



**HAL**  
open science

# Tubes estimates for diffusion processes under a local Hörmander condition of order one

Vlad Bally, Lucia Caramellino

► **To cite this version:**

Vlad Bally, Lucia Caramellino. Tubes estimates for diffusion processes under a local Hörmander condition of order one. 2015. hal-01104873

**HAL Id: hal-01104873**

**<https://hal.science/hal-01104873v1>**

Preprint submitted on 19 Jan 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Tubes estimates for diffusion processes under a local Hörmander condition of order one

VLAD BALLY\*

LUCIA CARAMELLINO†

February 23, 2012

**Abstract.** We consider a diffusion process  $X_t$  and a skeleton curve  $x_t(\phi)$  and we give a lower bound for  $P(\sup_{t \leq T} d(X_t, x_t(\phi)) \leq R)$ . This result is obtained under the hypothesis that the strong Hörmander condition of order one (which involves the diffusion vector fields and the first Lie brackets) holds in every point  $x_t(\phi)$ ,  $0 \leq t \leq T$ . Here  $d$  is a distance which reflects the non isotropic behavior of the diffusion process which moves with speed  $\sqrt{t}$  in the directions of the diffusion vector fields but with speed  $t$  in the directions of the first order Lie brackets. We prove that  $d$  is locally equivalent with the standard control metric  $d_c$  and that our estimates hold for  $d_c$  as well.

**Keywords:** Hörmander condition, Tube estimates, Diffusion processes, Caratheodory metric.

**2000 MSC:** 60H07, 60H30.

---

\*Laboratoire d'Analyse et de Mathématiques Appliquées, UMR 8050, Université Paris-Est Marne-la-Vallée, 5 Bld Descartes, Champs-sur-Marne, 77454 Marne-la-Vallée Cedex 2, 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

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Notations and main results</b>	<b>4</b>
<b>3</b>	<b>Multiple stochastic integrals</b>	<b>8</b>
3.1	Decomposition . . . . .	8
3.2	Main estimates . . . . .	12
<b>4</b>	<b>Diffusion processes</b>	<b>17</b>
4.1	Short time behavior . . . . .	17
4.2	Chain argument . . . . .	21
<b>5</b>	<b>Appendix 1. Exponential decay for multiple stochastic integrals</b>	<b>24</b>
<b>6</b>	<b>Appendix 2. Small perturbations of Gaussian random variables</b>	<b>25</b>
6.1	The inverse function theorem . . . . .	25
6.2	Estimates of the density . . . . .	28
<b>7</b>	<b>Appendix 3. Support Property</b>	<b>30</b>
<b>8</b>	<b>Appendix 4. Norms and distances</b>	<b>33</b>

## 1 Introduction

We consider the diffusion process solution of  $dX_t = \sum_{j=1}^d \sigma_j(t, X_t) \circ dW_t^j + b(t, X_t)dt$  where the coefficients  $\sigma_j, b$  are three times differentiable and verify the strong Hörmander condition on order one (involving  $\sigma_j$  and the first order Lie brackets  $[\sigma_i, \sigma_j]$ ) locally around a skeleton path  $dx_t(\phi) = \sum_{j=1}^d \sigma_j(t, x_t(\phi)) \phi_t^j dt + b(t, x_t(\phi))dt$ . The aim of this paper is to give a lower bound for the probability that  $X_t$  remains in a tube around  $x_t(\phi)$  for  $t \leq T$ . This problem has already been addressed in the literature. The first result was given by Stroock and Varadhan in their celebrated paper [15]. They obtain a lower bound for  $P(\sup_{t \leq T} \|X_t - x_t(\phi)\| \leq R)$  and use it in order to prove the support theorem for diffusion processes. Here  $\|X_t - x_t(\phi)\|$  is the Euclidian norm. Later, one has considered other norms which reflect the degree of regularity of the trajectories of the diffusion process  $X_t$ : Ben Arous and Gradinaru [4] and Ben Arous, Gradinaru and Ledoux [5] obtained similar results for the Hölder norm. And more recently Friz, Lyons and Stroock [10] use a norm related to the rough path theory. All these results hold without any non degeneracy assumption.

Tubes estimates has also been considered in connection with the Onsager-Machlup functional for diffusion processes. There is an abundant literature on this subject: see e.g. [7], [8], [11], [12], [16]. In this case one considers strong ellipticity conditions and the norm which describes the tube is the Euclidian norm or some Hölder norm. Notice that these are asymptotic results whether in our paper we give estimates which are non asymptotic.

Finally, in [1] and [3] one obtains similar lower bounds for general Itô processes under an ellipticity assumption.

The specific point in our paper is that we use a distance which reflects the non isotropic structure of the problem: the diffusion process  $X_t$  moves with speed  $\sqrt{t}$  in the direction of the diffusion vector fields  $\sigma_j$  and with speed  $t = \sqrt{t} \times \sqrt{t}$  in the direction of  $[\sigma_i, \sigma_j]$ . Let us be more precise. For  $R > 0$  and  $x \in R^n$  we construct the matrix  $A_R(t, x)$  with columns  $\sqrt{R}\sigma_i(t, x), [\sqrt{R}\sigma_j, \sqrt{R}\sigma_p](t, x), 1 \leq i, j, p \leq d$ . If the above vectors span  $R^n$  the matrix  $A_R A_R^*(t, x)$  is invertible, so we are able to define the norm

$$|y|_{A_R(t, x)}^2 = \langle (A_R A_R^*)^{-1}(t, x)y, y \rangle.$$

Our main result is the following (see Theorem 3 for a precise statement): we assume that the non-degeneracy condition holds along the curve  $x_t(\phi), 0 \leq t \leq T$  and we prove

$$P(\sup_{t \leq T} |X_t - x_t(\phi)|_{A_R(t, x_t(\phi))} \leq 1) \geq \exp\left(-C\left(\frac{1}{R} + \int_0^T |\phi_t|^2 dt\right)\right).$$

Computations involving the above norms are generally not easy - so we give some estimates which seem to be more explicit. In Proposition 1 we prove that  $|y|_{A_R(t, x)}$  describes (roughly speaking) ellipsoids with semi-axes of length  $\sqrt{R}$  in the directions of  $\sigma_j(t, x)$  and of length  $R$  in the directions of  $[\sigma_i, \sigma_j](t, x)$ . Moreover we associate to the above norms the following semi-distance:  $d(x, y) < R$  if and only if  $|y|_{A_R(x)} < 1$ . With this definition we have  $\{\sup_{t \leq T} |X_t - x_t(\phi)|_{A_R(t, x_t(\phi))} \leq 1\} = \{\sup_{t \leq T} d(x_t(\phi), X_t) \leq R\}$ . In Proposition 28 we prove that the semi-distance  $d$  is equivalent with the standard control metric  $d_c$  (see (11) for the definition) so the estimates of the tubes hold in the control metric as well. In Proposition 6 we give local lower and upper bounds for  $d$  and  $d_c$  in terms of some semi-distances which describe in a more explicit way the ellipsoid structure we mentioned above.

The paper is organized as follows. In Section 2 we give the statements of the main results. In Section 3 we consider a process  $Z_t$  which is a linear combination of  $W_t^j, j = 1, \dots, d$  and of  $\int_0^t W_s^i dW_s^j, 1 \leq i, j \leq d$ . And we give a decomposition of such a process - this decomposition represents the main ingredient in our approach. Roughly speaking the idea is the following: we consider a small interval of time  $[0, \delta]$  and we split it in  $d$  subintervals  $I_i = (t_{i-1}, t_i]$  with  $t_i = \frac{i}{d}\delta$ . We fix  $i$  and for  $t \in I_i$  we take conditional expectation with respect to  $W_t^j, j \neq i$  so all these processes appear as “controls”. And the only process which is at work is  $W_t^i$ . Then the vector  $(W_{t_i}^i - W_{t_{i-1}}^i), \int_{t_{i-1}}^{t_i} (W_s^j - W_{t_{i-1}}^j) dW_s^i, j \neq i$  is Gaussian (with respect to the above mentioned conditional probability). And we may choose the trajectories (controls)  $(W_s^j - W_{t_{i-1}}^j)_{s \in I_i}, j \neq i$  in such a way that the covariance matrix of the above Gaussian vector is non degenerated (this is a support property proven in Section 7). Then we are able to use estimates for non degenerated Gaussian random variables. The process  $Z_t$  appears as the principal part in the development in stochastic series of order two of the diffusion process  $X_t$ . In Section 4 we use the estimates for  $Z_t$  in order to obtain estimates for  $X_t$  and so to finish the proof of the main theorem stated in Section 2.

The fact that one may choose  $(W_s^j - W_{t_{i-1}}^j)_{s \in I_i}, j \neq i$  in an appropriate way is due to the support theorem for the Brownian motion. But the quantitative property that we use

employs in a crucial way the estimates of the variance (with respect to the time) of the Brownian motion obtained in [9].

**Acknowledgments.** We are grateful to Arturo Kohatsu-Higa and to Peter Friz for useful discussions on this topic.

## 2 Notations and main results

We consider the  $n$  dimensional diffusion process

$$dX_t = \sum_{j=1}^d \sigma_j(t, X_t) \circ dW_t^j + b(t, X_t)dt \quad (1)$$

where  $W = (W^1, \dots, W^d)$  is a standard Brownian motion,  $\circ dW_t^j$  denotes the Stratonovich integral and  $\sigma_j, b : R_+ \times R^n \rightarrow R^n$  are three time differentiable in  $x \in R^n$  and one time differentiable with respect to the time  $t \in R_+$ . We also assume that the derivatives with respect to the space  $x \in R^n$  are one time differentiable with respect to  $t$ . And for  $(t, x) \in R_+ \times R^n$  we denote by  $n(t, x)$  a constant such that for every  $s \in [(t-1) \vee 0, t+1]$ ,  $y \in B(x, 1)$  and for every multi index  $\alpha$  of length less or equal to three

$$|\partial_x^\alpha b(s, y)| + |\partial_t \partial_x^\alpha b(s, y)| + \sum_{j=1}^d |\partial_x^\alpha \sigma_j(s, y)| + |\partial_t \partial_x^\alpha \sigma_j(s, y)| \leq n(t, x). \quad (2)$$

Here,  $\alpha = (\alpha_1, \dots, \alpha_k) \in \{1, \dots, n\}^k$  is a multi index and  $|\alpha| = k$  is the length of  $\alpha$  and  $\partial_x^\alpha = \partial_{x_{\alpha_1}} \dots \partial_{x_{\alpha_k}}$ .

In the following we assume that for external reasons one produces a continuous adapted process  $X$  which solves equation (1) on the time interval  $[0, T]$  and we give estimates for this process. More precisely, for  $\phi \in L^2([0, T]; R^d)$ , we assume there exists a solution of

$$dx_t(\phi) = \sum_{j=1}^d \sigma_j(t, x_t(\phi)) \phi_t^j dt + b(t, x_t(\phi))dt \quad (3)$$

and we want to estimate the probability that  $X_t$  remains in a tube around the deterministic curve  $x_t = x_t(\phi)$ .

We need some more notations. First, we use the following notation of directional derivatives: for  $f, g : R_+ \times R^n \rightarrow R^n$  we define  $\partial_g f(t, x) = \sum_{i=1}^n g^i(t, x) \partial_{x_i} f(t, x)$  and we recall that the Lie bracket (with respect to the space variable  $x$ ) is defined as  $[f, g](t, x) = \partial_g f(t, x) - \partial_f g(t, x)$ . Moreover, let  $M \in \mathcal{M}_{n \times m}$  be a matrix (which generally may be not square) such that  $MM^*$  is invertible ( $M^*$  denotes the transposed matrix). We denote by  $\lambda_*(M)$  (respectively  $\lambda^*(M)$ ) the smaller (respectively the larger) eigenvalue of  $MM^*$  and we consider the norm on  $R^n$

$$|y|_M = \sqrt{\langle (MM^*)^{-1}y, y \rangle}. \quad (4)$$

We are concerned with the matrix  $A(t, x) \in \mathcal{M}_{n \times m}$  with columns  $\sigma_i(t, x), [\sigma_j, \sigma_p](t, x), 1 \leq i, j, p \leq d, j \neq p$ . Here and all along the paper

$$m = d^2.$$

We will write

$$A(t, x) = (\sigma_i(x), [\sigma_j, \sigma_p](t, x))_{i,j,p=1,\dots,d,j \neq p}. \quad (5)$$

We denote by  $\lambda(t, x)$  the lower eigenvalue of  $A(t, x)$  that is

$$\lambda(t, x) = \inf_{|\xi|=1} \sum_{i=1}^m \langle A_i(t, x), \xi \rangle^2, \quad (6)$$

$A_i(t, x), i = 1, \dots, m$ , denoting the columns of  $A(t, x)$ . Moreover for  $R > 0$  we define

$$A_R(t, x) = (\sqrt{R}\sigma_i(t, x), [\sqrt{R}\sigma_j, \sqrt{R}\sigma_p](t, x))_{i,j,p=1,\dots,d,j \neq p}.$$

Consider now some  $x \in R^n, t \geq 0$  such that  $(\sigma_i(t, x), [\sigma_j, \sigma_p](t, x))_{i,j,p=1,\dots,d,j \neq p}$  span  $R^n$ . Then  $A_R A_R^*(t, x)$  is invertible and we may define  $|y|_{A_R(t, x)}$ . We give some lower and upper bounds for  $|y|_{A_R(t, x)}$ . We denote by  $S(t, x)$  the space spanned by  $\sigma_1(t, x), \dots, \sigma_d(t, x)$  and by  $S^\perp(t, x)$  the orthogonal of  $S(t, x)$ . We also denote by  $\Pi_{t,x}$  the projection on  $S(t, x)$  and by  $\Pi_{t,x}^\perp$  the projection on  $S^\perp(t, x)$ . Moreover we denote

$$\lambda_{t,x} = \inf_{\xi \in S(t,x), |\xi|=1} \sum_{i=1}^d \langle \sigma_i(t, x), \xi \rangle^2, \quad \lambda_{t,x}^\perp = \inf_{\xi \in S^\perp(t,x), |\xi|=1} \sum_{i < j} \langle [\sigma_i, \sigma_j](t, x), \xi \rangle^2. \quad (7)$$

By the very definition  $\lambda_{t,x} > 0$  (which is different from  $\lambda(t, x)$ ) and under our hypothesis  $\lambda_{t,x}^\perp > 0$  also. Then Proposition 26 gives:

**Proposition 1** *If  $R \leq \lambda_{t,x}/(4m \times n^4(t, x))$  then*

$$\frac{1}{4Rn^2(t, x)} |\Pi_{t,x} y|^2 + \frac{1}{4R^2n^2(t, x)} |\Pi_{t,x}^\perp y|^2 \leq |y|_{A_R(t, x)}^2 \leq \frac{4}{R\lambda_{t,x}} |\Pi_{t,x} y|^2 + \frac{4}{R^2\lambda_{t,x}^\perp} |\Pi_{t,x}^\perp y|^2. \quad (8)$$

For  $\mu \geq 1$  and  $0 < h \leq 1$  we denote by  $L(\mu, h)$  the class of non negative functions  $f : R_+ \rightarrow R_+$  which have the property

$$f(t) \leq \mu f(s) \quad \text{for} \quad |t - s| \leq h.$$

We will make the following hypothesis: there exists some functions  $n : [0, T] \rightarrow [1, \infty)$  and  $\lambda : [0, T] \rightarrow (0, 1]$  such that for some  $\mu \geq 1$  and  $0 < h \leq 1$  we have

$$\begin{aligned} (H_1) \quad & n(t, x_t(\phi)) \leq n_t, \forall t \in [0, T], \\ (H_2) \quad & \lambda(t, x_t(\phi)) \geq \lambda_t > 0, \forall t \in [0, T], \\ (H_3) \quad & n, \lambda \in L(\mu, h). \end{aligned} \quad (9)$$

**Remark 2** *The hypothesis  $(H_2)$  implies that for each  $t \in (0, T)$ , the space  $R^n$  is spanned by the vectors  $(\sigma_i(t, x_t), [\sigma_j, \sigma_p](t, x_t))_{i,j,p=1,\dots,d,j < p}$ , so the Hörmander condition holds along the curve  $x_t(\phi)$ .*

The main result in this paper is the following.

**Theorem 3** *Suppose that  $(H_1)$ ,  $(H_2)$  and  $(H_3)$  hold and that  $X_0 = x_0(\phi)$ . Let  $\rho \in (0, 1)$ . There exists a universal constant  $C$  (depending on  $d$  and  $\rho$  only) such that for every  $R \in (0, 1)$  one has*

$$P(\sup_{t \leq T} |X_t - x_t(\phi)|_{A_R(t, x_t(\phi))} \leq 1) \geq \exp\left(-C\mu^9\left(\frac{T}{h} + \int_0^T \frac{n_t^{6(1+d\rho)}}{\lambda_t^{1+2d\rho}}\left(\frac{1}{R} + |\phi_t|^2\right)dt\right)\right). \quad (10)$$

**Remark 4** *Suppose that  $X_t = W_t$  is just the Brownian motion and that  $x_t = 0$ , so that  $n_t = 1$ ,  $\lambda_t = 1$ ,  $\mu = 1$  and  $\phi_t = 0$ . Then  $|X_t - x_t|_{A_R(x_t(\phi))} = R^{-1/2}W_t$  so we obtain  $P(\sup_{t \leq T} |W_t| \leq \sqrt{R}) \geq \exp(-CT/R)$  which is coherent with the standard estimate (see [12]).*

**Remark 5** *Since  $\partial_t x_t(\phi) - b(t, x_t(\phi)) = \sigma(t, x_t(\phi))\phi(t)$  we immediately obtain*

$$\frac{1}{dn(t, x_t(\phi))} |\partial_t x_t(\phi) - b(t, x_t(\phi))| \leq |\phi(t)| \leq \frac{1}{\sqrt{\lambda_{t, x_t(\phi)}}} |\partial_t x_t(\phi) - b(t, x_t(\phi))|$$

with  $\lambda_{t, x_t(\phi)}$  given in (7).

We establish now the link between the norm  $|z|_{A_R(t, x)}$  and the control (Caratheodory) distance. We will use in a crucial way the alternative characterizations given in [14] for this distance - and these results hold in the homogeneous case: the coefficients of the equations do not depend on time:  $\sigma_j(t, x) = \sigma_j(x)$  and  $b(t, x) = b(x)$ . Consequently now on we have a matrix  $A_R(x)$  instead of  $A_R(t, x)$ . We define the semi-distance  $d : R^n \times R^n \rightarrow R_+$  by  $d(x, y) < \sqrt{R}$  if and only if  $|y|_{A_R(x)} < 1$  (see page 37 for the definition of a semi-distance). We also consider the standard control distance  $d_c$  (Caratheodory distance) associated to  $\sigma_1, \dots, \sigma_d$  in the following way. Let  $y_t(\phi)$  be the solution of the equation  $dy_t(\phi) = \sum_{j=1}^d \sigma_j(y_t(\phi))\phi_t^j dt$  (notice that here  $b = 0$ ). We denote  $C(x, y) = \{\phi \in L^2(0, 1) : y_0(\phi) = x, y_1(\phi) = y\}$  and we define

$$d_c(x, y) = \inf \left\{ \left( \int_0^1 |\phi_s|^2 ds \right)^{1/2} : \phi \in C(x, y) \right\}. \quad (11)$$

In Section 8 Theorem 28 we prove that  $d$  is locally equivalent with  $d_c$ . Moreover we obtain the following bounds for them. We define  $\bar{d}(x, y)$  and  $\underline{d}(x, y)$  as follows:

- $\bar{d}(x, y) < \sqrt{R}$  if and only if

$$\frac{4}{R\lambda_x} |\Pi_x(y - x)|^2 + \frac{4}{R^2\lambda_x^\perp} |\Pi_x^\perp(y - x)|^2 < 1;$$

- $\underline{d}(x, y) < \sqrt{R}$  if and only if

$$\frac{1}{4Rn_x^2} |\Pi_x(y - x)|^2 + \frac{1}{4R^2n_x^2} |\Pi_x^\perp(y - x)|^2 < 1.$$

Then as an immediate consequence (we give a detailed proof at the end of Appendix 4) of Proposition 1 and Theorem 28 we obtain:

**Proposition 6** *Let  $x, y \in R^n$  be such that*

$$|y - x| \leq \frac{\lambda_x \sqrt{\lambda_*(A(x))}}{(4m)n^4(x)}. \quad (12)$$

*Then*

$$\underline{d}(x, y) \leq d(x, y) \leq \bar{d}(x, y). \quad (13)$$

*Moreover for every compact set  $K \subset R^n$  there exists some constants  $C_K, r_K$  such that for ever  $x, y \in K$  which satisfy (12) and such that  $\bar{d}(x, y) \leq r_K$  one has*

$$\frac{1}{C_K} \underline{d}(x, y) \leq d_c(x, y) \leq C_K \bar{d}(x, y). \quad (14)$$

As an immediate consequence of the definition of  $d$  and of the local equivalence of  $d_c$  with  $d$  we obtain the following:

**Proposition 7** *Suppose that  $(H_i), i = 1, 2, 3$  hold and  $X_0 = x_0(\phi)$ . Let  $\rho \in (0, 1)$ . There exists a universal constant  $C$  (depending on  $d$  and  $\rho$  only) such that for every  $R \in (0, 1)$  one has*

$$P(\sup_{t \leq T} d(x_t(\phi), X_t) \leq R) \geq \exp(-C\mu^9(\frac{T}{h} + \int_0^T \frac{n_t^{6(1+d\rho)}}{\lambda_t^{1+2d\rho}} (\frac{1}{R} + |\phi_t|^2) dt)).$$

*Moreover there exists a constant  $C$  (depending on  $d$  and  $\rho$  but also on  $x_t(\phi)$  and on the coefficients  $\sigma_i(x_t(\phi)), b(x_t(\phi))$  and on their derivatives up to order three) such that*

$$P(\sup_{t \leq T} d_c(x_t(\phi), X_t) \leq R) \geq \exp(-C\mu^9(\frac{T}{h} + \int_0^T \frac{n_t^{6(1+d\rho)}}{\lambda_t^{1+2d\rho}} (\frac{1}{R} + |\phi_t|^2) dt)). \quad (15)$$

We finish this section with two simple examples.

**Example 1.** We consider the two dimensional diffusion process

$$X_t^1 = x_1 + W_t^1, \quad X_t^2 = x_2 + \int_0^t X_s^1 dW_s^2.$$

Straightforward computations give

$$|\xi|_{A_\delta(x)}^2 = |T_{x,\delta}\xi|^2 \quad \text{with} \quad T_{x,\delta}\xi = \left( \frac{1}{\sqrt{\delta}}\xi_1, \frac{1}{\sqrt{\delta(\delta + x_1^2)}}\xi_2 \right).$$



In particular, if  $x_1 = 0$  then  $T_{0,\delta}\xi = (\frac{1}{\sqrt{\delta}}\xi_1, \frac{1}{\delta}\xi_2)$  and consequently  $\{\xi : |\xi|_{A_\delta(x)} \leq 1\}$  is an ellipsoid. But if  $x_1 \neq 0$  and  $\delta$  is small, then the distance given by  $|\xi|_{A_\delta(x)}$  is equivalent with the Euclidian one.

If we take a path  $x_t$  which keeps far from zero then we have ellipticity along the path and so we may use estimates for elliptic processes (see [1] and [3]). But if  $x_1(t) = 0$  for some  $t \in [0, T]$  then we may no more use them. Let us compare the norm here and the norm in the elliptic case: if  $x_1 > 0$  the diffusion matrix is not degenerated so we may consider the norm  $|\xi|_{B_\delta(x)}$  with  $B_\delta(x) = \delta\sigma\sigma^*(x)$ . We have

$$|\xi|_{B_\delta(x)}^2 = \frac{1}{\delta}\xi_1^2 + \frac{1}{\delta x_1^2}\xi_2^2 \geq \frac{1}{\delta}\xi_1^2 + \frac{1}{\delta(\delta + x_1)}\xi_2^2 = |\xi|_{A_\delta(x)}^2.$$

So the estimates obtained using the Lie brackets are sharper even if ellipticity holds. Let us now take  $x_1 = x_2 = 0$ ,  $x_t(\phi) = (0, 0)$ . We have  $n_s = 1$  and  $\lambda_s = 1$  and  $X_t - x_t = (W_t^1, \int_0^t W_s^1 dW_s^2)$ . And we obtain

$$P(\sup_{t \leq T} \left( \frac{1}{\delta} |W_t^1|^2 + \frac{1}{\delta^2} \left| \int_0^t W_s^1 dW_s^2 \right|^2 \right) \leq 1) = P(\sup_{t \leq T} (|X_t - x_t|_{A_\delta(0)}^2 \leq 1) \geq e^{-C/\delta}.$$

**Example 2. The principal invariant diffusion on the Heisenberg group.** We consider the diffusion process

$$X_t^1 = x_1 + W_t^1, \quad X_t^2 = x_2 + W_t^2, \quad X_t^3 = x_3 + \frac{1}{2} \int_0^t X_s^1 dW_s^2 - \frac{1}{2} \int_0^t X_s^2 dW_s^1.$$

Direct computations give

$$|\xi|_{A_\delta(x)}^2 = |A_\delta^{-1}(x)\xi|^2 = \frac{1}{\delta} \left( \xi_1 - \xi_3 \times \frac{x_2}{2\sqrt{\delta}} \right)^2 + \frac{1}{\delta} \left( \xi_2 - \xi_3 \times \frac{x_1}{2\sqrt{\delta}} \right)^2 + \frac{\xi_3^2}{\delta^2}.$$

In particular for  $x = 0$  we obtain

$$\begin{aligned} P\left( \sup_{t \leq T/\delta} \left( |W_t^1|^2 + |W_t^2|^2 + A_t^2(W) \right) \leq 1 \right) \\ = P\left( \sup_{t \leq T} \left( \frac{1}{\delta} |W_t^1|^2 + \frac{1}{\delta} |W_t^2|^2 + \frac{1}{\delta^2} A_t^2(W) \right) \leq 1 \right) \geq e^{-\frac{CT}{\delta}} \end{aligned}$$

where  $A_t(W) = \int_0^t W_s^1 dW_s^2 - \int_0^t W_s^2 dW_s^1$ .

## 3 Multiple stochastic integrals

### 3.1 Decomposition

We consider the stochastic process

$$Z(t) = \sum_{i=1}^d a_i W_t^i + \sum_{i,j=1}^d a_{i,j} \int_0^t W_s^i \circ dW_s^j \tag{16}$$

with  $a_i, a_{i,j} \in R^n$ . Our aim is to give a decomposition for this process. In order to do it we have to introduce some notation. We fix  $\delta > 0$  and we denote  $s_k(\delta) = \frac{k}{d}\delta$  and

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

Notice that  $\Delta_k^{i,j}(\delta, W)$  is the Stratonovich integral, but for  $i \neq j$  it coincides with the Ito integral. When now confusion is possible we use the short notation  $s_k = s_k(\delta)$ ,  $\Delta_k^i = \Delta_k^i(\delta, W)$ ,  $\Delta_k^{i,j} = \Delta_k^{i,j}(\delta, W)$ . Moreover for  $p = 1, \dots, d$  we define

$$\begin{aligned} \mu_p(\delta, W) &= \sum_{i \neq p} \Delta_i^p \\ \psi_p(\delta, W) &= \sum_{i \neq j, i \neq p, j \neq p} a_{i,j} \Delta_p^{i,j} + \sum_{l=p+1}^d \sum_{i \neq p} \sum_{j \neq l} a_{i,j} \Delta_l^j \Delta_p^i + \frac{1}{2} \sum_{i \neq p} a_{i,i} |\Delta_p^i|^2 \\ \varepsilon_p(\delta, W) &= \sum_{l > p} \sum_{j \neq l} a_{p,j} \Delta_l^j + \sum_{p > l} \sum_{j \neq l} a_{j,p} \Delta_l^j + \sum_{j \neq p} a_{p,j} \Delta_p^j \\ \eta_p(\delta, W) &= \frac{1}{2} a_{p,p} |\Delta_p^p|^2 + \sum_{l > p} a_{p,l} \Delta_l^l \Delta_p^p + \Delta_p^p \varepsilon_p. \end{aligned} \tag{17}$$

We denote  $\eta(\delta, W) = \sum_{p=1}^d \eta_p(\delta, W)$  and  $\psi(\delta, W) = \sum_{p=1}^d \psi_p(\delta, W)$  and

$$[a]_{i,p} = a_{i,p} - a_{p,i}. \tag{18}$$

Our aim is to prove the following decomposition.

### Proposition 8

$$Z(\delta) = \sum_{p=1}^d a_p(\Delta_p^p(\delta, W) + \mu_p(\delta, W)) + \sum_{p=1}^d \sum_{i \neq p} [a]_{i,p} \Delta_p^{i,p}(\delta, W) + \eta(\delta, W) + \psi(\delta, W) \tag{19}$$

**Remark 9** *The reason of being of this decomposition is the following. We split the time interval  $(0, \delta)$  in  $d$  sub intervals of length  $\delta/d$ . And we also split the Brownian motion in corresponding pieces:  $(W_s^i - W_{s_{p-1}}^i)_{s_{p-1} \leq s \leq s_p}$ ,  $i = 1, \dots, d$ . Let us fix  $i$ . For  $s \in (s_{i-1}, s_i)$  we have the processes  $(W_s^j - W_{s_{i-1}}^j)_{s_{i-1} \leq s \leq s_i}$ ,  $j = 1, \dots, d$ . Our idea is to settle a calculus which is based on  $W^i$  and to take conditional expectation with respect to  $W^j$ ,  $j \neq i$ . So  $(W_s^j - W_{s_{i-1}}^j)_{s_{i-1} \leq s \leq s_i}$ ,  $j \neq i$  will appear as parameters (or controls) which we may choose in an appropriate way. And the random variables on which the calculus is based are  $\Delta_i^i = W_{s_i}^i - W_{s_{i-1}}^i$  and  $\Delta_i^{j,i} = \int_{s_{i-1}}^{s_i} (W_s^j - W_{s_{i-1}}^j) dW_s^i$ ,  $j \neq i$ . These are the random variables that we have emphasized in the decomposition of  $Z(\delta)$ . Notice that, conditionally to the controls  $(W_s^j - W_{s_{i-1}}^j)_{s_{i-1} \leq s \leq s_i}$ ,  $j \neq i$ , this is a centered Gaussian vector and, under appropriate hypothesis on the controls this Gaussian vector is non degenerated (we treat in the Appendix 3 the problem of the choice of the controls). But there is another term which appear and which is difficult to handle by a choice of the controls  $W^j$ : this is  $\Delta_i^{i,j} = \int_{s_{i-1}}^{s_i} (W_s^i - W_{s_{i-1}}^i) dW_s^j$ . So we use the identity  $\Delta_i^{i,j} = \Delta_i^j \Delta_i^i - \Delta_i^{j,i}$  in order to eliminate this term - and this is the reason for which  $(a_{i,j} - a_{j,i}) = [a]_{i,j}$  appears.*

**Proof.** We decompose

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

and we write

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

Then

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

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

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

Then

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

We treat now  $S_2$ . We want to emphasis terms which contain  $\Delta_l^i$ . 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  $\sum_{l>p}^d = \sum_{p=1}^d \sum_{l=p+1}^d$  and

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

We have

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

and

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

so that

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

Finally

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

We want to compute the coefficient of  $\Delta_p^p$ : this term appears in

$$\begin{aligned} &\sum_{p=1}^d \Delta_p^p (a_p + \varepsilon_p) \quad \text{with} \\ \varepsilon_p &= \sum_{l>p}^d \sum_{j \neq l}^d a_{p,j} \Delta_l^j + \sum_{p>l}^d \sum_{j \neq l}^d a_{j,p} \Delta_l^j + \sum_{j \neq p}^d a_{p,j} \Delta_p^j. \end{aligned}$$

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

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

The other terms are

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

We put everything together and (19) is proved.  $\square$

### 3.2 Main estimates

Throughout this section we will assume that

$$\text{Span}\{a_i, [a]_{j,p}, i, j, p = 1, \dots, d, j \neq p\} = R^n. \quad (20)$$

Let us introduce some notation. We consider the matrix  $A = (a_i, [a]_{j,p}, i, j, p = 1, \dots, d, j \neq p)$  to be the matrix with columns  $a_i$  and  $[a]_{j,p}$ . For  $R \in (0, 1]$  we define the matrix  $A_R = (\sqrt{R}a_i, R[a]_{j,p}, i, j, p = 1, \dots, d, j \neq p)$  and we denote  $\lambda_*(A_R), \lambda^*(A_R)$  the lower and the larger eigenvalue of  $A_R A_R^*$ . We just write  $\lambda_*(A), \lambda^*(A)$  if  $R = 1$ . We associate the norms  $|y|_{A_R}^2 = \langle (A_R A_R)^{-1} y, y \rangle$ .

In Proposition 25 from the Appendix 4 we prove the following basic properties. For every  $0 < R \leq R' \leq 1$

$$\sqrt{\frac{R}{R'}} |y|_{A_R} \geq |y|_{A_{R'}} \geq \frac{R}{R'} |y|_{A_R} \quad (21)$$

and

$$\frac{1}{\sqrt{R} \sqrt{\lambda^*(A)}} |y| \leq |y|_{A_R} \leq \frac{1}{R \sqrt{\lambda_*(A)}} |y|. \quad (22)$$

Finally

$$|A_R y|_{A_R} \leq |y|. \quad (23)$$

**Lemma 10** *Suppose that (20) holds. There exists an universal constant  $C_0$  such that for every  $R \geq \delta > 0$  and  $r > 0$*

$$P(\sup_{t \leq \delta} |Z_t|_{A_R} \geq r) \leq \exp\left(-\frac{rR}{C_0 \delta} \left(r \wedge \frac{\sqrt{\lambda_*(A)}}{\bar{a}}\right)\right) \quad (24)$$

with

$$\bar{a} = 1 \vee \max_{i,j} |a_{i,j}|. \quad (25)$$

**Remark 11** *One might think to use directly Bernstein's inequality in order to estimate  $P(\sup_{t \leq \delta} |Z_t|_{A_R} \geq r)$  but then one would not obtain the right inequality. Indeed one writes  $|Z_t|_{A_R} \leq (R \sqrt{\lambda_*(A)})^{-1} |Z_t|$  and then the above probability is bounded by*

$$P(\sup_{t \leq \delta} |Z_t| \geq rR \sqrt{\lambda_*(A)}) \leq \exp\left(-\frac{r^2 R^2 \lambda_*(A)}{\delta}\right).$$

So one obtains  $\frac{R^2}{\delta}$  instead of  $\frac{R}{\delta}$  and this is not in the right scale. The reason is that in the above argument we just use the lower eigenvalue  $\lambda_*(A)$  in order to upper bound  $|Z_t|_{A_R}$  since in the proof of our lemma we use the more subtle inequality  $|A_R y|_{A_R} \leq |y|$ .

**Proof.** Let  $t \leq \delta$ . We decompose  $Z(t)$  instead of  $Z(\delta)$  and similarly to (19) we obtain

$$Z(t) = \sum_{p=1}^d a_p (\Delta_p^p(t, W) + \mu_p(t, W)) + \sum_{p=1}^d \sum_{i \neq p} [a]_{i,p} \Delta_p^{i,p}(t, W) + \eta(t, W) + \psi(t, W),$$

in which  $\eta(t, W)$  and  $\psi(t, W)$  are defined as in (17) with  $\Delta_p^i$  and  $\Delta_p^{ij}$  replaced by  $\Delta_p^i(t, W)$  and  $\Delta_p^{ij}(t, W)$  respectively, and these last quantities are defined as follows: for  $t \in [0, T]$ ,  $\Delta_p^i(t, W) = W_{s_p \wedge t}^i - W_{s_{p-1} \wedge t}^i$  and  $\Delta_p^{ij}(t, W) = \int_{s_{p-1} \wedge t}^{s_p \wedge t} (W_s^i - W_{s_{p-1} \wedge t}^i) dW_s^j$ . We denote by  $u(t) \in R^m$  the vector with component  $u_p(t) = t^{-1/2}(\Delta_p^p(t, W) + \mu_p(t, W)) = t^{-1/2}W_t^p, p = 1, \dots, d$  and  $u_{i,j}(t) = 0, i \neq j$  and we also denote

$$U(t) = \sum_{p=1}^d \sum_{i \neq p} [a]_{i,p} \Delta_p^{i,p}(t, W) + \eta(t, W) + \psi(t, W).$$

Then we have

$$Z(t) = \sum_{p=1}^d t^{1/2} a_p u_p(t) + \sum_{p=1}^d \sum_{i \neq p} t [a]_{i,p} \times 0 + U(t) = A_t u(t) + U(t).$$

Using the norm inequalities given above

$$|U(t)|_{A_R} \leq \frac{1}{R\sqrt{\lambda_*(A)}} |U(t)| \leq \frac{C\bar{a}}{R\sqrt{\lambda_*(A)}} \sum_{i,j=1}^d (|\Delta_j^i(t, W)|^2 + \sum_{p=1}^d |\Delta_p^{i,j}(t, W)|)$$

so that

$$\begin{aligned} P\left(\sup_{t \leq \delta} |U(t)|_{A_R} \geq \frac{r}{2}\right) &\leq \sum_{i,j=1}^d P\left(\sup_{t \leq \delta} |\Delta_j^i(t, W)|^2 \geq \frac{rR\sqrt{\lambda_*(A)}}{C\bar{a}}\right) \\ &+ \sum_{i,j,p=1}^d P\left(\sup_{t \leq \delta} |\Delta_p^{i,j}(t, W)| \geq \frac{rR\sqrt{\lambda_*(A)}}{C\bar{a}}\right). \end{aligned}$$

It is easy to check that

$$P\left(\sup_{t \leq \delta} |\Delta_p^p(t, W)|^2 \geq \frac{rR\sqrt{\lambda_*(A)}}{C\bar{a}}\right) \leq C' \exp\left(-\frac{rR\sqrt{\lambda_*(A)}}{C'\bar{a}\delta}\right).$$

Moreover,

$$\sup_{t \leq \delta} |\Delta_p^{i,j}(t, W)| \leq 2 \sup_{t \leq \delta} \left| \int_0^t W_s^i dW_s^j \right| + 2 \sup_{t \leq \delta} (|W_t^i|^2 + |W_t^j|^2).$$

Using (43) from the Appendix 1 we obtain

$$P\left(\sup_{t \leq \delta} \left| \int_0^t W_s^i dW_s^j \right| \geq \frac{rR\sqrt{\lambda_*(A)}}{C\bar{a}}\right) \leq C \exp\left(-\frac{rR\sqrt{\lambda_*(A)}}{C\bar{a}\delta}\right).$$

So we have proved that

$$P\left(\sup_{t \leq \delta} |U(t)|_{A_R} \geq \frac{r}{2}\right) \leq C \exp\left(-\frac{rR\sqrt{\lambda_*(A)}}{C\bar{a}\delta}\right).$$

Using (21) (recall that  $t \leq \delta \leq R$ ) and (23)

$$|A_t u(t)|_{A_R} \leq \sqrt{\frac{t}{R}} |A_t u(t)|_{A_t} \leq \sqrt{\frac{t}{R}} |u(t)| \leq \frac{C}{\sqrt{R}} \sup_{t \leq \delta} |W_t|.$$

It follows that

$$P\left(\sup_{t \leq \delta} |A_t u(t)|_{A_R} \geq \frac{r}{2}\right) \leq P\left(\sup_{t \leq \delta} |W_t| \geq \frac{r\sqrt{R}}{C}\right) \leq C \exp\left(-\frac{r^2 R}{C\delta}\right).$$

□

We give the main result in this section.

**Proposition 12** *Suppose that  $\lambda_*(A) > 0$ . Let  $\rho \in (0, 1)$  be fixed. There exists an universal constant  $C_*$  (depending on  $d$  and on  $\rho$  only) such that for every*

$$r \leq \frac{\lambda_*^{1/2}(A)}{C_* \bar{a}} \quad (26)$$

one has

$$P(|Z_\delta|_{A_\delta} \leq r) \geq \frac{r^m}{C_*} \times \frac{\lambda_*^{2d^3}(A)}{\bar{a}^{d^3}} \times \exp\left(-\frac{C_* \lambda_*^{d^2 \rho}(A)}{\bar{a}^2}\right). \quad (27)$$

**Proof. Step 1. Scaling.** Let  $B_t = \delta^{-1/2} W_{t\delta}$ . Then  $B$  is a standard Brownian motion and we denote

$$\Delta_i^j(B) = B_i^j - B_{i-1}^j, \quad \Delta_p^{i,j}(B) = \int_{p-1}^p (B_s^j - B_p^j) dB_s^i, \quad i \neq j.$$

We also denote by  $\Delta(B)$  the vector  $(\Delta_i^j(B), \Delta_p^{i,j}(B), i, j, p = 1, \dots, d)$  and we define  $\Theta(B) = (\Theta_1(B), \dots, \Theta_d(B))$  with  $\Theta_p(B) = (\Delta_p^p(B), \Delta_p^{j,p}(B), j \neq p)$ . We consider the  $\sigma$  field

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

Conditionally to  $\mathcal{G}$  the random variable  $\Theta_p(B)$  is Gaussian with covariance matrix  $Q_p(B)$  given by

$$\begin{aligned} Q_p^{p,j}(B) &= \int_{p-1}^p (B_s^j - B_{i-1}^j) ds, \quad j \neq p, \\ Q_p^{i,j}(B) &= \int_{p-1}^p (B_s^j - B_{i-1}^j)(B_s^i - B_{i-1}^i) ds, \quad j \neq p, i \neq p, \\ Q_p^{p,p}(B) &= 1. \end{aligned}$$

Since the random variables  $\Theta_1(B), \dots, \Theta_d(B)$  are independent  $\Theta(B)$  is a Gaussian random variable. We denote by  $Q(B)$  the covariance matrix of  $\Theta(B)$  and by  $\lambda_*(B), \lambda^*(B)$  the smaller and the larger eigenvalues of  $Q(B)$ . Since this matrix is built with the blocks  $Q_p(B), p = 1, \dots, d$  we have

$$\lambda_*(B) = \prod_{p=1}^d \lambda_{*,p}(B) \quad \text{and} \quad \lambda^*(B) = \prod_{p=1}^d \lambda_p^*(B)$$

where  $\lambda_{*,p}(B), \lambda_p^*(B)$  are the smaller and the larger eigenvalues of  $Q_p(B)$ .

We come now back to our problem. Let  $\eta(\Delta(B)), \psi(\Delta(B)), \varepsilon(\Delta(B)), \mu(\Delta(B))$  be the quantities defined in (17) with  $\Delta = \Delta(\delta, W)$  replaced by  $\Delta(B)$ . Then  $\delta\eta(\Delta(B)) = \eta(\delta, W)$ . The same is true for  $\psi$  and  $\varepsilon$  and finally  $\sqrt{\delta}\mu(\Delta(B)) = \mu(\delta, W)$ . So using (19)

$$Z_\delta = \sum_{p=1}^d \sqrt{\delta} a_p (\Delta_p^p(B) + \mu_p(\Delta(B))) + \sum_{p=1}^d \sum_{i \neq p} \delta [a]_{ip} \Delta_p^{i,p}(B) + \delta\eta(\Delta) + \delta\psi(\Delta).$$

We define now the vector  $\mu(\Delta(B)) = (\mu_p(\Delta(B)), \mu_{i,j}(\Delta(B)) \in R^m, i \neq j)$  by  $\mu_{i,j}(\Delta(B)) = 0$  and then we may write the above decomposition in matrix notation

$$\begin{aligned} Z_\delta &= A_\delta(\Theta(B) + \mu(\Delta(B))) + \delta\eta(\Theta(B)) + \delta\psi(\Delta(B)) \\ &= y + A_\delta\Theta(B) + \eta_\delta(\Theta(B)) \end{aligned} \quad (28)$$

with

$$y = A_\delta\mu(\Delta(B)) + \delta\psi(\Delta(B)), \quad \eta_\delta(\theta) = \delta\eta(\theta).$$

**Step 2. Localization.** We take

$$\varepsilon \leq \frac{\lambda_*(A)}{C_1 \bar{a}^2} \quad (29)$$

where  $C_1$  is an universal constant to be chosen in the sequel. For each  $p = 1, \dots, d$  we define the sets

$$\Lambda_{\rho, \varepsilon, p} = \left\{ \det Q_p(B) \geq \varepsilon^\rho, \sup_{p-1 \leq t \leq p} \sum_{j \neq p} |B_t^j - B_{p-1}^j| \leq \varepsilon^{-\rho}, q_p(B) \leq \varepsilon \right\}$$

with

$$q_p(B) = \sum_{j \neq p} |B_p^j - B_{p-1}^j| + \sum_{j \neq p, i \neq p} \left| \int_{p-1}^p (B_s^j - B_{i-1}^j) dB_s^i \right|.$$

By (61) in Appendix 3 we may find some constants  $c$  and  $\varepsilon_*$  depending on  $d$  and  $\rho$  only such that

$$P(\Lambda_{\rho, \varepsilon, p}) \geq c \varepsilon^{\frac{1}{2}d(d+1)} \quad \text{for} \quad \varepsilon \leq \varepsilon_* \quad (30)$$

And using the independence we obtain

$$P\left(\bigcap_{p=1}^d \Lambda_{\rho, \varepsilon, p}\right) \geq c^d \times \varepsilon^{\frac{1}{2}d^2(d+1)}. \quad (31)$$



On the set  $\cap_{p=1}^d \Lambda_{\rho, \varepsilon, p}$  we have  $\det Q_p(B) \geq \varepsilon^\rho$  so that  $\det Q(B) \geq \varepsilon^{d\rho}$ . We also have  $\lambda^*(B) \leq \varepsilon^{-\rho}$  and this gives  $\lambda_*(B) \geq \varepsilon^{d^2\rho}$ . And we also have  $\det Q(B) \leq \varepsilon^{-d\rho}$  so

$$\cap_{p=1}^d \Lambda_{\rho, \varepsilon, p} \subset \left\{ \det Q(B) \leq \varepsilon^{-d\rho}, \lambda_*(B) \geq \varepsilon^{d^2\rho}, \sum_{p=1}^d q_p(B) \leq d\varepsilon \right\} \quad (32)$$

**Step 3. Inverse function theorem.** We will use (55) with  $G = Z_\delta$  so we have to estimate the parameters associated to  $\eta_\delta$  and  $A_\delta$ . Notice first that  $\lambda_*(A_\delta) \geq \delta^2 \lambda_*(A)$ ,  $c_{3, \eta_\delta} = 0$  and  $c_{2, \eta_\delta} \leq C_2 \bar{a} \delta$ . So the first inequality in (54) reads

$$r \leq \frac{\lambda_*^{1/2}(A)}{C_2 \bar{a}} \leq \frac{\lambda_*^{1/2}(A_\delta)}{16(c_{2, \eta_\delta} + c_{3, \eta_\delta})}.$$

And this is verified by our hypothesis. Moreover

$$c_*(\eta_\delta, r) \leq C_3 \bar{a} (|\theta| + \sum_{p=1}^d |\varepsilon_p(\Delta(B))|) \leq C_4 \bar{a} (r + \sum_{p=1}^d q_p(B)) \leq C_4 \bar{a} \left( \frac{\lambda_*^{1/2}(A)}{C_2 \bar{a}} + d\varepsilon \right).$$

If we choose  $C_1$  in (29) sufficiently large and  $C_2$  large also we obtain  $c_*(\eta_\delta, r) \leq \frac{1}{2}$  which is the second restriction in (54). Let  $p_{\mathcal{G}, Z_\delta}(z)$  be the density of  $Z_\delta$  conditionally to  $\mathcal{G}$ . Then, using (55), if  $|z - y|_{A_\delta} \leq r \leq 1$  we obtain

$$\begin{aligned} p_{\mathcal{G}, Z_\delta}(z) &\geq \frac{(4\lambda_*(B))^{(m-n)/2}}{(8\pi)^{m/2} \sqrt{\det Q(B)} \sqrt{\det A_\delta A_\delta^*}} \exp\left(-\frac{1}{4\lambda_*(Q(B))} |z - y|_{A_\delta}^2\right) \\ &\geq \frac{\varepsilon^{d^3\rho}}{(8\pi)^{m/2} \sqrt{\det A_\delta A_\delta^*}} \exp\left(-\frac{1}{4\varepsilon^{d^2\rho}}\right) \end{aligned}$$

the second inequality being true on  $\cap_{p=1}^d \Lambda_{\rho, \varepsilon, p}$ . On this set we also have

$$|\mu(\Delta(B))| + |\psi(\Delta(B))| \leq C_5 \bar{a} \sum_{p=1}^d q_p(B) \leq C_6 \bar{a} \varepsilon$$

so that

$$\begin{aligned} |y|_{A_\delta} &\leq |A_\delta \mu(\Delta(B))|_{A_\delta} + \delta |\psi(\Delta(B))|_{A_\delta} \leq |\mu(\Delta(B))| + \frac{1}{\sqrt{\lambda_*(A)}} |\psi(\Delta(B))| \\ &\leq \frac{C_7 \bar{a}}{\sqrt{\lambda_*(A)}} \varepsilon \leq \frac{r}{2}. \end{aligned}$$

So, if  $|z|_{A_\delta} \leq \frac{r}{2}$  then  $|z - y|_{A_\delta} \leq r$ . It follows that

$$\begin{aligned} P_{\mathcal{G}}(|Z_\delta|_{A_\delta} \leq \frac{r}{2}) &= \int_{\{|z|_{A_\delta} \leq \frac{r}{2}\}} p_{\mathcal{G}, Z_\delta}(z) dz \geq \frac{\varepsilon^{d^3\rho}}{(8\pi)^{m/2}} \exp\left(-\frac{1}{4\varepsilon^{d^2\rho}}\right) \int_{\{|z|_{A_\delta} \leq \frac{r}{2}\}} \frac{1}{\sqrt{\det A_\delta A_\delta^*}} dz \\ &= \frac{\varepsilon^{d^3\rho}}{(8\pi)^{m/2}} \exp\left(-\frac{1}{4\varepsilon^{d^2\rho}}\right) \times \frac{r^m}{2^m} \end{aligned}$$

the last equality being obtained by a change of variable. Finally using (31)

$$P(|Z_\delta|_{A_\delta} \leq \frac{r}{2}) \geq P(P_G(|Z_\delta|_{A_\delta} \leq r), \cap_{p=1}^d \Lambda_{\rho, \varepsilon, p}) \geq \frac{r^m \varepsilon^{2d^3}}{C_8} \exp(-\frac{1}{4\varepsilon^{d^2\rho}}).$$

We replace now  $\varepsilon$  by the expression in the RHS of (29) and we obtain (27).  $\square$

**Corollary 13** *Suppose that  $\lambda_*(A) > 0$ . Let  $\rho \in (0, 1)$  be fixed. There exists some universal constant  $C$  (depending on  $d$  and on  $\rho$  only) such that for every  $r, R > 0$  the following holds. Suppose that*

$$\delta \leq \frac{rR}{C \ln \frac{1}{r}} \left( r \wedge \frac{\sqrt{\lambda_*(A)}}{\bar{a}} \right) \times \frac{\lambda_*^{d\rho}(A)}{\bar{a}^{2d\lambda}}. \quad (33)$$

Then

$$P(\sup_{t \leq \delta} |Z_t|_{A_R} \leq r, |Z_\delta|_{A_\delta} \leq r) \geq \frac{r^m}{2C_*} \exp(-\frac{C_* \bar{a}^{2d\rho}}{\lambda_*^{d\lambda}(A)}) \quad (34)$$

with  $C_*$  the constant from (27).

**Proof.** We use (24) and (27) in order to obtain

$$\begin{aligned} P(\sup_{t \leq \delta} |Z_t|_{A_R} \leq r, |Z_\delta|_{A_\delta} \leq r) &\geq P(|Z_\delta|_{A_\delta} \leq r) - P(\sup_{t \leq \delta} |Z_t|_{A_R} > r) \\ &\geq \frac{r^m}{C_3} \exp(-\frac{C_3 \bar{a}^{2d\rho}}{\lambda_*^{d\lambda}(A)}) - \exp(-\frac{rR}{C_0 \delta} \left( r \wedge \frac{\sqrt{\lambda_*(A)}}{\bar{a}} \right)) \\ &\geq \frac{r^m}{2C_3} \exp(-\frac{C_3 \bar{a}^{2d\rho}}{\lambda_*^{d\lambda}(A)}) \end{aligned}$$

the last inequality being a consequence of our restriction on  $\delta$ .  $\square$

## 4 Diffusion processes

### 4.1 Short time behavior

We consider the diffusion process  $X_t$  solution of (1) and the skeleton  $x_t = x_t(\phi)$  solution of (3) and we give for them an estimate which is analogous to (34). Using a development in stochastic Taylor series of order two we write

$$X_t = X_0 + Z_t + b(0, X_0)t + R_t$$

where  $Z_t$  is defined in (16) with  $a_i = \sigma_i(0, X_0)$ ,  $a_{i,j} = \partial_{\sigma_i} \sigma_j(0, X_0)$  so that  $[a]_{i,j} = [\sigma_i, \sigma_j](0, X_0)$ , and

$$\begin{aligned} R_t &= \sum_{j,i=1}^d \int_0^t \int_0^s (\partial_{\sigma_i} \sigma_j(u, X_u) - \partial_{\sigma_i} \sigma_j(0, X_0)) \circ dW_u^i \circ dW_s^j \\ &\quad + \sum_{i=1}^d \int_0^t \int_0^s \partial_b \sigma_i(u, X_u) du \circ dW_s^i + \sum_{i=1}^d \int_0^t \int_0^s \partial_u \sigma_j(u, X_u) du \circ dW_s^j \\ &\quad + \sum_{i=1}^d \int_0^t \int_0^s \partial_{\sigma_i} b(u, X_u) \circ dW_u^i ds + \int_0^t \int_0^s \partial_b b(u, X_u) dud s. \end{aligned}$$

We denote

$$\begin{aligned} A(t, x) &= (\sigma_i(t, x), [\sigma_j, \sigma_p](t, x))_{i,j,p=1,\dots,d,j \neq p} \quad \text{and} \\ A_\delta(t, x) &= (\sqrt{\delta}\sigma_i(t, x), [\sqrt{\delta}\sigma_j, \sqrt{\delta}\sigma_p](t, x))_{i,j,p=1,\dots,d,j \neq p}. \end{aligned}$$

In particular  $\lambda_*(A(t, x)) = \lambda(t, x)$ .

We will need the following estimate for the skeleton  $x_t = x_t(\phi)$  as in (3). And for  $\phi \in L^2([0, T], R^d)$ , we set

$$\varepsilon_\phi(\delta) = \left( \int_0^\delta |\phi_s|^2 ds \right)^{1/2}. \quad (35)$$

**Lemma 14** *Let  $\delta$  be such that  $\varepsilon_\phi(\delta) + \sqrt{\delta} \leq 1$ ,  $\delta < \frac{1}{4n(0, x_0)}$  and*

$$n(0, x_0)(\varepsilon_\phi(\delta) + \sqrt{\delta}) + \sqrt{\delta} \leq \frac{\sqrt{\lambda(0, x_0)}}{8d^3n^2(0, x_0)}. \quad (36)$$

*Then for every  $0 \leq t \leq \delta$  and  $z \in R^n$ ,*

$$|z|_{A_\delta(0, x_0)}^2 \leq 4 |z|_{A_\delta(t, x_t)}^2 \leq 16 |z|_{A_\delta(0, x_0)}^2. \quad (37)$$

*Moreover,*

$$\sup_{t \leq \delta} |x_t - x_0 - b(0, x_0)t|_{A_\delta(0, x_0)} \leq 4\varepsilon_\phi(\delta) + \frac{1}{n(0, x_0)}\delta. \quad (38)$$

**Proof.** First, one has  $x_s \in B(x_0, 1)$  for every  $s \leq \delta$ . In fact, setting  $\tau = \inf\{t > 0 : |x_t - x_0| > 1\}$ , for  $s \leq \delta \wedge \tau$  one has

$$|x_s - x_0| \leq n(0, x_0)\sqrt{\delta}(\varepsilon_\phi(\delta) + \sqrt{\delta}) \leq \frac{1}{2}$$

because  $\varepsilon_\phi(\delta) + \sqrt{\delta} \leq 1$  and  $\delta < \frac{1}{4n(0, x_0)}$ . This gives  $s < \tau$ . This means that  $\delta < \tau$ , so that  $|x_s - x_0| < 1$  for every  $s \leq \delta$ . Moreover, by using (36),

$$|x_s - x_0| + |s| \leq n(0, x_0)\sqrt{\delta}(\varepsilon_\phi(\delta) + \sqrt{\delta}) + \delta \leq \frac{\sqrt{\lambda(0, x_0)}}{8d^3n^2(0, x_0)} \times \sqrt{\delta}. \quad (39)$$

Now, (37) follows immediately from Proposition 27 in Appendix 4 (see page 36).

We prove now (38). For  $t \leq \delta$ , we write now

$$J_t := x_t - x_0 - b(0, x_0)t = \int_0^t (\partial_s x_s - b(s, x_s))ds + \int_0^t (b(s, x_s) - b(0, x_0))ds.$$

By using inequality (65) in Lemma 25 from Appendix 4 (see page 33), we get

$$\begin{aligned} |J_t|_{A_\delta(0, x_0)}^2 &\leq 2t \int_0^t |\partial_s x_s - b(s, x_s)|_{A_\delta(0, x_0)}^2 ds + 2t \int_0^t |b(s, x_s) - b(0, x_0)|_{A_\delta(0, x_0)}^2 ds \\ &=: I'_t + I''_t \end{aligned}$$

As for  $I'_t$ , we use (37): for  $s \leq t \leq \delta$  we have

$$|\partial_s x_s - b(s, x_s)|_{A_\delta(0, x_0)}^2 \leq 4|\partial_s x_s - b(s, x_s)|_{A_\delta(s, x_s)}^2.$$

Moreover, we can write

$$\partial_s x_s - b(s, x_s) = \sum_{j=1}^d \sigma_j(s, x_s) \phi_j(s) = A_\delta(s, x_s) \psi(s), \quad \text{with } \psi_j(s) = \frac{1}{\sqrt{\delta}} \phi_j, \quad \psi_{i,j}(s) = 0$$

so that

$$|\partial_s x_s - b(s, x_s)|_{A_\delta(s, x_s)} = |A_\delta(s, x_s) \psi(s)|_{A_\delta(s, x_s)} \leq |\psi(s)| = \frac{1}{\sqrt{\delta}} |\phi(s)|.$$

Then, for  $t \leq \delta$  we can write

$$I'_t \leq 8\delta \int_0^\delta |\partial_s x_s - b(s, x_s)|_{A_\delta(s, x_s)}^2 ds \leq 8 \int_0^\delta |\phi(s)|^2 ds = 8\varepsilon_\phi(\delta)^2.$$

We estimate now  $I''_t$ : by using (39),

$$\begin{aligned} I''_t &\leq 2\delta \int_0^\delta \frac{1}{\lambda_*(A_\delta(0, x_0))} |b(s, x_s) - b(0, x_0)|^2 ds \\ &\leq 2 \frac{n^2(0, x_0)}{\lambda(0, x_0)} \int_0^t (|s| + |x_s - x_0|)^2 ds \leq \frac{1}{n^2(0, x_0)} \times \delta^2. \end{aligned}$$

By inserting the estimates for  $I'_t$  and  $I''_t$ , we get

$$\sup_{t \leq \delta} |J_t|_{A_\delta(0, x_0)} \leq \left( 8\varepsilon_\phi(\delta)^2 + \frac{1}{n^2(0, x_0)} \delta^2 \right)^{1/2} \leq 4\varepsilon_\phi(\delta) + \frac{1}{n(0, x_0)} \delta.$$

□

The main estimate in this section is the following proposition.

**Proposition 15** *Let (9) hold and let  $\rho \in (0, 1)$  be fixed. Then there exist some universal constants  $C_1, C_2$  (depending on  $d$  and  $\rho$  only) such that the following holds. Let  $0 < \delta \leq R \leq 1$  and  $r \in (0, 1)$  be such that*

$$\varepsilon_\phi(\delta) \leq \frac{r \wedge \sqrt{\lambda(0, x_0)}}{C_1 n^3(0, x_0)}, \quad \delta \leq \frac{r^5 R}{C_1} \times \frac{\lambda^{1+3d\rho}(0, x_0)}{n^{6+6d\rho}(0, x_0)} \quad (40)$$

and suppose that

$$|X_0 - x_0|_{A_\delta(0, x_0)} \leq \frac{r}{8}. \quad (41)$$

Then

$$P\left(\sup_{t \leq \delta} |X_t - x_t|_{A_R(t, x_t)} \leq 2r, |X_\delta - x_\delta|_{A_\delta(\delta, x_\delta)} \leq r\right) \geq \frac{r^m}{C_2} \exp\left(-\frac{C_2 n^{2d\rho}(0, x_0)}{\lambda^{d\rho}(0, x_0)}\right). \quad (42)$$

**Proof.** For  $t \leq \delta$ , by using (37) we obtain

$$|X_t - x_t|_{A_\delta(t, x_t)} \leq 4 |X_t - x_t|_{A_\delta(0, x_0)} \leq 4 \sum_{j=1}^6 |I_j|_{A_\delta(0, x_0)}$$

with

$$\begin{aligned} I_1 &= X_0 - x_0, & I_2 &= Z_t, & I_3 &= R_t \\ I_4 &= x_t - x_0 - b(0, x_0)t, & I_5 &= (b(0, X_0) - b(0, x_0))t, \end{aligned}$$

We have to estimate the above terms for  $t \leq \delta$ . First

$$|I_5|_{A_\delta(0, x_0)} \leq \frac{n(0, x_0)}{\sqrt{\lambda(0, x_0)}} |X_0 - x_0| \leq \frac{n^2(0, x_0)}{\sqrt{\lambda(0, x_0)}} |X_0 - x_0|_{A_\delta(0, x_0)} \times \sqrt{\delta} \leq \frac{r}{8}$$

and by (38)

$$|I_4|_{A_\delta(0, x_0)} \leq 4\varepsilon_\phi(\delta) + \frac{1}{n(0, x_0)}\delta \leq \frac{r}{8}$$

And by our assumption  $|I_1|_{A_\delta(0, x_0)} \leq \frac{r}{8}$ . So we have

$$|X_t - x_t|_{A_\delta(t, x_t)} \leq \frac{r}{2} + 4(|Z_t|_{A_\delta(0, x_0)} + |R_t|_{A_\delta(0, x_0)}).$$

Since  $R \geq \delta$ , by (62) in Lemma 25 from Appendix 4 (see page 33) we have  $|y|_{A_R(0, x_0)} \leq |y|_{A_\delta(0, x_0)}$  so  $|I_i|_{A_R(0, x_0)} \leq \frac{r}{8}$  for  $i = 1, 4, 5$ . And this gives

$$|X_t - x_t|_{A_R(t, x_t)} \leq \frac{r}{2} + 4(|Z_t|_{A_R(0, x_0)} + |R_t|_{A_R(0, x_0)}).$$

Using the above inequalities we easily obtain

$$\begin{aligned} &P\left(\sup_{t \leq \delta} |X_t - x_t|_{A_R(t, x_t)} \leq 2r, |X_\delta - x_\delta|_{A_\delta(\delta, x_\delta)} \leq r\right) \\ &\geq P\left(\sup_{t \leq \delta} |Z_t|_{A_R(0, x_0)} + \sup_{t \leq \delta} |R_t|_{A_R(0, x_0)} \leq \frac{r}{4}, |Z_\delta|_{A_\delta(0, x_0)} + |R_\delta|_{A_\delta(0, x_0)} \leq \frac{r}{8}\right) \\ &\geq P\left(\sup_{t \leq \delta} |Z_t|_{A_R(0, x_0)} \leq \frac{r}{8}, |Z_\delta|_{A_\delta(0, x_0)} \leq \frac{r}{16}\right) - P\left(\sup_{t \leq \delta} |R_t|_{A_\delta(0, x_0)} > \frac{r}{8}\right). \end{aligned}$$

We upper bound now the last term. First, using the norms inequalities

$$P\left(\sup_{t \leq \delta} |R_t|_{A_\delta(0, x_0)} > \frac{r}{8}\right) \leq P\left(\sup_{t \leq \delta} |R_t| > K\right)$$

with  $K = \frac{r\delta\sqrt{\lambda(0, x_0)}}{8}$ . We define now  $\tau = \inf\{t : |X_t - X_0| \geq \frac{1}{2}\}$ . Using the norms inequalities, (40) and (41) we obtain  $|x_0 - X_0| \leq \frac{1}{2}$  so that for  $t \leq \tau$  we have  $|X_t - x_0| \leq 1$ . It follows that up to  $\tau$  the diffusion process  $X$  coincides with a diffusion process  $\bar{X}$  which

has the coefficients and their derivatives up to order three bounded by  $n(x_0)$ . We denote by  $\overline{R}$  the reminder in which  $X$  is replace with  $\overline{X}$  and we write

$$P\left(\sup_{t \leq \delta} |R_t| > K\right) \leq P\left(\sup_{t \leq \delta} |\overline{R}_t| > K\right) + P(\tau \leq \delta).$$

Since  $\tau = \overline{\tau} := \inf\{t : |\overline{X}_t - X_0| \geq \frac{1}{2}\}$  a standard reasoning based on Bernstin's inequality gives  $P(\tau \leq \delta) = P(\overline{\tau} \leq \delta) \leq \exp(-1/C\delta n^2(x_0))$ .

In order to estimate the last first we use (43) from the Appendix 1 (see Lemma 18 at page 24) with  $k = 3, p_3 = \frac{2}{3}$  and with  $k = 1, p_1 = 2$ , and  $K = \frac{r\delta}{4}\sqrt{\lambda(0, x_0)}$ . A straightforward computation gives

$$\begin{aligned} P\left(\sup_{t \leq \delta} |\overline{R}_t| > \frac{r\delta\sqrt{\lambda(0, x_0)}}{8}\right) &\leq C \exp\left(-\frac{r^2\lambda(0, x_0)}{C\delta n^4(0, x_0)}\right) + C \exp\left(-\frac{r^{2/3}\lambda^{1/3}(0, x_0)}{C\delta^{1/3}n^2(0, x_0)}\right) \\ &\leq C \exp\left(-\frac{r^{2/3}\lambda^{1/3}(0, x_0)}{C\delta^{1/3}n^2(0, x_0)}\right) \end{aligned}$$

the last inequality being a consequence of (40).

Using (34)

$$P\left(\sup_{t \leq \delta} |Z_t|_{A_R(0, x_t)} \leq \frac{r}{8}, |Z_\delta|_{A_\delta(0, x_0)} \leq \frac{r}{16}\right) \geq \frac{r^m}{2C_*} \exp\left(-\frac{C_*n^{2d\lambda}(0, x_0)}{\lambda^{d\lambda}(0, x_0)}\right)$$

with  $C_*$  the universal constant in (34). Our assumption on  $\delta$  gives

$$C \exp\left(-\frac{r^{2/3}\lambda^{1/3}(0, x_0)}{C\delta^{1/3}n^2(0, x_0)}\right) \leq \frac{1}{2} \times \frac{r^m}{2C_*} \exp\left(-\frac{C_*n^{2d\rho}(0, x_0)}{\lambda^{d\lambda}(0, x_0)}\right)$$

so we have proved that

$$P\left(\sup_{t \leq \delta} |X_t - x_t|_{A_R(t, x_t)} \leq 2r, |X_\delta - x_\delta|_{A_\delta(\delta, x_\delta)} \leq r\right) \geq \frac{r^m}{4C_*} \exp\left(-\frac{C_*n^{2d\rho}(0, x_0)}{\lambda^{d\lambda}(0, x_0)}\right).$$

□

## 4.2 Chain argument

We recall that, by the hypothesis (9) we have some functions  $\lambda, n \in L(\mu, h)$  such that  $\lambda(t) \leq 1 \wedge \lambda(t, x_t)$  and  $n_t \geq 1 \vee n(t, x_t)$  such that  $\lambda, n \in L(\mu, h)$  for some  $h > 0$  and  $\mu \geq 1$ . We also consider some  $R, r, \rho \in (0, 1)$  and we define (with  $C_1$  the constant in (40))

$$f_h(t) = \frac{2}{h} + \frac{C_1(\ln \frac{1}{r})^3 n_t^{\delta+4d\rho}}{Rr^2 \lambda_t^{1+d\rho}} + \frac{C_1^2 n_t^6}{r^2 \wedge \lambda_t} |\phi_t|^2.$$

Notice that, if  $d\rho \leq \frac{1}{5}$  then  $f_h \in L(\mu^8, h)$ . We define

$$\delta(t) = \inf \left\{ \delta > 0 : \int_t^{t+\delta} f_h(s) ds \geq \frac{1}{\mu^8} \right\}$$

**Lemma 16** *i) One has*

$$\delta(t) \leq \frac{h}{2} \wedge \frac{Rr_t^2 \lambda_t^{1+d\rho}}{C_1 (\ln \frac{1}{r})^3 n_t^{6+4d\rho}}, \quad \varepsilon_\phi(\delta(t)) \leq \frac{r \wedge \lambda_t^{1/2}}{C_1 n_t^3}.$$

*ii) If  $|t - t'| \leq \delta(t)$  then*

$$\frac{1}{4\mu^{16}} |y|_{A_{\delta(t)}(t, x_t)} \leq |y|_{A_{\delta(t')}(t', x_{t'})} \leq 4\mu^8 |y|_{A_{\delta(t)}(t, x_t)}.$$

**Proof.** *i)* Since  $\int_t^{t+h/2} \frac{2}{h} ds = 1 \geq 1/\mu^8$  we have  $\delta(t) \leq \frac{1}{2}h$ . So we may use the properties  $L(\mu, h)$  for  $t \leq s \leq t + \delta(t)$ . Consequently, for  $0 < \delta \leq \delta(t)$

$$\frac{1}{\mu^8} \geq \int_t^{t+\delta} \frac{C_1 (\ln \frac{1}{r})^3 n_s^{6+4d\rho}}{Rr^2 \lambda_s^{1+d\rho}} ds \geq \frac{1}{\mu^8} \times \frac{C_1 (\ln \frac{1}{r})^3 n_t^{6+4d\rho}}{Rr^2 \lambda_t^{1+d\rho}} \times \delta$$

which gives

$$\delta(t) \leq \frac{Rr^2 \lambda_t^{1+d\rho}}{C_1 (\ln \frac{1}{r})^3 n_t^{6+4d\rho}}.$$

We also have

$$\frac{1}{\mu^8} \geq \int_t^{t+\delta} \frac{C_1^2 n_s^6}{r^2 \wedge \lambda_s} |\phi_s|^2 ds \geq \frac{1}{\mu^8} \times \frac{C_1^2 n_t^6}{r^2 \wedge \lambda_t} \int_t^{t+\delta} |\phi_s|^2 ds$$

so that

$$\varepsilon_\phi^2(t) \leq \frac{r^2 \wedge \lambda_t}{C_1^2 n_t^6}.$$

This proves *i)*.

*ii)* We use here next Proposition 27 from Appendix 4 (see page 36).

If  $|t - t'| \leq \delta(t)$ , then  $|x_t - x_{t'}| \leq \delta^{1/2}(t)(d\varepsilon_\phi(\delta(t)) + \delta^{1/2}(t))n_t$  so (73) is verified and we may use (74) to obtain

$$\frac{1}{4} |y|_{A_{\delta(t)}(t, x_t)} \leq |y|_{A_{\delta(t')}(t', x_{t'})} \leq 4 |y|_{A_{\delta(t)}(t, x_t)}.$$

It remains to compare  $|y|_{A_{\delta(t)}(t, x_t)}$  with  $|y|_{A_{\delta(t')}(t, x_t)}$ . Since  $\delta(t') \leq \frac{1}{2}h$  and  $|t - t'| \leq \frac{1}{2}h$  we have  $|t - s| \leq h$  for every  $s \in (t', t' + \delta(t'))$ . We use the property  $L(\mu^8, h)$  for  $f_h$  and we obtain

$$\mu^8 f_h(t) \delta(t) \geq \int_t^{t+\delta(t)} f_h(s) ds = \frac{1}{\mu^8} = \int_{t'}^{t'+\delta(t')} f_h(s) ds \geq \mu^{-8} f_h(t) \delta(t').$$

So  $(\delta(t)/\delta(t'))^{1/2} \geq \mu^{-8}$ . Suppose now that  $\delta(t) \leq \delta(t')$ . We use then (21) and we obtain

$$\frac{1}{\mu^{16}} |y|_{A_{\delta(t)}(t, x_t)} \leq |y|_{A_{\delta(t')}(t, x_t)} \leq \frac{1}{\mu^8} |y|_{A_{\delta(t)}(t, x_t)}.$$

□

We construct now a time grid in the following way. We put  $t_0 = 0$  and

$$t_k = t_{k-1} + \delta(t_{k-1})$$

and we denote

$$\Theta_k = \left\{ \sup_{t_{k-1} \leq t \leq t_k} |X_t - x_t|_{A_R(t, x_t)} \leq r \right\}, \quad \Gamma_k = \left\{ |X_{t_k} - x_{t_k}|_{A_{\delta(t_k)}(t_k, x_{t_k})} \leq \frac{r}{8} \right\}.$$

**Proposition 17** *i) Suppose that (9) holds and let  $R, r \in (0, 1)$  and  $\rho \in (0, \frac{1}{5d})$ . There exists a universal constant  $C$  (depending on  $d$  and on  $\rho$ ) such that*

$$P(\cap_{i=1}^k \Theta_i \cap \Gamma_i) \geq P(\cap_{i=1}^{k-1} \Theta_i \cap \Gamma_i) \exp\left(-\frac{C n_{t_{k-1}}^{2d\rho}}{\lambda_{t_{k-1}}^{d\rho}}\right).$$

*ii) Moreover there exists an universal constant  $C$  such that*

$$\begin{aligned} P\left(\sup_{0 \leq s \leq T} |X_s - x_s|_{A_R(s, x(s))} \leq r\right) &\geq \exp\left(-C\mu^9 \int_0^T f_h(t) \frac{n_s^{2d\rho}}{\lambda_t^{d\rho}} dt\right) \\ &\geq \exp\left(-C\mu^9 \left(\frac{T}{h} + \frac{1}{r^2} \int_0^T \frac{n_t^{6+6d\rho}}{\lambda_t^{1+2d\rho}} \left(\frac{(\ln \frac{1}{r})^3}{R} + |\phi_t|^2\right) dt\right)\right). \end{aligned}$$

**Proof** *i)* Let

$$\tilde{\Gamma}_k = \left\{ |X_{t_k} - x_{t_k}|_{A_{\delta(t_{k-1})}(t_{k-1}, x_{t_{k-1}})} \leq \frac{1}{32\mu^8} r \right\}.$$

Using *ii)* from the previous lemma we obtain  $\tilde{\Gamma}_k \subset \Gamma_k$  so by (42)

$$P_{t_{k-1}}(\Theta_k \cap \Gamma_k) \geq P_{t_{k-1}}(\Theta_k \cap \tilde{\Gamma}_k) \geq \exp\left(-C \frac{n_{t_{k-1}}^{2d\rho}(x_0)}{\lambda_{t_{k-1}}^{d\rho}(A(x_0))}\right).$$

The above inequality holds if  $|X_{t_{k-1}} - x_{t_{k-1}}|_{A_{\delta(t_{k-1})}(t_{k-1}, x_{t_{k-1}})} \leq \frac{r}{8}$  and this is true on the set  $\Gamma_{k-1}$ .

*ii)* Let  $N_T = \min\{k : t_k > T\}$ . Since  $X_0 = x_0$  we may use the recursively the inequality from *i)* and we obtain

$$P\left(\sup_{t \leq T} |X_t - x_t|_{A_R(x_t)} \leq r\right) \geq P\left(\cap_{i=1}^{N_T} \Theta_i \cap \Gamma_i\right) \geq \exp\left(-C \sum_{k=1}^{N_T} \frac{n_{t_{k-1}}^{2d\rho}}{\lambda_{t_{k-1}}^{d\rho}}\right).$$

We write

$$\begin{aligned} \int_0^T f_h(s) \frac{n_s^{2d\rho}}{\lambda_s^{d\rho}} ds &\geq \sum_{i=1}^{N_T-1} \int_{t_{i-1}}^{t_i} f_h(s) \frac{n_s^{2d\rho}}{\lambda_s^{d\rho}} ds \geq \frac{1}{\mu^{3d\rho}} \sum_{i=1}^{N_T-1} \frac{n_{t_{i-1}}^{2d\rho}}{\lambda_{t_{i-1}}^{d\rho}} \int_{t_{i-1}}^{t_i} f_h(s) ds \\ &= \frac{1}{\mu^{8+3d\rho}} \sum_{i=1}^{N_T-1} \frac{n_{t_{i-1}}^{2d\rho}}{\lambda_{t_{i-1}}^{d\rho}} \end{aligned}$$

the last equality being a consequence of the definition of  $\delta(t_k)$ .  $\square$



## 5 Appendix 1. Exponential decay for multiple stochastic integrals

In this section  $W = (W^1, \dots, W^d)$  is a standard Brownian motion and  $\alpha = (\alpha_1, \dots, \alpha_k) \in \{1, \dots, d\}^k$  denotes a multi index. We use the notation  $\bar{\alpha} = (\alpha_1, \dots, \alpha_{k-1})$ . We consider an adapted and bounded stochastic process  $a$  and we denote by  $\|a\|_\infty$  a constant such that  $\sup_{t \leq T} |a(t, \omega)| \leq \|a\|_\infty$  almost surely. Then we define the iterated stochastic integrals

$$I_0(a, W)(t) = a(t), \quad I_k^\alpha(a, W)(t) = \int_0^t I_{k-1}^{\bar{\alpha}}(a, W)(s) dW_s^{\alpha_k}.$$

**Lemma 18** *There exist some universal constants  $C_k, C'_k$  such that for each  $T, K \geq 0$  and every multi-index  $\alpha = (\alpha_1, \dots, \alpha_k)$  one has*

$$P\left(\sup_{t \leq T} |I_k^\alpha(a, W)(t)| \geq K\right) \leq C_k \exp\left(-C'_k \left(\frac{K}{T^{k/2} \|a\|_\infty}\right)^{p_k}\right) \quad (43)$$

with

$$p_1 = 2, \quad p_{k+1} = \frac{2p_k}{2 + p_k}.$$

**Proof.** We assume that  $\|a\|_\infty = 1$  almost surely (if not we normalize with  $\|a\|_\infty$ ) and  $T = 1$  (if not we use a scaling argument). We proceed by recurrence. We take some  $Q \geq 0$  and we write

$$\begin{aligned} P(\sup_{t \leq 1} |I_k^\alpha(a, W)(t)| \geq K) &\leq I + J \quad \text{with} \\ I &= P(\sup_{t \leq 1} |I_k^\alpha(a, W)(t)| \geq K, \sup_{t \leq 1} |I_{k-1}^{\bar{\alpha}}(a, W)(t)| \leq Q) \\ J &= P(\sup_{t \leq 1} |I_{k-1}^{\bar{\alpha}}(a, W)(t)| \geq Q). \end{aligned}$$

Using the recurrence hypothesis

$$J \leq C_{k-1} \exp(-C'_{k-1} Q^{p_{k-1}}).$$

We set  $h(t) = \int_0^t |I_{k-1}^{\bar{\alpha}}(a, W)(s)|^2 ds$  and we write  $I_k^\alpha(a, W)(t) = b(h_t)$  where  $b$  is a standard Brownian motion. So, we obtain

$$I \leq P(\sup_{t \leq h_1} |b(t)| \geq K, h(1) \leq Q^2) \leq P(\sup_{t \leq Q^2} |b(t)| \geq K) \leq C \exp\left(-\frac{C' K^2}{Q^2}\right).$$

We choose  $Q$  solution of  $Q^{p_{k-1}} = K^2/Q^2$  that is  $Q = K^{\frac{2}{2+p_{k-1}}}$ . Then we obtain

$$P(\sup_{t \leq 1} |I_k^\alpha(a, W)(t)| \geq K) \leq C_k \exp(-C'_k K^{\frac{2p_{k-1}}{2+p_{k-1}}})$$

with  $C_k = C \vee C_{k-1}, C'_k = C' \wedge C'_{k-1}$ .  $\square$

## 6 Appendix 2. Small perturbations of Gaussian random variables

### 6.1 The inverse function theorem

We give first a quantitative version of the inverse function theorem. We consider a three time differentiable function

$$\eta : R^d \rightarrow R^d \quad \text{and} \quad \Phi(\theta) := \theta + \eta(\theta).$$

We assume that

$$\eta(0) = 0 \quad \text{and} \quad \nabla\eta(0) \leq \frac{1}{2}.$$

In particular this implies that  $\nabla\Phi(0)$  is invertible and

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

We also have  $|\nabla\Phi(0)x| \leq \sqrt{3}|x|$  so

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

We denote

$$c_2(\eta) = \max_{i,j=1,d} \sup_{|x| \leq 1} |\partial_{ij}^2 \eta(x)|, \quad c_3(\eta) = \max_{i,j,k=1,d} \sup_{|x| \leq 1} |\partial_{ijk}^3 \eta(x)|$$

and we take  $h_\eta > 0$  such that

$$h_\eta \leq \frac{1}{2} \quad \text{and} \quad h_\eta \leq \frac{1}{4d^3(c_2(\eta) + c_3(\eta))}. \quad (44)$$

**Proposition 19** *Suppose that  $\eta \in C^3(R^d, R^d)$ ,  $\eta(0) = 0$  and  $\nabla\eta(0) \leq \frac{1}{2}$ . Then there exists a neighborhood  $V_{(h_\eta)} \subset B(0, 2h_\eta)$  of zero such that  $\Phi : V_{(h_\eta)} \rightarrow B(0, \frac{1}{2}h_\eta)$  is a diffeomorphism. In particular, one has*

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

and for every  $y \in B(0, \frac{1}{2}h_\eta)$  the following estimates hold:

$$\frac{1}{4}|\Phi^{-1}(y)| \leq |y| \leq 4|\Phi^{-1}(y)|. \quad (45)$$

**Proof.** The existence and the differentiability property of the inverse function  $\Phi^{-1}$  in a neighborhood of the origin is a well known result from the Inverse Function Theorem. What we aim to prove is that  $\Phi^{-1} : B(0, \frac{1}{2}h_\eta) \rightarrow B(0, 2h_\eta)$  and the estimates in (45). Since  $\eta(0) = 0$  we have

$$\eta(\theta) = \nabla\eta(0)\theta + \int_0^1 (1-t) \langle \nabla^2\eta(t\theta)\theta, \theta \rangle dt$$

with  $\nabla^2 \eta^k = (\partial_{ij}^2 \eta^k)_{i,j=1,d}$ ,  $k = 1, \dots, d$ . So, given  $y \in R^d$  and recalling that  $\nabla \Phi(0) = I + \nabla \eta(0)$ , the equation  $\Phi(\theta) = y$  reads

$$\theta = U_y(\theta), \quad \text{with } U_y(\theta) := (\nabla \Phi(0))^{-1} \left( y - \int_0^1 (1-t) \langle \nabla^2 \eta(t\theta)\theta, \theta \rangle dt \right).$$

Recall that  $\frac{1}{2}|x| \leq |\nabla \Phi(0)x|$ . Then, for  $\theta_1, \theta_2 \in B(0, 2h_\eta)$  we have

$$\begin{aligned} |U_y(\theta_1) - U_y(\theta_2)| &= \left| (\nabla \Phi(0))^{-1} \int_0^1 (1-t) (\langle \nabla^2 \eta(t\theta_1)\theta_1, \theta_1 \rangle - \langle \nabla^2 \eta(t\theta_2)\theta_2, \theta_2 \rangle) dt \right| \\ &\leq 2 \int_0^1 (1-t) |\langle \nabla^2 \eta(t\theta_1)\theta_1, \theta_1 \rangle - \langle \nabla^2 \eta(t\theta_2)\theta_2, \theta_2 \rangle| dt \\ &\leq 2d^3 h_\eta (c_2(\eta) + c_3(\eta)) |\theta_1 - \theta_2| \leq \frac{1}{2} |\theta_1 - \theta_2|, \end{aligned}$$

so that

$$|U_y(\theta_1) - U_y(\theta_2)| \leq \frac{1}{2} |\theta_1 - \theta_2|. \quad (46)$$

Notice also that for  $y \in B(0, \frac{1}{2}h_\eta)$  and  $\theta \in B(0, 2h_\eta)$  the above inequality gives

$$|U_y(\theta)| \leq |U_y(\theta) - U_y(0)| + |U_y(0)| \leq \frac{1}{2} |\theta| + 2|y| \leq h_\eta + h_\eta = 2h_\eta. \quad (47)$$

We define now

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

From (47) we know that  $\theta_k \in B(0, 2h_\eta)$ ,  $k \in N$  and consequently

$$|U_y(\theta_{k+1}) - U_y(\theta_k)| \leq \frac{1}{2} |\theta_k - \theta_{k-1}|.$$

So the sequence  $\theta_k, k \in N$  converges to the solution of the equation  $\theta = U_y(\theta)$ , that is  $\Phi(y) = \theta$ . We have thus proved that for any  $y \in B(0, \frac{1}{2}h_\eta)$  there exists a unique  $\theta \in B(0, 2h_\eta)$  such that  $\Phi(\theta) = y$ , that is  $\Phi^{-1} : B(0, \frac{1}{2}h_\eta) \rightarrow B(0, 2h_\eta)$  is well defined.

Finally, for  $y \in B(0, \frac{1}{2}h_\eta)$  let  $\theta = \Phi^{-1}(y)$ . Then  $\theta = U_y(\theta)$  so, using (47)  $|\theta| = |U_y(\theta)| \leq \frac{1}{2} |\theta| + 2|y|$  which gives  $|\theta| \leq 4|y|$ . Moreover, again by (47),

$$|\theta| = |U_y(\theta)| \geq |U_y(0)| - |U_y(\theta) - U_y(0)| \geq \frac{1}{2} |y| - \frac{1}{2} |\theta|$$

which proves that  $|\theta| \geq \frac{1}{3} |y| \geq \frac{1}{4} |y|$ .  $\square$

Let us consider a more specific variant of the local inversion theorem we will need in next Section 8. We consider a matrix  $B \in \mathcal{M}_{d \times d}$  with columns  $B_i \in R^d, i = 1, \dots, d$  and we suppose that  $B$  is invertible. Then we consider the equation

$$y = B\theta + r(\theta) \quad (48)$$

where  $r \in C^3(R^d, R^d)$ . Our aim is to prove that for small  $y$  the above equation has a unique solution and to obtain some precise estimates for  $\theta$  and its projection on a suitable subspace of  $R^d$  in terms of  $y$ . In order to do it we have to introduce some more notations.

We fix  $d' \in \{1, \dots, d-1\}$  and we denote  $d'' = d - d'$ . For  $x = (x_1, \dots, x_d) \in R^d$  we denote  $\overleftarrow{x} = (x_1, \dots, x_{d'})$  and  $\overrightarrow{x} = (x_{d'+1}, \dots, x_d)$ . We denote by  $\overleftarrow{B} \in \mathcal{M}_{d \times d'}$  (respectively by  $\overrightarrow{B} \in \mathcal{M}_{d \times d''}$ ) the matrix with columns  $B_1, \dots, B_{d'}$  (respectively the matrix with columns  $B_{d'+1}, \dots, B_d$ ). Let  $S = \text{Vect}\{B_1, \dots, B_{d'}\}$ . Since  $B$  is invertible the columns  $B_i, i = 1, \dots, d$  are linearly independent and so  $\dim S = d'$ . We denote by  $S^\perp$  the orthogonal of  $S$  by  $\Pi$  the projection on  $S$  and by  $\Pi^\perp$  the projection on  $S^\perp$ . We define  $\overrightarrow{B}^\perp$  to be the matrix with columns  $B_i^\perp := \Pi^\perp B_i, i = d'+1, \dots, d$ . Since  $B_1, \dots, B_d$  span  $R^d$  it follows that  $B_i^\perp, i = d'+1, \dots, d$  span  $S^\perp$  which has dimension  $d''$ . So  $B_i^\perp, i = d'+1, \dots, d$  are linearly independent. We conclude that the matrices  $B^*B, \overleftarrow{B}^* \overleftarrow{B} \in \mathcal{M}_{d' \times d'}$  and  $\overrightarrow{B}^{\perp,*} \overrightarrow{B}^\perp \in \mathcal{M}_{d'' \times d''}$  are all invertible, and as usual we denote by  $\lambda_*(B), \lambda_*(\overleftarrow{B})$  and  $\lambda_*(\overrightarrow{B}^\perp)$  the smaller eigenvalue of  $B^*B, \overleftarrow{B}^* \overleftarrow{B}$  and  $\overrightarrow{B}^{\perp,*} \overrightarrow{B}^\perp$  respectively.

**Theorem 20** *We assume that the matrix  $B$  is invertible and that  $r(0) = \nabla r(0) = 0$ . Set  $|B|_\infty = \sup_{i,j=1,\dots,d} |B_{ij}|$ . Then for every  $y \in R^d$  such that*

$$|y| < \frac{\lambda_*(B)^{1/2}}{4} \quad \text{and} \quad |y| < \frac{\lambda_*(B)}{8d^3(c_2(r) + c_2(r))}, \quad (49)$$

the equation (48) has a unique solution  $\theta$  and

$$|\theta| \leq \frac{4}{\lambda_*^{1/2}(B)} |y|, \quad \left| \overrightarrow{\theta} \right| \leq \frac{|B|_\infty}{\lambda_*(\overrightarrow{B}^\perp)} |\Pi^\perp y| + \frac{16c_2(r) |B|_\infty}{\lambda_*(\overrightarrow{B}^\perp) \alpha_*(B)} |y|^2. \quad (50)$$

In particular if  $|\Pi^\perp y| \leq |\Pi y|$  then  $|y| \leq 2 |\Pi y|$  so

$$|\theta| \leq \frac{8}{\lambda_*^{1/2}(B)} |\Pi y|, \quad \left| \overrightarrow{\theta} \right| \leq \frac{|B|_\infty}{\lambda_*(\overrightarrow{B}^\perp)} |\Pi^\perp y| + \frac{64c_2(r) |B|_\infty}{\lambda_*(\overrightarrow{B}^\perp) \lambda_*(B)} |\Pi y|^2. \quad (51)$$

**Proof.** We write the equation (48) as  $B^{-1}y = \theta + B^{-1}r(\theta)$  and we use Proposition 19 with  $\eta(\theta) = B^{-1}r(\theta)$ . Since  $\partial_\alpha B^{-1}r(\theta) = B^{-1}\partial_\alpha r(\theta)$  we have  $c_2(\eta) + c_3(\eta) \leq \lambda_*(B)^{-1/2}(c_2(r) + c_3(r))$ . So our assumption (49) ensures that for some  $h_\eta$  fulfilling (44), one has  $|B^{-1}y| \leq \frac{1}{2}h_\eta$  and we may use Proposition 19 in order to produce the solution  $\theta$  of our equation. And moreover, by (45) one has

$$|\theta| \leq 4 |B^{-1}y| \leq \frac{4}{\lambda_*^{1/2}(B)} |y|.$$

In particular this proves the first inequality in (50). Using (49) we also have  $|\theta| \leq 1$ . Since  $r(0) = \nabla r(0) = 0$  we obtain

$$|r(\theta)| \leq \max_{|\alpha|=2} \sup_{|\theta'| \leq 1} |\partial_\theta^\alpha r(\theta')| \times |\theta|^2 = c_2(r) |\theta|^2 \leq \frac{16c_2(r)}{\lambda_*(B)} |y|^2.$$

We multiply our equation with  $(\overrightarrow{B}^{\perp,*} \overrightarrow{B}^\perp)^{-1} \overrightarrow{B}^{\perp,*}$  and we obtain

$$(\overrightarrow{B}^{\perp,*} \overrightarrow{B}^\perp)^{-1} \overrightarrow{B}^{\perp,*} y = \overrightarrow{\theta} + (\overrightarrow{B}^{\perp,*} \overrightarrow{B}^\perp)^{-1} \overrightarrow{B}^{\perp,*} r(\theta).$$

Notice that  $\vec{B}^{\perp,*}y = \vec{B}^{\perp,*}\Pi^{\perp}y$  so  $\left|(\vec{B}^{\perp,*}\vec{B}^{\perp})^{-1}\vec{B}^{\perp,*}y\right| \leq \lambda_*^{-1}(\vec{B}^{\perp})|B|_{\infty}|\Pi^{\perp}y|$  and this gives

$$\left|\vec{\theta}\right| \leq \lambda_*^{-1}(\vec{B}^{\perp})|B|_{\infty}|\Pi^{\perp}y| + \lambda_*^{-1}(\vec{B}^{\perp})|B|_{\infty}|r(\theta)| \leq \frac{|B|_{\infty}}{\lambda_*(\vec{B}^{\perp})}|\Pi^{\perp}y| + \frac{16c_2(r)|B|_{\infty}}{\lambda_*(\vec{B}^{\perp})\lambda_*(B)}|y|^2.$$

□

## 6.2 Estimates of the density

For  $h > 0$  we denote

$$c_*(\eta, h) = \sup_{|x| \leq 2h} \max_{i,j} |\partial_i \eta^j(x)|. \quad (52)$$

Let  $\Theta$  be a  $m$  dimensional centered Gaussian random variable with covariance matrix  $Q$ . We assume that  $Q$  is invertible and we denote by  $\underline{\lambda}(Q)$  and  $\bar{\lambda}(Q)$  the lower and the upper eigenvalue of  $Q$  respectively. We also consider a matrix  $\Gamma \in \mathcal{M}_{n \times m}$  with  $n \leq m$  and we recall that  $|x|_{\Gamma}^2 = \langle \Gamma \Gamma^* x, x \rangle$ ,  $\lambda_*(\Gamma)$  is the smaller eigenvalue of  $\Gamma \Gamma^*$  and  $B_{\Gamma}(y, r) = \{z : |y - z|_{\Gamma} < r\}$ .

**Lemma 21** *Suppose that  $\Gamma \Gamma^*$  is invertible. Let  $\eta \in C_b^3(R^m, R^n)$  such that  $\eta(0) = 0$ . Set*

$$G = y + \Gamma \Theta + \eta(\Theta) \quad (53)$$

and assume there exists  $r > 0$  such that

$$r \leq \frac{1}{2} \lambda_*(\Gamma)^{1/2} h_{\eta} \quad \text{and} \quad \frac{c_*(\eta, 4r)}{\lambda_*(\Gamma)^{1/2}} \leq \frac{1}{2m}, \quad (54)$$

$h_{\eta}$  being defined in (44). Then the law of  $G$  has a density  $p_G$  on  $B_{\Gamma}(y, r)$  and for  $z \in B_{\Gamma}(y, r)$  one has

$$p_G(z) \geq \frac{\underline{\lambda}(Q)^{(m-n)/2}}{\pi^{n/2} 8^{m/2} 2^{(m-n)/2} (\det Q \det \Gamma \Gamma^*)^{1/2}} \exp\left(-\frac{2}{\underline{\lambda}(Q)}|z - y|_{\Gamma}^2\right) \quad (55)$$

$$p_G(z) \leq \frac{\bar{\lambda}(Q)^{(m-n)/2} 8^{(m-n)/2} 2^{m/2}}{\pi^{n/2} (\det Q \det \Gamma \Gamma^*)^{1/2}} \exp\left(-\frac{1}{8\bar{\lambda}(Q)}|z - y|_{\Gamma}^2\right) \quad (56)$$

In particular, (55) and (56) imply that, for  $z \in B_{\Gamma}(y, r)$ ,

$$\left(\frac{\underline{\lambda}(Q)}{16\bar{\lambda}(Q)}\right)^{m/2} p_{N(y, \frac{1}{4}\underline{\lambda}(Q)\Gamma\Gamma^*)}(z) \leq p_G(z) \leq \left(\frac{16\bar{\lambda}(Q)}{\underline{\lambda}(Q)}\right)^{m/2} p_{N(y, 4\bar{\lambda}(Q)\Gamma\Gamma^*)}(z)$$

where  $p_{N(y, BB^*)}$  denotes the Gaussian density with mean  $y$  and covariance matrix  $BB^*$ .

**Proof. Step 1.** We assume first that  $n = m, y = 0$  and  $\Gamma$  is the identity matrix. We denote  $\Phi(\theta) = \theta + \eta(\theta)$ , so that  $\Phi(\Theta) = G$ . Let  $f : R^m \rightarrow R$  be a non negative measurable

function with the support included in the (Euclidian) ball  $B(0, r)$ , with  $r$  fulfilling (54). Using a change of variable and Proposition 19, we obtain

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

where we have set, for  $z \in B(0, r)$ ,

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

Since  $r \leq h_\eta$ , if  $z \in B(0, r)$  one has  $\theta = \Phi^{-1}(z) \in B(0, 4r)$  and for  $x \in B(0, 4r)$  we have

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

because  $c_*(\eta, 4r) \leq \frac{1}{2m}$ . Therefore, if  $z \in B(0, r)$  then

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

Moreover, using (45) we obtain

$$\begin{aligned} \langle Q^{-1}\Phi^{-1}(z), \Phi^{-1}(z) \rangle &\leq \frac{1}{\underline{\lambda}(Q)} |\Phi^{-1}(z)|^2 \leq \frac{4}{\underline{\lambda}(Q)} |z|^2 \quad \text{and} \\ \langle Q^{-1}\Phi^{-1}(z), \Phi^{-1}(z) \rangle &\geq \frac{1}{\overline{\lambda}(Q)} |\Phi^{-1}(z)|^2 \geq \frac{1}{4\overline{\lambda}(Q)} |z|^2 \end{aligned}$$

So, as  $z \in B(0, r)$  we get

$$\frac{1}{(8\pi)^{m/2} \sqrt{\det Q}} \exp\left(-\frac{2}{\underline{\lambda}(Q)} |z|^2\right) \leq p_{\Phi(\Theta)}(z) \leq \frac{2^{m/2}}{\pi^{m/2} \sqrt{\det Q}} \exp\left(-\frac{1}{8\overline{\lambda}(Q)} |z|^2\right) \quad (57)$$

**Step 2.** We still assume that  $n = m$  but now  $y$  and  $\Gamma$  are general, with  $\Gamma$  invertible. We write  $G = y + \Gamma(\theta + \eta_\Gamma(\theta))$  with  $\eta_\Gamma(\theta) = \Gamma^{-1}\eta(\theta)$  and denote  $\Phi_\Gamma(\theta) = \theta + \eta_\Gamma(\theta)$ . One has  $c_2(\eta_\Gamma) + c_3(\eta_\Gamma) \leq \lambda_*(\Gamma)^{-1/2}(c_2(\eta) + c_3(\eta))$ , so that  $h_{\eta_\Gamma} \geq \lambda_*(\Gamma)^{1/2}h_\eta$  and then (54) gives  $r \leq \frac{1}{2}h_{\eta_\Gamma}$ . Moreover, since  $c_*(\eta_\Gamma, 4r) \leq \lambda_*(\Gamma)^{-1/2}c_*(\eta, 4r)$ , (54) gives also  $c_*(\eta_\Gamma, 4r) \leq \frac{1}{2m}$ . And since  $|\Gamma x|_\Gamma = |x|$ , one has  $G \in B_\Gamma(y, r)$  iff  $\Phi_\Gamma(\Theta) \in B(0, r)$ . Then by a change of variable, for  $z \in B_\Gamma(y, r)$  we have

$$p_G(z) = \frac{1}{|\det \Gamma|} p_{\Phi_\Gamma(\Theta)}(\Gamma^{-1}(z - y)).$$

Since  $|\Gamma^{-1}(z - y)| = |z - y|_\Gamma$  we use (57) and we obtain

$$\begin{aligned} p_G(z) &\geq \frac{1}{(8\pi)^{m/2} \sqrt{\det Q} |\det \Gamma|} \exp\left(-\frac{2}{\underline{\lambda}(Q)} |z - y|_\Gamma^2\right) \\ p_G(z) &\leq \frac{2^{m/2}}{\pi^{m/2} \sqrt{\det Q} |\det \Gamma|} \exp\left(-\frac{1}{8\overline{\lambda}(Q)} |z - y|_\Gamma^2\right). \end{aligned}$$

**Step 3.** Now we allow  $n$  to be strictly smaller than  $m$ . Since  $\Gamma\Gamma^*$  is invertible the lines  $\Gamma^1, \dots, \Gamma^n \in R^m$  of  $\Gamma$  are linearly independent. We denote  $S = Vect\{\Gamma^1, \dots, \Gamma^n\}$  and we take  $\Gamma^{n+1}, \dots, \Gamma^m$  to be an orthonormal basis in the orthogonal of  $S$ . Then we define  $\tilde{\Gamma} \in \mathcal{M}_{m \times m}$  to be the matrix with lines  $\Gamma^1, \dots, \Gamma^n, \Gamma^{n+1}, \dots, \Gamma^m$  and we notice that

$$\tilde{\Gamma}\tilde{\Gamma}^* = \begin{pmatrix} \Gamma\Gamma^* & 0 \\ 0 & I_{m-n} \end{pmatrix}$$

where  $I_{m-n} \in \mathcal{M}_{(m-n) \times (m-n)}$  is the identity matrix. In particular  $\det \tilde{\Gamma} = \sqrt{\det \Gamma\Gamma^*}$  and for  $z = (z_1, z_2), z_1 \in R^n, z_2 \in R^{m-n}$  we have  $|z|_{\tilde{\Gamma}}^2 = |z_1|_{\Gamma}^2 + |z_2|^2$ . Moreover, for  $y \in R^n$  we denote  $\tilde{y} = (y, 0)$  and we also set  $\tilde{\eta}(\theta) = (\eta(\theta), 0)$ . So, we define  $H = \tilde{y} + \tilde{\Gamma}\Theta + \tilde{\eta}(\Theta)$ , and we notice that  $h_{\tilde{\eta}} = h_{\eta}$  and  $c_*(\tilde{\eta}, 4r) = c_*(\eta, 4r)$ . For the density of  $H = (H_1, H_2)$  we can use the estimate from the previous step. Notice that since  $H_2$  is an orthogonal transformation of a Gaussian random variable, one easily gets that the estimates hold for  $(z, u) \in R^m$  such that  $z \in B(0, r)$  and  $u \in R^{m-n}$ . Now, since  $H_1 = G$  we obtain

$$\begin{aligned} p_G(z) &= \int_{R^{m-n}} p_H(z, u) du \\ &\geq \int_{R^{m-n}} \frac{1}{(8\pi)^{d/2} \sqrt{\det Q} |\det \tilde{\Gamma}|} \exp\left(-\frac{2}{\underline{\lambda}(Q)} \left(|z - y|_{\Gamma}^2 + |u|^2\right)\right) du \\ &= \frac{(4\lambda_*(Q))^{\frac{1}{2}(m-n)}}{(8\pi)^{m/2} \sqrt{\det Q} \sqrt{|\det \Gamma\Gamma^*|}} \exp\left(-\frac{2}{\underline{\lambda}(Q)} |z - y|_{\Gamma}^2\right). \end{aligned}$$

The proof of the other inequality is the same.  $\square$

## 7 Appendix 3. Support Property

In this section we prove (30). Let  $B = (B^1, \dots, B^{d-1})$  be a standard Brownian motion. We consider the analogues of the covariance matrix  $Q_i(B)$  considered in the previous sections: we define a symmetric square matrix of dimension  $d \times d$  by

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

and we denote by  $\underline{\lambda}(Q)$  (respectively by  $\bar{\lambda}(Q)$ ) the lower (respectively larger) eigenvalue of  $Q$ .

For a measurable function  $g : [0, 1] \rightarrow R^{d-1}$  we denote

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

We need the following two preliminary lemmas.

**Lemma 22** 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

$$\underline{\lambda}(Q) = \inf_{|\xi|=1} (\alpha_B^2(\xi) + \beta_B(\xi)) \quad \text{and} \quad \bar{\lambda}(Q) \leq \sup_{|\xi|=1} (\alpha_B^2(\xi) + \beta_B(\xi)) \leq \left(1 + \sup_{t \leq 1} |B_t|\right)^2.$$

Taking  $\xi_* = 0$  and  $\xi_d = 1$  we obtain  $\langle Q\xi, \xi \rangle = 1$  so that  $\underline{\lambda}(Q) \leq 1 \leq \bar{\lambda}(Q)$ .

**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 23** For each  $p \geq 1$  one has

$$E(|\det Q|^{-p}) \leq C_{p,d} < \infty \tag{58}$$

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

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

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

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

We integrate and we use Schwartz inequality in order to obtain

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



For each fixed  $\xi_*$  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  $V_{\xi_*}$  is the variance of  $b_{\xi_*}$  with respect to the time. Then it is proved in [9] (see (1.f), p. 183) that

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

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

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

$$q(B) = \sum_{i=1}^{d-1} |B_1^i| + \sum_{j \neq p} \left| \int_0^1 B_s^j dB_s^p \right| \quad (59)$$

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

$$\Lambda_{\rho, \varepsilon}(B) = \{ \det Q(B) \geq \varepsilon^\rho, \sup_{t \leq 1} |B_t| \leq \varepsilon^{-\rho}, q(B) \leq \varepsilon \}. \quad (60)$$

**Proposition 24** *There exist some universal constants  $c_{\rho, d}, \varepsilon_{\rho, d} \in (0, 1)$  (depending on  $\rho$  and  $d$  only) such that for every  $\varepsilon \in (0, \varepsilon_{\rho, d})$  one has*

$$P(\Lambda_{\rho, \varepsilon}(B)) \geq c_{\rho, d} \times \varepsilon^{\frac{1}{2}d(d+1)}. \quad (61)$$

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

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

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  $\left| \int_0^1 B_s^j dB_s^p \right| \leq |B_1^j| |B_1^p| + \left| \int_0^1 B_s^j dB_s^j \right|$  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  $X = (X^i, X^{j, p}, i = 1, \dots, d, 1 \leq j < p \leq d)$  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 [13] or [2]). 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

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

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

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

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

## 8 Appendix 4. Norms and distances

In this section we use the notation from Section 3 and 4. We consider the matrix  $A$  with columns  $a_i, [a]_{j,p} = a_{j,p} - a_{p,j}, i = 1, \dots, d, j \neq p$  defined in Section 3.2 and we assume that the non degeneracy condition (20) holds. For notational convenience we denote  $A_i = a_i, i = 1, \dots, d$  and  $A_i, i = d + 1, \dots, m$  will be an enumeration of  $[a]_{j,p}, 1 \leq j, p \leq d, j \neq p$ . We work with the norm  $|y|_{A_R}^2 = \langle (A_R A_R^*)^{-1} y, y \rangle, y \in R^n$ . We have the following simple properties:

**Lemma 25** *i) For every  $y \in R^n$  and  $0 < R \leq R' \leq 1$  one has*

$$\sqrt{\frac{R}{R'}} |y|_{A_R} \geq |y|_{A_{R'}} \geq \frac{R}{R'} |y|_{A_R} \quad \text{and} \quad (62)$$

$$\frac{1}{\sqrt{R\sqrt{\lambda^*(A)}}} |y| \leq |y|_{A_R} \leq \frac{1}{R\sqrt{\lambda_*(A)}} |y|. \quad (63)$$

*ii) For every  $z \in R^m$  and  $R > 0$  one has*

$$|A_R z|_{A_R} \leq |z|. \quad (64)$$

*iii) For every  $\mu \in L^2([0, T]; R^m)$  and  $R > 0$  one has*

$$\left| \int_0^t \mu_s ds \right|_{A_R}^2 \leq t \int_0^t |\mu_s|_{A_R}^2 ds, \quad t \in [0, T]. \quad (65)$$

**Proof.** *i)* It is easy to check that

$$\frac{R'}{R} A_R A_R^* \leq A_{R'} A_{R'}^* \leq \left( \frac{R'}{R} \right)^2 A_R A_R^*$$

which is equivalent with (62). This also implies (one takes  $R' = 1$  so  $A_{R'} = A$ ) that

$$\frac{1}{R} \lambda_*(A_R) \leq \lambda_*(A) \leq \frac{1}{R^2} \lambda_*(A_R) \quad \text{and} \quad \frac{1}{R} \lambda^*(A_R) \leq \lambda^*(A) \leq \frac{1}{R^2} \lambda^*(A_R)$$

which immediately gives (63).

*ii)* For  $z \in R^m$ , we write  $z = A_R^* y + w$  with  $y \in R^n$  and  $w \in (\text{Im} A_R^*)^\perp = \text{Ker} A_R$ . Then  $A_R z = A_R A_R^* y$  so that

$$\begin{aligned} |A_R z|_{A_R}^2 &= |A_R A_R^* y|_{A_R}^2 = \langle (A_R A_R^*)^{-1} A_R A_R^* y, A_R A_R^* y \rangle \\ &= \langle z, A_R A_R^* z \rangle = \langle A_R^* y, A_R^* y \rangle = |A_R^* y|^2 \leq |z|^2 \end{aligned}$$

and (64) holds.

iii) For  $\mu \in L^2([0, T]; R^m)$  and  $t \in [0, T]$  one has

$$\begin{aligned}
\left| \int_0^t \mu_s ds \right|_{A_R}^2 &= \langle A_R^{-1} \int_0^t \mu_s ds, \int_0^t \mu_s ds \rangle = \int_0^t \int_0^t \langle A_R^{-1} \mu_s, \mu_u \rangle ds du \\
&= \frac{1}{2} \int_0^t \int_0^t \left( \langle A_R^{-1}(\mu_s - \mu_u), \mu_s - \mu_u \rangle - \langle A_R^{-1} \mu_s, \mu_s \rangle - \langle A_R^{-1} \mu_u, \mu_u \rangle \right) ds du \\
&= \frac{1}{2} \int_0^t \int_0^t \left( |\mu_s - \mu_u|_{A_R}^2 - 2|\mu_s|_{A_R}^2 \right) ds du \\
&\leq \int_0^t \int_0^t |\mu_u|_{A_R}^2 ds du = t \int_0^t |\mu_u|_{A_R}^2 du.
\end{aligned}$$

□

We give now some lower and upper bounds for  $|y|_{A_R}$ . We denote  $S = \text{Vect}\{A_i, i = 1, \dots, d\}$  and  $\Pi_S$  is the projection on  $S$ .  $S^\perp$  is the orthogonal of  $S$  and  $\Pi_{S^\perp}$  is the projection on  $S^\perp$ . Moreover we denote

$$\begin{aligned}
\underline{\lambda}_S &= \inf_{\xi \in S, |\xi|=1} \sum_{i=1}^d \langle A_i, \xi \rangle^2, & \bar{\lambda}_S &= \sup_{\xi \in S, |\xi|=1} \sum_{i=1}^d \langle A_i, \xi \rangle^2 \\
\underline{\lambda}_{S^\perp} &= \inf_{\xi \in S^\perp, |\xi|=1} \sum_{i=d+1}^m \langle A_i, \xi \rangle^2, & \bar{\lambda}_{S^\perp} &= \sup_{\xi \in S^\perp, |\xi|=1} \sum_{i=d+1}^m \langle A_i, \xi \rangle^2.
\end{aligned} \tag{66}$$

By the very definition  $\underline{\lambda}_S > 0$  and under assumption (20) we also have  $\underline{\lambda}_{S^\perp} > 0$ . And  $\bar{\lambda}_S \leq \lambda^*(A)$ ,  $\bar{\lambda}_{S^\perp} \leq \lambda^*(A)$ .

**Proposition 26** *Suppose that (20) holds and let*

$$R < \frac{\underline{\lambda}_S}{4 \sum_{i=d+1}^m |\Pi_S A_i|^2}. \tag{67}$$

Then for every  $y \in R^n$

$$\frac{1}{4R \bar{\lambda}_S} |\Pi_S y|^2 + \frac{1}{4R^2 \bar{\lambda}_{S^\perp}} |\Pi_{S^\perp} y|^2 \leq |y|_{A_R}^2 \leq \frac{4}{R \underline{\lambda}_S} |\Pi_S y|^2 + \frac{4}{R^2 \underline{\lambda}_{S^\perp}} |\Pi_{S^\perp} y|^2. \tag{68}$$

In particular, if  $|A|_\infty = \max_{i=1, \dots, m} |A_i|$  and  $R \leq \underline{\lambda}_S / 4m |A|_\infty$  then

$$\frac{1}{4R |A|_\infty} |\Pi_S y|^2 + \frac{1}{4R^2 |A|_\infty} |\Pi_{S^\perp} y|^2 \leq |y|_{A_R}^2 \leq \frac{4}{R \underline{\lambda}_S} |\Pi_S y|^2 + \frac{4}{R^2 \underline{\lambda}_{S^\perp}} |\Pi_{S^\perp} y|^2. \tag{69}$$

**Proof.** In a first stage we assume that  $A_i \perp A_j$  for  $i \leq d < j$ . We will drop out this restriction in the second part of the proof. Let  $T_S$  and  $T_{S^\perp}$  be the restriction of  $y \mapsto A_R A_R^* y$  to  $S$  and to  $S^\perp$  respectively. Since

$$A_R A_R^* y = R \sum_{i=1}^d \langle A_i, y \rangle A_i + R^2 \sum_{i=d+1}^m \langle A_i, y \rangle A_i$$

our orthogonality hypothesis implies that  $T_S y = R \sum_{i=1}^d \langle A_i, y \rangle A_i \in S$  for  $y \in S$  and  $T_{S^\perp} y = R^2 \sum_{i=d+1}^m \langle A_i, y \rangle A_i \in S^\perp$  for  $y \in S^\perp$ . Since  $A_R A_R^*$  is invertible it follows that  $T_S$  (respectively  $T_{S^\perp}$ ) is an invertible operator from  $S$  (respectively from  $S^\perp$ ) into itself. For  $y \in S$  we have

$$R \bar{\lambda}_S |y|^2 \geq \langle T_S y, y \rangle = R \sum_{i=1}^d \langle A_i, y \rangle^2 \geq R \underline{\lambda}_S |y|^2$$

and since  $|y|_{A_R}^2 = \langle T_S^{-1} y, y \rangle$  for  $y \in S$ , we obtain

$$\frac{1}{R \bar{\lambda}_S} |y|^2 \leq |y|_{A_R}^2 \leq \frac{1}{R \underline{\lambda}_S} |y|^2, \quad \text{if } y \in S.$$

Similarly, we get

$$\frac{1}{R^2 \bar{\lambda}_{S^\perp}} |y|^2 \leq |y|_{A_R}^2 \leq \frac{1}{R^2 \underline{\lambda}_{S^\perp}} |y|^2, \quad \text{if } y \in S^\perp.$$

Let  $y \in R^n$ . Since  $(A_R A_R^*)^{-1} \Pi_S y \in S$  we have  $\langle (A_R A_R^*)^{-1} \Pi_S y, \Pi_{S^\perp} y \rangle = 0$  so that  $|y|_{A_R}^2 = |\Pi_S y|_{A_R}^2 + |\Pi_{S^\perp} y|_{A_R}^2$ . We obtain

$$\frac{1}{R \bar{\lambda}_S} |\Pi_S y|^2 + \frac{1}{R^2 \bar{\lambda}_{S^\perp}} |\Pi_{S^\perp} y|^2 \leq |y|_{A_R}^2 \leq \frac{1}{R \underline{\lambda}_S} |\Pi_S y|^2 + \frac{1}{R^2 \underline{\lambda}_{S^\perp}} |\Pi_{S^\perp} y|^2. \quad (70)$$

We drop now out the orthogonality assumption. For  $j > d$  we consider the decomposition  $A_j = \Pi_S A_j + \Pi_{S^\perp} A_j$  and we define the matrices  $\bar{A}_R = (\sqrt{R} A_1, \dots, \sqrt{R} A_d, R \Pi_{S^\perp} A_{d+1}, \dots, R \Pi_{S^\perp} A_m)$  and  $\hat{A}_R = (0, \dots, 0, R \Pi_S A_{d+1}, \dots, R \Pi_S A_m)$  so that  $A_R = \bar{A}_R + \hat{A}_R$ . We will check that under the restriction (67) we have

$$4 \left| \bar{A}_R^* y \right|^2 \geq |A_R^* y|^2 \geq \frac{1}{4} \left| \bar{A}_R^* y \right|^2 \quad \forall y \in R^n. \quad (71)$$

We suppose for the moment that the above inequality is true and we prove (68). Since  $|A_R^* y|^2 = \langle A_R A_R^* y, y \rangle$  the above inequality means that  $4 \bar{A}_R \bar{A}_R^* \geq A_R A_R^* \geq \frac{1}{4} \bar{A}_R \bar{A}_R^*$  and this gives

$$\frac{1}{4} |y|_{\bar{A}_R}^2 \leq |y|_{A_R}^2 \leq 4 |y|_{\bar{A}_R}^2. \quad (72)$$

Since the columns of  $\bar{A}_R$  verify the orthogonality assumption we may use the result from the first step with  $A$  replaced with  $\bar{A} = (\bar{A}_1, \dots, \bar{A}_d, \bar{A}_{d+1}, \dots, \bar{A}_m)$ , with  $\bar{A}_j = A_j$  if  $j \leq d$  and  $\bar{A}_j = \Pi_{S^\perp} A_j$  for  $j > d$ . Here, we have  $\bar{S} = \text{Vect}\{\bar{A}_1, \dots, \bar{A}_d\} = \text{Vect}\{A_1, \dots, A_d\} = S$ , so that  $\underline{\lambda}_{\bar{S}} = \underline{\lambda}_S$  and  $\bar{\lambda}_{\bar{S}} = \bar{\lambda}_S$ . Moreover, since  $\bar{S}^\perp = S^\perp$ , the computations in (66) are actually done with  $\xi \in S^\perp$ , and thus we obtain  $\underline{\lambda}_{\bar{S}^\perp} = \underline{\lambda}_{S^\perp}$  and  $\bar{\lambda}_{\bar{S}^\perp} = \bar{\lambda}_{S^\perp}$ . So (70) gives

$$\frac{1}{R \bar{\lambda}_S} |\Pi_S y|^2 + \frac{1}{R^2 \bar{\lambda}_{S^\perp}} |\Pi_{S^\perp} y|^2 \leq |y|_{\bar{A}_R}^2 \leq \frac{1}{R \underline{\lambda}_S} |\Pi_S y|^2 + \frac{1}{R^2 \underline{\lambda}_{S^\perp}} |\Pi_{S^\perp} y|^2$$

which together with (72) imply (68).

It remains to prove (71). We have

$$\left| \overline{A_R^* y} \right|^2 \geq R \sum_{j=1}^d \langle A_j, y \rangle^2 = R \sum_{j=1}^d \langle A_j, \Pi_S y \rangle^2 \geq R \underline{\lambda}_S |\Pi_S y|^2$$

and

$$\left| \widehat{A_R^* y} \right|^2 = R^2 \sum_{j=d+1}^m \langle \Pi_S A_j, y \rangle^2 = R^2 \sum_{j=d+1}^m \langle \Pi_S A_j, \Pi_S y \rangle^2 \leq R^2 |\Pi_S y|^2 \sum_{j=d+1}^m |\Pi_S A_j|^2.$$

Then (67) gives  $\left| \overline{A_R^* y} \right|^2 \geq 4 \left| \widehat{A_R^* y} \right|^2$ . Using the inequality  $(a+b)^2 \geq \frac{1}{2}a^2 - b^2$  we obtain

$$|A_R^* y|^2 = \left| \overline{A_R^* y} + \widehat{A_R^* y} \right|^2 \geq \frac{1}{2} \left| \overline{A_R^* y} \right|^2 - \left| \widehat{A_R^* y} \right|^2 \geq \frac{1}{4} \left| \overline{A_R^* y} \right|^2$$

and using  $(a+b)^2 \leq 2a^2 + 2b^2$  we get  $|A_R^* y|^2 \leq 4 \left| \overline{A_R^* y} \right|^2$ .  $\square$

From now on we consider the specific situation when  $a_i = \sigma_i(t, x)$ ,  $[a]_{i,j} = [\sigma_i, \sigma_j](t, x)$  and we denote by  $A(t, x)$  respectively  $A_R(t, x)$  the matrices associated to these coefficients. We will need the following

**Lemma 27** *Let  $x, y \in R^n$  be such that  $|x - y| \leq 1$  and let  $s, t \in [0, 1]$ . Assume that*

$$|x - y| + |t - s| \leq \frac{\sqrt{\lambda_*(A(t, x))}}{(8dm)n^2(t, x)} \times \sqrt{\delta} \quad (73)$$

*Then for every  $z \in R^n$  and  $\delta \leq 1$  one has*

$$\frac{1}{4} |z|_{A_\delta(t, x)}^2 \leq |z|_{A_\delta(s, y)}^2 \leq 4 |z|_{A_\delta(t, x)}^2. \quad (74)$$

**Proof.** The inequality (74) is equivalent to

$$4(A_\delta A_\delta^*)(t, x) \geq (A_\delta A_\delta^*)(s, y) \geq \frac{1}{4}(A_\delta A_\delta^*)(t, x).$$

We use the numerical inequality  $(a+b)^2 \geq \frac{1}{2}a^2 - b^2$ , the hypothesis (73) and we obtain

$$\begin{aligned} \langle (A_\delta A_\delta^*)(s, y) z, z \rangle &= \sum_{k=1}^m \langle A_{\delta, k}(s, y), z \rangle^2 \\ &= \sum_{k=1}^m (\langle A_{\delta, k}(t, x), z \rangle + \langle A_{\delta, k}(s, y) - A_{\delta, k}(t, x), z \rangle)^2 \\ &\geq \frac{1}{2} \sum_{k=1}^m \langle A_{\delta, k}(t, x), z \rangle^2 - \sum_{k=1}^m (\langle A_{\delta, k}(s, y) - A_{\delta, k}(t, x), z \rangle)^2 \\ &\geq \frac{1}{2} \sum_{k=1}^m \langle A_{\delta, k}(t, x), z \rangle^2 - (2dm)^2 n^4(t, x) \delta (|x - y|^2 + |t - s|^2) \times |z|^2. \end{aligned}$$

Since  $\lambda_*(A_\delta(t, x)) \geq \delta^2 \lambda_*(A(t, x))$  our hypothesis says that

$$\begin{aligned} (2dm)^2 n^4(t, x) \delta(|x - y|^2 + |t - s|^2) \times |z|^2 &\leq \frac{1}{4} \delta^2 \lambda_*(A(t, x)) \times |z|^2 \\ &\leq \frac{1}{4} \lambda_*(A_\delta(t, x)) \times |z|^2 \\ &\leq \frac{1}{4} \sum_{k=1}^m \langle A_{\delta, k}(t, x), z \rangle^2 \end{aligned}$$

so that

$$\langle (A_\delta A_\delta^*)(s, y)z, z \rangle \geq \frac{1}{4} \sum_{k=1}^m \langle A_{\delta, k}(t, x), z \rangle^2 = \frac{1}{4} \langle (A_\delta A_\delta^*)(t, x)z, z \rangle.$$

Using  $(a + b)^2 \leq 2a^2 + 2b^2$  one proves the other inequality.  $\square$

In the last part of this section we establish the link between the norm  $|z|_{A_R(t, x)}$  and the control (Caratheodory) distance. We will use in a crucial way the alternative characterizations given in [14] for this distance - and these results hold in the homogeneous case: the coefficients of the equations do not depend on time any more, so that we suppose now  $\sigma_j(t, x) = \sigma_j(x)$ . Consequently, we handle the matrix  $A_R(x)$  instead of  $A_R(t, x)$ .

We first introduce a semi-distance  $d$  on an open set  $\Omega \subset R^n$  which is naturally associated to the family of norms  $|y|_{A_R(x)}$ .

We set  $\Omega = \{x \in R^n : \lambda_*(A(x)) > 0\} = \{x : \det(AA^*(x)) \neq 0\}$ , which is open because  $x \mapsto \det AA^*(x)$  is continuous. Notice that if  $x \in \Omega$  then  $\det A_R A_R^*(x) > 0$  for every  $R > 0$ . For  $x, y \in \Omega$ , we define  $d(x, y)$  by  $d(x, y) < \sqrt{R}$  if and only if  $|y - x|_{A_R(x)} < 1$ . The motivation of taking  $\sqrt{R}$  is the following: if we are in the elliptic case then  $|y - x|_{A_R(x)} \sim R^{-1/2} |y - x|$  so  $|y - x|_{A_R(x)} \leq 1$  amounts to  $|y - x| \leq \sqrt{R}$ .

It is straightforward to see that  $d$  is a semi-distance on  $\Omega$ , in the sense that  $d$  verifies the following three properties (see [14]):

- i) for every  $r > 0$ , the set  $\{y \in \Omega : d(x, y) < r\}$  is open;
- ii)  $d(x, y) = 0$  if and only if  $x = y$ ;
- iii) for every compact set  $K \Subset \Omega$  there exists  $C > 0$  such that for every  $x, y, z \in K$  one has  $d(x, y) \leq C(d(x, z) + d(z, y))$ .

Moreover, one says that  $d_1 : \Omega \times \Omega \rightarrow R_+$  and  $d_2 : \Omega \times \Omega \rightarrow R_+$  are equivalent if for every compact set  $K \Subset \Omega$  there exists a constant  $C$  such that for every  $x, y \in K$

$$\frac{1}{C} d_1(x, y) \leq d_2(x, y) \leq C d_1(x, y). \quad (75)$$

In particular if  $d_1$  is a distance and  $d_2$  is equivalent with  $d_1$  then  $d_2$  is a semi-distance. And one says that  $d_1$  is locally equivalent with  $d_2$  if for every  $x_0 \in \Omega$  there exists a neighborhood  $V$  of  $x_0$  and a constant  $C$  such that (75) holds for every  $x, y \in V$ .

We introduce now the control metric. For  $x, y \in R^n$  we denote by  $C(x, y)$  the set of controls  $\psi \in L^2([0, 1]; R^n)$  such that the corresponding skeleton  $du_t(\psi) = \sum_{j=1}^d \sigma_j(u_t(\psi)) \psi_t^j dt$

with  $u_0(\psi) = x$  satisfies  $u_1(\psi) = y$ . Notice that the drift  $b$  does not appear in the equation of  $u_t(\psi)$ . Then we define

$$d_c(x, y) = \inf \left\{ \left( \int_0^1 |\psi_s|^2 ds \right)^{1/2} : \psi \in C(x, y) \right\}.$$

**Theorem 28 A.** *Let*

$$\alpha(x) = \frac{\lambda_*^{1/2}(A(x))}{256d^6n^6(x)}.$$

*Then for every  $x, y \in \Omega$  such that  $d_c(x, y) \leq \frac{1}{4}\alpha^2(x)$  one has  $d(x, y) \leq 9\alpha^2(x)d_c(x, y)$ .*

**B.**  *$d$  is locally equivalent to  $d_c$  on  $\Omega$ .*

**C.** *In particular for every compact set  $K \Subset \Omega$  there exists  $r_K$  and  $C_K$  such that for every  $x, y \in K$  with  $d(x, y) \leq r_K$  one has  $d_c(x, y) \leq C_K d(x, y)$ .*

**Proof A.** Let  $\delta > 0$ ,  $x, y \in \Omega$  and  $\psi \in C(x, y)$ . Setting  $x_t = u_{t/\delta}(\psi)$ , we obtain  $dx_t = \sum_{j=1}^d \sigma_j(x_t) \phi_t^j dt$  with  $\phi(t) = \delta^{-1}\psi(t\delta^{-1})$ , which means that  $x_t = u_t(\phi)$ . Notice also that  $\int_0^1 |\psi_s|^2 ds = \delta \int_0^\delta |\phi_s|^2 ds$ .

We denote now  $C_\delta(x, y)$  the set of controls  $\phi \in L^2([0, \delta]; R^n)$  such that the corresponding skeleton  $u_t(\phi)$  with  $u_0(\phi) = x$  verifies  $u_\delta(\phi) = y$ . As a consequence of the previous computations, one has

$$d_c(x, y) = \sqrt{\delta} \times \inf \left\{ \left( \int_0^\delta |\phi_s|^2 ds \right)^{1/2} : \phi \in C_\delta(x, y) \right\} \equiv \sqrt{\delta} \times \inf \{ \varepsilon_\phi(\delta) : \phi \in C_\delta(x, y) \}.$$

Suppose now that  $d_c(x, y) \leq \frac{1}{4}\alpha^2(x)$ . We take  $\delta = \frac{1}{4}\alpha^2(x)$  so that  $d_c(x, y) \leq \frac{1}{2}\alpha(x)\sqrt{\delta}$ . Then one may find a control  $\phi \in C_\delta(x, y)$  such that  $\varepsilon_\phi(\delta) \leq \frac{1}{2}\alpha(x)$  and  $y = x_\delta(\phi)$ . Since  $\varepsilon_\phi(\delta) + \sqrt{\delta} \leq \alpha(x)$  we may use (38) and we obtain  $|y - x|_{A_\delta(x)} \leq 4\varepsilon_\phi(\delta) + \sqrt{\delta} \leq 3\alpha(x)$ . It follows that

$$|y - x|_{A_{9\alpha^2(x)\delta}(x)} \leq \frac{1}{3\alpha(x)} |y - x|_{A_\delta(x)} \leq 1$$

and this gives  $d(x, y) \leq 9\alpha^2(x)\delta = 9\alpha^2(x) \times \frac{1}{4}\alpha^2(x)$ . And this guarantees that  $d(x, y) \leq 9\alpha^2(x) \times d_c(x, y)$ .

**B.** We prove now the converse inequality. We use the results from [14] so we recall the definition of the semi-distance  $d_*$  (which is denoted by  $\rho_2$  in [14]). Given  $\phi^i, \phi^{k,j} \in R$ ,  $i = 0, \dots, d, 1 \leq k < j \leq d$  we consider the equation

$$v_t(\phi) = x + \int_0^t \left( \sum_{j=1}^d \phi^j \sigma_j(v_s(\phi)) + \sum_{i \neq j} \phi^{i,j} [\sigma_i, \sigma_j](v_s(\phi)) \right) ds. \quad (76)$$

Notice that  $\phi^j, \phi^{i,j}$  are now real constants (in contrast with the time depending controls in the standard skeleton) and we have added the vector fields  $[\sigma_i, \sigma_j]$  which does not appear in skeletons. And the drift term  $b$  does not appear. We denote by  $P_*(x, y)$  the family of paths  $v_t(\phi)$  which satisfy (76) and such that  $v_1(\phi) = y$ . We define  $d_*$  by:  $d_*(x, y) < \delta$  if and only if one may find  $\phi^i, \phi^{k,j} \in R$ ,  $1 \leq i, k, j \leq d, j \neq k$  such that  $v_t(\phi) \in P_*(x, y)$  and  $|\phi^i| < \delta, |\phi^{i,j}| < \delta^2$ . As a consequence of Theorem 2 and Theorem 4 from [14]  $d_*$  is

locally equivalent with  $d_c$ . So our aim is to prove that  $d_*$  is locally dominated by  $d$ . Let us be more precise: we fix  $x \in \Omega$  and we look for two constants  $C_x, \delta_x > 0$  such that the following holds: if  $0 < \delta \leq \delta_x$  and  $d(x, y) \leq \sqrt{\delta}$  then one may construct a control  $\phi \in R^m$  such that  $v(\phi) \in P_*(x, y)$  and  $|\phi^i| < C_x \sqrt{\delta}, |\phi^{i,j}| < C_x^2 \delta$ . This implies  $d_*(x, y) \leq C_x \sqrt{\delta}$ , and the statement will hold. Notice that we discuss local equivalence, that is why we may take  $C_x, \delta_x$  depending on  $x$ .

We recall that  $A_i(x), i = 1, \dots, m$  is an enumeration of  $\sigma_i(x), [\sigma_j, \sigma_p](x), i, j, p = 1, \dots, d$  and that they span  $R^n$  because  $x \in \Omega$ . So, we choose  $i_1 < \dots < i_{d'} \leq d < i_{d'+1} < \dots < i_n \leq m$  such that  $A_{i_k}(x), k = 1, \dots, d'$  span  $Vect\{A_1(x), \dots, A_d(x)\}$  and  $A_{i_k}(x), k = 1, \dots, n$  span  $R^n$ . In particular all of them are linearly independent. We denote  $B_k(x) = A_{i_k}(x)$  and we want to use Theorem 20 for them. Notice that  $Vect\{A_1(x), \dots, A_d(x)\} = Vect\{B_1(x), \dots, B_{d'}(x)\}$  so the projections  $\Pi$  and  $\Pi^\perp$  considered in Theorem 20 and in Proposition 62 coincide. In particular if  $d(x, y) \leq \sqrt{\delta}$  then  $|\Pi(y - x)| \leq |A(x)|_\infty \sqrt{\delta}$  and  $|\Pi^\perp(y - x)| \leq |A(x)|_\infty \delta$ . And this also implies that  $|y - x| \leq 2|A(x)|_\infty \sqrt{\delta}$ .

As  $\theta \in R^n$ , we look for a solution to the equation

$$y = \xi_1(\theta), \quad \text{with} \quad \xi_t(\theta) = x + \sum_{k=1}^n \theta_k \int_0^t B_k(\xi_s(\theta)) ds = x + \sum_{k=1}^n \theta_k \int_0^t A_{i_k}(\xi_s(\theta)) ds.$$

So, we write it as

$$y = x + B(x)\theta + r(\theta) \quad \text{with} \quad r(\theta) = \sum_{k=1}^n \theta_k \int_0^t (B_k(\xi_s(\theta)) - B_k(x)) ds.$$

Clearly  $r \in C^3(R^n, R^n)$  and  $r(0) = \nabla r(0) = 0$ . Then,

$$|y - x| \leq 2|A(x)|_\infty \sqrt{\delta} \leq 2|A(x)|_\infty \sqrt{\delta_x}$$

and we suppose that  $\delta_x$  is sufficiently small in order that  $|y - x|$  satisfies (49), that is

$$|y - x| < \frac{\lambda_*(B(x))^{1/2}}{4} \quad \text{and} \quad |y - x| < \frac{\lambda_*(B(x))}{8d^3(c_2(r) + c_2(r))}.$$

We use then (51) and we obtain  $|\theta_i| \leq C_x \sqrt{\delta}, i = 1, \dots, d'$  and  $|\theta_i| \leq C_x^2 \delta, i = d' + 1, \dots, n$ . This proves that  $d_*(x, y) \leq C_x \sqrt{\delta}$ .

**C.** For  $x \in \Omega$ , we denote  $B_d(x, r) := \{y \in \Omega : d(x, y) < r\}$  and this is an open set. Since  $d$  and  $d_c$  are locally equivalent for every compact  $K \Subset \Omega$  and for every  $x \in K$  there exists  $C_x, \varepsilon_x > 0$  such that for  $y \in B_d(x, \varepsilon_x)$  we have  $d_c(x, y) \leq C_x d(x, y)$ . Since the set  $K$  is compact we may find  $x_1, \dots, x_N \in K$  such that  $K \subset \cup_{i=1}^N B_d(x_i, \varepsilon_{x_i})$ . We denote  $C_{\max} = \max_{i=1, \dots, N} C_{x_i}$ . Let us prove that there exists  $r_* > 0$  such that for every  $x \in K$  and every  $y \in B_d(x, r_*)$  we have  $d_c(x, y) \leq C_{\max} d(x, y)$ .

For  $x \in K$  one may find  $i$  such that  $x \in B_d(x_i, \varepsilon_{x_i})$  and  $r > 0$  such that  $B_d(x, r) \subset B_d(x_i, \varepsilon_{x_i})$ . We define  $r_x = \sup\{r > 0 : \exists i \in \{1, \dots, N\} \text{ such that } B_d(x, r) \subset B_d(x_i, \varepsilon_{x_i})\}$ . We claim that  $r_* := \inf_{x \in K} r_x > 0$ . Indeed suppose that this is not true. Then one may find a sequence  $y_n \rightarrow y_0$  such that  $r_{y_n} \rightarrow 0$ . Since  $r_{y_0} > 0$  there exists  $n_*$  such that for  $n \geq n_*$  one has  $B_d(y_n, \frac{1}{2C_K} r_{y_0}) \subset B_d(y_0, r_{y_0}) \subset B_d(x_i, \varepsilon_{x_i})$  for some  $i$ . Here  $C_K$  is the



constant in the triangle inequality *iii*) at page 37. And this means that  $r_{y_n} \geq \frac{1}{2}r_{y_0} > 0$  which is in contradiction with our hypothesis. So we have proved that  $r_* > 0$ .

Consider now  $y \in B_d(x, r_*)$ . There exists  $i$  such that  $B_d(x, r_*) \subset B_d(x_i, \varepsilon_{x_i})$  and this means that  $y, x \in B_d(x_i, \varepsilon_{x_i})$  and consequently  $d_c(x, y) \leq C_{x_i}d(x, y) \leq C_{\max}d(x, y)$ .  $\square$

Finally we give:

**Proof of Proposition 6.** We will first prove that under our hypothesis  $d(x, y) \leq \sqrt{\lambda_x/(4m)n^4(x)}$ . Let  $R$  be such that  $d(x, y) \geq \sqrt{R}$  so that  $|y - x|_{A_R(x)} \geq 1$ . Then by (63)

$$\frac{\lambda_x \sqrt{\lambda_*(A)}}{(4m)n^4(x)} \geq |y - x| \geq R \sqrt{\lambda_*(A)} |y - x|_{A_R} \geq R \sqrt{\lambda_*(A)}.$$

It follows that  $R \leq \lambda_x/(4m)n^4(x)$  which proves our assertion.

We suppose now that  $d(x, y) \geq \sqrt{R}$ . Since  $R \leq \lambda_x/(4m)n^4(x)$  we may use (8) and we obtain

$$\frac{4}{R\lambda_x} |\Pi_x(y - x)|^2 + \frac{4}{R^2\lambda_x^\perp} |\Pi_x^\perp(y - x)|^2 \geq |y - x|_{A_R(x)} \geq 1$$

which gives  $\bar{d}(x, y) \geq \sqrt{R}$ . So  $\bar{d}(x, y) \geq d(x, y)$ . Suppose that  $d(x, y) < \sqrt{R}$ . Then

$$\frac{1}{4Rn^2(x)} |\Pi_x(y - x)|^2 + \frac{1}{4R^2n^2(x)} |\Pi_x^\perp(y - x)|^2 \leq |y - x|_{A_R(x)}^2 < 1$$

and this reads  $\underline{d}(x, y) < \sqrt{R}$  which gives  $\underline{d}(x, y) \leq d(x, y)$ .  $\square$

## References

- [1] BALLY V., FERNANDEZ B. AND MEDA A. (2011). Estimates for the probability that a Itô process remains near a path. *Stochastic Processes and their Applications* **121**, 2087-2113.
- [2] BALLY V. AND CARAMELLINO L. (2011) Positivity and lower bounds for the density of Wiener functionals. arXiv.org math arXiv:1004.5269
- [3] BALLY V. AND DE MARCO S. (2011) Tube estimates for Itô processes and density minorations for log-normal type diffusions. Working paper.
- [4] BEN AROUS G. AND GRADINARU M. (1993). Normes Hölderiennes et support des diffusions. *C.R. Acad. Sci. Paris Sér. I Math.* **16** (3), 283-286.
- [5] BEN AROUS G., GRADINARU M. AND LEDOUX M. (1994). Hölder norms and the support theorem for diffusions. *Ann. Inst. H. Poincaré Probab. Statist.* **30** (3), 415-436.
- [6] BICHTLER K., GRAVEREAUX J-B. AND JACOD J. (1987). *Malliavin Calculus for Processes with Jumps*. Stochastic monographs Volume 2, Gordon and Breach Science Publishers.

- [7] CAPITAINE M. (1995). Onsager-Machlup functional for some smooth norms on Wiener space. *Prob. Th. Rel. Fields* **102**, 189-201.
- [8] CAPITAINE M. (2000). On the Onsager-Machlup functional for elliptic diffusion processes. *Séminaires de Probabilités (Strasbourg)* **34**, 313-328.
- [9] DONATI-MARTIN C. AND YOR M. (1991). Fubini's theorem for double Wiener integrals and the variance of the Brownian path. *Ann. Inst. Henri Poincaré* **27**, 181-200.
- [10] FRIZ P., LYONS T. AND STROOCK D. (2006). Lévy's area under conditioning. *Ann. Inst. H. Poincaré Probab. Statist.* **42**, 89-101.
- [11] HARA K. AND TAKAHASHI Y. (1996). Lagrangian for pinned diffusion processes. *Itô's Stochastic Calculus and Probability Theory*, 117-128, Springer.
- [12] IKEDA N. AND WATANABE S. (1981). *Stochastic Differential Equations and Diffusion Processes*, North Holland.
- [13] KUSUOKA S. AND STROOCK D. (1987). Applications of the Malliavin calculus. III. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **34**, 391-442.
- [14] NAGEL A., STEIN E.M. AND WEINGER S. (1985). Balls and metrics defined by vector fields, I. Basic properties. *Acta Math.* **155**, 103-147.
- [15] STROOCK D. AND VARADHAN S. (1970). On the support of diffusion processes with applications to the strong maximum principle. In: *Proceedings of the Sixth Berkeley Symposium of Mathematical Statistics and Probability, vol III, Probability Theory*, 333-359.
- [16] ZEITOUNI O. (1989). On the Onsager-Machlup functional for diffusion processes around non  $C^2$  curves. *Ann. Probab.* **17**, 1037-1054.