vergata, Via della Ricerca Scientifica 1, I-00133 Roma, Italy. Email: caramell@mat.uniroma2.it.
‡INRIA, Villers-lès-Nancy, F-54600, France Université de Lorraine, IECL, UMR 7502, Vandoeuvre-lès-Nancy, F-54600, France paolo.pigato@inria.fr.

1 Introduction

In this paper we study bounds for the density of a diffusion process at a small time under a local strong Hörmander condition. To be more precise, let X denote the process in $\mathbb{R}^n$ solution to

$$dX_t = \sum_{j=1}^{d} \sigma_j(t, X_t) \circ dW_t^j + b(t, X_t)dt, \qquad X_0 = x_0. \tag{1.1}$$

where $W = (W^1, ..., W^d)$ is a standard Brownian motion and $\circ dW_t^j$ denotes the Stratonovich integral. We assume nice differentiability and boundedness or sublinearity properties for the diffusion coefficients b and σ_j, $j = 1, \ldots, d$, and we consider a degenerate case:

$$\dim \sigma(0, x_0) = \dim \mathrm{Span}\{\sigma_1(0, x_0), \ldots, \sigma_d(0, x_0)\} < n, \tag{1.2}$$

$\dim S$ denoting the dimension of the vector space S. Our aim is to study lower and upper bounds for the density of the solution to (1.1) at a small (but not asymptotic) time, say δ, under the following local strong Hörmander condition:

$$\mathrm{Span}\{\sigma_i(0, x_0), [\sigma_p, \sigma_j](0, x_0),\ i, p, j = 1, \ldots, d\} = \mathbb{R}^n \tag{1.3}$$

in which $[\cdot, \cdot]$ denotes the standard Lie bracket vector field. Notice that we ask for a Hörmander condition at time 0. Our estimates are written in terms of a norm which reflects the non-isotropic structure of the problem: roughly speaking, in a small time interval of length δ, the diffusion process moves with speed $\sqrt{\delta}$ in the direction of the diffusion vector fields σ_j and with speed $\delta = \sqrt{\delta} \times \sqrt{\delta}$ in the direction of $[\sigma_i, \sigma_j]$. In order to catch this behavior we introduce the following norms. Let $A_\delta(0, x_0)$ denote the $n \times d^2$ matrix

$$A_\delta(0, x_0) = [A_{1,\delta}(0, x_0), \ldots, A_{d^2,\delta}(0, x_0)]$$

where the general column $A_{l,\delta}(0, x_0)$, $l = 1, \ldots, d^2$, is defined as follows:

- for $l = (p-1)d + i$ with $p, i \in \{1, \ldots, d\}$ and $p \neq i$ then

$$A_{l,\delta}(0, x_0) = [\sqrt{\delta}\sigma_i, \sqrt{\delta}\sigma_p](0, x_0) = \delta[\sigma_i, \sigma_p](0, x_0);$$

- for $l = (p-1)d + i$ with $p, i \in \{1, \ldots, d\}$ and $p = i$ then

$$A_{l,\delta}(0, x_0) = \sqrt{\delta}\, \sigma_i(0, x_0).$$

Under (1.3), the rank of $A_\delta(0,x_0)$ is equal to n, hence the following norm is well defined:

$$|\xi|_{A_\delta(0,x_0)} = \langle (A_\delta(0,x_0)A_\delta(0,x_0)^T)^{-1}\xi,\xi\rangle^{1/2}, \quad \xi\in\mathbb{R}^n,$$

where the supscript T denotes the transpose and $\langle\cdot,\cdot\rangle$ stands for the standard scalar product. We prove in [3] that the metric given by this norm is locally equivalent with the control distance d_c (the Carathéodory distance) which is usually used in this framework. We denote by $p_\delta(x_0,\cdot)$ the density of the solution to (1.1) at time δ. Under (1.3) and assuming suitable hypotheses on the boundedness and sublinearity of the coefficients b and σ_j, $j=1,\ldots,d$ (see Assumption 2.1 for details), we prove the following result (recall $\dim\sigma(0,x_0)$ given in (1.2)):

[LOWER BOUND] there exist positive constants r,δ_*,C such that for every $\delta\le\delta_*$ and for every y with $|y-x_0-b(0,x_0)\delta|_{A_\delta(0,x_0)}\le r$ one has

$$p_\delta(x_0,y)\ge \frac{1}{C\delta^{n-\frac{\dim\sigma(0,x_0)}{2}}};$$

[UPPER BOUND] for any $p>1$, there exists a positive constant C such that for every $\delta\le 1$ and for every $y\in\mathbb{R}^n$ one has

$$p_\delta(x_0,y)\le \frac{1}{\delta^{n-\frac{\dim\sigma(0,x_0)}{2}}}\frac{C}{1+|y-x_0|^p_{A_\delta(0,x_0)}}.$$

This is stated in Theorem 2.4, where an exponential upper bound is achieved as well, provided that stronger boundedness assumptions on the diffusion coefficients hold (see Assumption 2.3).

In the context of a degenerate diffusion coefficient which fulfills a strong Hörmander condition, the main result in this direction is due to Kusuoka and Stroock. In the celebrated paper [12], they prove the following two-sided Gaussian bounds: there exists a constant $M\ge 1$ such that

$$\begin{aligned}\frac{1}{M|B_{d_c}(x_0,\delta^{1/2})|}&\exp\left(-\frac{Md_c(x_0,y)^2}{\delta}\right)\\ &\le p_\delta(x_0,y)\le \frac{M}{|B_{d_c}(x_0,\delta^{1/2})|}\exp\left(-\frac{d_c(x_0,y)^2}{M\delta}\right)\end{aligned}\tag{1.4}$$

where $\delta\in(0,1]$, $x_0,y\in\mathbb{R}^n$, $B_{d_c}(x,r)=\{y\in\mathbb{R}^n : d_c(x,y)<r\}$, d_c denoting the control (Carathéodory) distance, and $|B_{d_c}(x,r)|$ stands for the Lebesgue measure of $B_{d_c}(x,r)$. It is worth to be said that (1.4) holds under special hypotheses: in [12] it is assumed that the coefficients do not depend on the time variable and that $b(x)=\sum_{j=1}^d\alpha_i\sigma_i(x)$, with $\alpha_i\in C_b^\infty(\mathbb{R}^n)$ (i.e. the drift is generated by the vector fields of the diffusive part, which is a quite restrictive hypothesis). Other celebrated estimates for the heat kernel under strong Hörmander condition are provided in [4, 5]. The subject has also been widely studied by analytical methods - see for example [10] and [17]. We stress that these are asymptotic results, whereas we prove estimate for a finite, positive and fixed time. In [15], [8], non-isotropic norms similar to $|\cdot|_{A_\delta(0,x_0)}$ are used to provide density estimates for SDEs under Hörmander conditions of weak type. We also refer to [7], which considers the existence of the density for SDEs with time dependent coefficients, under very weak regularity assumption.

The paper [3] will follow the present work, considering related results from a "control" point of view, discussing in particular tube estimates and the connection with the control-Carathéodory distance. Tube estimates are estimates on the probability that an Itô process remains around a deterministic path up to a given time. These will be obtained from a concatenation of the short-time density estimates presented here. Then we will consider, as in [12], $b(t,x) = b(x)$ and $\sigma(t,x) = \sigma(x)$. Defining the semi distance d via: $d(x,y) < \sqrt{\delta}$ if and only if $|x-y|_{A_\delta(x)} < 1$, we will prove in [3] the local equivalence of d and d_c. This will give a rewriting of the upper/lower estimates of the density in terms of the control distance as well.

The paper is organized as follows. In Section 2 we set-up the framework and give the precise statement of our main result (Theorem 2.4). The proof is split in two sections: Section 3, which is devoted to the lower bound, and Section 4, in which we deal with the upper bound. The main tool we are going to use is given by the estimates of localized densities which have been developed in [2]. These need to use techniques from Malliavin calculus, so we briefly report in Appendix D all these arguments. But in order to set-up our program we also require some other facts, which have been collected in other appendices. First, we use a key decomposition of the solution X_δ to (1.1) at a small (but not asymptotic) time δ (see Section 3.1), and we postpone the proof in Appendix A. This decomposition allows us to work with a random variable whose law, conditional to a suitable σ-algebra, is Gaussian, and in Appendix B we study some useful support properties that are applied to our case. Moreover, since the key-decomposition brings to handle a perturbed Gaussian random variable, in Appendix C we prove density estimates via local inversion for such kind of random variables.

2 Notations and main results

We need to recall some notations. For $f, g : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^n$ we define the directional derivative (w.r.t. the space variable x) $\partial_g f(t,x) = \sum_{i=1}^n g^i(t,x)\partial_{x_i} f(t,x)$, and we recall that the Lie bracket (again w.r.t. the space variable) is defined as $[g,f](t,x) = \partial_g f(t,x) - \partial_f g(t,x)$. Let $M \in \mathcal{M}_{n\times m}$ be a matrix with full row rank. We write M^T for the transposed matrix, and MM^T is invertible. We denote by $\lambda_*(M)$ (respectively $\lambda^*(M)$) the smallest (respectively the largest) singular value of M. We recall that singular values are the square roots of the eigenvalues of MM^T, and that, when M is symmetric, singular values coincide with the absolute values of the eigenvalues of M. In particular, when M is a covariance matrix, $\lambda_*(M)$ and $\lambda^*(M)$ are the smallest and the largest eigenvalues of M.

We consider the following norm on $\mathbb{R}^n$:

$$|y|_M = \sqrt{\langle (MM^T)^{-1}y, y\rangle}. \tag{2.1}$$

Hereafter, $\alpha = (\alpha_1, ..., \alpha_k) \in \{1, ..., n\}^k$ represents a multi-index with length $|\alpha| = k$ and $\partial_x^\alpha = \partial_{x_{\alpha_1}} ... \partial_{x_{\alpha_k}}$. We allow the case $k = 0$, giving $\alpha = \emptyset$, and $\partial_x^\alpha f = f$. Finally, for given vectors $v_1, \dots, v_n \in \mathbb{R}^m$, we define $\langle v_1, \dots, v_n\rangle \subset \mathbb{R}^m$ the vector space spanned by $v_1, \dots, v_n$. Let X denote the process in $\mathbb{R}^n$ already introduced in (1.1), that is

$$dX_t = \sum_{j=1}^d \sigma_j(t, X_t) \circ dW_t^j + b(t, X_t)dt, \qquad X_0 = x_0, \tag{2.2}$$

W being a standard Brownian motion in $\mathbb{R}^d$. We suppose the diffusion coefficients fulfill the following requests:

Assumption 2.1. *There exists a constant $\kappa > 0$ such that, $\forall t \in [0,1]$, $\forall x \in \mathbb{R}^n$:*

$$\sum_{j=1}^{d} |\sigma_j(t,x)| + |b(t,x)| + \sum_{j=1}^{d} \sum_{0\leq|\alpha|\leq 2} |\partial_x^\alpha \partial_t \sigma_j(t,x)| \leq \kappa(1+|x|)$$
$$\sum_{j=1}^{d} \sum_{1\leq|\alpha|\leq 4} |\partial_x^\alpha \sigma_j(t,x)| + \sum_{1\leq|\alpha|\leq 3} |\partial_x^\alpha b(t,x)| \leq \kappa$$

Remark that Assumption 2.1 ensures the strong existence and uniqueness of the solution to (2.2). We do not assume here ellipticity but a non degeneracy of Hörmander type. In order to do this, we need to introduce the $n \times d^2$ matrix $A(t,x)$ defined as follows. We set $m = d^2$ and define the function

$$l(i,p) = (p-1)d + i \in \{1,\dots,m\}, \quad p,i \in \{1,\dots,d\}. \tag{2.3}$$

Notice that $l(i,p)$ is invertible. For $l = 1,\dots,m$, we set the (column) vector field $A_l(t,x)$ in $\mathbb{R}^n$ as follows:

$$\begin{aligned} A_l(t,x) &= [\sigma_i, \sigma_p](t,x) \quad \text{if} \quad l = l(i,p) \quad \text{with} \quad i \neq p, \\ &= \sigma_i(t,x) \quad \text{if} \quad l = l(i,p) \quad \text{with} \quad i = p \end{aligned} \tag{2.4}$$

and we set $A(t,x)$ to be the $n \times m$ matrix whose columns are given by $A_1(t,x),\dots,A_m(t,x)$:

$$A(t,x) = [A_1(t,x),\dots,A_m(t,x)]. \tag{2.5}$$

$A(t,x)$ can be interpreted as a directional matrix. We denote by $\lambda(t,x)$ the smallest singular value of $A(t,x)$, i.e.

$$\lambda(t,x)^2 = \lambda_*(A(t,x))^2 = \inf_{|\xi|=1} \sum_{i=1}^{m} \langle A_i(t,x), \xi \rangle^2 . \tag{2.6}$$

In this paper, we assume the following non degeneracy condition. We write it in a "time dependent way" because this is useful in [3], which represents the second part of the present article. In fact, we use here just $A(0,x_0)$ and $\lambda(0,x_0)$, whereas in [3] we consider $A(t,x_t)$ and $\lambda(t,x_t)$, x_t denoting a skeleton path.

Assumption 2.2. *Let x_0 denote the starting point of the diffusion X solving* (2.2)*. We suppose that*

$$\lambda(0,x_0) > 0.$$

Notice that Assumption 2.2 is actually equivalent to require that the first order Hörmander condition holds in the starting point x_0, i.e. the vector fields $\sigma_i(0,x_0)$, $[\sigma_j,\sigma_p](0,x_0)$, as $i,j,p = 1,...,d$, span the whole $\mathbb{R}^n$.

We define now the $m \times m$ diagonal scaling matrix $\mathcal{D}_\delta$ as

$$\begin{aligned} (\mathcal{D}_\delta)_{l,l} &= \delta \quad \text{if} \quad l = l(i,p) \quad \text{with} \quad i \neq p, \\ &= \sqrt{\delta} \quad \text{if} \quad l = l(i,p) \quad \text{with} \quad i = p \end{aligned}$$

and the scaled directional matrix

$$A_\delta(t,x) = A(t,x)\mathcal{D}_\delta. \tag{2.7}$$

Notice that the lth column of the matrix $A_\delta(t,x)$ is given by $\sqrt{\delta}\sigma_i(t,x)$ if $l = l(i,p)$ with $i = p$ and by $\delta[\sigma_i, \sigma_p](t,x) = [\sqrt{\delta}\sigma_i, \sqrt{\delta}\sigma_p](t,x)$ if $i \neq p$. Therefore, $A_\delta(t,x)$ is the matrix given in (2.5) when the original diffusion coefficients $\sigma_j(t,x)$, $j = 1,\ldots,d$, are replaced by $\sqrt{\delta}\sigma_j(t,x)$, $j = 1,\ldots,d$.

This matrix and the associated norm $|\cdot|_{A_\delta(0,x_0)}$ are the tools that allow us to account of the different speeds of propagation of the diffusion: $\sqrt{\delta}$ (diffusive scaling) in the direction of σ and δ in the direction of the first order Lie Brackets. In particuar, straightforward computations easily give that

$$\frac{1}{\sqrt{\delta}\lambda^*(A(t,x))}|y| \leq |y|_{A_\delta(t,x)} \leq \frac{1}{\delta\lambda_*(A(t,x))}|y|, \tag{2.8}$$

We also consider the following assumption, as a stronger version of Assumption 2.1 (morally we ask for boundedness instead of sublinearity of the coefficients, in the spirit of Kusuoka-Stroock estimates in [12]).

Assumption 2.3. *There exists a constant $\kappa > 0$ such that for every $t \in [0,1]$ and $x \in \mathbb{R}^n$ one has*

$$\sum_{0\leq|\alpha|\leq 4}\Big[\sum_{j=1}^d |\partial_x^\alpha \sigma_j(t,x)| + |\partial_x^\alpha b(t,x)| + |\partial_x^\alpha \partial_t \sigma_j(t,x)|\Big] \leq \kappa.$$

The aim of this paper is to prove the following result:

Theorem 2.4. *Let Assumption 2.1 and 2.2 hold. Let p_{X_t} denote the density of X_t, $t > 0$, with the starting condition $X_0 = x_0$. Then the following holds.*

(1) *There exist positive constants r, δ_*, C such that for every $\delta \leq \delta_*$ and for every y such that $|y - x_0 - b(0,x_0)\delta|_{A_\delta(0,x_0)} \leq r$,*

$$\frac{1}{C\delta^{n-\frac{\dim\langle\sigma(0,x_0)\rangle}{2}}} \leq p_{X_\delta}(y).$$

(2) *For any $p > 1$, there exists a positive constant C such that for every $\delta \leq 1$ and for every $y \in \mathbb{R}^n$*

$$p_{X_\delta}(y) \leq \frac{1}{\delta^{n-\frac{\dim\langle\sigma(0,x_0)\rangle}{2}}}\frac{C}{1 + |y - x_0|^p_{A_\delta(0,x_0)}}.$$

(3) *If also Assumption 2.3 holds (boundedness of coefficients) there exists a constant C such that for every $\delta \leq 1$ and for every $y \in \mathbb{R}^n$.*

$$p_{X_\delta}(y) \leq \frac{C}{\delta^{n-\frac{dim\langle\sigma(0,x_0)\rangle}{2}}}\exp\big(-\frac{1}{C}|y - x_0|_{A_\delta(0,x_0)}\big).$$

Here $\dim\langle\sigma(0,x_0)\rangle$ *denotes the dimension of the vector space spanned by* $\sigma_1(0,x_0), \ldots,$ $\sigma_d(0,x_0)$.

Remark 2.5. It might appear contradictory that the lower estimate (1) in Theorem 2.4 is centered in $x_0 + \delta b(x_0)$, whereas the upper estimates are centered in x_0. In fact, this is important only for the lower bound, the upper bounds (2) and (3) holding true either if we write $|y - x_0 - \delta b(x_0)|_{A_\delta(0,x_0)}$ or $|y - x_0|_{A_\delta(0,x_0)}$ (see next Remark 4.8).

Remark 2.6. As already mentioned, the two sided bound (1.4) by Kusuoka and Stroock [12] is proved under a strong Hörmander condition of any order, but the drift coefficient must be generated by the vector fields of the diffusive part, and the diffusion coefficients b and σ_j, $j = 1, \dots, d$, must not depend on time. Here, on the contrary, we allow for a general drift and time dependence in the coefficients, but we consider only first order Lie Brackets. Moreover, in assumption 2.1, we also relax the hypothesis of bounded coefficients. Anyways, the two estimates are strictly related, since our matrix norm is locally equivalent to the Carathéodory distance – this is proved [3] (see Section 4 therein).

Remark 2.7. Our main application is developed in [3], which is the second part of this paper and concerns tube estimates. To this aim, we are mostly interested in the diagonal estimates, that is, around $x_0 + \delta b(0, x_0)$. In particular, what we need is the precise exponent $n - \dim\langle\sigma(0, x_0)\rangle/2$, which accounts for the time-scale of the heat kernel when δ goes to zero. However, our results are not asymptotic, but hold uniformly for δ small enough. This is crucial for our application to tube estimates, and this is also a main difference with the estimates in [4, 5].

Remark 2.8. The upper bounds in (2) and (3) of Theorem 2.4 give also the tail estimates, which are exponential if we assume the boundedness of the coefficients, polynomial otherwise.

The proof of Theorem 2.4 is long, also different according to the lower or upper estimate, and we proceed by organizing two sections where such results will be separately proved. So, the lower estimate will be discussed in Section 3 and proved in Theorem 3.7, whereas Section 4 and Theorem 4.6 will be devoted to the upper estimate.

3 Lower bound

We study here the lower bound for the density of X_δ.

3.1 The key-decomposition

We start with the decomposition of the process that will allow us to produce the lower bound in short (but not asymptotic) time.
We first use a development in stochastic Taylor series of order two of the diffusion process X defined through (2.2). This gives

$$X_t = x_0 + Z_t + b(0, x_0)t + R_t \tag{3.1}$$

where

$$Z_t = \sum_{i=1}^{d} a_i W_t^i + \sum_{i,j=1}^{d} a_{i,j} \int_0^t W_s^i \circ dW_s^j \qquad \text{with } a_i = \sigma_i(0, x_0),\ a_{i,j} = \partial_{\sigma_i}\sigma_j(0, x_0) \tag{3.2}$$

and

$$R_t = \sum_{j,i=1}^d \int_0^t \int_0^s (\partial_{\sigma_i}\sigma_j(u,X_u) - \partial_{\sigma_i}\sigma_j(0,x_0)) \circ dW_u^i \circ dW_s^j \tag{3.3}$$
$$+ \sum_{i=1}^d \int_0^t \int_0^s \partial_b \sigma_i(u,X_u) du \circ dW_s^i + \sum_{i=1}^d \int_0^t \int_0^s \partial_u \sigma_j(u,X_u) du \circ dW_s^i$$
$$+ \sum_{i=1}^d \int_0^t \int_0^s \partial_{\sigma_i} b(u,X_u) \circ dW_u^i ds + \int_0^t \int_0^s \partial_b b(u,X_u) du ds.$$

Since $R_t = \mathcal{O}(t^{3/2})$, we expect the behavior of X_t and Z_t to be somehow close for small values of t. Our first goal is to give a decomposition for Z_t in (3.2). We start introducing some notation. We fix $\delta > 0$ and set

$$s_k(\delta) = \frac{k}{d}\delta, \quad k = 1, \dots, d.$$

We now consider the following random variables: for $i, k = 1, \dots, d$,

$$\Delta_k^i(\delta, W) = W^i_{s_k(\delta)} - W^i_{s_{k-1}(\delta)}, \quad \Delta_k^{i,j}(\delta, W) = \int_{s_{k-1}(\delta)}^{s_k(\delta)} (W_s^i - W^i_{s_{k-1}}) \circ dW_s^j. \tag{3.4}$$

Notice that $\Delta_k^{i,j}(\delta, W)$ is the Stratonovich integral, but for $i \neq j$ it coincides with the Itô integral. When no confusion is possible we use the short notation $s_k = s_k(\delta), \Delta_k^i = \Delta_k^i(\delta, W), \Delta_k^{i,j} = \Delta_k^{i,j}(\delta, W)$. We also denote the random vector $\Delta(\delta, W)$ in $\mathbb{R}^m$

$$\begin{aligned} \Delta_l(\delta, W) &= \Delta_p^{i,p}(\delta, W) \quad \text{if } l = l(i,p) \text{ with } i \neq p, \\ &= \Delta_p^p(\delta, W) \quad \text{if } l = l(i,p) \text{ with } i = p. \end{aligned} \tag{3.5}$$

(recall $l(i,p)$ in (2.3)). Moreover, with $\sum_{l>p}^d = \sum_{p=1}^d \sum_{l=p+1}^d$, we define

$$\begin{aligned} V(\delta, W) &= \sum_{p=1}^d \Big[\sum_{i\neq p} \Delta_p^i + \sum_{i\neq j, i\neq p, j\neq p} a_{i,j}\Delta_p^{i,j} + \sum_{l=p+1}^d \sum_{i\neq p} \sum_{j\neq l} a_{i,j} \Delta_l^j \Delta_p^i + \frac{1}{2} \sum_{i\neq p} a_{i,i} \left|\Delta_p^i\right|^2 \Big]; \\ \varepsilon_p(\delta, W) &= \sum_{l>p}^d \sum_{j\neq l} a_{p,j}\Delta_l^j + \sum_{p>l}^d \sum_{j\neq l} a_{j,p}\Delta_l^j + \sum_{j\neq p} a_{p,j}\Delta_p^j, \quad p = 1, ..., d; \\ \eta_p(\delta, W) &= \frac{1}{2} a_{p,p} \left|\Delta_p^p\right|^2 + \sum_{l>p}^d a_{p,l}\Delta_l^l \Delta_p^p + \Delta_p^p \varepsilon_p(\delta, W), \quad p = 1, ..., d. \end{aligned} \tag{3.6}$$

We have the following decomposition:

Lemma 3.1. *Let $\Delta(\delta, W)$ and $A(0, x_0)$ be given in (3.5) and (2.5) respectively. One has*

$$Z_\delta = V(\delta, W) + A(0, x_0)\Delta(\delta, W) + \eta(\delta, W), \tag{3.7}$$

where $V(\delta, W)$ is given in (3.6) and $\eta(\delta, W) = \sum_{p=1}^d \eta_p(\delta, W)$, $\eta_p(\delta, W)$ being given in (3.6).

The proof of Lemma 3.1 is quite long, so it is postponed to Appendix A.

Remark 3.2. The reason of this decomposition is the following. We split the time interval $(0,\delta)$ in d sub intervals of length δ/d. We also split the Brownian motion in corresponding increments: $(W_s^p - W_{s_{k-1}}^p)_{s_{k-1}\leq s\leq s_k}, p = 1,...,d$. Let us fix p. For $s \in (s_{p-1}, s_p)$ we have the processes $(W_s^i - W_{s_{p-1}}^i)_{s_{p-1}\leq s\leq s_p}, i = 1,...,d$. Our idea is to settle a calculus which is based on W^p and to take conditional expectation with respect to $W^i, i \neq p$. So $(W_s^i - W_{s_{p-1}}^i)_{s_{p-1}\leq s\leq s_p}, i \neq p$ will appear as parameters (or controls) which we may choose in an appropriate way. The random variables on which the calculus is based are $\Delta_p^p = W_{s_p}^p - W_{s_{p-1}}^p$ and $\Delta_p^{i,p} = \int_{s_{p-1}}^{s_p}(W_s^i - W_{s_{p-1}}^i)dW_s^p, j \neq p$. These are the r.v. that we have emphasized in the decomposition of Z_δ. Notice that, conditionally to the controls $(W_s^i - W_{s_{p-1}}^i)_{s_{p-1}\leq s\leq s_p}, i \neq p$, this is a centered Gaussian vector and, under appropriate hypothesis on the controls this Gaussian vector is non degenerate (we treat in section B the problem of the choice of the controls). In order to handle the term $\Delta_p^{p,i} = \int_{s_{p-1}}^{s_p}(W_s^p - W_{s_{p-1}}^p)dW_s^i$. we use the identity $\Delta_p^{p,i} = \Delta_p^i\Delta_p^p - \Delta_p^{i,p}$.

We now emphasize the scaling in δ in the random vector $\Delta(\delta, W)$. We define $B_t = \delta^{-1/2}W_{t\delta}$ and denote

$$\begin{aligned}\Theta_l = \frac{1}{\delta}\Delta_p^{i,p} &= \int_{\frac{p-1}{d}}^{\frac{p}{d}}(B_s^i - B_{\frac{p-1}{d}}^i)dB_s^p \quad \text{if } l = l(i,p) \text{ with } i \neq p,\\ &= \frac{1}{\sqrt{\delta}}\Delta_p^p = B_{\frac{p}{d}}^p - B_{\frac{p-1}{d}}^p \quad \text{if } l = l(i,p) \text{ with } i = p,\end{aligned} \tag{3.8}$$

$l(i,p)$ being given in (2.3). For $p = 1,...,d$ we denote with $\Theta_{(p)}$ the pth block of Θ with length d, that is

$$\Theta_{(p)} = (\Theta_{(p-1)d+1}, ..., \Theta_{pd}),$$

so that $\Theta = (\Theta_{(1)}, \ldots, \Theta_{(d)})$. We will also denote

$$l(p) = l(p,p) = (p-1)d + p \quad \text{and} \quad \Theta_{l(p)} = \frac{1}{\sqrt{\delta}}\Delta_p^p. \tag{3.9}$$

Consider now the σ field

$$\mathcal{G} := \sigma(W_s^j - W_{s_{p-1}(\delta)}^j, s_{p-1}(\delta) \leq s \leq s_p(\delta), p = 1, ...d, j \neq p). \tag{3.10}$$

Then conditionally to $\mathcal{G}$ the random variables $\Theta_{(p)}, p = 1,...,d$ are independent centered Gaussian d dimensional vectors and the covariance matrix Q_p of $\Theta_{(p)}$ is given by

$$\begin{aligned}Q_p^{p,j} &= Q_p^{j,p} = \int_{\frac{p-1}{d}}^{\frac{p}{d}}\left(B_s^j - B_{\frac{p-1}{d}}^j\right)ds, \quad j \neq p,\\ Q_p^{i,j} &= \int_{\frac{p-1}{d}}^{\frac{p}{d}}\left(B_s^j - B_{\frac{p-1}{d}}^j\right)\left(B_s^i - B_{\frac{p-1}{d}}^i\right)ds, \quad j \neq p, i \neq p,\\ Q_p^{p,p} &= \tfrac{1}{d}.\end{aligned} \tag{3.11}$$

It is easy to see that $\det Q_p \neq 0$ almost surely. It follows that conditionally to $\mathcal{G}$ the random variable $\Theta = (\Theta_{(1)}, ..., \Theta_{(d)})$ is a centered $m = d^2$ dimensional Gaussian vector. Its covariance matrix Q is a block-diagonal matrix built with Q_p, $p=1,\ldots,d$:

$$Q=\begin{pmatrix} Q_1 & & \\ & \ddots & \\ & & Q_d \end{pmatrix} \tag{3.12}$$

In particular $\det Q=\prod_{p=1}^d \det Q_p\neq 0$ almost surely, and $\lambda_*(Q)=\min_{p=1,\ldots,d}\lambda_*(Q_p)$. We also have $\lambda^*(Q)=\max_{p=1,\ldots,d}\lambda^*(Q_p)$. We will need to work on subsets where we have a quantitative control of this quantities, so we will come back soon on these covariance matrices. But let us show now how one can rewrite decomposition (3.7) in terms of the random vector Θ. As a consequence, the scaled matrix $A_\delta=A_\delta(0,x_0)$ in (2.7) will appear. We denote by $A_\delta^i\in\mathbb{R}^m, i=1,...,n$ the rows of the matrix A_δ. We also denote $S=\langle A_\delta^1,...,A_\delta^n\rangle\subset\mathbb{R}^m$ and $S^\perp$ its orthogonal. Under Assumption 2.2 the columns of A_δ span $\mathbb{R}^n$ so the rows $A_\delta^1,...,A_\delta^n$ are linearly independent and $S^\perp$ has dimension $m-n$. We take $\Gamma_\delta^i, i=n+1,...,m$ to be an orthonormal basis in $S^\perp$ and we denote $\Gamma_\delta^i=A_\delta^i(0,x_0)$ for $i=1,...,n$. We also denote $\underline{\Gamma}_\delta$ the $(m-n)\times m$ matrix with rows $\Gamma_\delta^i, i=n+1,\ldots,m$. Finally we denote by Γ_δ the $m\times m$ dimensional matrix with rows $\Gamma_\delta^i, i=1,...,m$. Notice that

$$\Gamma_\delta\Gamma_\delta^T=\begin{pmatrix} A_\delta A_\delta^T(0,x_0) & 0 \\ 0 & \mathrm{Id}_{m-n}\end{pmatrix} \tag{3.13}$$

where Id_{m-n} is the identity matrix in $\mathbb{R}^{m-n}$. It follows that for a point $y=(y_{(1)},y_{(2)})\in\mathbb{R}^m$ with $y_{(1)}\in\mathbb{R}^n, y_{(2)}\in\mathbb{R}^{m-n}$ we have

$$|y|_{\Gamma_\delta}^2=\left|y_{(1)}\right|_{A_\delta(0,x_0)}^2+\left|y_{(2)}\right|^2 \tag{3.14}$$

where we recall that $|y|_{\Gamma_\delta}^2=\left\langle(\Gamma_\delta\Gamma_\delta^T)^{-1}y,y\right\rangle$. For $a\in\mathbb{R}^m$ we define the immersion

$$J_a:\mathbb{R}^n\to\mathbb{R}^m,\quad (J_a(z))_i=z_i, i=1,...,n\quad\text{and}\quad (J_a(z))_i=\left\langle\Gamma_\delta^i,a\right\rangle, i=n+1,...,m. \tag{3.15}$$

In particular $J_0(z)=(z,0,...,0)$ and

$$|J_0z|_{\Gamma_\delta}=|z|_{A_\delta(0,x_0)}\,. \tag{3.16}$$

Finally we denote

$$\begin{aligned} V_\omega &= V(\delta,W) \\ \eta_\omega(\Theta) &= \sum_{p=1}^d\left(\frac{a_{p,p}}{2}\delta\Theta_{l(p)}^2+\delta^{1/2}\Theta_{l(p)}\varepsilon_p(\delta,W)+\sum_{q>p}^d a_{p,q}\delta\Theta_{l(q)}\Theta_{l(p)}\right)\end{aligned} \tag{3.17}$$

where $V(\delta,W)$ and $\varepsilon_p(\delta,W)$ are defined in (3.6) and $\Theta_{l(p)}$ is given in (3.9). We notice that $\eta_\omega(\Theta)=\sum_{p=1}^d\eta_p(\delta,W)$, $\eta_p(\delta,W)$ being defined in (3.6). We also remark that both $V(\delta,W)$ and $\varepsilon_p(\delta,W)$ are $\mathcal{G}$-measurable, so (3.17) stresses a dependence on ω which is $\mathcal{G}$-measurable and a dependence on the random vector Θ whose conditional law w.r.t. $\mathcal{G}$ is Gaussian.
Now the decomposition (3.7) may be written as

$$Z_\delta=V_\omega+A_\delta(0,x_0)\Theta+\eta_\omega(\Theta).$$

We embed this relation in $\mathbb{R}^m$ and obtain

$$J_\Theta(Z_\delta) = J_0(V_\omega) + \Gamma_\delta \Theta + J_0(\eta_\omega(\Theta)).$$

We now multiply with Γ_δ^{-1}: setting

$$\widetilde{Z} = \Gamma_\delta^{-1} J_\Theta(Z_\delta), \quad \widetilde{V}_\omega = \Gamma_\delta^{-1} J_0(V_\omega), \quad \widetilde{\eta}_\omega(\Theta) = \Gamma_\delta^{-1} J_0(\eta_\omega(\Theta)) \tag{3.18}$$

and

$$G = \Theta + \widetilde{\eta}_\omega(\Theta), \tag{3.19}$$

we get

$$\widetilde{Z} = \widetilde{V}_\omega + G. \tag{3.20}$$

Notice that, conditionally to $\mathcal{G}$, $\widetilde{Z}$ is a translation of the random variable $G = \Theta + \widetilde{\eta}_\omega(\Theta)$ which is a perturbation of a centred Gaussian random variable. Thanks to this fact, we can to apply the results in Appendix C: we use a local inversion argument in order to give bounds for the conditional density of $\widetilde{Z}$, which will be used in order to get bounds for the non conditional density. As a consequence, we will get density estimates for Z_δ.

3.2 Localized density for the principal term $\widetilde{Z}$

We study here the density of $\widetilde{Z}$ in (3.20), "around" (that is, localized on) a suitable set of Brownian trajectories, where we have a quantitative control on the "non-degeneracy" (conditionally to $\mathcal{G}$) of the main Gaussian Θ.
We denote

$$q_p(B) = \sum_{j\neq p} \left| B^j_{\frac{p}{d}} - B^j_{\frac{p-1}{d}} \right| + \sum_{j\neq p, i\neq p, i\neq j} \left| \int_{\frac{p-1}{d}}^{\frac{p}{d}} (B^j_s - B^j_{\frac{i-1}{d}}) dB^i_s \right|. \tag{3.21}$$

For fixed $\varepsilon, \rho > 0$, we define

$$\begin{aligned} &\Lambda_{\rho,\varepsilon,p} = \Big\{ \det Q_p \geq \varepsilon^\rho, \sup_{\frac{p-1}{d} \leq t \leq \frac{p}{d}} \textstyle\sum_{j\neq p} |B^j_t - B^j_{\frac{p-1}{d}}| \leq \varepsilon^{-\rho}, q_p(B) \leq \varepsilon \Big\}, \ p = 1, \dots, d \\ &\Lambda_{\rho,\varepsilon} = \cap_{p=1}^d \Lambda_{\rho,\varepsilon,p}. \end{aligned} \tag{3.22}$$

Notice that $\Lambda_{\rho,\varepsilon,p} \in \mathcal{G}$ for every $p = 1, \dots, d$. By using some results in Appendix B, we get the following.

Lemma 3.3. *Let $\Lambda_{\rho,\varepsilon}$ be as in (3.22). There exist c and ε_* such that for every $\varepsilon \leq \varepsilon_*$ one has*

$$\mathbb{P}\big(\Lambda_{\rho,\varepsilon}\big) \geq c \times \varepsilon^{\frac{1}{2}m(d+1)}. \tag{3.23}$$

Proof. We apply here Proposition B.3. Let $p \in \{1, \dots d\}$ be fixed and consider the Brownian motion $\widehat{B}_t = \sqrt{d}(B_{\frac{p-1+t}{d}} - B_{\frac{p-1}{d}})$. Let $Q(\widehat{B})$ be the matrix in (B.1). Up to a permutation of the components of $\widehat{B}$, we easily get $Q^{p,p}(\widehat{B}) = d \times Q_p^{p,p}$, $Q^{p,j}(\widehat{B}) = d^{3/2} \times Q_p^{p,j}$ for $j \neq p$, $Q^{i,j}(\widehat{B}) = d^2 \times Q_p^{i,j}$ for $i \neq p$ and $j \neq p$. Therefore,

$$\det Q_p = d^{2d-1} \det Q(\widehat{B}) \geq \det Q(\widehat{B}).$$

Let now $q(\widehat{B})$ be the quantity defined in (B.3). With $q_p(B)$ as in (3.21), it easily follows that

$$q_p(B) \leq q(\widehat{B}).$$

Moreover, $\sup_{t\leq 1}|\widehat{B}_t| = \sqrt{d}\sup_{\frac{p-1}{d}\leq s\leq\frac{p}{d}}|B_s - B_{\frac{p-1}{d}}| \geq \sup_{\frac{p-1}{d}\leq s\leq\frac{p}{d}}|B_s - B_{\frac{p-1}{d}}|$. As a consequence, with $\Upsilon_{\rho,\varepsilon}$ the set defined in (B.4), we get

$$\Upsilon_{\rho,\varepsilon}(\widehat{B}) \subset \Lambda_{\rho,\varepsilon,p}$$

and by using (B.5), we may find some constants c and ε_* such that $\mathbb{P}(\Lambda_{\rho,\varepsilon,p}) \geq c\varepsilon^{\frac{1}{2}d(d+1)}$, for $\varepsilon \leq \varepsilon_*$. This holds for every p. Since $\Lambda_{\rho,\varepsilon} = \cap_{p=1}^d \Lambda_{\rho,\varepsilon,p}$, by using the independence property we get (3.23). $\square$

Let Q be the matrix in (3.12). On the set $\Lambda_{\rho,\varepsilon} \in \mathcal{G}$ we have $\det Q = \prod_{p=1}^d \det Q_p \geq \varepsilon^{d\rho}$. Remark that

$$\frac{\lambda^*(Q)}{\sqrt{m}} \leq |Q|_l := \left(\frac{1}{m}\sum_{1\leq i,j\leq m} Q_{i,j}^2\right)^{1/2} \leq \lambda^*(Q). \tag{3.24}$$

For $a > 0$ we introduce the following function,

$$\psi_a(x) = 1_{|x|\leq a} + \exp\left(1 - \frac{a^2}{a^2 - (x-a)^2}\right) 1_{a<|x|<2a},$$

which is a mollified version of $1_{[0,a]}$. We can now define our *localization variables.*

$$\widetilde{U}_\varepsilon = (\psi_{a_1}(1/\det Q))\psi_{a_2}(|Q|_l)\psi_{a_3}(q(B)), \quad \text{with} \quad a_1 = \varepsilon^{-d\rho},\ a_2 = \varepsilon^{-2\rho},\ a_3 = d\varepsilon \tag{3.25}$$

in which we have set

$$q(B) = \sum_{p=1}^d q_p(B).$$

Remark that $\widetilde{U}_\varepsilon$ is measurable w.r.t. $\mathcal{G}$. The following inclusions hold: for every ε small enough,

$$\Lambda_{\rho,\varepsilon} \subset \left\{\det Q \geq \varepsilon^{d\rho}, |Q|_l \leq \varepsilon^{-2\rho}, q(B) \leq d\varepsilon\right\} = \{\widetilde{U}_\varepsilon = 1\} \subset \{\widetilde{U}_\varepsilon \neq 0\}. \tag{3.26}$$

We can consider $\widetilde{U}_\varepsilon$ as a smooth version of the indicator function of $\Lambda_{\rho,\varepsilon}$. We also define, for fixed $r > 0$,

$$\bar{U}_r = \prod_{i=1}^n \psi_r(\Theta_i). \tag{3.27}$$

In order to state a lower estimate for the (localized) density of $\widetilde{Z}$ in (3.20), we define the following set of constants:

$$\mathcal{C} = \left\{C > 0 \,:\, C = \exp\left(c\Big(\frac{\kappa}{\lambda(0,x_0)}\Big)^q\right),\ \exists\, c,q > 0\right\} \tag{3.28}$$

and we set

$$1/\mathcal{C} = \{C > 0 \,:\, 1/C \in \mathcal{C}\}.$$

Lemma 3.4. *Suppose Assumption 2.1 and 2.2 both hold. Let* $U_{\varepsilon,r} = \widetilde{U}_\varepsilon \bar{U}_r$*,* $\widetilde{U}_\varepsilon$ *and* $\bar{U}_r$ *being defined in* (3.25) *and* (3.27) *respectively, with* $\rho = \frac{1}{8m}$*. Set* $d\mathbb{P}_{U_{\varepsilon,r}} = U_{\varepsilon,r} d\mathbb{P}$ *and let* $p_{\widetilde{Z},U_{\varepsilon,r}}$ *denote the density of* $\widetilde{Z}$ *in* (3.20) *when we endow* Ω *with the measure* $\mathbb{P}_{U_{\varepsilon,r}}$*. Then there exist* $C \in \mathcal{C}$*,* $\varepsilon, r \in 1/\mathcal{C}$ *such that for* $|z| \le r/2$*,*

$$p_{\widetilde{Z},U_{\varepsilon,r}}(z) \ge \frac{1}{C}. \tag{3.29}$$

Proof. STEP 1: lower bound for the localized conditional density given $\mathcal{G}$.
Let $p_{\widetilde{Z},\bar{U}_r|\mathcal{G}}$ denote the localized density w.r.t. the localization $\bar{U}_r$ of $\widetilde{Z}$ conditioned to $\mathcal{G}$, i.e.

$$\mathbb{E}[f(\widetilde{Z})\bar{U}_r|\mathcal{G}] = \int f(z) p_{\widetilde{Z},\bar{U}_r|\mathcal{G}}(z)dz, \tag{3.30}$$

for f positive, measurable, with support included in $B(0, r/2)$. We start proving that there exist $C \in \mathcal{C}$, $\varepsilon, r \in 1/\mathcal{C}$ such that, on $\widetilde{U}_\varepsilon \neq 0$, for $|z| \le r/2$

$$p_{\widetilde{Z},\bar{U}_r|\mathcal{G}}(z) \ge \frac{1}{C}.$$

We recall (3.20): $\widetilde{Z} = \widetilde{V}_\omega + \Theta + \widetilde{\eta}_\omega(\Theta)$, where $\omega \mapsto \widetilde{V}_\omega$ and $\omega \mapsto \widetilde{\eta}_\omega(\cdot)$ are both $\mathcal{G}$-measurable and the conditional law of Θ given $\mathcal{G}$ is Gaussian. This allows us to use the results in Appendix C. In particular, we are interested in working on the set $\{\widetilde{U}_\varepsilon \neq 0\} \in \mathcal{G}$, so one has to keep in mind that $\omega \in \{\widetilde{U}_\varepsilon \neq 0\}$.
On $\widetilde{U}_\varepsilon \neq 0$, by (3.25) and (3.24) one has $\lambda^*(Q) \le 2\sqrt{m}\varepsilon^{-2\rho}$, and

$$\frac{\varepsilon^{d\rho}}{2} \le \det Q \le \lambda_*(Q)\lambda^*(Q)^{m-1} \le \lambda_*(Q)(2\sqrt{m})^{m-1}\varepsilon^{-2\rho(m-1)},$$

and this gives $\lambda_*(Q) \ge \frac{\varepsilon^{3m\rho}}{(2\sqrt{m})^m}$. So, fixing $\rho = 1/(8m)$, for $\varepsilon \le \varepsilon^*$,

$$\frac{1}{16m^2}\frac{\lambda_*(Q)}{\lambda^*(Q)} \ge C_m \varepsilon^{3m\rho+2\rho} \ge \varepsilon. \tag{3.31}$$

To apply (C.8) to $G = \Theta + \widetilde{\eta}_\omega(\Theta)$ we need to check the hypothesis of Lemma C.3. We are going to use the notation of Appendix C, in particular for $c_*(\widetilde{\eta}_\omega, r)$ in (C.5) and $c_i(\widetilde{\eta}_\omega)$, $i = 2, 3$, in (C.1). Recall that $\widetilde{\eta}_\omega$ is defined in (3.18) through η_ω given in (3.17). Since the third order derivatives of η_ω are null, we have $c_3(\widetilde{\eta}_\omega) = 0$. Also, for $i = l(p)$ and $j = l(q)$ we have $\partial_{i,j}\eta_\omega(\Theta) = \delta a_{ij}$, otherwise we get $\partial_{i,j}\eta_\omega(\Theta) = 0$. So $|\partial_{i,j}\eta_\omega(\Theta)| \le \delta \sum_{i,j} |a_{i,j}|$. Using (2.8) we obtain

$$|\partial_{i,j}\widetilde{\eta}_\omega(\Theta)| = |J_0(\partial_{i,j}\eta_\omega(\Theta))|_{\Gamma_\delta} = |\partial_{i,j}\eta_\omega(\Theta)|_{A_\delta} \le \frac{\sum_{i,j}|a_{i,j}|}{\lambda(0, x_0)} \le C \in \mathcal{C}.$$

So, with h_{η_ω} as in (C.2), we get

$$h_{\eta_\omega} = \frac{1}{16m^2(c_2(\widetilde{\eta}_\omega) + \sqrt{c_3(\widetilde{\eta}_\omega)})} \ge \frac{1}{C_1}, \quad \exists C_1 \in \mathcal{C} \tag{3.32}$$

We compute now the first order derivatives. For $j \notin \{l(p) : p = 1, ..., d\}$ we have $\partial_j \eta_\omega = 0$ and for $j = l(p)$ we have

$$\partial_j \eta_\omega(\Theta) = \delta \sum_{q=p}^{d} a_{p\wedge q, p\vee q} \Theta_{l(q)} + \sqrt{\delta}\varepsilon_j(\delta, W).$$

So, as above, we obtain $|\partial_j \widetilde{\eta}_\omega(\Theta)| \leq C(|\Theta| + |\varepsilon_j(\delta, W)|/\sqrt{\delta})$. Remark now that on $\{\bar{U}_r \neq 0\}$ we have $|\Theta| \leq Cr$, and on $\{\widetilde{U}_\varepsilon \neq 0\}$ we have $q(B) \leq 2d\varepsilon$, so

$$\sum_{j=1}^{d} |\varepsilon_j(\delta, W)| \leq C\sqrt{\delta} q(B) \leq C\sqrt{\delta}\varepsilon. \tag{3.33}$$

Therefore

$$c_*(\widetilde{\eta}_\omega, 16r) \leq C_2(r + \varepsilon), \quad \exists C_2 \in \mathcal{C}. \tag{3.34}$$

We also consider the following estimate of $|\widetilde{V}_\omega| = |V_\omega|_{A_\delta}$. First, we rewrite V_ω as follows:

$$V_\omega = \sum_p a_p \mu_p(\delta, W) + \sum_p \psi_p(\delta, W), \quad \text{with}$$

$$\mu_p(\delta, W) = \sum_{i\neq p} \Delta_i^p \text{ and } \psi_p(\delta, W) = \sum_{i\neq j, i\neq p, j\neq p} a_{i,j}\Delta_p^{i,j} + \sum_{l=p+1}^{d} \sum_{i\neq p} \sum_{j\neq l} a_{i,j}\Delta_l^j \Delta_p^i + \frac{1}{2}\sum_{i\neq p} a_{i,i} \left|\Delta_p^i\right|^2$$

Using again (2.8) we have

$$\Big|\sum_{p=1}^{d} a_p \mu_p(\delta, W)\Big|_{A_\delta} = \frac{1}{\sqrt{\delta}} \Big|A_\delta J_0\Big(\sum_{p=1}^{d} \mu_p(\delta, W)\Big)\Big|_{A_\delta} \leq \sum_{p=1}^{d} \frac{1}{\sqrt{\delta}} |\mu_p(\delta, W))| \leq Cq(B)$$

and

$$|\psi(\delta, W)|_{A_\delta} \leq \frac{|\psi(\delta, W)|}{\delta\sqrt{\lambda(0, x_0)}} \leq Cq(B).$$

Since $\omega \in \{\widetilde{U}_\varepsilon \neq 0\}$ we get

$$|\widetilde{V}_\omega| \leq Cq(B) \leq C_3\varepsilon, \quad \exists C_3 \in \mathcal{C}. \tag{3.35}$$

We consider (3.35), and fix $\frac{r}{\varepsilon} = 2C_3 \in \mathcal{C}$, so $|\widetilde{V}_\omega| \leq r/2$. Then we consider (3.34) and we obtain

$$c_*(\widetilde{\eta}_\omega, 4r) \leq C_2(2C_3 + 1)\varepsilon \leq \varepsilon^{1/2}, \quad \text{for } \varepsilon \leq \frac{1}{(4C_2C_3)^2} \in \frac{1}{\mathcal{C}}.$$

Moreover, looking at (3.32)

$$r = 2C_3\varepsilon \leq \frac{1}{C_1} \quad \text{for } \varepsilon \leq \frac{1}{2C_1C_3} \in \frac{1}{\mathcal{C}}.$$

So, with

$$\varepsilon = \varepsilon^* \wedge \frac{1}{(4C_2C_3)^2} \wedge \frac{1}{2C_1C_3} \in \frac{1}{\mathcal{C}},$$

and $r = 2C_3\varepsilon$ we have

$$|\widetilde{V}_\omega| \leq \frac{r}{2}, \qquad c_*(\widetilde{\eta}_\omega, 4r) \leq \varepsilon^{1/2}, \qquad r \leq \frac{1}{C_1}.$$

Now, by using also (3.31) and (3.32), it follows that (C.6) holds, and we can apply Lemma C.3. We obtain

$$\frac{1}{K \det Q^{1/2}} \exp\left(-\frac{K}{\lambda_*(Q)}|z|^2\right) \leq p_{G,\bar{U}_r|\mathcal{G}}(z)$$

for $|z| \leq r$, where K does not depend on σ, b. Remark that, using $\lambda_*(Q) \geq \frac{\varepsilon^{3m\rho}}{(2\sqrt{m})^m}$, $\rho = \frac{1}{8m}$, $r/\varepsilon = 2C_1$ and $\varepsilon \leq 1/(4C_2C_1)^2$,

$$\frac{|z|^2}{\lambda_*(Q)} \leq \frac{(2\sqrt{m})^m r^2}{\varepsilon^{3m\rho}} \leq (2\sqrt{m})^m \frac{r^2}{\varepsilon} \leq (2\sqrt{m})^m \frac{r^2}{\varepsilon^2}\varepsilon \leq (2\sqrt{m})^m (2C_1)^2 \varepsilon \leq \bar{K} \tag{3.36}$$

where $\bar{K}$ does not depend on σ, b. Therefore $p_{G,\bar{U}_r|\mathcal{G}}(z) \geq \frac{1}{C}$, for $|z| \leq r$, for some $C \in \mathcal{C}$, on $\widetilde{U}_\varepsilon \neq 0$. Now, by recalling that $|\widetilde{V}_\omega| \leq r/2$ and (3.20), we have

$$p_{\widetilde{Z},\bar{U}_r|\mathcal{G}}(z) \geq \frac{1}{C}, \quad \text{for } |z| \leq r/2 \text{ on the set } \{\widetilde{U}_\varepsilon \neq 0\}. \tag{3.37}$$

STEP 2: we get rid of the conditioning on $\mathcal{G}$, to have non-conditional bound for $p_{\widetilde{Z},U_{\varepsilon,r}}$. Since $\widetilde{U}_\varepsilon$ is $\mathcal{G}$ measurable, for every non-negative and measurable function f with support included in $B(0, r/2)$ we have

$$\mathbb{E}(f(\widetilde{Z})U_{\varepsilon,r}) = \mathbb{E}\big(\widetilde{U}_\varepsilon \mathbb{E}(f(\widetilde{Z})\bar{U}_r|\mathcal{G})\big).$$

By (3.30) and (3.37), we obtain

$$\mathbb{E}(f(\widetilde{Z})U_{\varepsilon,r}) \geq \frac{1}{C}\mathbb{E}(\widetilde{U}_\varepsilon)\int f(z)dz$$

Since $\Lambda_{\rho,\varepsilon} \subset \{\widetilde{U}_\varepsilon = 1\}$, $\mathbb{E}(\widetilde{U}_\varepsilon) \geq \mathbb{P}(\Lambda_{\rho,\varepsilon})$, so by using (3.23) and $\varepsilon \in 1/\mathcal{C}$ we finally get that $\mathbb{E}(\widetilde{U}_\varepsilon) \geq \frac{1}{C}$, so (3.29) is proved. $\square$

3.3 Lower bound for the density of X_δ

We study here a lower bound for the density of X_δ, X being the solution to (2.2). Recall decomposition (3.1):

$$X_\delta - x_0 - b(0, x_0)\delta = Z_\delta + R_\delta.$$

Our aim is to "transfer" the lower bound for $\widetilde{Z} = \Gamma_\delta^{-1} J_\Theta(Z_\delta)$ already studied in Lemma 3.4 to a lower bound for X_δ. In order to set up this program, we use results on the distance between probability densities which have been developed in [1]. In particular, we are going to use now Malliavin calculus. Appendix D is devoted to a recall of all the results and notation the present section is based on. In particular, we denote with D the Malliavin derivative with respect to W, the Brownian motion driving the original equation (2.2).
But first of all, we need some properties of the matrix Γ_δ, which can be resumed as follows. We set $SO(d)$ the set of the $d \times d$ orthogonal matrices and we denote with Id_d the $d \times d$ identity matrix.

Lemma 3.5. *Set for simplicity $A_\delta = A_\delta(0,x_0)$ and let Γ_δ be as in* (3.13)*. There exist $\mathcal{U} \in SO(n)$, $\underline{\mathcal{U}} \in SO(m-n)$ and $\mathcal{V} \in SO(m)$ such that*

$$\Gamma_\delta = \begin{pmatrix} \mathcal{U} & 0 \\ 0^T & \underline{\mathcal{U}} \end{pmatrix} \begin{pmatrix} \bar\Sigma & 0 \\ 0^T & \mathrm{Id}_{m-n} \end{pmatrix} \mathcal{V}^T$$

where 0 *denotes a null $n \times (m-n)$ matrix and $\bar\Sigma = \bar\Sigma = \mathrm{Diag}(\lambda_1(A_\delta), \dots, \lambda_n(A_\delta))$, $\lambda_i(A_\delta)$, $i = 1, \dots, n$, being the singular values of A_δ (which are strictly positive because A_δ has full rank).*

Proof. We recall that

$$\Gamma_\delta = \begin{pmatrix} A_\delta \\ \underline{\Gamma_\delta} \end{pmatrix},$$

where $\underline{\Gamma_\delta}$ is a $(m-n) \times n$ matrix whose rows are vectors of $\mathbb{R}^m$ which are orthonormal and orthogonal with the rows of A_δ. We take a singular value decomposition for A_δ and for $\underline{\Gamma_\delta}$. So, there exist $\mathcal{U} \in SO(n)$ and $\bar{\mathcal{V}} \in SO(m)$ such that

$$A_\delta = \mathcal{U}(\bar\Sigma \ 0)\bar{\mathcal{V}}^T,$$

0 denoting the $n \times (m-n)$ null matrix. Similarly, there exist $\underline{\mathcal{U}} \in SO(m-n)$ and $\underline{\mathcal{V}} \in SO(m)$ such that

$$\underline{\Gamma_\delta} = \underline{\mathcal{U}}(0^T \ \mathrm{Id}_{m-n})\underline{\mathcal{V}}^T,$$

the diagonal matrix being Id_{m-n} because the rows of $\underline{\Gamma_\delta}$ are orthonormal. Therefore, we get

$$\Gamma_\delta = \begin{pmatrix} \mathcal{U} & 0 \\ 0^T & \underline{\mathcal{U}} \end{pmatrix} \begin{pmatrix} \bar\Sigma & 0 \\ 0^T & \mathrm{Id}_{m-n} \end{pmatrix} \mathcal{V}^T$$

where $\mathcal{V}$ is a $m \times m$ matrix whose first n columns are given by the first n columns of $\overline{\mathcal{V}}$ and the remaining $m-n$ columns are given by the last $m-n$ columns of $\underline{\mathcal{V}}$. Moreover, since each row of A_δ is orthogonal to any row of $\underline{\Gamma_\delta}$, it immediately follows that all columns of $\mathcal{V}$ are orthogonal. This proves that $\mathcal{V} \in SO(m)$, and the statement follows. ☐

Then we have

Lemma 3.6. *Suppose Assumption 2.1 and 2.2 both hold. Let $U_{\varepsilon,r}$ denote the localization in Lemma 3.4 and let $\mathcal{U}$ and $\bar\Sigma$ be the matrices in Lemma 3.5. Set*

$$\alpha = \mathcal{U}\bar\Sigma \quad \textit{and} \quad \widehat{X}_\delta = \alpha^{-1}(X_\delta - x_0 - b(0,x_0)\delta).$$

Then there exist $C \in \mathcal{C}$, $\delta_*, r \in 1/\mathcal{C}$ such that for $\delta \le \delta_*$, $|z| \le r/2$,

$$p_{\widehat{X}_\delta, U_{\varepsilon,r}}(z) \ge \frac{1}{C}, \tag{3.38}$$

$p_{\widehat{X}_\delta, U_{\varepsilon,r}}$ denoting the density of $\widehat{X}_\delta$ with respect to the measure $\mathbb{P}_{U_{\varepsilon,r}}$.

Proof. We set $\widehat{Z}_\delta = \alpha^{-1}Z_\delta$ and we use Proposition D.1, with the localization $U = U_{\varepsilon,r}$, applied to $F = \widehat{X}_\delta$ and $G = \widehat{Z}_\delta$. Recall that the requests in (1) of Proposition D.1 involve several quantities: the lowest singular value (that in this case coincides with the lowest eigenvalue) $\lambda_*(\gamma_{\widehat{X}_\delta})$ and $\lambda_*(\gamma_{\widehat{Z}_\delta})$ of the Malliavin covariance matrix of $\widehat{X}_\delta$ and $\widehat{Z}_\delta$ respectively, as well as $m_{U_{\varepsilon,r}}(p)$ in (D.2), the Sobolev-Malliavin norms $\|\widehat{X}_\delta\|_{2,p,U_{\varepsilon,r}}$, $\|\widehat{Z}_\delta\|_{2,p,U_{\varepsilon,r}}$, and $\|\widehat{X}_\delta - \widehat{Z}_\delta\|_{2,p,U_{\varepsilon,r}} = \|\alpha^{-1}R_\delta\|_{2,p,U_{\varepsilon,r}}$. First of all, by using Assumption 2.1, one easily gets

$$\|\alpha^{-1}R_\delta\|_{2,p} \leq C\delta^{-1}\delta^{3/2} = C\sqrt{\delta} \quad \text{and} \quad \|\widehat{X}_\delta\|_{2,p} + \|\widehat{Z}_\delta\|_{2,p} \leq C, \quad \exists C \in \mathcal{C}. \tag{3.39}$$

We now check that $m_{U_{\varepsilon,r}}(p) < \infty$ for every p. Standard computations and (B.2) give, for every p,

$$\|1/\det Q\|_{2,p} + \|\,|Q|_l\|_{2,p} + \|q(B)\|_{2,p} + \|\Theta\|_{2,p} \leq C,$$

so we can apply (D.4) and conclude

$$m_{U_{\varepsilon,r}}(p) \leq C \in \mathcal{C}. \tag{3.40}$$

We now study the lower eigenvalue of the Malliavin covariance matrix of $\widehat{Z}_\delta$. From the definition of $\widehat{Z}_\delta$, we have

$$\widetilde{Z} = \mathcal{V}\begin{pmatrix} \alpha^{-1}Z_\delta \\ \underline{\mathcal{U}}^T\underline{\Gamma}_\delta\Theta \end{pmatrix} = \mathcal{V}\begin{pmatrix} \widehat{Z}_\delta \\ \underline{\mathcal{U}}^T\underline{\Gamma}_\delta\Theta \end{pmatrix}, \tag{3.41}$$

(see the proof of Lemma 3.5 for the definition of $\underline{\Gamma}_\delta$). As an immediate consequence, one has $\lambda_*(\gamma_{\widehat{Z}_\delta}) \geq \lambda_*(\gamma_{\widetilde{Z}})$, and it suffices to study the lower eigenvalue of $\widetilde{Z}$. By using (3.20), we have

$$\begin{aligned}
\langle \gamma_{\widetilde{Z}}\xi, \xi\rangle &= \sum_{i=1}^d \int_0^\delta \langle D^i_s\widetilde{Z}, \xi\rangle^2 = \sum_{i=1}^d \int_{s_{i-1}(\delta)}^{s_i(\delta)} \langle D^i_s\widetilde{Z}, \xi\rangle^2 = \sum_{i=1}^d \int_{s_{i-1}(\delta)}^{s_i(\delta)} \langle D^i_s(\Theta + \widetilde{\eta}(\Theta)), \xi\rangle^2 \\
&\geq \sum_{i=1}^d \int_{s_{i-1}(\delta)}^{s_i(\delta)} \Big(\frac{1}{2}\langle D^i_s\Theta, \xi\rangle^2 - \langle D^i_s\eta(\Theta), \xi\rangle^2\Big)ds \\
&= S_1 + S_2.
\end{aligned}$$

We write

$$\begin{aligned}
S_1 &= \sum_{i=1}^d \int_{s_{i-1}(\delta)}^{s_i(\delta)} \frac{1}{2}\langle D^i_s\Theta, \xi\rangle^2 \geq \frac{\lambda_*(Q)}{2} \\
S_2 &= \sum_{i=1}^d \int_{s_{i-1}(\delta)}^{s_i(\delta)} \langle \nabla\eta(\Theta)D^i_s\Theta, \xi\rangle^2 ds = \sum_{i=1}^d \int_{s_{i-1}(\delta)}^{s_i(\delta)} \langle D^i_s\Theta, \nabla\eta(\Theta)^T\xi\rangle^2 ds \leq \lambda^*(Q)|\nabla\eta(\Theta)|^2|\xi|^2,
\end{aligned}$$

so that

$$\lambda_*(\gamma_{\widehat{Z}_\delta}) \geq \lambda_*(\gamma_{\widetilde{Z}}) \geq \lambda_*(Q)\left(\frac{1}{2} - \frac{\lambda^*(Q)}{\lambda_*(Q)}|\nabla\eta(\Theta)|^2\right).$$

On $\{\widetilde{U}_\varepsilon \neq 0\}$, we have already proved in Lemma 3.4 that $c_*(\eta,\Theta) \leq \frac{\sqrt{\lambda_*(Q)/\lambda^*(Q)}}{2m}$. Since $|\nabla\eta(\Theta)| \leq mc_*(\eta,\Theta)$, we obtain

$$|\nabla\eta(\Theta)| \leq \frac{1}{2}\sqrt{\frac{\lambda_*(Q)}{\lambda^*(Q)}},$$

and therefore $\lambda_*(\gamma_{\widehat{Z}_\delta}) \geq 4\lambda_*(\gamma_{\widetilde{Z}}) \geq \lambda_*(Q) \geq \varepsilon^{3m\rho}$, which implies that $\mathbb{E}_{U_{\varepsilon,r}}(\lambda_*(\widehat{Z}_\delta)^{-p}) < \infty$ for every p. Let us study the lowest eigenvalue of $\gamma_{\widehat{X}_\delta}$. We use here some results from next Section 4, namely Lemma 4.5. There, we actually prove the desired bound for the Malliavin covariance matrix of $\alpha^{-1}(X_\delta - x_0)$. Here we are considering $\widehat{X}_\delta = \alpha^{-1}(X_\delta - x_0 - b(0,x_0)\delta)$, but their Malliavin covariance matrix is the clearly the same. Then, Lemma 4.5 gives that $\mathbb{E}(\lambda_*(\gamma_{\widehat{X}_\delta})^{-p}) < \infty$ for every p.
So, we have proved that all the requests in Proposition D.1 hold. Then, we can apply (D.6) and we get

$$p_{\widehat{X}_\delta, U_{\varepsilon,r}}(z) \geq p_{\widehat{Z}_\delta, U_{\varepsilon,r}}(z) - C'\sqrt{\delta}$$

with $C' \in \mathcal{C}$. Now, from (3.41) and (3.29), with a simple change of variables, we get

$$p_{\widehat{Z}_\delta, U_{\varepsilon,r}}(z) \geq \frac{1}{C}, \quad \text{for} \quad |z| \leq \frac{r}{2}. \tag{3.42}$$

We can assert the existence of $\delta_* \in 1/\mathcal{C}$ and $C \in \mathcal{C}$ such that for all $\delta \leq \delta_*$

$$p_{\widehat{X}_\delta, U_{\varepsilon,r}}(z) \geq \frac{1}{2C},$$

and the statement follows. □

We are now ready for the proof of the lower bound:

Theorem 3.7. *Let Assumption 2.1 and 2.2 hold. Let p_{X_t} denote the density of X_t, $t > 0$. Then there exist positive constants r, δ_*, C such that for every $\delta \leq \delta_*$ and for every y such that $|y - x_0 - b(0,x_0)\delta|_{A_\delta(0,x_0)} \leq r$,*

$$p_{X_\delta}(y) \geq \frac{1}{C\delta^{n - \frac{\dim\langle\sigma(0,x_0)\rangle}{2}}},$$

$\dim\langle\sigma(0,x_0)\rangle$ *denoting the dimension of the vector space spanned by $\sigma_1(0,x_0), \ldots, \sigma_d(0,x_0)$. Here, $C \in \mathcal{C}$ and $r, \delta_* \in 1/\mathcal{C}$.*

Proof. We take the same δ_*, r as in Lemma 3.6 and let $\widehat{X}_\delta$ denotes the r.v. handled in Lemma 3.6. By construction, we have $X_\delta = x_0 + b(0,x_0) + \alpha\widehat{X}_\delta$, so by applying Lemma 3.6 we get

$$\begin{aligned}
\mathbb{E}(f(X_\delta)) \geq \mathbb{E}_{U_{\varepsilon,r}}(f(X_\delta)) &= \mathbb{E}_{U_{\varepsilon,r}}(f(x_0 + b(0,x_0)\delta + \alpha\widehat{X}_\delta)) \\
&= \int f(x_0 + b(0,x_0)\delta + \alpha z) p_{\widehat{X}_\delta, U_{\varepsilon,r}}(z) dz \\
&\geq \frac{1}{C} \int_{\{|z| \leq r/2\}} f(x_0 + b(0,x_0)\delta + \alpha z) dz \\
&\geq \frac{1}{C|\det \alpha|} \int_{|y|_\alpha \leq r/2} f(x_0 + b(0,x_0)\delta + y) dy
\end{aligned}$$

From (2.7) and the Cauchy-Binet formula we obtain

$$\frac{1}{C}\delta^{n - \frac{\dim\langle\sigma\rangle}{2}} \leq \sqrt{|\det A_\delta A_\delta^T|} = \det(\alpha) \leq C\delta^{n - \frac{\dim\langle\sigma\rangle}{2}} \tag{3.43}$$

and the statement follows. □

Remark 3.8. We observe that if the diffusion coefficients are bounded, that is Assumption 2.3 holds, then the class $\mathcal{C}$ in (3.28) of the constants can be replaced by

$$\mathcal{D}_0 = \Big\{ C > 0 \,:\, C = c\Big(\frac{\kappa}{\lambda(0,x_0)}\Big)^q, \ \exists\, c, q > 0 \Big\}$$

and, as before, $1/\mathcal{D}_0 = \{C > 0 \,:\, 1/C \in \mathcal{D}_0\}$. This is because in the estimates for $\|\widehat{X}_\delta - \widehat{Z}_\delta\|_{2,p}$ and $\|\widehat{X}_\delta\|_{2,p}$ one does not need any more to use the Gronwall's Lemma but it suffices to use the boundedness of the coefficients and the Burkholder inequality.

4 Upper bound

We study here the upper bound for the density of X_δ.

4.1 Rescaling of the diffusion

As for the lower bound, we again scale X_δ. We recall the results and the notation in Lemma 3.5 and we define the change of variable

$$T_\alpha : \mathbb{R}^n \to \mathbb{R}^n, \qquad T_\alpha(y) = \alpha^{-1} y, \quad \text{where} \quad \alpha = \mathcal{U}\bar{\Sigma} \tag{4.1}$$

and its adjoint $T^*_\alpha(v) = \alpha^{-1,T} v$. Note that α is a $n \times n$ matrix. We write $A_{\delta,j}$, for $j = 1, \dots, m$, for the columns of A_δ (which can be $\sqrt{\delta}\sigma_i$ or $\delta[\sigma_i, \sigma_p]$). The following properties hold:

Lemma 4.1. *Let T_α be defined in (4.1). Then one has:*

$$|y|_{A_\delta} = |T_\alpha y| = |y|_\alpha, \quad \forall y \in \mathbb{R}^n, \qquad \textit{and} \ \det \alpha = \sqrt{\det A_\delta A_\delta^T} \tag{4.2}$$

$$\forall v \in \mathbb{R}^n \ \textit{with} \ |v| = 1, \quad \exists j = 1, \dots, m: \quad |T^*_\alpha v \cdot A_{\delta,j}| \geq \frac{1}{m} \tag{4.3}$$

$$\forall j = 1, \dots, d, \quad \sqrt{\delta}|T_\alpha \sigma_j| \leq 1 \tag{4.4}$$

Proof. (4.2) follows easily from $\alpha = \mathcal{U}\bar{\Sigma}$ and the definition (2.1) of $|\cdot|_M$. Now, $(T^*_\alpha v)^T A_\delta = v^T \alpha^{-1} A_\delta = [v^T 0]\mathcal{V}^T$. So $|(T^*_\alpha v)^T A_\delta| = |[v^T 0]\mathcal{V}^T| = 1$. Recall that $A_{\delta,j}$ are the columns of A_δ, therefore $\exists j = 1, \dots, m: \quad |(T^*_\alpha v)^T A_{\delta,j}| \geq \frac{1}{m}$, which is equivalent to (4.3). Moreover, $T_\alpha A_\delta = [\mathrm{Id}_n \, 0]\mathcal{V}^T$. This easily implies that $\forall i = 1, \dots, m, \quad |T_\alpha A_{\delta,i}| \leq 1$. For $A_{\delta,i} = \sigma_j(0, x_0)\sqrt{\delta}$ we have (4.4). □

We define now

$$F = \alpha^{-1}(X_\delta - x_0) = T_\alpha(X_\delta - x_0). \tag{4.5}$$

As for the lower bound, we first estimate the density of F, using the results in Appendix D (specifically, (D.7) in Proposition D.1), and then recover the estimates for the density of X_δ via a change of variable.

4.2 Malliavin Covariance Matrix

Let F be as in (4.5). To prove the upper bound for its density p_F we need a quantitative control on the Malliavin covariance matrix γ_F of F. We start with some preliminary results. The following lemma is a slight modification of Lemma 2.3.1. in [14].

Lemma 4.2. *Let γ be a symmetric nonnegative definite $n \times n$ random matrix. Denoting $|\gamma| = \sum_{1\leq i,j\leq n} |\gamma^{i,j}|^2)^{1/2}$, we assume that, for $p \geq 2$, $\mathbb{E}|\gamma|^{p+1} < \infty$, and that there exists $\varepsilon_0 > 0$ such that for $\varepsilon \leq \varepsilon_0$,*

$$\sup_{|\xi|=1} \mathbb{P}[\langle \gamma\xi, \xi\rangle < \varepsilon] \leq \varepsilon^{p+2n}$$

Then there exists a constant C depending only on the dimension n such that

$$\mathbb{E}\lambda_*(\gamma)^{-p} \leq C\mathbb{E}|\gamma|^{1+p}\varepsilon_0^{-p}$$

We also need the following technical result.

Lemma 4.3. *Let $\delta \in (0, 1]$ and let a_t, b_t, $t \in [0, \delta]$ be stochastic processes which are a.s. increasing. Assume that $b_0 = 0$. Suppose for fixed $p \geq 1$ and for all $t \in [0, \delta]$ one has*

$$\mathbb{E}[b_t^p] \leq C_p t^{2p} \quad \text{and} \quad a_t \geq \frac{t - b_t}{\delta}.$$

Then for all $\varepsilon > 0$

$$\mathbb{P}(a_\delta \leq \varepsilon) \leq 4^p C_p \varepsilon^p.$$

Proof. Set

$$S_\varepsilon = \inf\left\{ s \geq 0 : b_s \geq \frac{\delta\varepsilon}{2} \right\} \wedge \delta,$$

Remark that for any $p > 0$

$$\mathbb{P}(S_\varepsilon < \delta\varepsilon) = \mathbb{P}\left(b_{\delta\varepsilon}^p \geq \left(\frac{\delta\varepsilon}{2}\right)^p \right) \leq 2^p \frac{\mathbb{E}b_{\delta\varepsilon}^p}{(\delta\varepsilon)^p} \leq 2^p C_p (\delta\varepsilon)^p.$$

On the other hand, on $S_\varepsilon \geq \delta\varepsilon$,

$$a_{S_\varepsilon} \geq a_{\delta\varepsilon} \geq \frac{\delta\varepsilon - \delta\varepsilon/2}{\delta} \geq \varepsilon/2.$$

Therefore

$$\mathbb{P}(a_\delta < \varepsilon/2) \leq \mathbb{P}(a_\delta < \varepsilon/2, S_\varepsilon < \delta\varepsilon) + \mathbb{P}(a_\delta < \varepsilon/2, S_\varepsilon \geq \delta\varepsilon) \leq \mathbb{P}(S_\varepsilon < \delta\varepsilon) \leq 2^p C_p \varepsilon^p.$$

This implies that $\mathbb{P}(a_\delta < \varepsilon) \leq 4^p C_p \varepsilon^p$. $\square$

The following Lemma 4.4 is a refinement of what was proved by Norris in [13], in the sense that we take care of the same quantities, but handling more carefully the dependence on the final time t_0. This is a key estimate for the proof of next Theorem 4.6.

Lemma 4.4. *Suppose $u(t) = (u_1(t), \ldots, u_d(t))$ and $a(t)$ are a.s. continuous and adapted processes such that for some $p \geq 1$, $C > 0$ and for every $t_0 \leq 1$ one has*

$$\mathbb{E}\left[\sup_{0\leq s\leq t_0} |u_s|^p\right] \leq \frac{C}{t_0^p}, \qquad \mathbb{E}\left[\sup_{0\leq s\leq t_0} |a_s|^p\right] \leq \frac{C}{t_0^p}. \tag{4.6}$$

Set

$$Y(t) = y + \int_0^t a(s)ds + \sum_{k=1}^d \int_0^t u_k(s)dW_s^k.$$

Then, for any $q > 4$ and $r > 0$ such that $6r + 4 < q$, there exists $\varepsilon_0(q, r, p)$ such that for every $t_0 \le 1$ and $\varepsilon \le \varepsilon_0(q, r, p)$ one has

$$\mathbb{P}\left\{\int_0^{t_0} Y_t^2 dt < \varepsilon^q,\ \int_0^{t_0} |u(t)|^2 dt \ge \frac{6\varepsilon}{t_0}\right\} \le (2^p C + 1)\varepsilon^{rp}.$$

Proof. Set $\theta_t = |a_t| + |u_t|$, and

$$\tau = \inf\left\{s \ge 0 : \sup_{0 \le u \le s} \theta_u > \frac{\varepsilon^{-r}}{t_0}\right\} \wedge t_0.$$

We have

$$\mathbb{P}\left\{\int_0^{t_0} Y_t^2 dt < \varepsilon^q,\ \int_0^{t_0} |u(t)|^2 dt \ge \frac{\varepsilon}{t_0}\right\} \le A_1 + A_2.$$

where $A_1 = \mathbb{P}[\tau < t_0]$ and

$$A_2 = \mathbb{P}\left\{\int_0^{t_0} Y_t^2 dt < \varepsilon^q,\ \int_0^{t_0} |u_t|^2 dt \ge \frac{\varepsilon}{t_0}, \tau = t_0\right\}$$

An upper bound for A_1 easily follows from (4.6). Indeed

$$\mathbb{P}[\tau < t_0] \le \mathbb{P}\left[\sup_{0 \le s \le t_0} \theta_u > \frac{\varepsilon^{-r}}{t_0}\right] \le t_0^p \varepsilon^{rp} \mathbb{E}\left[\sup_{0 \le s \le t_0} \theta_s^p\right] \le 2^p C \varepsilon^{rp}.$$

for $\varepsilon \le \varepsilon_0$. To estimate A_2 we introduce

$$N_t = \int_0^t Y_s \sum_{k=1}^d u_s^k dW_s^k \quad \text{and}$$

$$B = \left\{\langle N\rangle_\tau < \rho, \sup_{0 \le s \le \tau} |N_s| \ge \delta\right\}, \quad \text{with} \quad \delta = \frac{\varepsilon^{2r+2}}{t_0} \quad \text{and} \quad \rho = \frac{\varepsilon^{q-2r}}{t_0^2}.$$

By the exponential martingale inequality,

$$\mathbb{P}(B) \le \exp\left(\frac{-\delta^2}{2\rho}\right) \le \exp\left(-\varepsilon^{6r+4-q}\right).$$

So, in order to conclude the proof, it suffices to show that

$$\left\{\int_0^{t_0} Y_t^2 dt < \varepsilon^q,\ \int_0^{t_0} |u_t|^2 dt \ge \frac{6\varepsilon}{t_0}, \tau = t_0\right\} \subset B, \tag{4.7}$$

We suppose $\omega \notin B$, $\int_0^{t_0} Y_t^2 dt < \varepsilon^q$ and $\tau = t_0$ and show $\int_0^{t_0} |u_t|^2 dt < 6\varepsilon/t_0$. With these assumptions,

$$\langle N\rangle_\tau = \int_0^\tau Y_t^2 |u_t|^2 dt \le \int_0^{t_0} Y_t^2 dt \sup_{0 \le t \le \tau} |u_t|^2 < \frac{\varepsilon^{q-2r}}{t_0^2} = \rho.$$

So, for $\omega \notin B$ then one necessarily has $\sup_{0 \le t \le \tau} |\int_0^t Y_s \sum_{k=1}^d u_s^k dW_s^k| < \delta = \varepsilon^{2r+2}/t_0$. From $6r + 4 \le q$, if $\tau = t_0$ then

$$\sup_{0 \le t \le \tau} \left|\int_0^t Y_s a_s ds\right| \le \left(t_0 \int_0^\tau Y_s^2 a_s^2 ds\right)^{1/2} \le t_0 \Big(\int_0^{t_0} Y_s^2 ds \sup_{0 \le s \le \tau} |a_s|^2\Big)^{1/2}$$
$$\le t_0 \Big(\varepsilon^q \frac{\varepsilon^{-2r}}{t_0^2}\Big)^{1/2} \le \frac{\varepsilon^{2r+2}}{t_0}.$$

Thus

$$\sup_{0\le t\le\tau}\left|\int_0^t Y_s dY_s\right| \le \sup_{0\le t\le\tau}\left|\int_0^t Y_s a_s ds + \int_0^t Y_s u_s dW_s\right| \le \frac{2\varepsilon^{2r+2}}{t_0}.$$

By Itô's formula, $Y_t^2 = y^2 + 2\int_0^t Y_s dY_s + \langle M\rangle_t$ with $\langle M\rangle_t = \int_0^t |u_s|^2 ds$. So, recalling that $q > 2r+2$,

$$\int_0^\tau \langle M\rangle_t dt = \int_0^\tau Y_t^2 dt - \tau y^2 - 2\int_0^\tau\int_0^t Y_s dY_s dt < \varepsilon^q + 4\frac{\varepsilon^{2r+2}}{t_0} < 5\frac{\varepsilon^{2r+2}}{t_0}.$$

Since $t\mapsto\langle M\rangle_t$ is non negative and increasing, for $0<\gamma<\tau$ we have

$$\gamma\langle M\rangle_{\tau-\gamma} \le \int_{\tau-\gamma}^\tau \langle M\rangle_t dt \le 5\frac{\varepsilon^{2r+2}}{t_0}.$$

Using also the fact that

$$\langle M\rangle_\tau - \langle M\rangle_{\tau-\gamma} = \int_{\tau-\gamma}^\tau |u_s|^2 ds \le \gamma\frac{\varepsilon^{-2r}}{t_0^2},$$

we have

$$\langle M\rangle_\tau < \frac{5\varepsilon^{2r+2}}{\gamma} + \gamma\frac{\varepsilon^{-2r}}{t_0^2}.$$

With $\gamma = t_0\varepsilon^{2r+1}$, this gives $\int_0^{t_0}|u_s|^2 ds = \langle M\rangle_\tau < \frac{6\varepsilon}{t_0}$. □

We are now ready to prove the non degeneracy of the Malliavin covariance matrix. More precisely, we prove a quantitative version of this property: the L^p norm of the inverse of the Malliavin covariance matrix of F is upper bounded by a constant in $\mathcal{C}$, $\mathcal{C}$ being defined in (3.28).

Lemma 4.5. *Let α, T_α and $F = T_\alpha(X_\delta - x_0)$ be defined as in* (4.1) *and* (4.5)*. Let γ_F denote the Malliavin covariance matrix of F. Then for any $p>1$ there exists $C\in\mathcal{C}$ such that, for $\delta\le 1$, $\mathbb{E}|\lambda_*(\gamma_F)|^{-p}\le C$.*

Proof. We need a bound for the moments of the inverse of

$$\gamma_F = \sum_{k=1}^d \int_0^\delta D_s^k F D_s^k F^T ds.$$

Following [14] we define the tangent flow Y of X as the derivative with respect to the initial condition of X: $Y_t := \partial_x X_t$. We also denote its inverse $Z_t = Y_t^{-1}$. Then one has (remark that the equations we consider for X, Y and Z are all in Stratonovich form):

$$\begin{aligned} Y_t &= \mathrm{Id} + \sum_{k=1}^d \int_0^t \nabla_x\sigma_k(s,X_s)Y_s \circ dW_s^k + \int_0^t \nabla_x b(s,X_s)Y_s ds \\ Z_t &= \mathrm{Id} - \sum_{k=1}^d \int_0^t Z_s\nabla_x\sigma_k(s,X_s)\circ dW_s^k - \int_0^t Z_s\nabla_x b(s,X_s)ds, \end{aligned} \tag{4.8}$$

where $\nabla_x \sigma_k$ and $\nabla_x b$ are the Jacobian matrix with respect to the space variable. It holds

$$D_s X_\delta = Y_\delta Z_s \sigma(s, X_s), \quad s < \delta.$$

By applying Itô's formula we have the following representation, for $\phi \in C^{1,2}$:

$$\begin{aligned} Z_t \phi(t, X_t) = \phi(0, x_0) &+ \int_0^t Z_s \sum_{k=1}^d [\sigma_k, \phi](s, X_s) dW_s^k \\ &+ \int_0^t Z_s \left\{ [b, \phi] + \frac{1}{2} \sum_{k=1}^d [\sigma_k, [\sigma_k, \phi]] + \frac{\partial \phi}{\partial s} \right\} (s, X_s)\, ds \end{aligned} \tag{4.9}$$

(details are given in [14], remark that in the r.h.s. above we are taking into account an Itô integral). We now compute

$$D_s F = \alpha^{-1} D_s X_\delta = \alpha^{-1} Y_\delta Z_s \sigma(s, X_s) = \alpha^{-1} Y_\delta \alpha \alpha^{-1} Z_s \sigma(s, X_s)$$

so

$$\gamma_F = \alpha^{-1} Y_\delta \alpha\, \bar{\gamma}_F\, (\alpha^{-1} Y_\delta \alpha)^T \quad \text{where} \quad \bar{\gamma}_F = \alpha^{-1} \int_0^\delta Z_s \sigma(s, X_s) \sigma(s, X_s)^T Z_s^T ds\, \alpha^{-1,T},$$

and

$$\gamma_F^{-1} = (\alpha^{-1} Y_\delta \alpha)^{-1,T}\, \bar{\gamma}_F^{-1}\, (\alpha^{-1} Y_\delta \alpha)^{-1}.$$

Now,

$$(\alpha^{-1} Y_\delta \alpha)^{-1} = \alpha^{-1} Z_\delta \alpha = \mathrm{Id}_n + \alpha^{-1}(Z_\delta - \mathrm{Id}_n)\alpha$$

Using the fact that $\lambda^*(\cdot)$ is a norm on the set of matrices, and that for two $n \times n$ matrices A, B, $\lambda^*(AB) \le n \lambda^*(A) \lambda^*(B)$, we have

$$\lambda_*\big(\gamma_F\big)^{-1} = \lambda^*\big(\gamma_F^{-1}\big) \le n^2 \lambda^*\big(\bar{\gamma}_F^{-1}\big)\, \lambda^*\big((\alpha^{-1} Y_\delta \alpha)^{-1}\big)^2$$

and

$$\lambda^*\big((\alpha^{-1} Y_\delta \alpha)^{-1}\big) \le 1 + n^2 \lambda^*(\alpha^{-1}) \lambda^*(Z_\delta - \mathrm{Id}_n) \lambda^*(\alpha).$$

Standard estimates (see also (4.8)) give $\lambda^*(Z_\delta - \mathrm{Id}_n) \le C_1 \sqrt{\delta}$ for some $C_1 \in \mathcal{C}$. Moreover

$$\begin{aligned} &\lambda^*(\alpha) = \lambda^*(A_\delta) = \lambda^*(A\mathcal{D}_\delta) \le n \lambda^*(A) \lambda^*(\mathcal{D}_\delta) \le C_2 \sqrt{\delta}, \quad C_2 \in \mathcal{C} \\ &\lambda^*(\alpha^{-1}) \le \frac{1}{\lambda_*(A\mathcal{D}_\delta)} \le \frac{C_3}{\delta}, \quad C_3 \in \mathcal{C} \end{aligned}$$

and so for all $q > 1$ exists $C \in \mathcal{C}$ such that

$$\mathbb{E} \lambda^* \left((\alpha^{-1} Y_\delta \alpha)^{-1} \right)^q \le C$$

We now need to estimate the reduced matrix, i.e. prove that for all $q > 1$ exists $C \in \mathcal{C}$ such that

$$\mathbb{E} \lambda^*(\bar{\gamma}_F^{-1})^q = \mathbb{E} \lambda_*(\bar{\gamma}_F)^{-q} \le C \tag{4.10}$$

We show now that for any $p > 0$, $\sup_{|v|=1} \mathbb{P}\left(\langle \bar{\gamma}_F v, v \rangle \le \varepsilon \right) \le \varepsilon^p$, for $\delta \le 1$ for $\varepsilon \le \varepsilon_0 \in 1/\mathcal{C}$ not depending on δ. Together with lemma 4.2 this implies (4.10).

Denote $\xi = T_\alpha^* v = \alpha^{-1,T} v$. From (4.3) and the definition (2.7) of A_δ we have two possible cases: A) $|\xi \cdot \sigma_j(0, x_0)| \geq \frac{1}{m\delta^{1/2}}$ for some $j = 1, \ldots, d$, or B) $|\xi \cdot [\sigma_j, \sigma_l](0, x_0)| \geq \frac{1}{m\delta}$ for some $j, l = 1, \ldots, d$, $j \neq l$. Moreover

$$\alpha \bar{\gamma}_F \alpha^T = \int_0^\delta Z_s \sigma(s, X_s) \sigma(s, X_s)^T Z_s^T ds. \tag{4.11}$$

Therefore, with $\xi = T_\alpha^* v$, we have for any $q > 1$

$$\begin{aligned}
\mathbb{P}(\langle \bar{\gamma}_F v, v \rangle \leq \varepsilon^q) &= \mathbb{P}\left(\xi^T \int_0^\delta Z_s \sigma(s, X_s) \sigma(s, X_s)^T Z_s^T ds \, \xi \leq \varepsilon^q\right) \\
&= \mathbb{P}\left(\sum_{i=1}^d \int_0^\delta |\xi^T Z_s \sigma_i(s, X_s)|^2 ds \leq \varepsilon^q\right)
\end{aligned}$$

We decompose this probability:

$$\begin{aligned}
\mathbb{P}(\langle \bar{\gamma}_F v, v \rangle \leq \varepsilon^q) &= \mathbb{P}\left(\sum_{i=1}^d \int_0^\delta |\xi^T Z_t \sigma_i(t, X_t)|^2 dt \leq \varepsilon^q\right) \\
&\leq \mathbb{P}\left(\sum_{i=1}^d \int_0^\delta |\xi^T Z_t \sigma_i(t, X_t)|^2 dt \leq \varepsilon^q, \sum_{i,k=1}^d \int_0^\delta |\xi^T Z_t [\sigma_i, \sigma_k](t, X_t)|^2 dt \leq \frac{\varepsilon}{\delta}\right) \\
&+ \mathbb{P}\left(\sum_{i=1}^d \int_0^\delta |\xi^T Z_t \sigma_i(t, X_t)|^2 dt \leq \varepsilon^q, \sum_{i,k=1}^d \int_0^\delta |\xi^T Z_t [\sigma_i, \sigma_k](t, X_t)|^2 dt > \frac{\varepsilon}{\delta}\right) \\
&=: I_1 + I_2
\end{aligned}$$

To estimate I_1 we distinguish the two cases A) and B) above.
Case A): $|\xi \cdot \sigma_j(0, x_0)| \geq \frac{1}{m\delta^{1/2}}$ for some $j = 1, \ldots, d$. We fix this j. Then,

$$\begin{aligned}
I_1 &\leq \mathbb{P}\Big(\int_0^\delta |\xi^T Z_t \sigma_j(t, X_t)|^2 dt \leq \varepsilon^q, \int_0^\delta \Big|\xi^T \sum_{k=1}^d Z_t [\sigma_k, \sigma_j](t, X_t)\Big|^2 dt < \frac{\varepsilon}{\delta}\Big) \\
&\leq \mathbb{P}\Big(\int_0^\delta |\xi^T Z_t \sigma_j(t, X_t)|^2 dt \leq \varepsilon^q, \sup_{0 \leq t \leq \delta} \Big|\int_0^t \xi^T \sum_{k=1}^d Z_s [\sigma_k, \sigma_j](s, X_s) dW_s^k\Big|^2 < \frac{1}{12m^2\delta}\Big) \\
&+ \mathbb{P}\Big(\sup_{0 \leq t \leq \delta} |\int_0^t \xi^T \sum_{k=1}^d Z_s [\sigma_k, \sigma_j](s, X_s) dW_s^k|^2 \geq \frac{1}{12m^2\delta}, \int_0^\delta |\xi^T \sum_{k=1}^d Z_t [\sigma_k, \sigma_j](t, X_t)|^2 dt < \frac{\varepsilon}{\delta}\Big)
\end{aligned}$$

Set $u_s = (\xi^T Z_s [\sigma_k, \sigma_j](s, X_s))_{k=1,\ldots,d}$. From the exponential martingale inequality we have

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{0 \leq t \leq \delta} |\sum_{k=1}^d \int_0^t u_s^k dW_s^k|^2 \geq \frac{1}{12m^2\delta}, \quad \int_0^\delta |u_t|^2 dt < \frac{\varepsilon}{\delta}\Big) \\
&\leq 2\exp\left(-\frac{1}{12m^2\delta} \times \frac{\delta}{2\varepsilon}\right) = 2\exp\left(-\frac{1}{24m^2\varepsilon}\right) < \varepsilon^p,
\end{aligned} \tag{4.12}$$

the latter inequality holding for every $p > 1$ and $\varepsilon \leq \varepsilon_0$. We now define

$$D := \Big\{\sup_{0 \leq t \leq \delta} |\int_0^t \xi^T \sum_{k=1}^d Z_s [\sigma_k, \sigma_j](s, X_s) dW_s^k|^2 < \frac{1}{12m^2\delta}\Big\}$$

and prove

$$\mathbb{P}\left(\left\{\int_0^\delta |\xi^T Z_t\sigma_j(t,X_t)|^2 dt \le \frac{\varepsilon^q}{4m^2}\right\}\cap D\right)\le \varepsilon^p$$

which is equivalent to the desired estimate $\mathbb{P}\left(\left\{\int_0^\delta |\xi^T Z_t\sigma_j(t,X_t)|^2 dt \le \varepsilon^q\right\}\cap D\right)\le \varepsilon^p$. From representation (4.9), for $\phi=\sigma_j$ we find

$$Z_t\sigma_j(t,X_t)=\sigma_j(0,x_0)+\int_0^t\sum_{k=1}^d Z_s[\sigma_k,\sigma_j](s,X_s)dW_s^k+R_t,$$

with

$$R_t=\int_0^t Z_s\left\{[b,\sigma_j]+\frac{1}{2}\sum_{k=1}^d[\sigma_k,[\sigma_k,\sigma_j]]+\frac{\partial\sigma_j}{\partial s}\right\}(s,X_s)\,ds.$$

From $(a+b+c)^2\ge a^2/3-b^2-c^2$ and $|\xi\cdot\sigma_j(0,x_0)|\ge\frac{1}{m\delta^{1/2}}$, for $\bar t\le\delta$ we can write

$$\begin{aligned}
&\int_0^{\bar t}|\xi^T Z_t\sigma_j(t,X_t)|^2dt\\
&\quad\ge\frac{\bar t|\xi^T\sigma_j(0,x_0)|^2}{3}-\int_0^{\bar t}|\sum_{k=1}^d\int_0^t\xi^T Z_s[\sigma_k,\sigma_j](s,X_s)dW_s^k|^2dt-\int_0^{\bar t}|\xi^T R_t|^2dt\\
&\quad\ge\frac{\bar t}{3\delta m^2}-\int_0^{\bar t}|\sum_{k=1}^d\int_0^t\xi^T Z_s[\sigma_k,\sigma_j](s,X_s)dW_s^k|^2dt-\int_0^{\bar t}|\xi^T R_t|^2dt.
\end{aligned}$$

On the set D one has

$$\int_0^{\bar t}|\sum_{k=1}^d\int_0^t\xi^T Z_s[\sigma_k,\sigma_j](s,X_s)dW_s^k|^2dt\le\bar t\frac{1}{12m^2\delta},$$

so

$$\int_0^{\bar t}|\xi^T Z_t\sigma_j(t,X_t)|^2dt\ge\frac{\bar t}{4m^2\delta}-\int_0^{\bar t}|\xi^T R_t|^2dt,$$

that we rewrite as

$$4m^2\int_0^{\bar t}|\xi^T Z_t\sigma_j(t,X_t)|^2dt\ge\frac{\bar t-4m^2\delta\int_0^{\bar t}|\xi^T R_t|^2dt}{\delta}. \tag{4.13}$$

We now set

$$a_{\bar t}=4m^2\int_0^{\bar t}|\xi^T Z_t\sigma_j(t,X_t)|^2dt \text{ on the set } D \text{ and } a_{\bar t}=\bar t/\delta \text{ on the set } D^c,$$

D^c denoting the complement of D. Standard computations, considering also $|\xi|=|T_\alpha^* v|\le|v|C/\delta=C/\delta$, give $\mathbb{E}(\int_0^{\bar t}|\xi^T R_t|^2dt)^q\le C\bar t^{3q}/\delta^{2q}$, so $\mathbb{E}(4\delta m^2\int_0^{\bar t}|\xi^T R_t|^2dt)^q\le C\bar t^{2q}$, for $C\in\mathcal{C}$ (recall also $\bar t\le\delta$). This estimate and (4.13) allow us to apply lemma 4.3 with $a_{\bar t}$ defined above and

$$b_{\bar t}=4\delta m^2\int_0^{\bar t}|\xi^T R_t|^2dt.$$

We find

$$\left\{\int_0^\delta |\xi^T Z_t\sigma_j(t,X_t)|^2dt \le \frac{\varepsilon^q}{4m^2}\right\}\cap D = \{a_\delta \le \varepsilon^q\}\cap D$$

and we have

$$\mathbb{P}\Big(\Big\{\int_0^\delta |\xi^T Z_t\sigma_j(t,X_t)|^2dt \le \frac{\varepsilon^q}{4m^2}\Big\}\cap D\Big) = \mathbb{P}(\{a_\delta \le \varepsilon^q\}\cap D) \le \mathbb{P}(a_\delta \le \varepsilon^q) \le \varepsilon^p.$$

We the obtain $I_1 < \varepsilon^p$ for any $p > 1$, for $\delta \le 1$, $\varepsilon \le \varepsilon_0$.

Case B) $|\xi \cdot [\sigma_j, \sigma_l](t, x_0)| \ge \frac{1}{m\delta}$ for some $j, l = 1 \dots d$, $j \ne l$. In this case we write

$$I_1 \le \mathbb{P}\left(\int_0^\delta |\xi^T Z_t[\sigma_j,\sigma_l](t,X_t)|^2dt \le \frac{\varepsilon}{\delta}\right)$$

From representation (4.9) with $\phi = [\sigma_j, \sigma_l]$ we find

$$Z_t[\sigma_j,\sigma_l](t,X_t) = [\sigma_j,\sigma_l](0,x_0) + R_t,$$

with

$$\begin{aligned} R_t = &\int_0^t Z_s \sum_{k=1}^d [\sigma_k,[\sigma_j,\sigma_l]](s,X_s)dW_s^k \\ &+ \int_0^t Z_s \left\{[b,[\sigma_j,\sigma_l]] + \frac{1}{2}\sum_{k=1}^d [\sigma_k,[\sigma_k,[\sigma_j,\sigma_l]]] + \frac{\partial[\sigma_j,\sigma_l]}{\partial s}\right\}(s,X_s)\,ds. \end{aligned} \tag{4.14}$$

From $(a+b)^2 \ge a^2/2 - b^2$ and $|\xi \cdot [\sigma_j,\sigma_l](0,x_0)| \ge \frac{1}{m\delta}$, for $\bar t \le \delta$ we have

$$\int_0^{\bar t} |\xi^T Z_t[\sigma_j,\sigma_l](t,X_t)|^2dt \ge \frac{\bar t|\xi^T[\sigma_j,\sigma_l](0,x_0)|^2}{2} - \int_0^{\bar t}|\xi^T R_t|^2dt \ge \frac{\bar t}{2\delta^2m^2} - \int_0^{\bar t}|\xi^T R_t|^2dt. \tag{4.15}$$

We apply lemma 4.3 with

$$a_{\bar t} = 2m^2\delta\int_0^{\bar t}|\xi^T Z_s[\sigma_j,\sigma_l](s,X_s)|^2ds \quad\text{and}\quad b_{\bar t} = 2m^2\delta^2\int_0^{\bar t}|\xi^T R_t|^2dt$$

Indeed from (4.14) and $|\xi| \le C/\delta$,

$$\mathbb{E}|b_{\bar t}|^q \le C\bar t^2$$

and from (4.15) we have $a_{\bar t} \ge \frac{\bar t - b_{\bar t}}{\delta}$. So, we find $I_1 < \varepsilon^p$, for $\delta \le 1$, $\varepsilon \le \varepsilon_0$.

We estimate now

$$I_2 = \mathbb{P}\left(\sum_{i=1}^d\int_0^\delta |\xi^T Z_s\sigma_i(s,X_s)|^2ds \le \varepsilon^q, \sum_{i,j=1}^d\int_0^\delta|\xi^T Z_s[\sigma_i,\sigma_j](s,X_s)|^2ds > \frac{\varepsilon}{\delta}\right).$$

By using again (4.9), we find

$$\begin{aligned} \xi^T Z_t\sigma_i(t,X_t) = \sigma_i(0,x_0) &+ \sum_{j=1}^d\int_0^t \xi^T Z_s[\sigma_j,\sigma_i](s,X_s)dW_s^j \\ &+ \int_0^t \xi^T Z_s\left\{[b,\sigma_i] + \frac{1}{2}\sum_{j=1}^n[\sigma_j,[\sigma_j,\sigma_i]] + \frac{\partial\sigma_i}{\partial s}\right\}(s,X_s)ds. \end{aligned}$$

For $t_0 = \delta$ and from the fact that $|\xi| \leq \frac{C}{\delta}$, we have

$$\mathbb{E}[\sup_{0\leq s\leq\delta} |\xi^T Z_s[\sigma_j,\sigma_i](s,X_s)|^p] \leq \frac{C}{\delta^p}, \quad C\in\mathcal{C}, \quad \text{and}$$
$$\mathbb{E}\left[\sup_{0\leq s\leq\delta} |\xi^T Z_s\left\{[b,\sigma_i]+\frac{1}{2}\sum_{j=1}^d[\sigma_j,[\sigma_j,\sigma_i]]\right\}(s,X_s)|^p\right] \leq \frac{C}{\delta^p}, \quad C\in\mathcal{C}.$$

Thus we can apply Lemma 4.4 and we get

$$\mathbb{P}\left(\sum_{i=1}^d\int_0^\delta |\xi^T Z_s\sigma_i(s,X_s)|^2 ds \leq \varepsilon^q \text{ and } \sum_{i,j=1}^d\int_0^\delta |\xi^T Z_s[\sigma_i,\sigma_j](s,X_s)|^2 ds > \frac{\varepsilon}{\delta}\right) \leq \varepsilon^p$$

for any $p>1$, $\delta\leq 1$ for $\varepsilon\leq\varepsilon_0$. We have now both the estimates of I_1 and I_2, so we have $\sup_{|v|=1}\mathbb{P}(\langle\bar{\gamma}_F v,v\rangle\leq\varepsilon^q)\leq\varepsilon^p$ for $p>1$, $\delta\leq 1$ for $\varepsilon\leq\varepsilon_0$, and the statement holds. □

4.3 Upper bound for the density of X_δ

Theorem 4.6. *Let Assumption 2.1 and 2.2 hold. Let p_{X_t} denote the density of X_t, $t>0$. Then, for any $p>1$, there exists a positive constant $C\in\mathcal{C}$ such that for every $\delta\leq 1$ and for every $y\in\mathbb{R}^n$*

$$p_{X_\delta}(y) \leq \frac{1}{\delta^{n-\frac{\dim\langle\sigma(0,x_0)\rangle}{2}}}\frac{C}{1+|y-x_0|^p_{A_\delta(0,x_0)}}.$$

Again, $\dim\langle\sigma(0,x_0)\rangle$ *denotes the dimension of the vector space spanned by $\sigma_1(0,x_0),\ldots,$ $\sigma_d(0,x_0)$.*

Proof. Set $F = T_\alpha(X_\delta - x_0)$. We apply estimate (D.7): there exist constants p and a depending only on the dimension n, such that

$$p_F(z) \leq C\max\{1,\mathbb{E}|\lambda_*(\gamma_F)|^{-p}\|F\|_{2,p}\}\mathbb{P}(|F-z|<2)^a.$$

We first show that $\|F\|_{2,p}\leq C\in\mathcal{C}$, as a consequence of Assumption 2.1. We prove just that $\|F\|_p\leq C$ for every p, for the Malliavin derivatives the proof is heavier but analogous. We write

$$F = T_\alpha\Big(\sum_{j=1}^d\int_0^\delta\sigma_j(t,X_t)\circ dW_t^j+\int_0^\delta b(t,X_t)dt\Big) = T_\alpha\Big(\sum_{j=1}^d\sigma_j(0,x_0)W_\delta^j + B_\delta\Big),$$

where

$$B_\delta = \sum_{j=1}^d\int_0^\delta\big(\sigma_j(t,X_t)-\sigma_j(0,x_0)\big)\circ dW_t^j + \int_0^\delta b(t,X_t)dt.$$

Therefore

$$|F| \leq \sum_{j=1}^d |T_\alpha\sigma_j(0,x_0)W_\delta^j| + |T_\alpha B_\delta|. \tag{4.16}$$

(4.4) implies $|T_\alpha\sigma_j(0,x_0)W_\delta^j|\leq CW_\delta^j/\sqrt{\delta}$, for $j=1,\ldots,d$. Moreover $|T_\alpha B_\delta|\leq|B_\delta|_{A_\delta}\leq C|B_\delta|/\delta$. If assumption 2.1 holds we conclude that $\mathbb{E}|F|^p\leq C\in\mathcal{C}$.

As in [2], Remark 2.4, it is easy to reduce the estimate of $\mathbb{P}(|F-z|<2)$ to the tail estimate of F, and then to use Markov inequality to relate the estimate of the tails to the moments of F:

$$\mathbb{P}(|F-z|<2) \leq \mathbb{P}(|F|>|z|/2) \leq C\frac{1\vee \mathbb{E}|F|^p}{1+|z|^p}, \quad \forall z \in \mathbb{R}^n \tag{4.17}$$

Since, from Assumption 2.1, all the moments of F are bounded by constants in $\mathcal{C}$, we have that for any exponent $p>1$ this term decays faster than $|z|^{-p}$ for $|z|\to\infty$.
In Lemma 4.5 we have already proved that $\mathbb{E}|\lambda_*(\gamma_F)|^{-q} \leq C \in \mathcal{C}$, for $\delta \leq 1$. We conclude that $p_F(z) \leq \frac{C}{1+|z|^p}$. The upper bound for the density of X_δ comes from the simple change of variable $y = x_0 + \alpha z$. For a positive and bounded measurable function $f:\mathbb{R}^n \to \mathbb{R}$, we write

$$\mathbb{E}f(X_\delta) = \mathbb{E}f(x_0+\alpha F) = \int f(x_0+\alpha z)p_F(z)dz$$

and we apply our density estimate, so that

$$\mathbb{E}f(X_\delta) \leq \int \frac{Cf(x_0+\alpha z)}{1+|z|^p}dz \leq \frac{C}{|\det \alpha|}\int \frac{f(y)}{1+|x_0-y|^p_{A_\delta(0,x_0)}}dy,$$

in which we have used (4.2). Concerning $|\det\alpha|$, we recall (3.43) and we obtain

$$p_{X_\delta}(y) \leq \frac{1}{\delta^{n-\frac{dim\langle\sigma(0,x_0)\rangle}{2}}}\frac{C}{1+|x_0-y|^p_{A_\delta(0,x_0)}}.$$

$\square$

Remark 4.7. If Assumption 2.3 holds then the upper estimate in Theorem 4.6 is of exponential type: there exists a constant $C\in\mathcal{C}$ such that for every $\delta\leq 1$ and for every $y\in\mathbb{R}^n$

$$p_{X_\delta}(y) \leq \frac{C}{\delta^{n-\frac{dim\langle\sigma(0,x_0)\rangle}{2}}}\exp(-\frac{1}{C}|y-x_0|_{A_\delta(0,x_0)}).$$

The proof is identical to the previous one except for the last part. In fact, looking at (4.16), in this case the boundedness of the coefficients allows one to apply the exponential martingale inequality, so instead of (4.17) we obtain the exponential bound $\mathbb{P}(|F|>|y|/2)\leq C\exp(-|y|/C)$. This actually gives the proof of (3) in Theorem 2.4.

Remark 4.8. In Theorem 3.7 the lower bound is centered at $x_0+\delta b(x_0)$ but for the upper estimate in Theorem 4.6, one can choose to center at x_0 or at $x_0+\delta b(x_0)$. In fact, in this case we notice that

$$|\delta b(x_0)|_{A_\delta(0,x_0)} \leq \frac{C'}{\delta}|\delta b(x_0)| \leq C'',$$

so

$$\frac{C_1}{1+|x_0-y|_{A_\delta(0,x_0)}} \leq \frac{C_2}{1+|x_0+\delta b(x_0)-y|_{A_\delta(0,x_0)}} \leq \frac{C_3}{1+|x_0-y|_{A_\delta(0,x_0)}},$$

and the estimate of Theorem 4.6 can be equivalently written as

$$p_{X_\delta}(y) \leq \frac{1}{\delta^{n-\frac{\dim\langle\sigma(0,x_0)\rangle}{2}}}\frac{C}{1+|y-x_0-\delta b(x_0)|^p_{A_\delta(0,x_0)}}.$$

Remark 4.9. Theorem 4.6 can be seen as an improvement of the upper bound in [11] in the sense that it precisely identifies the exponent $n - \frac{\dim\langle\sigma(0,x_0)\rangle}{2}$, which accounts of the time-scale of the heat kernel when δ goes to zero. This is evident when we consider the diagonal estimate $y = x_0$, and the same consideration holds when y is close to x_0. When looking at the tails (y far from x_0), it is not clear which of the two upper bounds is more accurate, unless we further specify the model.

A Proof of Lemma 3.1

We prove the decomposition (3.7) in Lemma 3.1. We recall Z_t in (3.2):

$$Z_t = \sum_{i=1}^{d} a_i W_t^i + \sum_{i,j=1}^{d} a_{i,j} \int_0^t W_s^i \circ dW_s^j$$

with $a_i = \sigma_i(0, x_0)$, $a_{i,j} = \partial_{\sigma_i}\sigma_j(0, x_0)$. Setting $s_l = \frac{l}{d}\,\delta$, $l = 1, \ldots, d$, we have

$$Z_\delta = \sum_{l=1}^{d} Z(s_l) - Z(s_{l-1}) = \sum_{l=1}^{d} \left(\sum_{i=1}^{d} a_i \Delta_l^i + \sum_{i,j=1}^{d} a_{i,j} \int_{s_{l-1}}^{s_l} W_s^i \circ dW_s^j \right).$$

Recalling the quantities Δ_l^j and $\Delta_l^{i,j}$ in (3.4), we write

$$\int_{s_{l-1}}^{s_l} W_s^i \circ dW_s^j = W_{s_{l-1}}^i \Delta_l^j + \Delta_l^{i,j} = (\sum_{p=1}^{l-1} \Delta_p^i)\Delta_l^j + \Delta_l^{i,j}.$$

Then

$$Z_\delta = \sum_{l=1}^{d}\sum_{i=1}^{d} a_i \Delta_l^i + \sum_{l=1}^{d}\sum_{i,j=1}^{d} a_{i,j} (\sum_{p=1}^{l-1} \Delta_p^i)\Delta_l^j + \sum_{l=1}^{d}\sum_{i,j=1}^{d} a_{i,j}\Delta_l^{i,j} =: S_1 + S_2 + S_3.$$

Notice first that

$$S_1 = \sum_{l=1}^{d} a_l \Delta_l^l + \sum_{l=1}^{d}\sum_{i\neq l} a_i \Delta_l^i.$$

We treat now S_3. We will use the identities

$$\left|\Delta_l^i\right|^2 = 2\Delta_l^{i,i} \quad \text{and} \quad \Delta_l^i \Delta_l^j = \Delta_l^{i,j} + \Delta_l^{j,i}.$$

Then

$$\begin{aligned}
S_3 &= \sum_{l=1}^{d}\sum_{i=1}^{d} a_{i,i}\Delta_l^{i,i} + \sum_{l=1}^{d}\sum_{i\neq j} a_{i,j}\Delta_l^{i,j} \\
&= \sum_{l=1}^{d}\sum_{i=1}^{d} a_{i,i}\Delta_l^{i,i} + \sum_{l=1}^{d}\sum_{i\neq l} a_{i,l}\Delta_l^{i,l} + \sum_{l=1}^{d}\sum_{j\neq l} a_{l,j}\Delta_l^{l,j} + \sum_{l=1}^{d}\sum_{i\neq j, i\neq l j\neq l} a_{i,j}\Delta_l^{i,j} \\
&= \frac{1}{2}\sum_{l=1}^{d}\sum_{i=1}^{d} a_{i,i}\left|\Delta_l^i\right|^2 + \sum_{l=1}^{d}\sum_{i\neq l} a_{i,l}\Delta_l^{i,l} \\
&\quad + \sum_{l=1}^{d}\sum_{j\neq l} a_{l,j}\left(\Delta_l^j\Delta_l^l - \Delta_l^{j,l}\right) + \sum_{l=1}^{d}\sum_{i\neq j, i\neq l, j\neq l} a_{i,j}\Delta_l^{i,j} \\
&= \frac{1}{2}\sum_{i=1}^{d} a_{i,i}\left|\Delta_i^i\right|^2 + \frac{1}{2}\sum_{l=1}^{d}\sum_{i\neq l}^{d} a_{i,i}\left|\Delta_l^i\right|^2 + \sum_{l=1}^{d}\sum_{i\neq l}(a_{i,l}-a_{l,i})\Delta_l^{i,l} \\
&\quad + \sum_{l=1}^{d}\left(\sum_{j\neq l} a_{l,j}\Delta_l^j\right)\Delta_l^l + \sum_{l=1}^{d}\sum_{i\neq j, i\neq l, \neq j\neq} a_{i,j}\Delta_l^{i,j}.
\end{aligned}$$

We treat now S_2. We want to emphasize the terms containing Δ_i^i. We have

$$S_2 = \sum_{l>p}^{d}\sum_{i,j=1}^{d} a_{i,j}\Delta_p^i\Delta_l^j = S_2' + S_2'' + S_2''' + S_2^{iv}$$

with

$$\begin{aligned}
S_2' &= \sum_{l>p}^{d} a_{p,l}\Delta_p^p\Delta_l^l, \quad S_2'' = \sum_{l>p}^{d}\sum_{j\neq l} a_{p,j}\Delta_p^p\Delta_l^j \\
S_2''' &= \sum_{l>p}^{d}\sum_{i\neq p}^{d} a_{i,l}\Delta_p^i\Delta_l^l, \quad S_2^{iv} = \sum_{l>p}^{d}\sum_{i\neq p, j\neq l} a_{i,j}\Delta_p^i\Delta_l^j.
\end{aligned}$$

We have

$$S_2'' = \sum_{p=1}^{d}\Delta_p^p\left(\sum_{l=p+1}^{d}\sum_{j\neq l} a_{p,j}\Delta_l^j\right)$$

and

$$S_2''' = \sum_{l=1}^{d}\Delta_l^l\left(\sum_{p=1}^{l-1}\sum_{i\neq p} a_{i,l}\Delta_p^i\right) = \sum_{p=1}^{d}\Delta_p^p\left(\sum_{l=1}^{p-1}\sum_{j\neq l} a_{j,p}\Delta_l^j\right)$$

so that

$$S_2'' + S_2''' = \sum_{p=1}^{d}\Delta_p^p\left(\sum_{l=p+1}^{d}\sum_{j\neq l} a_{p,j}\Delta_l^j + \sum_{l=1}^{p-1}\sum_{j\neq l} a_{j,p}\Delta_l^j\right).$$

Finally

$$\begin{aligned}
Z_\delta &= \sum_{l=1}^{d} a_l \Delta_l^l + \sum_{l=1}^{d} \sum_{i \neq l} a_i \Delta_l^i \\
&\quad + \sum_{l>p}^{d} a_{p,l} \Delta_p^p \Delta_l^l + \sum_{p=1}^{d} \Delta_p^p \left(\sum_{l>p}^{d} \sum_{j \neq l} a_{p,j} \Delta_l^j + \sum_{p>l}^{d} \sum_{j \neq l} a_{j,p} \Delta_l^j \right) \\
&\quad + \sum_{l>p}^{d} \sum_{i \neq p, j \neq l} a_{i,j} \Delta_p^i \Delta_l^j + \frac{1}{2} \sum_{i=1}^{d} a_{i,i} \left| \Delta_i^i \right|^2 + \frac{1}{2} \sum_{l=1}^{d} \sum_{i \neq l} a_{i,i} \left| \Delta_l^i \right|^2 \\
&\quad + \sum_{l=1}^{d} \sum_{i \neq l} (a_{i,l} - a_{l,i}) \Delta_l^{i,l} + \sum_{l=1}^{d} \left(\sum_{j \neq l} a_{l,j} \Delta_l^j \right) \Delta_l^l + \sum_{l=1}^{d} \sum_{i \neq j, i \neq l, j \neq l} a_{i,j} \Delta_l^{i,j}.
\end{aligned}$$

We want to compute the coefficient of Δ_p^p : this term appears in $\sum_{p=1}^{d} \Delta_p^p (a_p + \varepsilon_p)$, with

$$\varepsilon_p = \sum_{l>p}^{d} \sum_{j \neq l} a_{p,j} \Delta_l^j + \sum_{p>l}^{d} \sum_{j \neq l} a_{j,p} \Delta_l^j + \sum_{j \neq p} a_{p,j} \Delta_p^j.$$

We consider now $\Delta_p^{i,p}$. It appears in

$$\sum_{p=1}^{d} \sum_{i \neq p} (a_{i,p} - a_{p,i}) \Delta_p^{i,p}$$

The vector $a_{i,p} - a_{p,i}$ corresponds to the bracket $[\sigma_i, \sigma_p](0,x)$. Notice that for $l = l(i,p)$ when $i \neq p$, then $[\sigma_i, \sigma_p](0,x) = A_l(0,x)$, $A_l(0,x)$ being the lth column of $A(0,x)$. The other terms are

$$\begin{aligned}
&\sum_{l=1}^{d} \sum_{i \neq l} a_i \Delta_l^i + \sum_{l>p}^{d} \sum_{i \neq p, j \neq l} a_{i,j} \Delta_p^i \Delta_l^j + \frac{1}{2} \sum_{i=1}^{d} a_{i,i} \left| \Delta_i^i \right|^2 + \frac{1}{2} \sum_{l=1}^{d} \sum_{i \neq l} a_{i,i} \left| \Delta_l^i \right|^2 \\
&+ \sum_{l=1}^{d} \sum_{i \neq j, i \neq l, j \neq l} a_{i,j} \Delta_l^{i,j} + \sum_{l=p+1}^{d} a_{p,l} \Delta_p^p \Delta_l^l.
\end{aligned}$$

We put everything together and (3.7) is proved.

B Support property

The aim of this section is the proof of the inequality in (B.5), which has been strongly used in Lemma 3.3.

Let $B = (B^1, ..., B^{d-1})$ be a standard Brownian motion. We consider the analogous of the covariance matrix $Q_i(B)$ considered in Section 3.1: we define a symmetric square matrix of dimension $d \times d$ by

$$\begin{aligned}
&Q^{d,d} = 1, \quad Q^{d,j} = Q^{j,d} = \textstyle\int_0^1 B_s^j ds, \quad j = 1, ..., d-1, \\
&Q^{j,p} = Q^{p,j} = \textstyle\int_0^1 B_s^j B_s^p ds, \quad j, p = 1, ..., d-1
\end{aligned} \tag{B.1}$$

and we denote by $\lambda_*(Q)$ (respectively by $\lambda^*(Q)$) the lowest (respectively largest) eigenvalue of Q.
For a measurable function $g : [0,1] \to R^{d-1}$ we denote

$$\begin{aligned}\alpha_g(\xi) &= \xi_d + \int_0^1 \langle g_s, \xi_* \rangle \, ds, \quad \beta_g(\xi) = \int_0^1 \langle g_s, \xi_* \rangle^2 \, ds - \left(\int_0^1 \langle g_s, \xi_* \rangle \, ds \right)^2 \quad \textit{with} \\ \xi &= (\xi_1, ..., \xi_d) \in \mathbb{R}^d \quad \textit{and} \quad \xi_* = (\xi_1, ..., \xi_{d-1}).\end{aligned}$$

We need the following two preliminary lemmas.

Lemma B.1. *With $g(s) = B_s, s \in [0,1]$ we have*

$$\langle Q\xi, \xi \rangle = \alpha_B^2(\xi) + \beta_B(\xi).$$

As a consequence, one has

$$\lambda_*(Q) = \inf_{|\xi|=1} (\alpha_B^2(\xi) + \beta_B(\xi)) \quad \textit{and} \quad \lambda^*(Q) \le \sup_{|\xi|=1} (\alpha_B^2(\xi) + \beta_B(\xi)) \le \big(1 + \sup_{t \le 1} |B_t|\big)^2.$$

Taking $\xi_* = 0$ and $\xi_d = 1$ we obtain $\langle Q\xi, \xi \rangle = 1$ so that $\lambda_*(Q) \le 1 \le \lambda^*(Q)$.

Proof. By direct computation

$$\begin{aligned}\langle Q\xi, \xi \rangle &= \xi_d^2 + 2\xi_d \int_0^1 \langle B_s, \xi_* \rangle \, ds + \left(\int_0^1 \langle B_s, \xi_* \rangle \, ds) \right)^2 \\ &\quad + \int_0^1 \langle B_s, \xi_* \rangle^2 \, ds - \left(\int_0^1 \langle B_s, \xi_* \rangle \, ds \right)^2 \\ &= \left(\xi_d + \int_0^1 \langle B_s, \xi_* \rangle \, ds \right)^2 + \int_0^1 \langle B_s, \xi_* \rangle^2 \, ds - \left(\int_0^1 \langle B_s, \xi_* \rangle \, ds \right)^2.\end{aligned}$$

The remaining statements follow straightforwardly. $\square$

Proposition B.2. *For each $p \ge 1$ one has*

$$\mathbb{E}(|\det Q|^{-p}) \le C_{p,d} < \infty \tag{B.2}$$

where $C_{p,d}$ is a constant depending on p, d only.

Proof. By Lemma 7-29, pg 92 in [6], for every $p \in (0, \infty)$ one has

$$\frac{1}{|\det Q|^p} \le \frac{1}{\Gamma(p)} \int_{R^d} |\xi|^{d(2p-1)} \, e^{-\langle Q\xi, \xi \rangle} d\xi.$$

Let $\theta(\xi_*) := \int_0^1 \langle B_s, \xi_* \rangle \, ds$. Using the previous lemma

$$\begin{aligned}\int_{R^d} |\xi|^{d(2p-1)} \, e^{-\langle Q\xi, \xi \rangle} d\xi &= \int_{R^d} (\xi_d^2 + |\xi_*|^2)^{d(2p-1)/2} e^{-(\xi_d + \theta(\xi_*))^2 - \beta_B(\xi_*)} d\xi \\ &\le C \int_{R^{d-1}} ((1 + \theta^2(\xi_*))^{d(2p-1)/2} + |\xi_*|^{d(2p-1)}) e^{-\beta_B(\xi_*)} d\xi_* \\ &\le C \int_{R^{d-1}} \sup_{t \le 1} 1 \vee |B_t|^{d(2p-1)} \, (1 + |\xi_*|^{d(2p-1)+1}) e^{-\beta_B(\xi_*)} d\xi_*.\end{aligned}$$

We integrate and we use Schwartz inequality in order to obtain

$$\mathbb{E}\Big(\frac{1}{|\det Q|^p}\Big) \leq C + C\int_{\{|\xi_*|\geq 1\}} (\mathbb{E}((1+|\xi_*|^{d(2p-1)+1})^2 e^{-2\beta_B(\xi_*)}))^{1/2} d\xi_*.$$

For each fixed ξ_* the process $b_{\xi_*}(t) := |\xi_*|^{-1}\langle B_t, \xi_*\rangle$ is a standard Brownian motion and $\beta_B(\xi_*) = |\xi_*|^2 \int_0^1 (b_{\xi_*}(t) - \int_0^1 b_{\xi_*}(s)ds)^2 dt =: |\xi_*|^2 V_{\xi_*}$ where V_{ξ_*} is the variance of b_{ξ_*} with respect to the time. Then it is proved in [9] (see (1.f), p. 183) that

$$\mathbb{E}(e^{-2\beta_B(\xi_*)}) = \mathbb{E}(e^{-2|\xi_*|^2 V_{\xi_*}}) = \frac{2|\xi_*|^2}{\sinh 2|\xi_*|^2}.$$

We insert this in the previous inequality and we obtain $\mathbb{E}(|\det Q|^{-p}) < \infty$. $\square$

We are now able to give the main result in this section. We define

$$q(B) = \sum_{i=1}^{d-1} |B_1^i| + \sum_{j\neq p} \left|\int_0^1 B_s^j dB_s^p\right| \tag{B.3}$$

and for $\varepsilon, \rho > 0$ we denote

$$\Upsilon_{\rho,\varepsilon}(B) = \{\det Q \geq \varepsilon^\rho, \sup_{t\leq 1}|B_t| \leq \varepsilon^{-\rho}, q(B) \leq \varepsilon\}. \tag{B.4}$$

Proposition B.3. *There exist some universal constants $c_{\rho,d}, \varepsilon_{\rho,d} \in (0,1)$ (depending on ρ and d only) such that for every $\varepsilon \in (0, \varepsilon_{\rho,d})$ one has*

$$\mathbb{P}(\Upsilon_{\rho,\varepsilon}(B)) \geq c_{\rho,d} \times \varepsilon^{\frac{1}{2}d(d+1)}. \tag{B.5}$$

Proof. Using the previous proposition and Chebyshev's inequality we get

$$\mathbb{P}(\det Q < \varepsilon^\rho) \leq \varepsilon^{p\rho}\mathbb{E}\,|\det Q|^{-p} \leq C_{p,d}\varepsilon^{p\rho} \quad \text{and} \quad \mathbb{P}(\sup_{t\leq 1}|B_t| > \varepsilon^{-\rho}) \leq \exp(-\frac{1}{C\varepsilon^{2\rho}}).$$

Let $q'(B) = \sum_{i=1}^{d-1}|B_1^i| + \sum_{j<p}\left|\int_0^1 B_s^j dB_s^p\right|$. Since $\left|\int_0^1 B_s^j dB_s^p\right| \leq \left|B_1^j\right| |B_1^p| + \left|\int_0^1 B_s^p dB_s^j\right|$ we have $q(B) \leq 2q'(B) + q'(B)^2$ so that $\{q'(B) \leq \frac{1}{3}\varepsilon\} \subset \{q(B) \leq \varepsilon\}$. We will now use the following fact: consider the diffusion process $X = (X^i, X^{j,p},\ i = 1,...,d, 1 \leq j < p \leq d)$ solution of the equation $dX_t^i = dB_t^i, dX_t^{j,p} = X_t^j dB_t^p$. The strong Hörmander condition holds for this process and the support of the law of X_1 is the whole space. So the law of X_1 is absolutely continuous with respect to the Lebesgue measure and has a continuous and strictly positive density p. This result is well known (see for example [12] or [1]). We denote $c_d := \inf_{|x|\leq 1} p(x) > 0$ and this is a constant which depends on d only. Then, by observing that $q'(B) \leq \sqrt{m}\,|X_1|$, where $m = \frac{1}{2}d(d+1)$ is the dimension of the diffusion X, we get

$$\mathbb{P}(q(B) \leq \varepsilon) \geq \mathbb{P}\Big(q'(B) \leq \frac{\varepsilon}{3}\Big) \geq \mathbb{P}\Big(|X_1| \leq \frac{\varepsilon}{3\sqrt{m}}\Big) \geq \frac{\varepsilon^m}{(3\sqrt{m})^m} \times \bar{c}_d,$$

with $\bar{c}_d > 0$. So finally we obtain

$$\mathbb{P}(\Upsilon_{\rho,\varepsilon}(B)) \geq \bar{c}_d \varepsilon^{\frac{1}{2}d(d+1)} - C_{p,d}\varepsilon^{p\rho} - \exp(-\frac{1}{C\varepsilon^{2\rho}}).$$

Choosing $p > \frac{1}{2\rho}d(d+1)$ and ε small we obtain our inequality. $\square$

C Density estimates via local inversion

In this section we see how to use the inverse function theorem to transfer a known estimate for a Gaussian random variable to its image via a certain function η. For a standard version of the inverse function theorem see [16].
We consider $\Phi(\theta) = \theta + \eta(\theta)$, for a three times differentiable function $\eta : \mathbb{R}^m \to \mathbb{R}^m$. Define

$$c_2(\eta) = \max_{i,j=1,..,m} \sup_{|x|\leq 1} |\partial^2_{ij}\eta(x)|, \quad c_3(\eta) = \max_{i,j,k=1,..,m} \sup_{|x|\leq 1} |\partial^3_{ijk}\eta(x)|, \tag{C.1}$$

and

$$h_\eta = \frac{1}{16m^2(c_2(\eta) + \sqrt{c_3(\eta)})} \tag{C.2}$$

Lemma C.1. *Take h_η as above. If the function η is such that*

$$\eta \in C^3(\mathbb{R}^m, \mathbb{R}^m), \quad \eta(0) = 0, \quad \nabla\eta(0) \leq \frac{1}{2},$$

then there exists a neighborhood of 0, that we denote with $V_{h_\eta} \subset B(0, 2h_\eta)$, such that $\Phi : V_{h_\eta} \to B\left(0, \frac{1}{2}h_\eta\right)$ is a diffeomorphism. In particular, if we denote with Φ^{-1} the local inverse of Φ, we have

$$\Phi^{-1} : B\left(0, \frac{1}{2}h_\eta\right) \to B\left(0, 2h_\eta\right),$$

and we have this quantitative estimate:

$$\forall y \in B\left(0, \frac{1}{2}h_\eta\right), \quad \frac{1}{4}|\Phi^{-1}(y)| \leq |y| \leq 4|\Phi^{-1}(y)|. \tag{C.3}$$

Remark C.2. Here we write Φ^{-1} for the inverse of the restriction of Φ to V_{h_η}, what is called a *local* inverse.

Proof. We have

$$\nabla\Phi(0) = \mathrm{Id} + \nabla\eta(0).$$

So

$$|\nabla\Phi(0)x|^2 \geq \frac{1}{2}|x|^2 - |\nabla\eta(0)x|^2 \geq \frac{1}{2}|x|^2 - \frac{1}{4}|x|^2 = \frac{1}{4}|x|^2.$$

and

$$|\nabla\Phi(0)x|^2 \leq 2|x|^2 + 2|\nabla\eta(0)x|^2 \leq 2|x|^2 + \frac{1}{2}|x|^2 \leq \frac{5}{2}|x|^2.$$

Therefore

$$\frac{1}{2}|x| \leq |\nabla\Phi(0)x| \leq \sqrt{3}|x|$$

This implies $\Phi(0)$ is invertible locally around 0, and the local inverse differentiable, using the classical inverse function theorem. We now look now at the image of the inverse, and at the estimates (C.3). We develop η around 0, writing $\nabla^2\eta(x)[u, v]$ to denote $\nabla^2\eta(x)$ computed in u and v.

$$\eta(\theta) = \nabla\eta(0)\theta + \int_0^1 (1-t)\nabla^2\eta(t\theta)[\theta, \theta]dt.$$

Fix $y \in \mathbb{R}^m$. Suppose $\Phi(\theta) = y$. Then

$$\begin{aligned}\theta &= (\nabla\Phi(0))^{-1}\nabla\Phi(0)\theta \\ &= (\nabla\Phi(0))^{-1}(\theta + \nabla\eta(0)\theta) \\ &= (\nabla\Phi(0))^{-1}\left(\theta + \eta(\theta) - \int_0^1 (1-t)\nabla^2\eta(t\theta)[\theta,\theta]dt\right) \\ &= (\nabla\Phi(0))^{-1}\left(y - \int_0^1 (1-t)\nabla^2\eta(t\theta)[\theta,\theta]dt\right).\end{aligned}$$

We define

$$U_y(\theta) = \left(y - \int_0^1 (1-t)\nabla^2\eta(t\theta)[\theta,\theta]dt\right),$$

so that θ can be seen as a fixed point for U_y. Recall that $|\frac{1}{2}x| \leq |\nabla\Phi(0)x|$.

$$\begin{aligned}|U_y(\theta_1) - U_y(\theta_2)| &= \left|(\nabla\Phi(0))^{-1}\left(\int_0^1 (1-t)(\nabla^2\eta(t\theta_2)[\theta_2,\theta_2] - \nabla^2\eta(t\theta_1)[\theta_1,\theta_1])dt\right)\right| \\ &\leq 2\left|\int_0^1 (1-t)(\nabla^2\eta(t\theta_2)[\theta_2,\theta_2] - \nabla^2\eta(t\theta_1)[\theta_1,\theta_1])dt\right| \\ &\leq 2\int_0^1 (1-t)(|\nabla^2\eta(t\theta_1)[\theta_1,\theta_1-\theta_2]| + |\nabla^2\eta(t\theta_1)[\theta_1-\theta_2,\theta_2]| \\ &+ |\nabla^2\eta(t\theta_1)[\theta_2,\theta_2] - \nabla^2\eta(t\theta_2)[\theta_2,\theta_2]|)dt.\end{aligned}$$

Now, form (C.2), for $\theta_1, \theta_2 \in B(0,h_\eta)$

$$|\nabla^2\eta(t\theta_1)[\theta_1,\theta_1-\theta_2]| \leq m^2 c_2(\eta)h_\eta|\theta_1-\theta_2| \leq \frac{1}{16}|\theta_1-\theta_2|,$$

and

$$|\nabla^2\eta(t\theta_1)[\theta_2,\theta_2] - \nabla^2\eta(t\theta_2)[\theta_2,\theta_2]| \leq m^3 c_3(\eta)|\theta_1-\theta_2|h_\eta^2 \leq \frac{1}{256}|\theta_1-\theta_2|,$$

and therefore

$$|U_y(\theta_1) - U_y(\theta_2)| \leq \frac{1}{4}|\theta_1-\theta_2|. \tag{C.4}$$

For $y \in B(0,\frac{1}{2}h_\eta)$ and $\theta \in B(0,2h_\eta)$ this implies

$$|U_y(\theta)| \leq |U_y(\theta) - U_y(0)| + |U_y(0)| \leq \frac{1}{4}|\theta| + 2y \leq 2h_\eta$$

Define now the sequence

$$\theta_0 = 0, \quad \theta_{k+1} = U_y(\theta_k).$$

We know that $\theta_k \in B(0,2h_\eta)$ for any $k \in \mathbb{N}$, and therefore inequality (C.4) implies

$$|U_y(\theta_k) - U_y(\theta_{k+1})| \leq \frac{1}{4}|\theta_k - \theta_{k+1}|.$$

The Banach fixed-point theorem tells us that θ_k converges to the unique solution of $\theta = U_y(\theta)$, which is $\theta = \Phi^{-1}(y)$, and $\theta \in B(0, 2h_\eta)$. So it is possible to define the local inverse Φ^{-1} on $B\left(0, \frac{1}{2}h_\eta\right)$, and

$$V_{h_\eta} := \Phi^{-1} B\left(0, \frac{1}{2}h_\eta\right) \subset B(0, 2h_\eta).$$

Now, for $y \in B(0, \frac{1}{2}h_\eta)$, let $\theta = \Phi^{-1}(y)$ and the following inequalities hold

$$\begin{aligned} |\theta| &= |U_y(\theta)| \leq \frac{1}{2}\theta + 2|y| && \Rightarrow |\theta| \leq 4|y| \\ |\theta| &= U_y(\theta) \geq |U_y(0)| - |U_y(\theta) - U_y(0)| \geq \frac{1}{2}|y| - \frac{1}{2}|\theta| && \Rightarrow |\theta| \geq \frac{1}{4}|y|. \end{aligned}$$

□

Let now Θ be a m-dimensional centered Gaussian variable with covariance matrix Q. Denote by $\underline{\lambda}$ and $\overline{\lambda}$ the lowest and the largest eigenvalues of Q. Keeping in mind the setting of the last subsection, we also introduce the notation

$$c_*(\eta, h) = \sup_{|x| \leq 2h} \max_{i,j} |\partial_i \eta^j(x)| \tag{C.5}$$

for $h > 0$. Recall we are supposing $\eta \in C^3(\mathbb{R}^m, \mathbb{R}^m)$ and $\eta(0) = 0$.
Take $r > 0$ such that

$$c_*(\eta, 16r) \leq \frac{1}{2m}\sqrt{\frac{\underline{\lambda}}{\overline{\lambda}}}, \qquad r \leq h_\eta = \frac{1}{16m^2(c_2(\eta) + \sqrt{c_3(\eta)})}. \tag{C.6}$$

We take a localizing function as in (D.3):

$$U = \prod_{i=1}^{m} \psi_r(\Theta_i). \tag{C.7}$$

Lemma C.3. *Let Q be non degenerate. Let r such that* (C.6) *holds and set U as in* (C.7)*. Then the density $p_{G,U}$ of*

$$G := \Phi(\Theta) = \Theta + \eta(\Theta)$$

under $\mathbb{P}_U$ has the following bounds on $B(0, r)$:

$$\frac{1}{C \det Q^{1/2}} \exp\left(-\frac{C}{\underline{\lambda}}|z|^2\right) \leq p_{G,U}(z) \leq \frac{C}{\det Q^{1/2}} \exp\left(-\frac{1}{C\overline{\lambda}}|z|^2\right) \tag{C.8}$$

Proof. For a general nonnegative, measurable function $f : \mathbb{R}^m \to \mathbb{R}$ with support included in $B(0, 4r)$, we compute $\mathbb{E}(f(G)1_{\{\Theta \in \Phi^{-1}B(0,4r)\}})$. Here Φ^{-1} is the local diffeomorphism of the inverse function theorem. After the multiplication with the characteristic function, on the support of the random variable that we are averaging, Φ is a diffeomorphism and the first equality holds. The second follows from the change of variable suggested by Lemma C.1 for $G = \Phi(\Theta)$

$$\begin{aligned} &\mathbb{E}\left(f(G)1_{\{\Theta \in \Phi^{-1}B(0,4r)\}}\right) = \\ &= \int_{\Phi^{-1}(B(0,4r))} f(\Phi(\theta)) \frac{1}{(2\pi)^{m/2} \det Q^{1/2}} \exp\left(-\frac{1}{2}\langle Q^{-1}\theta, \theta\rangle\right) d\theta \\ &= \int_{B(0,4r)} f(z)\bar{p}_G(z) dz, \end{aligned}$$

where for $z \in B(0, 4r)$

$$\bar{p}_G(z) = \frac{1}{(2\pi)^{m/2} \det Q^{1/2} |\det \nabla\Phi(\Phi^{-1}(z))|} \exp\left(-\frac{1}{2}\langle Q^{-1}\Phi^{-1}(z), \Phi^{-1}(z)\rangle\right).$$

Again from Lemma C.1, since $4r \leq \frac{h_\eta}{2}$, we have $z \in B(0,4r) \Rightarrow \theta \in B(0,16r)$. Using $c_*(\eta, 16r) \leq \frac{1}{2m}\sqrt{\frac{\underline{\lambda}}{\overline{\lambda}}}$,

$$\frac{1}{2}|x|^2 \leq (1 - m\, c_*(\eta, h_\eta))|x|^2 \leq |\langle \nabla\Phi(\theta)x, x\rangle| \leq (1 + m\, c_*(\eta, h_\eta))|x|^2 \leq 2|x|^2.$$

Therefore if $z \in B(0, 4r)$

$$2^{-m} \leq |\det \Phi(\Phi^{-1}(z))| \leq 2^m.$$

Moreover, using Lemma C.1

$$\begin{aligned}
\langle Q^{-1}\Phi^{-1}(z), \Phi^{-1}(z)\rangle &\leq \frac{1}{\underline{\lambda}}|\Phi^{-1}(z)|^2 \leq \frac{16}{\underline{\lambda}}|z|^2, \\
\langle Q^{-1}\Phi^{-1}(z), \Phi^{-1}(z)\rangle &\geq \frac{1}{\overline{\lambda}}|\Phi^{-1}(z)|^2 \geq \frac{1}{16\overline{\lambda}}|z|^2.
\end{aligned}$$

Therefore

$$\frac{1}{(8\pi)^{m/2} \det Q^{1/2}} \exp\left(-\frac{8}{\underline{\lambda}}|z|^2\right) \leq \bar{p}_G(z) \leq \frac{2^{m/2}}{\pi^{m/2} \det Q^{1/2}} \exp\left(-\frac{1}{32\overline{\lambda}}|z|^2\right).$$

Now we define, as in (D.3) the localization variables

$$U_1 = \prod_{i=1}^m \psi_{16r}(\Theta_i), \quad U_2 = \prod_{i=1}^m \psi_r(\Theta_i).$$

Notice that

$$U_2 \leq 1_{\{\Theta \in \Phi^{-1}B(0,4r)\}} \leq U_1,$$

so that we have

$$\mathbb{E}\left(f(G)U_2\right) \leq \mathbb{E}\left(f(G)1_{\{\Theta \in \Phi^{-1}B(0,4r)\}}\right) \leq \mathbb{E}\left(f(G)U_1\right).$$

The following bounds for the local densities follow:

$$\begin{aligned}
p_{G,U_1}(z) &\geq \frac{1}{(8\pi)^{m/2} \det Q^{1/2}} \exp\left(-\frac{8}{\underline{\lambda}}|z|^2\right), \\
p_{G,U_2}(z) &\leq \frac{2^{m/2}}{\pi^{m/2} \det Q^{1/2}} \exp\left(-\frac{1}{32\overline{\lambda}}|z|^2\right).
\end{aligned}$$

$U_1 \geq U = U_2$, so for the localization via U both bounds hold. $\square$

D Estimates of the distance between localized densities

D.1 Elements of Malliavin calculus

We recall some basic notions in Malliavin calculus. Our main reference is [14]. We consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a Brownian motion $W = (W_t^1, ..., W_t^d)_{t\geq 0}$ and the filtration $(\mathcal{F}_t)_{t\geq 0}$ generated by W. For fixed $T > 0$, we denote by $\mathcal{H}$ the Hilbert space $L^2([0,T], \mathbb{R}^d)$. For $h \in \mathcal{H}$ we introduce this notation for the Itô integral of h: $W(h) = \sum_{j=1}^d \int_0^T h^j(s) dW_s^j$. We denote by $C_p^\infty(\mathbb{R}^n)$ the set of all infinitely continuously differentiable functions $f : \mathbb{R}^n \to \mathbb{R}$ such that f and all of its partial derivatives have polynomial growth. We also denote by $\mathcal{S}$ the class of simple random variables of the form

$$F = f(W(h_1), ..., W(h_n)),$$

for some $f \in C_p^\infty(\mathbb{R}^n)$, $h_1, ..., h_n$ in $\mathcal{H}$, $n \geq 1$. The Malliavin derivative of $F \in \mathcal{S}$ is the $\mathcal{H}$ valued random variable given by

$$DF = (DF^1, \dots, DF^d)^T = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(W(h_1), ..., W(h_n))h_i. \tag{D.1}$$

We introduce the Sobolev norm of F:

$$\|F\|_{1,p} = [\mathbb{E}|F|^p + \mathbb{E}|DF|^p]^{\frac{1}{p}}$$

where

$$|DF| = \left(\int_0^T |D_s F|^2 ds\right)^{\frac{1}{2}}.$$

It is possible to prove that D is a closable operator and take the extension of D in the standard way. We can now define in the obvious way DF for any F in the closure of $\mathcal{S}$ with respect to this norm. Therefore, the domain of D will be the closure of $\mathcal{S}$.
The higher order derivative of F is obtained by iteration. For any $k \in \mathbb{N}$, for a multi-index $\alpha = (\alpha_1, ..., \alpha_k) \in \{1, ..., d\}^k$ and $(s_1, ..., s_k) \in [0,T]^k$, we can define

$$D^\alpha_{s_1,...,s_k} F := D^{\alpha_1}_{s_1} ... D^{\alpha_k}_{s_k} F.$$

We denote by $|\alpha| = k$ the length of the multi-index. Remark that $D^\alpha_{s_1,...,s_k} F$, is a random variable with values in $\mathcal{H}^{\otimes k}$, and so we define its Sobolev norm as

$$\|F\|_{k,p} = [\mathbb{E}|F|^p + \sum_{j=1}^k \mathbb{E}|D^{(j)} F|^p]^{\frac{1}{p}}$$

where

$$|D^{(j)} F| = \left(\sum_{|\alpha|=j} \int_{[0,T]^j} |D^\alpha_{s_1,...,s_j} F|^2 ds_1 ... ds_j\right)^{1/2}.$$

The extension to the closure of $\mathcal{S}$ with respect to this norm is analogous to the first order derivative. Notice that with this notation $|DF| = |D^{(1)} F|$. Also notice that $D^{(j)}$ means "derivative of order j" and D^j means "derivative with respect to W^j".

We denote by $\mathbb{D}^{k,p}$ the space of the random variables which are k times differentiable in the Malliavin sense in L^p, and $\mathbb{D}^{k,\infty} = \bigcap_{p=1}^{\infty} \mathbb{D}^{k,p}$. As usual, we also denote by L the Ornstein-Uhlenbeck operator, i.e. $L = -\delta \circ D$, where δ is the adjoint operator of D.
We consider random vector $F = (F_1, ..., F_n)$ in the domain of D. We define its *Malliavin covariance matrix* as follows:

$$\gamma_F^{i,j} = \langle DF_i, DF_j \rangle_{\mathcal{H}} = \sum_{k=1}^{d} \int_0^T D_s^k F_i \times D_s^k F_j ds.$$

D.2 Localization and density estimates

The following notion of localization is introduced in [1]. Consider a random variable $U \in [0, 1]$ and denote

$$d\mathbb{P}_U = U d\mathbb{P}.$$

$\mathbb{P}_U$ is a non-negative measure (not a probability measure, in general). We also set $\mathbb{E}_U$ the expectation (integral) w.r.t. $\mathbb{P}_U$, and denote

$$\|F\|_{p,U}^p = \mathbb{E}_U(|F|^p) = \mathbb{E}(|F|^p U)$$
$$\|F\|_{k,p,U}^p = \|F\|_{p,U}^p + \sum_{j=1}^{k} \mathbb{E}_U(|D^{(j)} F|^p).$$

We assume that $U \in \mathbb{D}^{1,\infty}$ and for every $p \geq 1$

$$m_U(p) := 1 + \mathbb{E}_U |D \ln U|^p < \infty. \tag{D.2}$$

The specific localizing function we will use is the following. Consider the function depending on a parameter $a > 0$:

$$\psi_a(x) = 1_{|x| \leq a} + \exp\left(1 - \frac{a^2}{a^2 - (x-a)^2}\right) 1_{a < |x| < 2a}.$$

For $\Theta_i \in \mathbb{D}^{2,\infty}$ and $a_i > 0$, $i = 1 \ldots, n$ we define the localization variable:

$$U = \prod_{i=1}^{n} \psi_{a_i}(\Theta_i) \tag{D.3}$$

For this choice of U we have that for any $p, k \in \mathbb{N}_0$

$$m_U(p) \leq C \frac{\|\Theta\|_{1,p}^p}{|a|} \tag{D.4}$$

The proof of (D.4) follows from standard computations and inequality

$$\sup_x |(\ln \psi_a)(x)|^p \psi_a(x) \leq \frac{C}{a^p} \tag{D.5}$$

In the following proposition we state the general lower and upper bound that we use in our density estimate. These results come from [1] and [2].

Proposition D.1. *Let $F \in (\mathbb{D}^{2,\infty})^d$.*

(1) *Suppose that for every $p \in \mathbb{N} : \mathbb{E}_U|\lambda_*(\gamma_F)|^{-p} < \infty$, $U \in \mathbb{D}^{1,\infty}$ and $m_U(p) < \infty$. Let $G \in (\mathbb{D}^{2,\infty})^d$ such that for every $p \in \mathbb{N}$*

$$\mathbb{E}_U|\lambda_*(\gamma_G)|^{-p} < \infty.$$

Then for every $p > d$

$$p_{F,U}(y) \geq p_{G,U}(y) - Cm_U(p)^b \max\left\{1, (\mathbb{E}_U|\lambda_*(\gamma_G)|^{-p})^b(\|F\|_{2,p,U} + \|G\|_{2,p,U})\right\}\|F - G\|_{2,p,U} \tag{D.6}$$

where C, b are constants depending only on d, p and $m_U(p)$ is given by (D.2).

(2) *Assume $\mathbb{E}|\lambda_*(\gamma_F)|^{-p} < \infty$, $\forall p$. Then $\exists C, p, b$ constants depending only on the dimension d such that*

$$|p_F(y)| \leq C \max\{1, \mathbb{E}|\lambda_*(\gamma_F)|^{-p}\|F\|_{2,p}\}\mathbb{P}(|F - y| < 2)^b \tag{D.7}$$

Proof. (1) The lower bound (D.6) for $p_{F,U}$ is a version of Proposition 2.5. in [1] with the lowest eigenvalue instead of the determinant.

(2) The upper bound (D.7) for p_F is a version of Theorem 2.14, point A., in [2]. We take therein $q = 0$, so there is no derivative, and $\Theta = 1$, that means that we are not localizing.

□

References

[1] V. Bally and L. Caramellino. Positivity and lower bounds for the density of Wiener functionals. *Potential Analysis*, 39:141–168, April 2013.

[2] V. Bally and L. Caramellino. On the distances between probability density functions. *Electron. J. Probab.*, 19:no. 110, 1–33, 2014.

[3] V. Bally, L. Caramellino, and P. Pigato. Diffusions under a local strong Hörmander condition. Part II: tube estimates. *ArXiv e-prints*, 1607.04544, July 2016.

[4] G. Ben Arous and R. Léandre. Décroissance exponentielle du noyau de la chaleur sur la diagonale. I. *Probab. Theory Related Fields*, 90(2):175–202, 1991.

[5] G. Ben Arous and R. Léandre. Décroissance exponentielle du noyau de la chaleur sur la diagonale. II. *Probab. Theory Related Fields*, 90(3):377–402, 1991.

[6] K. Bichteler, J-B. Gravereaux, and J. Jacod. *Malliavin Calculus for Processes with Jumps*, volume Stochastic mongraphs Volume 2. Gordon and Breach Science Publishers, 1987.

[7] P. Cattiaux and L. Mesnager. Hypoelliptic non-homogeneous diffusions. *Prob. Theory and Rel. Fields*, 123:453–483, 2002.

[8] F. Delarue and S. Menozzi. Density estimates for a random noise propagating through a chain of differential equations. *J. Funct. Anal.*, 259(6):1577–1630, 2010.

[9] C. Donati-Martin and M. Yor. Fubini's theorem for double wiener integrals and the variance of the brownian path. *Ann. Inst. Henri Poincaré*, 27:181–200, 1991.

[10] D. Jerison and A. Sánchez-Calle. *Subelliptic, second order differential operators*, pages 46–77. Springer Berlin Heidelberg, Berlin, Heidelberg, 1987.

[11] S. Kusuoka and D. Stroock. Applications of the Malliavin calculus. II. *Journal of the Faculty of Science of the Univ. of Tokyo Sect IA Math*, pages 1–76, 1985.

[12] S. Kusuoka and D. Stroock. Applications of the Malliavin calculus. III. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 34(2):391–442, 1987.

[13] J. Norris. Simplified Malliavin calculus. *Seminaire de Probabilitees XX: Lecture Notes in Math*, 1204:101130, 1986.

[14] D. Nualart. *Malliavin Calculus and Related Topics.* Springer, Berlin, 2006.

[15] P. Pigato. Tube estimates for diffusion processes under a weak Hörmander condition. *arXiv Preprint arXiv:1412.4917*, 2014.

[16] W. Rudin. *Principles of mathematical analysis.* McGraw-Hill Book Co., New York, third edition, 1976. International Series in Pure and Applied Mathematics.

[17] A. Sánchez-Calle. Fundamental solutions and geometry of the sum of squares of vector fields. *Inventiones mathematicae*, 78(1):143–160, 1984.