PhD-FSTC-2017-60
Faculty of Science, Technology and Communication

DISSERTATION

Defence held on the 7th of September in Esch-sur-Alzette
to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG EN MATHÉMATIQUES

by

ANTOINE LEDENT
Born on the 21st of June 1991 in Verviers, Belgium

KUSUOKA-STROOCK TYPE BOUNDS FOR DENSITIES RELATED TO LOW-DIMENSIONAL PROJECTIONS OF SOLUTIONS TO HIGH-DIMENSIONAL SDE

DISSERTATION DEFENCE COMMITTEE

Dr Giovanni PECCATI, chairman
Professeur, University of Luxembourg

Dr Ivan NOURDIN, vice chairman
Professeur, University of Luxembourg

Dr Anton THALMAIER, dissertation supervisor
Professeur, University of Luxembourg

Dr Vlad BALLY, member of the committee
Professeur, Université Paris-Est Marne-la-Vallée

Dr Eva LÖCHERBACH, member of the committee
Professeur, Université Cergy-Pontoise

Contents

Acknowledgements

At this point, I would like to thank all the people who made this thesis possible for me.

First of all, I would like to express my gratitude to my PhD supervisor, Prof. Anton Thalmaier, for giving me the opportunity to write this thesis at the University of Luxembourg. I am grateful for his patience, his proofreading, his engagement and attention to the issues I raised, and his relentless encouragement in this challenging four-year enterprise. None of this would have been possible without him.

I would like to warmly thank all the members of my thesis committee, who have accepted to judge my research work. I am truly grateful to Prof. Ivan Nourdin and Prof. Giovanni Peccati for accepting to be part of the Jury and to carefully examine my work. I also especially thank Prof. Vlad Bally and Prof. Eva Löcherbach, who are great experts on many aspects of the study of Stochastic Differential Equations, and whose articles have been a great source of inspiration for my work. I am truly honoured that they accepted take some of their time to review my thesis. I would also like to thank Paolo Pigato, for helpful discussions and explanations of his articles in the context of his visit to Luxembourg as well as for more recent discussion, and Erlend Grong, for advice, and discussions about Lie Algebra theory.

A very special thanks is due to Xi Geng, for stimulating discussions and comments about my work and his own as well as other mathematical topics, for informally teaching me certain fundamental aspects of Rough Path theory, and, most importantly, for communicating his sincere passion and beautiful enthusiasm for mathematics to me.

Last but not least, I would like to thank my parents, Peihsuan Tsai, all my friends at the university in general, and in particular Lijuan Cheng, for their daily moral support and guidance in those perforce, occasionally trying times.

Abstract

One of the purposes of this thesis is to use Malliavin calculus and Stochastic Taylor expansions to study the densities of interacting systems of stochastic differential equations (SDE), seen as projections of SDE onto a low-dimensional space, and to control the dependence of the constants on the dimension of the background space. The setting includes time-dependent SDE and a relatively large class of path-dependent SDE. Several results also shed light even on the classical theory of SDE in $\mathbb{R}^n$, independently of the control on the constants.

In Part 1, assuming the system satisfies suitably defined projected equivalents of the classic ellipticity or weak Hörmander conditions, we prove Gaussian estimates in terms of the Euclidean distance where, provided natural assumptions, *for a fixed target-space dimension, the constants depend polynomially on the background dimension, and, in the elliptic case, on the number of driving Brownian motions*.

In Part 2, we first define suitable generalisations of (time-dependent) control distances and prove Kusuoka-Stroock type results without control on the constants. In particular, we obtain a time-dependent extension of a result of Léandre about SDE with non-trivial drifts, i.e., drifts which are not uniformly contained in the span of the other vector fields.

Then, we introduce a condition which we call the 'Progressive Hörmander condition' and prove similar control-type estimates valid under this assumption, with polynomial control on the growth of the constants with background space dimension. The condition is of independent interest in the study of SDE on $\mathbb{R}^n$, and shows the connection between the classic works of Ben Arous, Kusuoka, Léandre and Stroock, and the more recent works of Bally, Caramellino, Delarue, Menozzi and Pigato. To main technique required is the study of density and scaling properties of some careful choice of linear combinations of terms of the signature of the driving path.

In Part 3, we introduce a stricter condition called the 'separated progressive Hörmander condition', and prove lower bounds and local strict positivity under this assumption. (By 'local' we mean local around the solution of the deterministic ODE driven by a null control, rather than local round the initial point.) The main technical difficulty is the identification of points contained in the interior of the support of the log-signature[1] of the path in $\mathbb{R}^{d+1}$ composed of d Brownian motions and a deterministic linear component.

The purpose of Part 4 is to use some results and techniques of the rest of the thesis to prove extensions of a theorem of Löcherbach about uniformly elliptic interacting branching diffusions.

[1] Ignoring the deterministic time component (of the log-signature).

Introduction and main results

The purpose of this thesis is to prove bounds for the densities of low-dimensional quantities relative to solutions to high-dimensional (possibly degenerate, non-Hörmander) SDE satisfying either a modified form of ellipticity defined directly on the target space or a modified form of the weak Hörmander condition defined directly on the target space.

More specifically, let $\sigma^0, \sigma^1, \ldots, \sigma^d$ be sufficiently smooth vector fields on $\mathbb{R}^m$, and let X be the solution to the following Stratonovich SDE:

$$dX_t = \sum_{i=1}^{d} \sigma^i(X_t) \circ dW_t^i + \sigma^0(X_t)dt,$$

with initial condition $X_0 = x$, where W_t is a d-dimensional Brownian motion. Let $F\colon \mathbb{R}^m \to \mathbb{R}^n$ be a function that sends X to some smaller-dimensional space $\mathbb{R}^n$ (the initial space $\mathbb{R}^m$ is called the *background space*, and $\mathbb{R}^n$ the *target space*). Write $Y_t = F(X_t)$. We call the ordered list (x, σ, F) a *random SDE system* or simply *system*. We are interested in bounds on the density of Y_t in the target space (when this density exists). For all of the introduction and most of the thesis, we assume that F is linear.

The above model includes interacting systems of SDE, SDE with time-dependent coefficients, as well as a large class of SDE with path-dependent coefficients. Indeed, it is now understood (cf. [**12, 27**]) that reasonable paths are determined up to tree-like equivalence by the integrals of arbitrary functionals along it. The above setting allows for the possibility of including a finite number of integrals of functionals as elements of the background space.

Several results appear to be new even in the traditional setting of an SDE on $\mathbb{R}^n$.

Literature review: what is known about the case of a single SDE in $\mathbb{R}^n$ (i.e. $F = \mathrm{Id}$).

In this case, local integrability in space-time has previously been achieved by Gaussian bounds of the form

$$p_t(x, y) \leq C \frac{e^{-\frac{|y-x|^2}{Mt}}}{t^{\frac{n}{2}}}$$

in the elliptic case (cf. [**36**]) and by bounds in terms of the control distance of the following form:

$$p_t(x, y) \leq C \frac{e^{-\frac{d(x,y)^2}{Mt}}}{|B_d(x, \sqrt{t})|}$$

in [**37**] (or see [**13**] for an analytic approach) in the strong Hörmander case *with a drift uniformly inside the span of the diffusion vector fields*. In the case of weak Hörmander with arbitrary drift, there is of course the following Gaussian bound (proved in [**36**]):

$$p_t(x, y) \leq C \frac{e^{-\frac{|y-x|^2}{Mt}}}{t^{\frac{\nu}{2}}},$$

where ν is some number greater than the dimension and the order of hypoellipticity. This bound is not sharp, and indeed not even integrable in space-time.

Some work has been done on providing sharper bounds in the context of weak Hörmander with non-trivial drift. It has been achieved *locally* by Léandre in [**39**], where a bound of the following form is shown

$$p_t(x,y) \leq \frac{C}{|B_{d_t}(x,\sqrt{t})|}$$

where d_t is some natural time-dependent control 'distance'. Before that, asymptotic estimates have been obtained by Ben Arous in [**1**]. More asymptotic results are provided in [**38**].

For lower bounds, the classical result is

$$p_t(x,y) \geq C\frac{e^{-d(x,y)^2/Mt}}{|B_d(x,\sqrt{t})|}$$

in [**37**], valid only for trivial drifts (i.e. assuming the drift is uniformly contained in the span of the diffusion vector fields): It is well known that the densities of solution to weak-Hörmander SDE with non-trivial drifts can fail to have full support (cf. example 10.1.1). Finding the precise, most general set of conditions on σ required to guarantee strict positivity in a weak Hörmander setting is a difficult open problem.

Other works that examine positivity for elliptic SDE from a theoretical point of view include [**5**], [**3**], etc. In [**34**], the author defines a notion of 'elliptic random variable' further than the concept of solutions to SDE, and proves the positivity of the density for such random variables. Other relevant work includes the analysis of the strict positivity of densities of solutions to equations of the form $dX_t = \sigma dW + B(X_t)dt$ where σ is constant and B is polynomial, satisfying modified versions of the Hörmander condition in [**28, 33**].

The strict positivity of the density is also known for certain specific classes weak Hörmander SDE, for instance, in [**30**], strict positivity is established for an SDE arising from the modeling of Neurons (see [**30**]).

There has been renewed interest in proving (non-Euclidean) upper and lower bounds for densities of SDE satisfying the weak Hörmander condition, with particular emphasis on expressing the distance in the exponential in concrete and simple ways that emphasise scaling properties: *two-sided* estimates have been obtained in [**50, 6, 20**] in the following particular cases:

(1) 'Weak Hörmander of order three' (with our terminology), with an additional 'geometric condition on the variance' assumption, locally ([**50**]).
(2) A chain of SDE with only the first SDE involving Brownian motions and the other equations only having a drift depending on the previous SDE (globally in [**20**] and then locally with sharper constants in [**50**] Chapter 3, section 3.5)
(3) Strong (time-dependent) Hörmander of order 2, locally, in [**8**] and [**50**] (Chapter 4).

For time-dependent coefficients, the existence and smoothness of densities, as well as bounds of the form

$$p_t(x,y) \leq C\frac{e^{-\frac{|y-x|^2}{Mt}}}{t^{\frac{\nu}{2}}}$$

are known, cf. [**15**], [**53**], [**17**], and integrable upper bounds are known under Hörmander of order 2 from [**8, 9**].

The Löcherbach problem

In [**42**], Löcherbach considers the following model (which is a generalisation of the model where particles are independent, considered by Bally and Löcherbach in [**10**]):

Particles are born in $\mathbb{R}^n$ following a Poisson process with known intensity measure, particles die or split into two at a Poisson rate that depends on the position of all particles, and for a fixed total number l of particles, the positions ξ^i of the l particles ($i = 1, \ldots, l$) evolve according to the following interacting system of SDE (here ξ is the total configuration of all particles, $\xi = (\xi^1, \xi^2, \ldots, \xi^l)$):

$$d\xi^i_t = \sigma^i(\xi^i, \xi)dW^i_t$$

where the W^i are (independent) d-dimensional Brownian motions. Under smoothness, boundedness, subcriticality, and *uniform ellipticity* conditions, the author obtains the existence and continuity of a density for the invariant measure. The main technical step of the proof is to obtain a bound of the form

$$p_t^i(\xi, y) \leq C \frac{e^{-\frac{|y-x|^2}{Mt}}}{t^{\frac{n}{2}}},$$

where $p_t^i(\xi, y)$ is the density of the position of particle ξ^i at time $T + t$ conditionally given that there is a jump (i.e. either an immigration, death or branching event) at time T and no jump between T and $T + t$, with the complete position of all particles at time T being given by ξ. The final argument then relies on the local space-time integrability of the estimate and on the fact that the constants involved depend polynomially on the number of particles.

Our contributions: Description of Part 1, and motivation for Parts 2 and 3

It is a fact not usually explicitly stated, but following easily by inspection of classic proofs, that the constants in traditional Gaussian upper bounds on solutions to SDE satisfying the Hörmander condition only depend on the following finite number of real quantities:

- the bounds on the derivatives of the driving vector fields;
- the order at which the Hörmander condition holds;
- the Hörmander constant;
- the number of driving vector fields and corresponding Brownian motions;
- the dimension of the ambient space,

i.e., the bounds do not actually depend on the fine structure of the driving vector fields or the sub-Riemannian metric which they induce.

For systems, another relevant quantity is the dimension of the background space. In this work, one particular aim is, in the framework of systems, to reproduce the SDE results under the weaker degeneracy assumptions defined directly on the target space, and to control the dependence of the constants on the dimension of the background space and (in the elliptic case) on the number of Brownian Motions. In many cases this can be a highly non-trivial task.

We will define the following notion of projected weak Hörmander constant H_L of order L directly on the target space:

DEFINITION. The weak Hörmander H_L constant at x of order L of the system (x, σ, F) is defined as

$$H_L := \min \left(1, \inf_{\substack{v \in \mathbb{R}^n \\ |v|=1}} \sum_{\substack{|\alpha| \leq L \\ \alpha_{\#(\alpha)} \neq 0}} \langle dF.\sigma^{[\alpha]}, v \rangle^2 \right).$$

(Here $\sigma^{[\alpha]}$ denotes the iterated brackets of σ following the multi-index α and $|\alpha|$ is the order of the multi-index α, where drift components are counted twice, see thesis for precise definitions.) The ellipticity constant of a system is the weak-Hörmander constant of order 1. A system is '(L, H_L)-weak Hörmander at x' if its weak Hörmander constant at x of order L is greater than H_L.

REMARK 0.0.1. It is possible for a system to be (L, H_L) weak Hörmander, whilst the background SDE is only (L', ϵ)-weak Hörmander, with L' arbitrarily larger than L and ϵ arbitrarily smaller than H_L. This is another reason why estimates with any control on the constants behave better than estimates obtained through the estimation of the density of the background process in its intrinsic space.

We also define the following quantity G (relative to the orders (L, g)) which we call the 'tension', and corresponds to boundedness assumptions:

DEFINITION. The tension of order (L, g) of a system (x, σ, F) is defined (when F is linear) as

$$G = \sup_{x \in \mathbb{R}^m} \sup_{\substack{v \in \mathbb{R}^m \\ |v|=1}} \sum_{|\alpha| \leq L} \langle \sigma^\alpha, v \rangle^2$$

$$+ \sup_{v\in\mathbb{R}^m,|v|=1} \sup_{o\leq g} \sup_{\substack{w\in(\mathbb{R}^m)^{\otimes o} \\ \forall i,|w^i|=1}} \sum_{k=0}^{d} \langle \frac{\partial^o \sigma^k}{\Pi \partial w^i}, v\rangle^2 + |dF|^2 + 1,$$

where $|dF|^2$ denotes the squared operator norm of dF, not the Frobenius norm, and σ^α denotes iterated derivatives of σ following the multi-index α, see thesis for precise definition.

A system is said to be (L, g, G)-tense if its (L, g)-order tension is less than G.

Constants which, for fixed n, L, g, do not depend at all on m or d, and depend polynomially on H_L, G, will be called *proper constants*.

Constants which, for fixed n, L, g, d, depend polynomially on m, G, H will be called *polynomial constants*.

Constants which, for fixed n, L, g, depend polynomially on d, m, G, H will be called *strongly polynomial constants*.

REMARK 0.0.2. It is now clear to us that this definition of G is not optimal *except for deterministic purposes*. Research on better control of the constants in probabilistic results with a different definition of G [2] is ongoing.

The main aims of Part 1 of this thesis are to show the following:

- If (x, σ, F) is a uniformly $(1, g, G)$-tense, uniformly H-elliptic system, and g is large enough ($g \geq n + N$), there exist *strongly polynomial* constants C, M, D such that for any $t \leq D$, Y_t admits a density $p_t(x, y)$ satisfying, for all $N \leq g - 3 - n$, and any unit $v_1, v_2, \dots, v_N$

$$|\frac{\partial^N p_t(x,y)}{\partial v_1 \partial v_2 \dots \partial v_N}| \leq C \frac{\exp(\frac{-|F.x-y|^2}{Mt})}{t^{\frac{n+N}{2}}},$$

  cf. Theorem 4.3.2. See also extension 4.3.3 for better control on the constants.
- If $\mathcal{A}$ is a uniformly (L, g, G)-tense, uniformly (L, H_L)-weak Hörmander system, and $g \geq n + 3$, there exist both *polynomial* constants C, M, D such that for any $t \leq D$, Y_t admits a density $p_t(x, y)$ satisfying, for all $N \leq g - 3 - n$, and any unit $v_1, v_2, \dots v_N$

$$|\frac{\partial^N p_t(x,y)}{\partial v_1 \partial v_2 \dots \partial v_N}| \leq C \frac{\exp(\frac{-|F.x-y|^2}{Mt})}{t^{(n+N)2^{4L}}},$$

  cf. Theorem 4.4.2.

The above theorems represent motivation for the rest of the thesis and show that defining hypoellipticity and ellipticity directly in the target space and in terms of the projected vector fields is enough to ensure the existence of a density worthy of further study.

Similarly to the classical situation, those bounds are also the first technical step towards attempts to show the more challenging estimates in Part 2.

As mentioned in the literature review, it was also proved in [**37**] that drift-free Stratonovich SDE satisfying the Hörmander condition uniformly admit bounds of the form

$$p_t(x, y) \leq C \frac{e^{-\frac{d(x,y)^2}{Mt}}}{|B_d(x, \sqrt{t})|},$$

where d is the Carnot-Carathéodory distance. This motivates the following question:

[2] A reasonable candidate is

$$G = \sup_{v\in\mathbb{R}^m,|v|=1} \sup_{o\leq g} \sum_{\substack{w\in(\mathbb{R}^m)^{\otimes o} \\ w^i\in B}} \sum_{k=0}^{d} \langle \frac{\partial^o \sigma^k}{\Pi \partial w^i}, v\rangle^2 + |dF|^2,$$

where B is an orthonormal basis of $(\mathbb{R}^m)^{\otimes o}$

QUESTION 0.0.3. *Define a function*

$$d : \mathbb{R}^m \otimes \mathbb{R}^n \to \mathbb{R},\ (x, y) \mapsto d(x, y) = \inf_{\gamma \in \mathcal{P}_1^d} \left(|\gamma|_{L^2} : y = F(X_1), X_1 = \mathrm{Sol}(\gamma, x) \right),$$

where $\mathrm{Sol}(\gamma, x)$ *denotes the solution to the ordinary differential equation*

$$dX_t = \sum_{i=1}^{d} \sigma^i(X_t) d\gamma_t^i$$

with initial condition $X_0 = x$*, evaluated at time 1, and* $|\gamma|_{L^2}^2$ *denotes the energy of the path* γ*, and* $\mathcal{P}_1^d$ *denotes the set of smooth paths starting at* 0*.*

(a) *Can one obtain estimates of the form*

$$p_t(x, y) \leq C \frac{e^{-\frac{d(x,y)^2}{Mt}}}{|B_d(x, \sqrt{t})|},$$

for a drift-free system satisfying the (projected equivalent of the) Hörmander condition?

(b) *If so, is it possible to ensure that the constants* M, C *are proper, or at least, (strongly) polynomial?*

(c) *Can results be extended to non-trivial drifts by changing the metric?*

The above bound will be proved globally, in a weak Hörmander context, with a time-dependent metric d_t in Theorem 8.2.1, using techniques inspired from [**37, 39**].[3]

However, the problem of ensuring that the constants are (strongly) polynomial proved to be extremely challenging.

We provide an alternative proof with polynomial constants valid only under a condition which we call the 'detailed-Progressive weak Hörmander condition' (which includes systems satisfying the weak-Hörmander condition of order 2) cf. Theorem 8.3.1, which is one of the main Theorems of this thesis. The condition and its implications in terms of equivalences of control distances is of independent interest, and shows links between [**37**] and some of the more modern part of the literature (cf.[**50, 8**]) Aside from not using the Fourier approximation argument of Kusuoka and Stroock, the proof presents the advantage that it would establish the *strong* polynomiality of the constants assuming a proof of the strong polynomiality of the Euclidean bound 4.4.1.

Any systems satisfying the Hörmander condition of order 2, and all the SDE treated in [**50, 20, 8**] satisfy the 'detailed-Progressive weak Hörmander condition'.

Because the works [**50, 20, 8**] include lower bounds, at this point, it is interesting to note that the classical example (see system (10.1.1) below, going back to [**39**], also quoted in [**50**]) of a weak Hörmander SDE whose solution fails to have full support does, in fact, satisfy our 'detailed-Progressive weak Hörmander condition'. This means that the 'detailed-Progressive Hörmander' condition is not enough to guarantee strict positivity. It is therefore natural to attempt to construct stricter sets of conditions that guarantee the strict positivity of the density. The investigation of that issue is the aim of Part 3, where we provide a class of systems the densities of whose solutions can be proved to have full support.

Our contributions: Description of Parts 2,3,4 and list of main results

In Parts 2 and 3 depending on the assumptions on the model and what kind of control we aim for on the dependence of the constants, we introduce various time-dependent functions of the form $d_t(x, y)$ on $\mathbb{R}^m \otimes \mathbb{R}^n$. To simplify the exposition, we call these objects 'distances' even though they clearly cannot satisfy the mathematical definition of a distance. In the most simple situation where the drift is null and we don't aim for control on the dependence of constants, $d_t(x, y)$ is usually the smallest possible energy of a driving path γ in $\mathbb{R}^d$ such that the solution to the ODE

[3] For a single SDE, the problem was proved globally for trivial drifts in [**37**] and locally for the weak Hörmander case in [**39**]. The local part of our estimate coincides with the estimate in [**39**] for SDE, and although our metric is not the 'optimal' metric from a theoretical point of view, our estimate is globally space-time integrable.

associated to the problem, $dX_t = \sigma(X)d\gamma_t$, with $x_0 = x$, satisfies $y = F(X_1)$. In other contexts d_t is a suitable generalisation of that concept.

We prove various upper bounds both in terms of the Euclidean metric, and bounds where those 'distances' play the same role as the control distance in Kusuoka-Stroock bounds.

We introduce the *'progressive Hörmander condition'*[4] and the *'separated progressive Hörmander condition'*. Roughly speaking, the progressive Hörmander condition states that for all $i \leq L$, arbitrary iterated derivatives (of the vector fields) of order i must be uniformly in the span of brackets of order i or less. The separated progressive Hörmander condition requires, additionally, that iterated brackets involving more than one drift- and more than two non-drift indices be uniformly contained in the span of brackets of *strictly* lower order.

We will see that the progressive Hörmander condition is precisely the condition required for statements such as 'the diffusion moves at speed $t^{k/2}$ in the direction of every k'th order bracket' to make sense:

In general, the control distance can be seen locally as the push-forward of a homogeneous norm on the log-signature space by the composition of the exponential map and a linear map. In some specific situations such as the ones treated in [**50**](Ch.3) and [**20**], this can be expressed as some homogeneous norm on $\mathbb{R}^n$. The progressive Hörmander condition is the condition required to make it possible to express the control distance as the push-forward of a homogeneous norm by a linear map only, without the exponential: Heuristically, we can equally think of the Carnot-Carathéodory distance as the minimum square root of the energy of a control γ in $\mathbb{R}^d$, satisfying $dz = \partial f.d\gamma$ for a curve z going from x to y, or as the minimum square root of the integrated homogeneous norm of a driving geometric rough path $\boldsymbol{\gamma}$ in $G^l(\mathbb{R}^d)$ such that $dz = \partial f.d\boldsymbol{\gamma} = \partial f.d\exp(\log(\boldsymbol{\gamma}))$, for a solution curve z leading to y. We will define the metric $d_{t,\log,\infty}(x,y)$ as the minimum square root of the integrated squared homogeneous norm of a driving control Γ in the free lie algebra $\mathcal{L}^l(\mathbb{R}^d)$ (seen simply as a vector space) such that $dz = \partial f.d\Gamma$ (without exponentials!) for a solution curve z leading to y. On the spectrum of 'increasing complexity of metrics and decreasing specificity of assumptions', $d_{t,\log,\infty}$ can be seen as 'in between' the metrics introduced in [**50, 8, 20**] and the general control distance from [**37**]. The control metric and $d_{t,\log,\infty}$ are equivalent if the 'progressive Hörmander' condition holds. The control distance, $d_{t,\log,\infty}$ and the metric d_{A_R} from [**8, 50**] are locally equivalent whenever the assumptions of those references are satisfied.

We can now list the main results of the thesis for a random SDE system:

- Integrable Gaussian upper bounds in the elliptic case with *strongly polynomial constants*; cf. Theorem 4.3.2 in Part 1, i.e. $\forall t \leq D, x \in \mathbb{R}^m, y \in \mathbb{R}^n$
$$\left|\frac{\partial^N p_t(x,y)}{\partial v_1 \partial v_2 \ldots \partial v_N}\right| \leq C\frac{\exp\left(\frac{-|Fx-y|^2}{Mt}\right)}{t^{\frac{n+N}{2}}},$$
where D, C and M are proper.
- Non-integrable Gaussian upper bounds in the Weak Hörmander case with *polynomial constants*; cf. Theorem 4.4.2 in Part 1. i.e. $\forall t \leq D, x \in \mathbb{R}^m, y \in \mathbb{R}^n$
$$|\frac{\partial^N p_t(x,y)}{\partial v_1 \partial v_2 \ldots \partial v_N}| \leq C\frac{\exp(\frac{-|F.x-y|^2}{Mt})}{t^{(n+N)2^{4L}}},$$
where D, C and M are polynomial.
- Integrable upper bounds, in terms of the 'log-homogeneous distance', under the detailed-Progressive Hörmander condition and the assumption that the background vector fields are uniformly progressively finitely generated; cf. Theorem 8.3.1 in Part 2, i.e. $\forall t \leq D,\ x \in \mathbb{R}^m,\ y \in \mathbb{R}^n$,
$$p_t(x,y) = \mathbb{E}_x(\delta(Y_t = y)) \leq C\frac{e^{-\frac{d_{t,\log,\infty}(x,y)^2}{Mt}}}{|B_{d_{t,\log,\infty}}(x,\sqrt{t})|},$$

[4]The 'detailed-Progressive Hörmander condition' is a slightly stricter version, the difference is mostly technical.

where D, C and M are polynomial, and $d_{t,\log,\infty}$ can be locally described as the smallest homogeneous norm of a tensor $s^\bullet \in \operatorname{span}_{|\alpha|\leq L} e^{[\alpha]}$ such that $\sum_{|\alpha|\leq L} s^\alpha \sigma^\alpha = y - dF.x$. Note that $d_{t,\log,\infty}$ locally generalises the norms introduced in [**50, 20, 8**].

REMARK 0.0.4. Our estimate is global and the constants C, M in particular do not depend on the initial point x. This includes situations where the dimension of $\operatorname{span}_{i=1}^d \sigma^i$ is not constant, whilst keeping uniformity over x. Such global estimates translate to estimates with a power of t in the denominator replacing the volume of the ball. If the power of t is required to be optimal (i.e. the homogeneous dimension at that point e.g. $2n - \dim(\operatorname{span}_k \sigma^k)$ when $L = 2$), then this latter estimate will only hold with constants depending on the initial point.

REMARK 0.0.5. For a zero drift SDE satisfying the progressive Hörmander condition, $d_{t,\log,\infty}$ is locally equivalent to the Carnot-Carathéodory distance (cf. 7.1.1), but neither the full equivalence of the metrics nor the above estimate hold in the more general setting. It is, however, true for drift-free SDE that *the volumes of* balls of radius $\sqrt{t}$ with respect to the distances d and $d_{t,\log,\infty}$ are uniformly comparable (cf. Lemma 5.1.12) even if the progressive Hörmander condition does not hold.

- Assuming the detailed-progressive Hörmander condition (on the target space), but not that the background vector fields are uniformly progressively finitely generated, cf. Theorem 8.4.1 there are *polynomial* constants D, M, C such that for $t \leq D, x \in \mathbb{R}^m, y \in \mathbb{R}^n$,
$$p_t(x, y) \leq C \frac{e^{-\frac{d_t(x,y)^2}{Mt}}}{|B_{d_t}(x, \sqrt{t})|}.$$

REMARK 0.0.6. This is almost the same as the above estimate, except for the removal of the background UPFG condition. The proof is only a slight variation. Nonetheless, removing this assumption could be seen as a significant difference.

- Uniform (and global) integrable upper bounds in the general case (including non-trivial drifts) [5]; cf. Theorem 8.2.1 in Part 2. i.e. for $t \leq D, x \in \mathbb{R}^m, y \in \mathbb{R}^n$,
$$p_t(x, y) \leq C \frac{e^{-\frac{d_t(x,y)^2}{Mt}}}{|B_{d_t}(x, \sqrt{t})|},$$
where D, C, M do not depend on the background space dimension, but depend uncontrollably on d. Only the most local part of the estimate, was proved in the SDE case before in [**39**] [6]
- Control-type lower bounds for systems with zero drift, without control on the constants; cf. Theorems 10.3.1 and 10.3.2 in Part 3.
$$p_t(x, y) \geq \frac{C}{|B_d(x, \sqrt{t})|},$$
for $t, d(x, y)^2 \leq D$ for some constants D, C.
- Under the Hörmander condition with zero drift, *diagonal* upper bounds with the 'log-homogeneous distance', with *polynomial constants*; cf. Theorem 8.6.1 in Part 2. This estimate is meant to show potential for better results rather than be particularly interesting in itself: like the proof of Theorem 8.3.1, it has the advantage that the constants would be strongly polynomial assuming the strongly polynomial Euclidean result.

[5] In fact, the constants can also be shown to be (weakly) polynomial

[6] Note the distance used in [**39**] is not the same as the one we use, though this does not affect the most local part of the estimate. As explained in the thesis, the distance used in [**39**], which we write $\tilde{d}_t$ really is the optimal one from a theoretical point of view, but (even in the SDE case) it is difficult to get a completely global estimate using this distance. Our global estimate is still space-time integrable. Even in the weak separated progressive Hörmander case, it is difficult to globalise the estimate, even though we have a lower bound.

- Under the condition which we call ‘separated progressive Hörmander’, we prove local lower bounds[7] and global strict positivity (if the condition holds uniformly), cf. Theorems 10.3.1 and 10.3.2 in Part 3. Local positivity if the condition holds locally best summarizes the progress represented by these estimates.

The ‘separated progressive Hörmander’ condition is satisfied by a set of non-trivial examples that strictly includes the examples treated in [**50, 20, 8**].

Warning *For a single SDE satisfying the conditions of the main model in Chapter 3 of* [**50**] *or those of* [**20**] *(also treated in* [**50**]*), the estimates we obtain (including the upper bounds) are not as good as the ones in those references for $t \lesssim d^2 \lesssim t\ln(t)$, because our ‘distance’ is slightly different, though balls of radius $\sqrt{t}$ are comparable (not just in volume), and (therefore) the estimates coincide for $d^2 \leq t$.*

The lower bound estimate implies the strict positivity of the density, which (even for an SDE in $\mathbb{R}^n$) under weak Hörmander does not follow from classical results and, as mentioned above, (other) diffusions under weak Hörmander can fail to have full support. Our estimate, like the corresponding one in [**50**] and [**8**], but unlike [**20**], is only local. Understanding better the tail behaviour of solutions to weak Hörmander SDE’s would be an interesting topic to study further.

Part 4 deals with interacting branching diffusions as considered in [**42**]: Using similar techniques to the ones used for the proofs of our estimates, we show a generalisation of the theorem of Löcherbach first to a time- and (partly) path-dependent setting in the elliptic case with slightly weaker assumptions than in the original paper. Then, conditionally given a proof of a strongly polynomial version of the Euclidean Result 4.4.1, we give an extension to a weak Hörmander situation, provided the ‘No degeneracy from interaction’ (NDI) condition, which we introduce. The condition is equivalent to keeping the driving vector fields independent of each other, but allowing the Brownian motions relative to different particles to be correlated with correlations that solve an interactive SDE depending on the state of the whole system. Control-type lower bounds applied to a non-NDI and drift-free Löcherbach system demonstrate that it would be very difficult to prove a generalisation of the Löcherbach theorem to non-NDI situations using the same or a similar method of proof.

Summary of key technical difficulties and solutions

The main difficulty in proving Kusuoka-Stroock type bounds is that it seems impossible to employ a truly direct integration by parts formula approach: the best applying standard Malliavin calculus techniques to that situation can do is obtain Euclidean bounds, which are not space time integrable in general. To obtain a sharp enough estimate, we must use more information about the likely shape of a sample path. This is a very fundamental difference with the far less ambitious goal of finding Euclidean bounds, and in particular, with the elliptic case.

The intuitive idea is to attempt to bound the ‘density in path space’ of the realisation of a d-dimensional Brownian path, and use a disintegration formula on the solution map associated to the SDE considered.

More precisely, the solution provided by Kusuoka and Stroock was to look at an *auxiliary object*, which contains more information about the sample path than the solution of the SDE (evaluated at time t) does, but is still finite-dimensional. This basic auxiliary object used in that paper is now much better understood, and used extensively in Rough Path Theory, where it is known as the ‘truncated signature’. This object is a tool to study a given SDE (or RDE), and it is itself the solution of an auxiliary SDE (or RDE)[8].

One of the reasons why obtaining polynomial control of the constants in the estimates in the answer to question 0.0.3 is that bounds on the density of the truncated log signature grow uncontrollably with the number of independent Brownian motions involved.

---

[7] In general, and similarly to [**50, 20**], ‘local’ means local around (the image of) the solution to the deterministic system driven by the constant driving path 0, rather than local around the image of the initial point $F(x)$.

[8] This auxiliary SDE is associated to the original problem only through how many Brownian motions are present: dependence on the initial point only appears later, in the finite dimensional analogue of the solution map.

One of the biggest technical difficulty challenge was to define auxiliary objects better suited to the problem of controlling the dimensional dependence of the constants. We call these auxiliary objects the *log-compensated signature* (used for the diagonal estimate), the *compensated signature* (used for proving proper upper bounds under the progressive Hörmander condition), denoted R, the *strictly compensated signature* $\bar{R}$ (used for the proofs of lower bounds under the separated progressive Hörmander condition) and (for Löcherbach systems under NDI) the *interactive signature*. Contrary to the truncated signature, the objects are not solutions of SDE's. Furthermore, both the compensated signature and the interactive signature also fail to satisfy as good scaling properties as the truncated signature does (and there is no group structure either, contrary to the classical situation). However, all objects are solutions of auxiliary Random SDE systems[9]. The auxiliary objects which don't satisfy scaling in the usual way do satisfy a weaker form of scaling that is enough to prove the theorems listed above.

The main technical difficulty in the proof of our lower bounds under the 'separated progressive Hörmander condition' was the identification of points contained in the interior of the support of the log-signature of the path in $\mathbb{R}^{d+1}$ composed of d Brownian motions and a time-linear path, coupled with some arguments developed in the rest of the thesis for general 'Progressive Hörmander' systems. To show that the points in question are in the interior of the support, first, explicit piecewise linear controls were built by solving systems of equations obtained via the Baker-Campbell-Hausdorff formula, then some perturbation of the construction was applied to ensure strict positivity.

[9] The association to the original problem takes more of the structure of the original problem into account. The auxiliary SDE system depends on the initial point, but the final estimate does not.

Part 1

Proof of Euclidean bounds

CHAPTER 1

Definitions and notations

DEFINITION 1.0.1. Let $X \in \mathbb{R}^m$ be driven by the Stratonovich SDE

$$dX_t = \sum_{i=1}^{d} \sigma^i(X_t) \circ dW_t^i + \sigma^0(X_t)dt,$$

with initial condition $X_0 = x$, where W_t is a d-dimensional Brownian motion; here $\sigma\colon \mathbb{R}^m \to \mathrm{Mat}(m, d+1)$ is a function with smoothness conditions which will be introduced later. Let $F\colon \mathbb{R}^m \to \mathbb{R}^n$, $X \mapsto F(X)$ be a function that sends X into the smaller dimensional space $\mathbb{R}^n$. Write $Y_t = F(X_t)$. We call the ordered list $\mathcal{A} = (x, \sigma, F)$ a *random SDE system* or *system*. The space $\mathbb{R}^m$ is called the *background space*, and $\mathbb{R}^n$ the *target space*.

REMARK 1.0.2. There are no non-degeneracy assumptions in the background space, so X_t need not have a density in $\mathbb{R}^m$. We thinking of $m, d \gg n$.

DEFINITION 1.0.3. Let $\mathcal{A} = (x, \sigma, F)$ be a system. Let v be a vector field in A. We denote by $*v$ the image of the vector field v by the differential of F:

$$*v\colon \mathbb{R}^m \to \mathbb{R}^n,\ x \mapsto dF_x.v.$$

We will also similarly write $*x$ for $F(x)$.

DEFINITION 1.0.4. If $v = (v^i)_{i=1,2,\ldots,d}$ is a collection of d vector fields in $\mathbb{R}^m$, for any multi-index α, we define higher order derivatives (v^α) and brackets ($v^{[\alpha]}$) by the induction relations

$$v^{(i)} = v^i,$$

$$v^{i,\alpha} = v^i(v^\alpha) = \sum_{k=1}^{m} \frac{\partial v^\alpha}{\partial x_k}(v^i)_k, \quad \text{and}$$

$$v^{[i,\alpha]} = [v^i, v^\alpha] = v^i(\alpha) - v^\alpha(v^i) = \sum_{k=1}^{m} \frac{\partial v^\alpha}{\partial x_k}(v^i)_k - \sum_{k=1}^{m} \frac{\partial v^i}{\partial x_k}(v^\alpha)_k.$$

The partial orders $o_0(\alpha)$ and $o_1(\alpha)$ of a multi-index $\alpha \in \mathrm{Multi}(\{0, 1, , \ldots, d\})$ are defined as

$$o_0(\alpha) = \#(\{k : \alpha_k = 0\})$$
$$o_1(\alpha) = \#(\{k : \alpha_k \neq 0\}).$$

The order $|\alpha|$ is defined as $2o_0(\alpha) + o_1(\alpha)$.

In fact, it will sometimes be useful to view $v^\bullet$ (evaluated at a given point of $\mathbb{R}^m$) as a linear map from the tensor space $T^l(\mathbb{R}^d)$ to $\mathbb{R}^d$ which makes the two operations of taking brackets correspond to each other (brackets on the tensor space are defined via the formula $[e_1, e_2] = e_1 \otimes e_2 - e_2 \otimes e_1$ and linear extension). From this perspective, the above definition of v^α corresponds to using the shorthand v^α for $v^{\bigotimes_{k=1}^{\#\alpha} e_{\alpha_k}}$. It will be shown that the iterative definition of $\sigma^{[\alpha]}$ coincides with the definition of $\sigma^{e^{[\alpha]}}$ where $e^{[\alpha]}$ is defined linearly on the tensor space.

DEFINITION 1.0.5. We will use the following notation: for any d dimensional bounded variation path $(\gamma_t^1, \ldots, \gamma_t^d)$, we define $\gamma_t^{\alpha,i}$ by the induction $\int_0^t \gamma_s^\alpha d\gamma_s^i$. In the case where γ is rough, we must assume we are already given a geometric rough path lift of γ. In particular, when we use a similar notation for Brownian motion, the integrals are to be interpreted in the Stratonovich sense.

We have the same remark as above: we can also view $\gamma^\bullet$ or $W^\bullet$ as a linear map from the tensor space to $\mathbb{R}^d$.

DEFINITION 1.0.6. If v is a vector field and $n \in \mathbb{N}$, we say that the C^n *constant of* v is

$$\|v\|_{\partial,n} = \sup_{m \le n} (\sup (|(v,w)^\alpha| \colon \#(\alpha) = m,\ w = (w_2, w_3, \ldots, w_m),\ |w_i| = 1 \text{ for all } i))$$

where w runs over all collections of vectors and α over all corresponding multi-indices.

DEFINITION 1.0.7. The tension of order (L, g) of a system (x, σ, F) is defined as

$$G = \sup_{x \in \mathbb{R}^m} \sup_{\substack{v \in \mathbb{R}^m \\ |v|=1}} \sum_{|\alpha| \le L} \langle \sigma^\alpha, v \rangle^2$$
$$+ \sup_{v \in \mathbb{R}^m, |v|=1} \sup_{o \le g} \sup_{\substack{w \in (\mathbb{R}^m)^{\otimes o} \\ \forall i, |w^i|=1}} \sum_{k=0}^{d} \langle \frac{\partial^o \sigma^k}{\Pi \partial w^i}, v \rangle^2 + \sum_{q \le g} |d^q F|^2,$$

where $|d^q F|$ denotes the operator norm of the q-th derivative of F as a function of q arguments. A system is said to be (L, g, G)-tense if its (L, g)-order tension is less than G.

REMARK 1.0.8. Because we aim to completely remove background dependence, it is very important to distinguish between the operator norm and the Frobenius norm for matrices (one must always work with the former).

DEFINITION 1.0.9. The weak Hörmander constant H_L at x of order L of the system (x, σ, F) is defined as

$$\min(1, \inf_{\substack{v \in \mathbb{R}^n \\ |v|=1}} \sum_{\substack{|\alpha| \le L \\ \alpha_{\#(\alpha)} \ne 0}} \langle dF.\sigma^{[\alpha]}, v \rangle^2).$$

The ellipticity constant of a system is the weak-Hörmander constant of order 1. A system is '(L, H_L)-weak Hörmander at x' if its weak Hörmander constant at x of order L is greater than H_L.

DEFINITION 1.0.10. We denote the set of smooth paths $\gamma \colon [0,1] \to \mathbb{R}^d$ by $\mathcal{P}^d$. As usual, the length of a path is defined by

$$|\gamma| = \int_0^1 \sqrt{\sum_{i=1}^{d} |\frac{d\gamma^i}{dt}|^2}\, dt.$$

DEFINITION. Let $\mathcal{P}_T^d$ be the set of smooth d-dimensional paths parametrised over $[0, T]$. For a path $\gamma \in \mathcal{P}_T^d$, we define the path $*\gamma \in \mathcal{P}_T^{d+1}$ by

$$\forall s \in [0, T], \qquad *\gamma_s = (s, \gamma_s^1, \gamma_s^2, \ldots, \gamma_s^d).$$

DEFINITION. If we are given a system (x, σ, F), with some fixed quantities such as H_L and G known, we call a quantity $C(n, m, d, L, g, G, H_L)$ a proper constant if it only depends on n, G, H_L, L, and for fixed n, g, L, is a polynomial function of H_L, G, i.e. there exist absolute constants $K_1, K_2 > 0$, $N \in \mathbb{Z}$ and $q \in \mathbb{R}^{[-N,\ldots,N] \times [-N,\ldots,N]}$ such that

$$C(n, m, d, L, g, G, H_L) < K_1 \left(\sum_{i,j \in [-N,\ldots,N]} q_{i,j} G^i H_L^j \right)$$

and

$$C(n, m, d, L, g, G, H_L)^{-1} < K_2 \left(\sum_{i,j \in [-N,\ldots,N]} q_{i,j} G^i H_L^j \right).$$

We call $C(n, m, d, L, g, G, H_L)$ a polynomial (resp. strongly polynomial) constant if it depends polynomially on G, H_L, m for fixed L, n, g, d (resp. polynomially on G, H_L, m, d for fixed

L, n, g). These concepts readily extend to situations where G or H_L are replaced by quantities with a similar role (such as the 'progressive Hörmander constant' or the 'mixed tension').

We also denote by $X_{0\to t}$ the Jacobian process in $\mathbb{R}^m \times \mathbb{R}^m$, which satisfies the following, where X_t is the solution to the original SDE 1.0.1, where $u, v \in \{1, 2, \ldots, m\}$:

$$X_{0\to 0} = \mathrm{Id}_{m\times m}$$

$$d(X_{0\to t})_{v,u} = \left(\sum_{i=1}^{d} \sum_{j=1}^{m} \frac{\partial \sigma^i(X_t)}{\partial x_j} (X_{0\to t})_{j,u} \circ dW_t^i + \sum_{j=1}^{m} \frac{\partial \sigma^0(X_t)}{\partial x_j} (X_{0\to t})_{j,u} dt \right)_v .$$

As usual, we also define $X_{t\to 0} = (X_{0\to t})^{-1}$ and $X_{s\to t} = X_{0\to t} X_{s\to 0}$.

CHAPTER 2

Technical Lemmas

Here we write the most fundamental lemmas that will need to be used during the proofs. This chapter could be considered an appendix.

2.1. On the exponential decay of iterated integrals (concentration results)

For the proof of Euclidean bounds, only the classic first result (the martingale inequality) is required. The other results will be used in the proof of the control bounds. The beginning of this section contains a version of the result of Appendix 1 (section 5, page 24) of [**6**] and inequality 4.5 page 422 in [**37**] (the former is a variation of the latter).

REMARK 2.1.1. In this section integrals are Itô integrals. We will use the notation $I^\alpha(A,T)$ and $J^\alpha(A,T)$ for Itô iterated integrals, and different notations, such as $(AW)^\alpha$ for Stratonovich ones.

The following can be found in the appendix of [**49**]

LEMMA 2.1.2 (Martingale inequality). *Let M_t be a martingale, we have*

$$\mathbb{P}(\sup_{s\le t}(|M_s|\ge\delta))\le 2e^{-\frac{\delta^2}{2\rho}}.$$

Now, let $\alpha\in\{0,1,\dots,d\}^k$ be a multi-index. We write as usual

$$|\alpha| = k+\#\{i\in\{1,2,..,k\},\ \alpha_i=0\}$$

for the order of α. Write $\bar\alpha=(\alpha_1,\alpha_2,\dots,\alpha_{k-1})$ We define the multiple stochastic integral of order $|\alpha|$, $I^\alpha(A,T)$ by induction:

$$\begin{aligned}
I^{(0)}(A,T) &= \int_0^T A_t dt,\\
I^{(k)}(A,T) &= \int_0^T A_t dW_t^{\alpha_k},\\
I^{\alpha}(A,T) &= \int_0^T I^{\bar\alpha}(A,t)dW_t^k, \qquad \text{if } \alpha_k\neq 0,\\
I^{\alpha}(A,T) &= \int_0^T I^{\bar\alpha}(A,t)dt, \qquad \text{if } \alpha_k= 0.
\end{aligned}$$

Recall the following theorem from [**6**].

THEOREM 2.1.1 (Exponential decay of iterated integrals). *There exist some universal constants $C_{|\alpha|}$ such that*

$$\mathbb{P}\left(\sup_{0\le t\le T} I^\alpha(A,t)>R\right)\le C_{|\alpha|}\exp\left(-\frac{(\frac{R}{a})^{\frac{2}{|\alpha|}}}{2T}\right).$$

With the exact values of the constants being

$$C_\alpha = 2^{|\alpha|}.$$

PROOF. We prove the result by induction.

1. *Initial case.* If $\alpha_1 \neq 0$, this is just the lemma above, with $C_\alpha = 2$. If $\alpha_1 = 0$, we have for all $\forall T > 0$,

$$\left|\int_0^T A_t dt\right| \leq \sup_{0\leq t\leq T} |A_t| T = aT,$$

therefore

$$\mathbb{P}\left(\sup_{0\leq t\leq T} |\int_0^t A_s ds| > R\right) = 0 \quad \text{if} \quad \frac{R}{aT} > 1.$$

This implies that we indeed have

$$\mathbb{P}\left(\sup_{0\leq t\leq T} I^\alpha(A,t) > R\right) \leq C_{|\alpha|} \exp\left(-\frac{(\frac{R}{a})^{\frac{2}{|\alpha|}}}{2T}\right)$$

for $c_\alpha = c_{(1)} = e^{1/2}$.

2. *Induction step from $\bar{\alpha}$ to α.* We suppose $\alpha = (\bar{\alpha}, \alpha_k)$ for some $\bar{\alpha}$.

If $\alpha_k \neq 0$, let us split the event into two as follows, with B a real number (corresponding to the one being called A in [**37**]). We have:

$$\begin{aligned}
\mathbb{P}\left(\sup_{0\leq t\leq T} |I^\alpha(A,t)| > R\right) &\leq \mathbb{P}\left(\sup_{0\leq t\leq T} |I^\alpha(A,t)| > R, \sup_{0\leq t\leq T} |I^{\bar{\alpha}}(A,t)| < B\right) \\
&\quad + \mathbb{P}\left(\sup_{0\leq t\leq T} |I^{\bar{\alpha}}(A,t)| > B\right) \\
&\leq 2\exp\left(-\frac{R^2}{2B^2 T}\right) + C_{\bar{\alpha}} \exp\left(-\frac{(\frac{B}{a})^{\frac{2}{|\bar{\alpha}|}}}{2T}\right)
\end{aligned}$$

(here we have used the martingale inequality and the induction hypothesis).

Now set $B = a^{\frac{1}{|\bar{\alpha}|+1}} R^{\frac{|\bar{\alpha}|}{|\bar{\alpha}|+1}}$, and we get the required inequality with $C_\alpha = C_{\bar{\alpha}} + 2$.

If $\alpha_k = 0$, proceeding as in the initial case above (for $\alpha_1 = 0$), we find that

$$\begin{aligned}
&\mathbb{P}\left(\sup_{0\leq t\leq T} |I^\alpha(A,t)| > R\right) \\
&\leq \mathbb{P}\left(\sup_{0\leq t\leq T} |I^{\bar{\alpha}}(A,t)| > \frac{R}{T}\right) \\
&\leq C_{\bar{\alpha}} \exp\left(-\frac{(\frac{R}{aT})^{\frac{2}{|\bar{\alpha}|}}}{2T}\right) \\
&\leq C_{\bar{\alpha}} \exp\left(-\frac{(\frac{R}{a})^{\frac{2}{|\alpha|}}}{2T} (\frac{R}{a})^{\frac{4}{|\alpha|(|\alpha|-2)}} (\frac{1}{T})^{\frac{2}{|\alpha|-2}}\right) \\
&\leq C_{\bar{\alpha}} \exp\left(-\frac{(\frac{R}{a})^{\frac{2}{|\alpha|}}}{2T} (\frac{(\frac{R}{a})^{\frac{2}{|\alpha|}}}{T})^{\frac{2}{|\alpha|-2}}\right) \\
&\leq C_{\bar{\alpha}} \exp\left(-\frac{(\frac{R}{a})^{\frac{2}{|\alpha|}}}{2T}\right) \mathrm{Id}_{\left\{\frac{(\frac{R}{a})^{\frac{2}{|\alpha|}}}{T} > 1\right\}} + e^{1/2} C_{\bar{\alpha}} \exp\left(-\frac{(\frac{R}{a})^{\frac{2}{|\alpha|}}}{2T}\right) \mathrm{Id}_{\left\{\frac{(\frac{R}{a})^{\frac{2}{|\alpha|}}}{T} \leq 1\right\}}
\end{aligned}$$

$$\leq C_{\bar{\alpha}}(1+e^{1/2})\exp\left(-\frac{(\frac{R}{a})^{\frac{2}{|\alpha|}}}{2T}\right)$$

where at the last two lines, we have split into two cases according to whether $\frac{(\frac{R}{a})^{\frac{2}{|\alpha|}}}{2T}$ is greater or smaller than 1. Therefore, the inequality indeed holds with $C_\alpha = C_{\bar{\alpha}}(1+e^{1/2})$. This completes the proof. □

2.2. Disintegration formula

The following result from real analysis is the key ingredient for the developments in Part 2.

THEOREM 2.2.1 (Disintegration formula from real analysis, see [**37**], Lemma 3.14). *Let $F : \mathbb{R}^m \to \mathbb{R}^n$ be a smooth function. Let μ be a measure on $\mathbb{R}^m$ with density μ. The pushforward of the measure is given by the formula*

$$\mu_{\mathbb{R}^n}(y) = \int_{F(x)=y} \mu(x)J(x)$$

where

$$J(x) = \sqrt{\frac{1}{\det(JFJF^t)}}$$

with JF being the Jacobian matrix ($JF_{i,j} = \frac{\partial f_i}{x_j}$) and t denoting transposition.

2.3. Gronwall-type Lemma

The following Gronwall-type theorem can be seen as a generalisation of Proposition A.1 in [**42**], merged with the technical point of Corollary A1 in [**42**] and adapted to our general setting, with sharper control on the growth of the constants:

THEOREM 2.3.1. *Let $S_t \in \mathbb{R}^m$ be a process driven by the following SDE:*

$$dS_t = B^0 dt + \sum_{k=1}^{d} B^k dW_t^k$$

For some processes $B_t^k \in \mathbb{R}^m \times \mathbb{R}^m$ for $k = 0, 1, .., d$, such that, for all k,

$$B_t^k = E_t^k S_t + F_t^k$$

with E_t^k, F_t^k processes in $\mathbb{R}^m \times \mathbb{R}^m$. Let $p \geq 1$ be fixed. Assume we have the following conditions, for some given constants C_0, C_1, C_2, and β (below, $\|\cdot\|$ denotes the operator norm):

(1) $\|\|\|S_0\|^2\|_p^p \leq C_0(p)$;
(2) *if $E_t : \mathbb{R}^m \to \mathbb{R}^m$, $y \mapsto E_t(y) = \sum_{k=1}^d \|E_t^k y\|^2 + |\langle y, E_t^0 y\rangle|$, the operator norm $\|E_t\|$ satisfies $\|E_t\|^2 \leq C_1$ almost surely and for all $t \leq 1$;*
(3) *if $F_t = \sum_{k=1}^d \|F^k\|^2 + |F^0|^2$, then for all $t \leq 1$, $\mathbb{E}(|F_t|^p)^{\frac{1}{p}} \leq C_2 e^{\beta t}$.*

Then, for all $x \in \mathbb{R}^m$ with $|x|_{\mathbb{R}^m} = 1$, there exist constants K_1 and K_2, depending only on C_0, C_1, C_2, p but not β, with K_2 not dependent on C_0, such that for all $t \leq 1$, we have:

$$\|(\sup_{s\leq t} |S_s x|^2_{\mathbb{R}^m})\|_p \leq K_1 e^{\max(K_2,\beta)t},$$

where $\|\|_p$ denotes the (expected) L^p norm.

If, additionally, we have that

$$S_0 x = 0, \quad \|E^0\|^2 \leq C_1,$$

then there exists a constant M depending only on C_0, C_1, C_2, p but not β or m, such that for all $p \geq 1$,

$$\left\| \left(\sup_{s \leq t} |S_s x|^2_{\mathbb{R}^m} \right) \right\|_p \leq t^{\frac{1}{2}} M e^{\max(K_2, \beta) t}.$$

Explicitly,

$$\begin{aligned}
K_1 &= \bar{K}_1^{\frac{1}{p}} = (C_0(p) + 2p(2p-1))^{\frac{1}{p}}, \\
K_2 &= \frac{\bar{K}_2}{p} = 2(C_1 + 1) + 8(p-1)(C_1^2 + 1) + C_2 2(2p-1), \quad \textit{and} \\
M &= 4^{1-\frac{1}{p}} (1 + \Lambda_p)^{\frac{1}{p}} (K_1 C_1^p + C_2^p)^{\frac{1}{p}},
\end{aligned}$$

where Λ_p is the constant in the Burkholder inequality.

PROOF. First , by Itô's formula, we have (with the usual convention that $W^0 = t$):

$$\begin{aligned}
|S_t x|^2 &= |S_0 x|^2 + 2 \sum_{k=0}^{d} \int_0^t \langle S_s x, B^k \rangle dW_s^k + \sum_{k=1}^{d} \int_0^t |B^k S_s x|^2 ds \\
&= |S_0 x|^2 + 2 \sum_{k=1}^{d} \int_0^t J_s^k dW_s^k + \int_0^t H_s ds,
\end{aligned}$$

for processes $J_s^k, H_s \in \mathbb{R}$ with:

$$\begin{aligned}
J_s^k &= 2 \langle S_s x, B^k \rangle \\
&= 2 \langle S_s x, (E^k S x + F^k) \rangle
\end{aligned}$$

and

$$\begin{aligned}
H_s &= 2 \langle S_s x, B^0 \rangle + \sum_{k=1}^{d} |B^k|^2 \\
&= 2 \langle S_s x, E^0 S_s x + F^0 \rangle + \sum_{k=1}^{d} |E^k S_s x|^2 \\
&\quad + \sum_{k=1}^{d} |F^k|^2 + \sum_{k=1}^{d} \langle F^k, E^k S_s x \rangle \\
&\leq 2 \langle S_s x, E^0 S_s x + F^0 \rangle + \sum_{k=1}^{d} |E^k S_s x|^2 \\
&\quad + \sum_{k=1}^{d} |F^k|^2 + \sum_{k=1}^{d} |F^k|^2 + \sum_{k=1}^{d} |E^k S_s x|^2 \\
&\leq 2 C_1 |S_s x|^2 + 2|F| + 2|S_s x|^2. \qquad (2.3.1)
\end{aligned}$$

Next, we note that

$$\begin{aligned}
\sum_{k=1}^{d} (J^k)^2 &= \sum_{k=1}^{d} 4 \langle S_s x, (E^k S x + F^k) \rangle^2 \\
&\leq 4 \sum_{k=1}^{d} (|S_s x| |E^k S x| + |S_s x| |F^k|)^2
\end{aligned}$$

$$\le 4|S_s x|^2 \sum_{k=1}^{d} (|E^k Sx| + |F^k|)^2$$

$$\le 8(\sum_{k=1}^{d} |E^k Sx|^2)|S_s x|^2 + 8(\sum_{k=1}^{d} |F^k|^2)|S_s x|^2$$

$$\le 16|S_s x|^4 + 16|E(S_s x)|^2 + 8|F||S_s x|^2. \tag{2.3.2}$$

We use the notation $\Lambda_t = |S_s x|^2$. Now, let us apply Itô's formula to Λ_t and the function $f(t,x) := e^{\gamma t} x^p$, where $\gamma < 0$ will be chosen later. We will also use the notation $\Phi_t = \mathbb{E}(f(t, \Lambda_t))$. We get:

$$e^{\gamma t}\Lambda_t^p = \Lambda_0^p + \gamma \int_0^t e^{\gamma s}\Lambda_s^p ds + p\int_0^t e^{\gamma s}\Lambda_s^{p-1} H_s ds$$
$$+ \frac{1}{2}p(p-1)\int_0^t e^{\gamma s}\Lambda_s^{p-2}(\sum_{k=1}^{d}(J_s^k)^2)ds + M_t,$$

where M_t is a martingale. It follows that

$$\Phi_t \le \Phi_0 + \gamma\int_0^t \Phi_s ds + p\int_0^t \mathbb{E}(e^{\gamma s}\Lambda_s^{p-1}H_s)ds$$
$$+ \frac{1}{2}p(p-1)\int_0^t \mathbb{E}(e^{\gamma s}\Lambda_s^{p-2}(\sum_{k=1}^{d}(J_s^k)^2))ds.$$

Now, using Eqs. (2.3.1) and (2.3.2), we can further reduce this to:

$$\Phi_t \le \Phi_0 + \gamma\int_0^t \Phi_s ds + p\int_0^t \mathbb{E}\left(e^{\gamma s}\Lambda_s^{p-1}(2C_1|S_s x|^2 + 2|F| + 2|S_s x|^2)\right)ds$$
$$+ \frac{1}{2}p(p-1)\int_0^t \mathbb{E}\left(e^{\gamma s}\Lambda_s^{p-2}(16|S_s x|^4 + 16|E(S_s x)|^2 + 8|F||S_s x|^2)\right)ds$$
$$\le \Phi_0 + (\gamma + 2pC_1 + 2p + 8p(p-1) + 8p(p-1)C_1^2)\int_0^t \Phi_s ds$$
$$+ (2p(2p-1))\int_0^t \mathbb{E}\left(e^{\gamma s}\Lambda_s^{p-1}|F|\right)ds$$
$$\le \Phi_0 + (\gamma + 2p(C_1+1) + 8p(p-1)(C_1^2+1))\int_0^t \Phi_s ds$$
$$+ (2p(2p-1))\int_0^t \mathbb{E}\left(e^{\gamma s}\Lambda_s^{p-1}|F|\right)ds. \tag{2.3.3}$$

Now, by Hölder's inequality,

$$e^{\gamma s}\mathbb{E}(\Lambda_s^{p-1}|F|) \le \Phi_s^{1-\frac{1}{p}} e^{\gamma s \frac{1}{p}} \||F_s|\|_p$$
$$\le \Phi_s^{1-\frac{1}{p}} e^{\gamma s\frac{1}{p}} C_2 e^{\beta s}.$$

Setting $\gamma < -p\beta$, and using $\forall u \in \mathbb{R}, u^{\frac{p-1}{p}} \le 1 + u$, we can write

$$e^{\gamma s}\mathbb{E}(\Lambda_s^{p-1}|F|) \le \Phi_s^{1-\frac{1}{p}} e^{\gamma s\frac{1}{p}} C_2 e^{\beta s}$$
$$\le (1+\Phi_s)e^{\gamma s\frac{1}{p}} C_2 e^{\beta s}$$
$$\le (1+\Phi_s)C_2.$$

Now, writing $\bar{K}_2$ for $2p(C_1+1)+8p(p-1)(C_1^2+1)+C_2 2p(2p-1)$ and $\bar{K}_1$ for $C_0(p)+2p(2p-1)$, we can rewrite (2.3.3) as:

$$\begin{aligned}\Phi_t &\leq \Phi_0 + (\gamma + 2p(C_1+1)+8p(p-1)(C_1^2+1)+C_2 2p(2p-1))\int_0^t \Phi_s ds \\ &+ (2p(2p-1))\int_0^t 1 ds \\ &\leq C_0(p) + t(2p(2p-1)) + (\gamma+\bar{K}_2)\int_0^t \Phi_s ds \\ &\leq \bar{K}_1 + (\gamma+\bar{K}_2)\int_0^t \Phi_s ds\end{aligned}$$

where at the third line, we have used the first assumption, and at the fourth line, we have used the fact that $t \leq 1$.

We would like to use Gronwall's identity. We can do that directly if $\gamma + \bar{K}_2 \geq 0$, otherwise, we need to use the fact that $\Phi_s \geq 0$ to transform the above into $\Phi_s \leq \bar{K}_1$. In all cases, we can write

$$\Phi_s \leq \bar{K}_1 + (\max(\gamma+\bar{K}_2, 0))\int_0^t \Phi_s ds,$$

and then use Gronwall's identity to obtain, for all $t \leq 1$:

$$\Phi_t \leq \bar{K}_1 e^{\max(\gamma+\bar{K}_2,0)t}.$$

Recall that by definition $\Phi_t = \mathbb{E}(e^{\gamma t}\Lambda_t^p)) = \mathbb{E}(e^{\gamma t}(|Sx|^2)^p))$, therefore,

$$\begin{aligned}\||Sx|^2\| &\leq \bar{K}_1^{\frac{1}{p}}(e^{\max(\gamma+\bar{K}_2,0)t-\gamma t})^{\frac{1}{p}} \\ &= \bar{K}_1^{\frac{1}{p}} e^{\max(\frac{\bar{K}_2}{p},\beta)t} \\ &= K_1 e^{\max(K_2,\beta)t},\end{aligned}$$

with

$$K_1 = \bar{K}_1^{\frac{1}{p}} = (C_0(p)+2p(2p-1))^{\frac{1}{p}}$$

and

$$K_2 = \frac{\bar{K}_2}{p} = 2(C_1+1)+8(p-1)(C_1^2+1)+C_2 2(2p-1).$$

This concludes the proof of the first statement.

Now, for the second part of the theorem, observe first that for $1 \leq p \leq 2$, $\|Sx\|_p \leq \|Sx\|_{2p}$, so we can suppose without loss of generality that $p \geq 2$. Now, using the Burkholder and Jensen's inequalities, we can write (with Λ_p being the constant appearing in Burkholder's inequality):

$$\begin{aligned}\mathbb{E}(|Sx|^p) &\leq 4^{p-1}\mathbb{E}\left(\left(\int_0^t (E^0 Sx)_s ds\right)^p\right) + 4^{p-1}\mathbb{E}\left(\left(\int_0^t (F^0)_s ds\right)^p\right) \\ &+ 4^{p-1}\mathbb{E}\left(\left(\int_0^t \sum_{k=1}^d (E^k Sx)_s dW_s\right)^p\right) + 4^{p-1}\mathbb{E}\left(\left(\int_0^t \sum_{k=1}^d (F^k)_s dW_s\right)^p\right) \\ &\text{(by Jensen's inequality)} \\ &\leq 4^{p-1}\mathbb{E}\left(\left(\int_0^t (E^0 Sx)_s ds\right)^p\right) + 4^{p-1}\mathbb{E}\left(\left(\int_0^t (F^0)_s ds\right)^p\right) \\ &+ 4^{p-1}\Lambda_p\mathbb{E}\left(\left(\int_0^t \sum_{k=1}^d (E^k Sx)_s^2 ds\right)^p\right) + 4^{p-1}\Lambda_p\mathbb{E}\left(\left(\int_0^t \sum_{k=1}^d (F^k)_s^2 ds\right)^p\right) \\ &\text{(by Burkholder's inequality)}\end{aligned}$$

$$
\begin{aligned}
&\leq 4^{p-1}\mathbb{E}\left(t^{p-1}\int_0^t (E^0Sx)_s^p ds\right) + 4^{p-1}\mathbb{E}\left(t^{p-1}\int_0^t (F^0)_s^p ds\right)\\
&\quad + 4^{p-1}\Lambda_p\mathbb{E}\left(t^{p-1}\int_0^t \left(\sum_{k=1}^d (E^kSx)_s^2\right)^p ds\right) + 4^{p-1}\Lambda_p\mathbb{E}(t^{p-1}\int_0^t \left(\sum_{k=1}^d (F^k)_s^2)^p ds\right)\\
&\quad\quad \text{(by Jensen's inequality again)}\\
&\leq 4^{p-1}C_1^{\frac{p}{2}}\mathbb{E}\left(t^{p-1}\int_0^t (|Sx|)_s^p ds\right) + 4^{p-1}\mathbb{E}\left(t^{p-1}\int_0^t (F^0)_s^p ds\right)\\
&\quad + 4^{p-1}\Lambda_p\mathbb{E}\left(t^{p-1}\int_0^t (C_1|Sx|^2)^p ds\right) + 4^{p-1}\Lambda_p\mathbb{E}\left(t^{p-1}\int_0^t \left(\sum_{k=1}^d (F^k)_s^2\right)^p ds\right)\\
&= t^{p-1}(4^{p-1}C_1^{\frac{p}{2}} + 4^{p-1}\Lambda_p C_1^p)\mathbb{E}\left(\int_0^t (|Sx|)_s^p ds\right)\\
&\quad + 4^{p-1}\mathbb{E}\left(t^{p-1}\int_0^t (F^0)_s^p ds\right) + 4^{p-1}\Lambda_p\mathbb{E}\left(t^{p-1}\int_0^t \left(\sum_{k=1}^d (F^k)_s^2\right)^p ds\right)\\
&= t^{p-1}(4^{p-1}C_1^{\frac{p}{2}} + 4^{p-1}\Lambda_p C_1^p)\int_0^t \mathbb{E}\left((|Sx|)_s^p\right) ds\\
&\quad + 4^{p-1}t^{p-1}\int_0^t \mathbb{E}\left((F^0)_s^p\right) ds + 4^{p-1}\Lambda_p t^{p-1}\int_0^t \mathbb{E}\left(\left(\sum_{k=1}^d (F^k)_s^2\right)^p\right) ds\\
&\leq t^{p-1}(4^{p-1}C_1^{\frac{p}{2}} + 4^{p-1}\Lambda_p C_1^p)(t.K_1 e^{\max(K_2,\beta)tp})\\
&\quad + 4^{p-1}t^{p-1}C_2^p e^{\beta tp}t + 4^{p-1}\Lambda_p t^{p-1}C_2^p e^{\beta tp}t\\
&\quad\quad \text{(by the first part of the theorem, and the third assumption)}\\
&\leq 4^{p-1}(1+\Lambda_p)(K_1C_1^p + C_2^p)t^p e^{\max(K_2,\beta)tp}\\
&= M^p t^p e^{\max(K_2,\beta)tp},
\end{aligned}
$$

where $M = 4^{1-\frac{1}{p}}(1+\Lambda_p)^{\frac{1}{p}}(K_1C_1^p + C_2^p)^{\frac{1}{p}}$. The theorem follows. □

2.4. Norris Lemmas

The following is a modification of Lemma 2.3.2 in [**49**]. We use the same notation and almost the same proof, the main difference with [**49**] is that we have to be careful about the exact value of ϵ_0, and how it depends on t_0. (This is required to obtain any upper bounds on the density, regardless of issues about growth of constants.)

THEOREM 2.4.1. *Let $\alpha, \beta(t), y \in \mathbb{R}^m$. Suppose that*

$$\gamma(t) = (\gamma_1(t), \gamma_2(t), \ldots, \gamma_d(t)) \in \mathbb{R}^d$$

and $u(t) = (u_1(t), \ldots, u_d(t))$ are adapted processes. Set

$$\alpha(t) = \alpha + \int_0^t \beta(s)ds + \sum_{i=1}^d \int_0^t \gamma_i(s)dW_s^i$$

$$Y(t) = y + \int_0^t a(s)ds + \sum_{i=1}^d \int_0^t u_i(s)dW_s^i,$$

and assume that there exists $0 < t_0 < 1$ and $p \geq 2$ such that

$$c = \mathbb{E}\left(\sup_{0\leq t\leq t_0,} |\beta(t)|\right) + \mathbb{E}\left(\sup_{0\leq t\leq t_0, v\in\mathbb{R}^m, |v|=1} (\sum_{k=1}^d |\gamma(t)_k|^2)^{\frac{1}{2}}\right) \tag{2.4.1}$$

$$+\mathbb{E}\left(\sup_{0\le t\le t_0}|a(t)|\right)+\mathbb{E}\left(\sup_{0\le t\le t_0}|u(t)|)\right)<\infty.$$

Then, for any $q>8$ and for any $r,\nu,\kappa>0$ such that $18r+9\nu+18\kappa<q-8$ and $\kappa<1$, there exists a constant $C(\nu,\kappa)\le 1$, depending only on ν and κ, such that for all $\epsilon\le\epsilon_0=C(\nu,\kappa)t_0$.

Let ω_1 denote the following event:

$$\omega_1=\left\{\int_0^{t_0}Y_t^2dt<\epsilon^q,\int_0^{t_0}(|a(t)|^2+|u(t)|^2)dt>\epsilon\right\}.$$

We have:

$$\begin{aligned}\mathbb{P}(\omega_1)&\le\mathbb{P}\left(\sup_{t\le t_0}\left(|\beta(t)|+(\sum_{k=1}^d|\gamma_k(t)|^2)^{\frac12}+|a(t)|+|u(t)|\right)\ge e^{-r}\right)+2e^{-\epsilon^{-\nu}}\\&\le c\epsilon^{rp}+2e^{-\epsilon^{-\nu}}.\end{aligned}$$

PROOF. Set

$$\theta_t=\sup_{t\le t_0}\left(|\beta(t)|+(\sum_{k=1}^d|\gamma_k(t)|^2)^{\frac12}+|a(t)|+|u(t)|\right).$$

Fix $q>8$ and r,ν,κ such that $\Delta:=q-8-18r-9\nu-18\kappa>0$. Suppose that $\nu'=\nu+2\kappa$. ν' satisfies satisfies $\Delta:=q-8-18r-9\nu'>0$. Then we define the bounded stopping time

$$T=\inf\left(s\ge 0:\sup_{0\le u\le s}\theta_u>\epsilon^{-r}\right)\wedge t_0.$$

We have

$$\mathbb{P}(\omega_1)\le A_1+A_2,$$

with $A_1=\mathbb{P}(T<t_0)$ and

$$A_2=\mathbb{P}(\omega_1,T=t_0).$$

By the definition of T and condition (2.4.1), we obtain

$$A_1\le\mathbb{P}\left(\sup_{0\le s\le t_0}\theta_s>\epsilon^{-r}\right)\le\epsilon^{rp}\mathbb{E}\left(\sup_{0\le s\le t_0}\theta_s^p\right)\le c\epsilon^{rp}$$

Let us introduce the following notation (summation over i is implied)

$$\begin{aligned}A_t&=\int_0^t a(s)ds, & M_t^v&=\int_0^t u_i(s)dW_s^i,\\N_t^v&=\int_0^t Y(s)u_i(s)dW_s^i, & Q_t^v&=\int_0^t A(s)\gamma_i(s)dW_s^i.\end{aligned}$$

Define for any $\rho_i>0,\delta_i>0,i=1,2,3,$

$$B_1^v=\left([N_T]<\rho_1,\sup_{0\le s\le T}|N_s|\ge\delta_1\right)$$

$$B_2^v=\left([M_T]<\rho_1,\sup_{0\le s\le T}|M_s|\ge\delta_1\right)$$

$$B_3^v=\left([Q_T]<\rho_1,\sup_{0\le s\le T}|Q_s|\ge\delta_1\right).$$

The usual exponential martingale inequality implies $\forall v\in\mathbb{R}^m:\mathbb{P}(B_i^v)\le 2e^{-\frac{\delta_i^2}{2\rho_i}}$. Our aim is to prove the following inclusion:

$$\{\omega_1,T=t_0\}\subset B_1\cup B_2\cup B_3\tag{2.4.2}$$

for the particular choices of ρ_i and δ_i:

$$\rho_1 = \epsilon^{-2r+q}, \quad \delta_1 = \epsilon^{q_1},\ q_1 = \frac{q}{2} - r - \frac{\nu'}{2},$$
$$\rho_2 = 2(2t_0+1)^{\frac{1}{2}}\epsilon^{-2r+\frac{q_1}{2}}, \quad \delta_2 = \epsilon^{q_2},\ q_2 = \frac{q}{8} - \frac{5r}{4} - \frac{5\nu'}{8},$$
$$\rho_3 = 36t_0\epsilon^{-2r+2q_2}, \quad \delta_3 = \epsilon^{q_3},\ q_3 = \frac{q}{8} - \frac{9r}{4} - \frac{9\nu'}{8}.$$

From the inclusion (2.4.2) we get

$$\begin{aligned}
A_2 &\leq 2\left(\exp(-\frac{\delta_1^2}{2\rho_1}) + \exp(-\frac{\delta_2^2}{2\rho_2}) + \exp(-\frac{\delta_3^2}{2\rho_3})\right)\\
&\leq 2\left(\exp(-\frac{1}{2}\epsilon^{-\nu'}) + \exp(-\frac{1}{4\sqrt{1+2t_0}}\epsilon^{-\nu'}) + \exp(-\frac{1}{72t_0}\epsilon^{-\nu'}\right)\\
&\leq 2\left(\exp(-\frac{1}{2}\epsilon^{-\nu-\kappa}\epsilon^{-\kappa}) + \exp(-\frac{1}{4\sqrt{1+2t_0}}\epsilon^{-\nu-\kappa}\epsilon^{-\kappa}) + \exp(-\frac{1}{72t_0}\epsilon^{-\nu-\kappa}\epsilon^{-\kappa}\right)\\
&\leq 6\left(\exp(-\epsilon^{-\nu-\kappa})\right)
\end{aligned}$$

as long as $\epsilon \leq 72^{\frac{-1}{\kappa}}$. Now, for $\epsilon \leq \bar{C}(\nu,\kappa)$, for some $\bar{C}(\nu,\kappa)$, we can also conclude

$$A_2 \leq \left(\exp(-\epsilon^{-\nu})\right)$$

To conclude the above, we have also used the following:

$$\begin{aligned}
2q_1 + 2r - q &= -\nu',\\
2q_2 + 2r - \frac{q_1}{2} &- \nu',\\
3q_3 + 2r - 2q_2 &= -\nu'.
\end{aligned}$$

It remains to show only the inclusion (2.4.2).

Suppose that $\omega \notin B_1 \cup B_2 \cup B_3$, $T(\omega) = t_0$, and $\int_0^T Y_t^2 dt < \epsilon^q$. Then

$$[N]_T = \int_0^T Y_t^2 |u_t|^2 dt < \epsilon^{-2r+q} = \rho_1.$$

Then, since $\omega \notin B_1$, $\sup_{0\leq s\leq T} |\int_0^T Y_s u_s^i dW_s^i| \leq \delta_1 = \epsilon^{q_1}$. We also have

$$\sup_{0\leq s\leq T}\left|\int_0^T Y_s a_s ds\right| \leq \left(t_0 \int_0^T Y_t^2 a_t^2 dt\right)^{\frac{1}{2}} < \sqrt{t_0}\epsilon^{-r+\frac{q}{2}}.$$

The above two points allow us to deduce that

$$\sup_{0\leq s\leq T}\left|\int_0^T Y_s dY_s\right| < \sqrt{t_0}\epsilon^{-r+\frac{q}{2}} + \epsilon^{q_1}.$$

Then by Itô's formula, $Y_t^2 = y^2 + 2\int_0^t Y_s dYs + [M]_t$, which gives the following control on the time integral of $[M]_t$:

$$\begin{aligned}
\int_0^T [M]_t dt &= \int_0^T Y_t^2 dt - Ty^2 - 2\int_0^T \left(\int_0^t Y_s dY_s\right) dt\\
&\leq \epsilon^q + 2t_0(\sqrt{t_0}\epsilon^{-r+\frac{q}{2}} + \epsilon^{q_1}).
\end{aligned}$$

Now, observe that $q > q_1 + \kappa$ and $-r + \frac{q}{2} > q_1 + \kappa$. Therefore, as long as $\epsilon \leq 3^{\frac{-1}{\kappa}} \leq (1 + 2t_0\sqrt{t_0})^{\frac{-1}{\kappa}}$ we get

$$\int_0^T [M]_t dt \leq \epsilon^{q_1}(2t_0 + 1).$$

Now, since $[M]_t$ is an increasing process, for any $\gamma \leq T$ we have

$$\gamma[M]_{T-\gamma} < (2t_0+1)\epsilon^{q_1},$$

and hence $[M]_T < \gamma^{-1}(2t_0+1)\epsilon^{q_1} + \gamma\epsilon^{-2r}$. Choosing $\gamma = (2t_0+1)^{\frac{1}{2}}\epsilon^{\frac{q_1}{2}}$, we obtain $[M]_T < \rho_2$ (since $\epsilon < 1$). Since $\omega \notin B_2$, we get

$$\sup_{0 \leq T} |M_t| < \delta_2 = \epsilon^{q_2}$$

Recall that $\int_0^T Y_t^2 < \epsilon^q$, so by Tchebychev's inequality,

$$\lambda^1\left(t \in [0,T] : |Y_s(\omega)| \geq \epsilon^{\frac{q}{3}}\right) \leq \epsilon^{\frac{q}{3}}$$

(here, λ^1 is the Lebesgue measure on $\mathbb{R}$.) Then

$$\lambda^1\left(t \in [0,T] : |y + A_t(\omega)| \geq \epsilon^{\frac{q}{3}} + \epsilon^{q_2}\right) \leq \epsilon^{\frac{q}{3}}.$$

Now, as long as $C(\kappa,\nu) \leq \frac{1}{2}$, we have $\epsilon^{\frac{q}{3}} \leq \epsilon \leq \frac{t_0}{2}$. This means that for each $t \in [0,T]$, there exists $s \in [0,T]$ such that $|s-t| \leq \epsilon^{\frac{q}{3}}$ and $|y + A_s| < \epsilon^{\frac{q}{3}} + \epsilon^{q_2}$. It follows that

$$|y + A_t| \leq |y + A_s| + |\int_s^t a_r dr| < (1+\epsilon^{-r})\epsilon^{\frac{q}{3}} + \epsilon^{q_2}.$$

In particular (setting $t = 0$), we have $|y| < (1+\epsilon^{-r})\epsilon^{\frac{q}{3}} + \epsilon^{q_2}$, and for all $t \in [0,T]$ we have

$$|A_t| \leq 2\left((1+\epsilon^{-r})\epsilon^{\frac{q}{3}} + \epsilon^{q_2}\right) \leq 6\epsilon^{q_2},$$

because $q_2 < \frac{q}{3} - r$. This implies that

$$[Q]_T = \int_0^T A_t^2|\gamma_t|^2 dt < 36t_0\epsilon^{2q_2-2r} = \rho_3.$$

So since $\omega \notin B_3$, we have

$$|Q_T| = |\int_0^T |A_t|^2|\gamma_t|^2 dt < 36t_0\epsilon^{2q_2-2r} = \rho_3.$$

Finally, by Itô's formula, we obtain:

$$\begin{aligned}
\int_0^T (a_t^2 + |u_t|^2)dt &= \int_0^T a_t dA_t + [M]_T \\
&= a_T A_T - \int_0^T A_t\beta_t dt - \int_0^T A_t\gamma_i(t)dW_t^i + [M]_T \\
&\leq (1+t_0)6\epsilon^{q_2-r} + \epsilon^{q_3} + 2\sqrt{2t_0+1}\epsilon^{-2r+\frac{q_1}{2}} < \epsilon,
\end{aligned}$$

as long as $\epsilon \leq 17^{-\frac{1}{\kappa}} \leq ((1+t_0)6 + 1 + 2\sqrt{2t_0+1})^{-\frac{1}{\kappa}}$, since $q_2 - r, q_3, -2r + \frac{q_1}{2} > 1 + \kappa$.

It follows that the theorem holds for

$$C(\nu,\kappa) = \min\{17^{-\frac{1}{\kappa}}, \bar{C}(\kappa,\nu), 3^{\frac{-1}{\kappa}}, 72^{\frac{-1}{\kappa}}, \frac{1}{2}\} = \min\{\bar{C}(\kappa,\nu), 72^{\frac{-1}{\kappa}}\}$$

i.e.

$$\epsilon \leq \epsilon_0 = \min\{\bar{C}(\kappa,\nu), 72^{\frac{-1}{\kappa}}\}t_0.$$

$\square$

CHAPTER 3

Foundational Lemmas

Here we introduce suitable versions of lemmas that are fundamental to the study the behaviour of SDE and multi-dimensional Brownian motion, which are the main building blocks of any proof of upper bounds.

3.1. Linearisation

Here we show why we can always assume that F is linear.

THEOREM 3.1.1. *Let (x, σ, F) be a system, with F, σ having derivatives up to order $o+1$ at x. We have, for all $\alpha \in \mathrm{Multi}(\{0, 1, \ldots, d\})$ with $\#(\alpha) \leq o$,*

$$(*\sigma)^{[\alpha]} = *\sigma^{[\alpha]}.$$

In particular, the weak or strong Hörmander constants of the systems

$$(x, \sigma, F) \quad \textit{and} \quad ((x, F(x)), (\sigma, F(\sigma)), \tilde{F}),$$

where $(x, F(x)), (\sigma, F(\sigma)) \in \mathbb{R}^{m+n} \simeq \mathbb{R}^m \otimes \mathbb{R}^n$ and $\tilde{F}$ denotes projection on the last n components, coincide at any point.

PROOF. This is a straightforward consequence of the Schwartz theorem. We prove the result by induction. The result is clear for $\#(\alpha) = 1$. Next, we suppose that $\alpha = (\bar{\alpha}, i)$ for some $\bar{\alpha} \in \mathrm{Multi}(\{0, 1, \ldots, d\})$ and some $i \in \{0, 1, \ldots, d\}$. We have

$$\begin{aligned}
(*\sigma)^{[\alpha]} &= \frac{\partial dF.\sigma^i}{\partial (*\sigma)^{[\bar{\alpha}]}} - \frac{\partial dF.(*\sigma)^{[\bar{\alpha}]}}{\partial \sigma^i} \\
&= dF.\frac{\partial \sigma^i}{\partial (*\sigma)^{[\bar{\alpha}]}} + d^2F(\sigma^i, (*\sigma)^{[\bar{\alpha}]}) - dF.\frac{\partial (*\sigma)^{[\bar{\alpha}]}}{\partial \sigma^i} - d^2F((*\sigma)^{[\bar{\alpha}]}, \sigma^i) \\
&= dF.\frac{\partial \sigma^i}{\partial (*\sigma)^{[\bar{\alpha}]}} + -dF.\frac{\partial (*\sigma)^{[\bar{\alpha}]}}{\partial \sigma^i} \quad \text{(by Schwartz's theorem)} \\
&= dF.\frac{\partial \sigma^i}{\partial * \sigma^{[\bar{\alpha}]}} + -dF.\frac{\partial * \sigma^{[\bar{\alpha}]}}{\partial \sigma^i} \quad \text{(by the induction hypothesis)} \\
&= *\sigma^{\alpha},
\end{aligned}$$

as expected. □

The following proposition shows that so far as the tension is concerned, when proving Gaussian bounds in terms of the Euclidean distance, we can suppose that F is linear:

PROPOSITION 3.1.1. *Let (x, σ, F) be an (L, g, G)-tense system. Let $\tilde{F} : \mathbb{R}^m \otimes \mathbb{R}^n \to \mathbb{R}^n$ be the projection onto the last n components. There exists a proper constant $\tilde{G}$ such that the system $((x, F(x)), (\sigma, dF.\sigma), \tilde{F})$ is $(L, g, \tilde{G})$-tense.*

PROOF. This is a straightforward verification. □

3.2. Bounds on the norms of the solution and the Jacobian process

Here we show bounds on the Malliavin derivatives of the solution process and the Jacobian process. The main ingredient is our Gronwall-type Lemma 2.3.1.

The following is a generalisation of Proposition A.2 in [**42**]:

THEOREM 3.2.1. *Let $\mathcal{A} = (x, F, \sigma)$ be a system such that*

$$K_N = \sup_{|v|=1} \sum_{k=0}^{d} \|\langle \sigma^k, v\rangle\|_{\partial,N}^2 \leq G$$

uniformly for some $G \geq 1$ and for all $N \leq o + 3$ for some $o \in \mathbb{N}$. We have the following bounds on the Malliavin derivatives of X and its Jacobian: There exist constants $C^1_{o,p}, C^2_{o,p}, C^3_{n,p}$ and $\beta^1_{o,p}, \beta^2_{o,p}, \beta^3_{o,p}$ (for $o \leq N - 3$) such that for any $s_1, s_2, \ldots, s_o \in \mathbb{R}$,

$$\sup_{|w_1|,|w_2|,\ldots,|v|=1} \|\sup_{s\leq t} D^{w_1}_{s_1} D^{w_2}_{s_2} \ldots D^{w_o}_{s_o}(X^v_s)\|_p \leq C^1_{o,p} e^{\beta^1_{o,p} t},$$

$$\sup_{|w_1|,|w_2|,\ldots,|v_1|,|v_2|=1} \|\sup_{s\leq t} D^{w_1}_{s_1} D^{w_2}_{s_2} \ldots D^{w_o}_{s_o}(X^{v_1,v_2}_{0\to s})\|_p \leq C^2_{o,p} e^{\beta^2_{o,p} t},$$

$$\sup_{|w_1|,|w_2|,\ldots,|v_1|,|v_2|=1} \|\sup_{s\leq t} D^{w_1}_{s_1} D^{w_2}_{s_2} \ldots D^{w_o}_{s_o}(X^{v_1,v_2}_{s\to 0})\|_p \leq C^3_{o,p} e^{\beta^3_{o,p} t},$$

$$\sup_{\substack{|v|=1\\ v\in\mathbb{R}^m}} \left(\sup_{s\leq t} \mathbb{E}(|v^T(X_{0\to s} - \mathrm{Id})|^p)\right)^{\frac{1}{p}} \leq M^1_{o,p} t^{\frac{1}{2}} e^{\beta^4_{o,p} t},$$

$$\sup_{\substack{|v|=1\\ v\in\mathbb{R}^m}} \left(\sup_{s\leq t} \mathbb{E}(|v^T(X_{s\to 0} - \mathrm{Id})|^p)\right)^{\frac{1}{p}} \leq M^2_{o,p} t^{\frac{1}{2}} e^{\beta^5_{o,p} t}.$$

Here the supremum is over any combination of unit vectors $w_1, w_2, \ldots, w_o \in \mathbb{R}^d$, and any unit $v, v_1, v_2 \in \mathbb{R}^m$, and the norm is the L^p norm in expectation. In the last two inequalities, the matrix norm is the operator norm. Furthermore, the constants $C^1_{o,p}, C^2_{o,p}, C^3_{o,p}$ and $\beta^1_{o,p}, \beta^2_{o,p}, \beta^3_{o,p}$ depend only on o, p and G, and not on m, n, d. Furthermore, for fixed o and p, the constants $C^1_{o,p}, C^2_{o,p}, C^3_{o,p}$ and $\beta^1_{o,p}, \beta^2_{o,p}, \beta^3_{o,p}$ are polynomial in G.

PROOF. The proof is mostly made of straightforward and 'classic' calculations. Because the assumption does not change up to a proper constant if we change the system into an Itô system rather than a Stratonovich one, we can work with Itô integrals.

First inequality. Differentiating the SDE formally and using standard Malliavin calculus, we have:

$$X^v_t = X^v_0 + \sum_{k=0}^{d} \int_0^t \sigma^{k,v}(X_s) dW^k_s,$$

$$\begin{aligned} D^{w_1}_{s_1} X^v_t &= \sum_{k=1}^{d} \sigma(X_{s_1})^{k,v}(w_1)_k + \sum_{k=0}^{d} \int_0^t \langle \partial\sigma(X_s)^{k,v}, D^{w_1}_{s_1} X_s\rangle_{\mathbb{R}^m} dW^k_s \\ &= \sigma^{w_1,v}_{s_1} + \sum_{k=0}^{d} \int_0^t \langle \partial\sigma^{k,v}_s, D^{w_1}_{s_1} X_s\rangle_{\mathbb{R}^m} dW^k_s, \end{aligned}$$

where we have used the natural notations

$$\sigma^{w_1,v}_{s_1} = \sigma(X_{s_1})^{w_1,v} = \sum_{k=1}^{d} \sigma(X_{s_1})^{k,v}(w_1)_k,$$

$$\sigma_s = \sigma(X_s),$$

$$\langle \partial\sigma^{k,v}, \bar{v}\rangle = \sum_{i=1}^{m} \frac{\partial \sigma^{k,v}}{\partial x_i}\bar{v}_i,$$

etc.

For the next two degrees, we also obtain similarly:

$$D^{w_1}_{s_1} D^{w_2}_{s_2} X^v_t = \langle \partial\sigma^{k,v}_{s_1}, D^{w_2}_{s_2} X_{s_1}\rangle + \int_0^t \sum_{k=0}^{d} \langle \partial\sigma^{k,v}_s, D^{w_1}_{s_1} D^{w_2}_{s_2} X_s\rangle dW^k_s$$
$$+ \int_0^t \sum_{k=0}^{d} \langle \partial^2\sigma^{k,v}_s, D^{w_1}_{s_1} X_s, D^{w_2}_{s_2} X_s\rangle dW^k_s \quad \text{and}$$
$$D^{w_1}_{s_1} D^{w_2}_{s_2} D^{w_3}_{s_3} X_t = \langle \partial\sigma^{w_1,v}_{s_1}, D^{w_2}_{s_2} D^{w_3}_{s_3} X_{s_1}\rangle + \langle \partial^2\sigma^{w_1,v}_{s_1}, D^{w_2}_{s_2} X_{s_1}, D^{w_3}_{s_3} X_{s_1}\rangle$$
$$+ \int_0^t \sum_{k=0}^{d} \langle \partial\sigma^{k,v}_s, D^{w_1}_{s_1} D^{w_2}_{s_2} D^{w_3}_{s_3} X_s\rangle dW^k_s$$
$$+ \int_0^t \sum_{k=0}^{d} \langle \partial^2\sigma^{k,v}_s, D^{w_1}_{s_1} X_s, D^{w_2}_{s_2} D^{w_3}_{s_3} X_s\rangle dW^k_s$$
$$+ \int_0^t \sum_{k=0}^{d} \langle \partial^2\sigma^{k,v}_s, D^{w_1}_{s_1} D^{w_2}_{s_2} X_s, D^{w_3}_{s_3} X_s\rangle dW^k_s$$
$$+ \int_0^t \sum_{k=0}^{d} \langle \partial^2\sigma^{k,v}_s, D^{w_2}_{s_2} X_s, D^{w_1}_{s_1} D^{w_3}_{s_3} X_s\rangle dW^k_s$$
$$+ \int_0^t \sum_{k=0}^{d} \langle \partial^3\sigma^{k,v}_s, D^{w_1}_{s_1} X_s, D^{w_2}_{s_2} X_s, D^{w_3}_{s_3} X_s\rangle dW^k_s.$$

For a subset H of $1, 2, \ldots, o$, we write $D^H X_s$ for $D^{w_{\tilde{H}_1}}_{s_{\tilde{H}_1}} D^{w_{\tilde{H}_2}}_{s_{\tilde{H}_2}} \ldots D^{w_{\tilde{H}_{\#H}}}_{s_{\tilde{H}_{\#H}}}$, where $\tilde{H}$ is the ordered set obtained by ordering H according to the natural order relation on $1, 2, \ldots, o$. Now, for the general order, we get:

$$D^{1,2,3,\ldots,o} X^v_t = \sum_{\substack{\cup_{j=1}^i I_j \\ = \{2,3,\ldots,o\}}} \langle \partial^i \sigma^{w_1,v}_{s_1}, D^{I_1} X_{s_1}, D^{I_2} X_{s_1} \ldots, D^{I_j} X_{s_1}\rangle$$
$$+ \int_0^t \sum_{k=0}^{d} \sum_{\substack{\cup_{j=1}^i I_j \\ = \{1,2,3,\ldots,o\}}} \langle \partial^i \sigma^{k,v}_s, D^{I_1} X_s, D^{I_2} X_s \ldots, D^{I_j} X_s\rangle dW^k_s. \tag{3.2.1}$$

We can now prove the (first) claim by induction using Theorem 2.3.1: For the initial case, recall that

$$X_t = X_0 + \sum_{k=0}^{d} \int_0^t \sigma^k(X_s) dW^k_s.$$

This equation is of the form of the assumptions of Theorem 2.3.1, with $S_0 = X_0$, $E = 0$, $F = \sigma(X_s)$, which satisfy the assumptions of Theorem 2.3.1 with $\beta = 0$, $C_2 = G$ and all other constants being null. Therefore we can deduce that

$$\|X\|_p \le C_{0,p} e^{\beta_{0,p} t},$$

with $C_{0,p} = (2p(2p-1))^{\frac{1}{p}}$ and $\beta_{0,p} = 0$.

For the induction step, suppose that the inequality holds for $o \le N - 1$, we must check that it holds for $o = N$. Separating the only term that involves N^{th} order derivatives, and observing that

$D^{\cdot}_{s_1} X_{s_2} = 0$ unless $s_2 \geq s_1$, Eq. (3.2.1) becomes:

$$\begin{aligned} D^{1,2,3,\dots,N} X_t = & \sum_{\substack{\cup_{j=1}^i I_j \\ =\{2,3,..N\}}} \langle \partial^i \sigma_{s_1}^{w_1}, D^{I_1} X_{s_1}, D^{I_2} X_{s_1} \dots, D^{I_j} X_{s_1} \rangle \\ & + \int_{\max(s_i)}^t \sum_{k=0}^d \sum_{\substack{\cup_{j=1}^i I_j = \{1,2,..N\} \\ i \geq 2}} \langle \partial^i \sigma_s^k, D^{I_1} X_s, D^{I_2} X_s, \dots, D^{I_j} X_s \rangle dW_s^k \\ & + \int_{\max(s_i)}^t \sum_{k=0}^d \langle \partial \sigma_s^k, D^{1,2,\dots,N} X_t \rangle. \end{aligned}$$

This is an equation of the form of the assumptions of Theorem 2.3.1, started at $\max(s_i)$ (instead of 0), with

$$\begin{aligned} E_s^k. &= \langle \partial \sigma_s^k, . \rangle, \\ F^k &= \sum_{\substack{\cup_{j=1}^i I_j = \{1,2,3,..N\} \\ i \geq 2}} \langle \partial^i \sigma_s^k, D^{I_1} X_s, D^{I_2} X_s, \dots, D^{I_i} X_s \rangle \quad \text{and} \\ S_{s_{N-1}} &= \sum_{\substack{\cup_{j=1}^i I_j \\ =\{2,3,..N\}}} \langle \partial^i \sigma_{s_1}^{w_1}, D^{I_1} X_{s_1}, D^{I_2} X_{s_1} \dots, D^{I_i} X_{s_1} \rangle. \end{aligned}$$

Now by the assumption on the tension, we have that $\sum_{k=0}^d \langle E^k, v \rangle^2 \leq \bar{G}$ for any $|v| = 1$ and for some (strongly) polynomial G. Furthermore, again by this assumption, we have:

$$\begin{aligned} |S_{s_{N-1}}|^2 &= \sum_{\substack{\cup_{j=1}^i I_j \\ =\{2,3,..N\}}} \langle \partial^i \sigma_{s_1}^{w_1}, D^{I_1} X_{s_1}, D^{I_2} X_{s_1} \dots, D^{I_i} X_{s_1} \rangle \\ &\leq G \sum_{\substack{\cup_{j=1}^i I_j \\ =\{2,3,..N\}}} \prod_{j=1}^i |D^{I_j} X_{s_1}|^2, \end{aligned}$$

using the induction hypothesis, we get

$$\begin{aligned} \| |S_{s_{N-1}}|^2 \|_p &\leq G \sum_{\substack{\cup_{j=1}^i I_j \\ =\{2,3,..N\}}} (C_{N-1,p})^i e^{\beta_{N-1,p} G^2 (\max_u s_u) i} \\ &\leq 2^{N-2} G (C_{N-1,p})^{N-1} e^{\beta_{N-1,p} G^2 (\max_u s_u)(N-1)}, \end{aligned}$$

and for $s \geq \max_u(s_u)$,

$$\|F_s\|_p \leq 2^{N-1} G (C_{N-1,p})^{N-1} e^{\beta_{N-1,p} G^2 N (s - \max_u(s_u))}.$$

We are now in a position to apply Theorem 2.3.1, to obtain directly:

$$\begin{aligned} \| |S|^2 \|_p &\leq C_{N,p} e^{\beta_{N-1,p} G^2 (\max_u s_u)(N-1)} e^{\max(K_2, \beta_{N-1,p} G^2 (N-1))(t - \max_u s_u)} \\ &\leq C_{N,p} e^{\max(K_2, \beta_{N-1,p} G^2 (N-1)) t} \end{aligned}$$

with

$$\begin{aligned} C_{N,p} &= \left((2^{N-2} G (C_{N-1,p})^{N-1})^p + 2p(2p-1) \right)^{\frac{1}{p}} \\ &\leq 2 \max(2^{N-2} G (C_{N-1,p})^{N-1}, (2p(2p-1))^{\frac{1}{p}} \quad \text{and} \\ K_2 &= 2(2G+1) + 8(p-1)(4G^2+1) + 2(2p-1) 2^{N-1} G (C_{N-1,p})^{N-1}. \end{aligned}$$

The first result follows at once.

Second inequality: We have

$$
\begin{aligned}
X^{v,\bullet}_{0\to t} &= v + \sum_{\substack{0\le k\le d\\ 1\le i\le m}} \int_0^t \left(\frac{\partial \sigma^k(X_s)}{\partial e^i}\right)\left(X^{v,i}_{0\to s}\right) dW^k_s \\
&= v + \sum_{0\le k\le d} \int_0^t \langle \partial\sigma^k(X_s), X^{v,\bullet}_{0\to s}\rangle dW^k_s.
\end{aligned}
$$

This is an equation of the form of the assumptions of Theorem 2.3.1, with $E^k = \langle \partial\sigma^k, \bullet\rangle$, $F^k = 0$ and $S_0 = v$. The second inequality follows for $o = 0$.

For $o = 1$, formally differentiating the above equation yields:

$$
\begin{aligned}
D^w_{s_1} X^{v,\bullet}_{0\to t} = \langle \sigma^w(X_{s_1}), X^{v,\bullet}_{0\to s_1}\rangle + \sum_{k=0}^d \int_{s_1}^t \langle \partial^2\sigma^k(X_s), D^w_{s_1} X_s, X^{v,\bullet}_{0\to s}\rangle dW^k_s \\
+ \int_{s_1}^t \langle \partial\sigma^k, D^w_{s_1}(X^{v,\bullet}_{0\to s})\rangle dW^k_s.
\end{aligned}
$$

Again, this satisfies the assumptions of Theorem 2.3.1, with:

$$
\begin{aligned}
E^k_s &= \langle \partial\sigma^k(X_s), \bullet\rangle \\
S_s &= \langle \sigma^w(X_s), X^{v,\bullet}_{0\to s}\rangle \\
F^k_s &= \sum_{k=0}^d \int_{s_1}^t \langle \partial^2\sigma^k(X_s), D^w_{s_1} X_s, X^{v,\bullet}_{0\to s}\rangle dW^k_s.
\end{aligned}
$$

Therefore, the result for $o = 1$ follows by using the first inequality for $o = 1$ and the second inequality for $o = 0$

For greater o, we also proceed by induction. First note that the development above for the first inequality actually shows that we an inequality of the same form for $|D^{1,2,\ldots,N}(\sigma^w(X_t))|_p$ for any $w \in \mathbb{R}^d$, any $(w_1, s_1), (w_2, s_2), \ldots (w_N, s_N)$.

Next, observe that

$$
\begin{aligned}
D^{1,2,\ldots,o}(X^{v,\bullet}_{0\to t}) &= D^{2,\ldots,N}(\langle \partial\sigma^{w_1}(X_{s_1}), X_{0\to s_1}\rangle) \\
&\quad + \int_{s_1}^t \sum_{k=0}^d D^{1,2,\ldots,N}(\langle \partial\sigma^k(X_s), X^{v,\bullet}_{0\to s}\rangle) dW^k_s \\
&= \sum_{I\cup J=\{2,\ldots,o\}} (\langle D^I\partial\sigma^{w_1}(X_{s_1}), D^J X_{0\to s_1}\rangle) \\
&\quad + \sum_{I\cup J=\{1,\ldots,o\}} \int_{s_1}^t \sum_{k=0}^d D^{1,2,\ldots,N}(\langle D^I\partial\sigma^k(X_s), D^J X^{v,\bullet}_{0\to s}\rangle) dW^k_s.
\end{aligned}
$$

This fits the setting of Theorem 2.3.1 with

$$
\begin{aligned}
S_{s_1} &= \sum_{I\cup J=\{2,\ldots,o\}} \langle D^I\partial\sigma^{w_1}(X_{s_1}), D^J X_{0\to s_1}\rangle \quad \text{(initial condition)} \\
F^k_s &= \sum_{I\cup J=\{1,\ldots,o\}, J\neq\varnothing} \langle D^I\partial\sigma^k(X_s), D^J X^{v,\bullet}_{0\to s}\rangle \\
E^k_s &= \langle D^{1,2,\ldots,o}\partial\sigma^k(X_s), \bullet\rangle.
\end{aligned}
$$

This allows us to perform the induction step by using both the first inequality for differentiation indices less than o and the second inequality for differentiation indices less than $o - 1$. It is clear that at each step, the constants stay polynomial (p,o-dependent) functions of G.

Third inequality: This is again very similar to the above. We have

$$\begin{aligned} X^{v,\bullet}_{t\to 0} = v &- \sum_{k=1}^{d}\int_0^t \langle \partial\sigma^k(X_s), X^{v,\bullet}_{s\to 0}\rangle dW^k_s \\ &- \int_0^t \langle \partial\sigma^0(X_s) - \sum_{k=1}^{d}\langle \partial\sigma^k(X_s), \partial\sigma^k(X_s)\rangle, X^{v,\bullet}_{s\to 0}\rangle ds. \end{aligned}$$

This is indeed a setting for applying Theorem 2.3.1 with

$$\begin{aligned} S_0 &= v, \\ E^k_s &= -\langle \partial\sigma^k(X_s), \cdot\rangle \quad \text{for} \quad k \neq 0, \\ E^0_s &= -\partial\sigma^0(X_s) + \sum_{k=1}^{d}\langle \partial\sigma^k(X_s), \partial\sigma^k(X_s)\rangle, \\ F &= 0. \end{aligned}$$

This allows us to conclude the case $o = 0$ for the third inequality.

For higher derivatives, the situation is the same as above with the matrix function of X_s given by $\partial\sigma^k(X_s)$ replaced by $-\partial\sigma^0(X_s) + \sum_{k=1}^{d}\langle \partial\sigma^k(X_s), \partial\sigma^k(X_s)\rangle$ for $k = 0$ and with just the sign changed for $k \neq 0$:

Writing

$$\begin{aligned} \mu^k : \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^m, \quad &\xi \mapsto \mu^k(\xi) = -\partial\sigma^k(\xi) \quad \text{if} \quad k \neq 0 \\ &\xi \mapsto \mu^0(\xi) = -\partial\sigma^0(\xi) + \sum_{k=1}^{d}\langle \partial\sigma^k(\xi), \partial\sigma^k(\xi)\rangle \quad \text{for} \quad k = 0, \end{aligned}$$

we can apply Theorem 2.3.1 with

$$\begin{aligned} S_{s_1} &= \sum_{I\cup J=\{2,\ldots,o\}} \langle D^I \mu^{\omega_1}(X_{s_1}), D^J X_{0\to s_1}\rangle \quad \text{(initial condition)} \\ F^k_s &= \sum_{I\cup J=\{1,\ldots,o\}, J\neq\varnothing} \langle D^I \mu^k(X_s), D^J X^{v,\bullet}_{0\to s}\rangle \\ E^k_s &= \langle D^{1,2,\ldots,o}\mu^k(X_s), \cdot\rangle. \end{aligned}$$

Because it is as easy to bound the sums over k of the operator norms of μ^k as those of σ^k, we can, again conclude using the induction hypothesis for step less than $o-1$ and using the first inequality for step less than o. Again, it is clear that at each step, the constants stay (p,o-dependent) polynomial functions of G. □

We can now prove a similar development for the Ornstein-Uhlenbeck operator L applied to X, and its Malliavin derivatives. We will need a stricter result, involving a factor of $t^{\frac{1}{2}}$, for the zeroth order derivative. This is the analogue of Corollary A.1 in [**42**], merged together with point '(4)' of the proof of Theorem 4.1 in the same paper.

THEOREM 3.2.2. *Let $\mathcal{A} = (x, F, \sigma)$ be a system such that*

$$K_N = \sup_{|v|=1}\sum_{k=0}^{d} \|\langle \sigma^k, v\rangle\|^2_{\partial,N} \leq G$$

uniformly for some $G \geq 1$ and for all $N \leq o+3$ for some $o \in \mathbb{N}$. Let L denote the Ornstein Uhlenbeck operator: for $v_1 \in \mathbb{R}^d$ and $v_2 \in \mathbb{R}^m$,

$$L^{v_1,v_2} = L^{v_1}(X^{v_2}) = -\sum_{i=1}^{d}(v_1)_i \delta^i(D^i(X^{v_2}))$$

$$L^{k,v} = L^{e^k, v_2}$$
$$L^v = \sum_k L^{k,v} = -\delta(D(X^v)),$$

for all $p \in \mathbb{N}$, there exist constants M,γ, C_o, and K_o, depending only on o, p, G, and polynomial in G for fixed o, p, such that for any combination of unit vectors $w_1, w_2, \ldots, w_o \in \mathbb{R}^d$, any $s_1, s_2, \ldots, s_o \in \mathbb{R}$, and any unit $v \in \mathbb{R}^m$,

$$\sup_{|w_1|,|w_2|,\ldots,|v|=1} \sup_{0\le s_1,s_2,\ldots,s_o\le t} \|\sup_{s\le t} D^{w_1}_{s_1} D^{w_2}_{s_2} \ldots D^{w_o}_{s_o}(L^v_s)\|_p \le C_o e^{K_o t}, \quad \textit{and}$$
$$\sup_{|v|=1} \|\sup_{s\le t} L^v_s\|_p \le M t^{\frac{1}{2}} e^{\gamma t}.$$

PROOF. By the same remark as in the Proof of 3.2.1, we can suppose $w_1, \ldots, v$ are fixed as long as all calculations on the norms of E, F are uniform over that choice. We have that

$$L_t = \sum_{k=1}^d \int_0^t \sigma^k(X_s) dW^k_s + \sum_{k=0}^d \int_0^t \Big(\langle \partial\sigma^k, L_s\rangle - \langle \partial^2\sigma^k(X_s), DX_s, DX_s\rangle\Big)\, dW^k_s.$$

This puts us in the situation of Theorem 2.3.1 with

$$S_0 = 0,$$
$$F^k_s = \sigma^k(X_s)1_{k\neq 0} - \langle \partial^2\sigma^k(X_s), DX_s, DX_s\rangle,$$
$$E^k = \partial\sigma^k(X_s).$$

The assumptions on the system, along with the first inequality of Theorem 3.2.1, shows that the assumptions of Theorem 2.3.1 are satisfied, and since $S_0 = 0$, we can even apply the second part of Theorem 2.3.1. It follows immediately that $\|L^v\|_p \le M t^{\frac{1}{2}} e^{\gamma t}$ for appropriate proper constants M, γ.

Now, for the induction case, differentiating Eq. (3.2.2), we get, using the same notations as in the proof of Theorem 2.3.1 (write $N = o$)

$$\begin{aligned}
&D^{w_1}_{s_1} D^{w_2}_{s_2} \ldots D^{w_o}_{s_o}(L^v_t)\\
&= D^{2,\ldots,N}\sigma^{\omega_1}(X_{s_1}) + \sum_{\substack{I\cup J=\\ \{2,\ldots,N\}}} \langle D^I\partial\sigma^{\omega_1}(X_{s_1}), D^J L_{s_1}\rangle\\
&\quad - \sum_{\substack{I\cup J\cup Q\\ =\{2,\ldots,N\}}} \langle D^I\partial^2\sigma^{\omega_1}(X_{s_1}), D^J(X_{s_1}), D^Q(X_{s_1})\rangle\\
&\quad + \int_{s_1}^t \sum_{k=0}^d D^{1,\ldots,N}\sigma^k(X_s)dW^k_s + \int_{s_1}^t \sum_{k=0}^d \sum_{\substack{I\cup J=\\ \{1,\ldots,N\}}} \langle D^I\partial\sigma^k(X_s), D^J L_s\rangle dW^k_s\\
&\quad - \int_{s_1}^t \sum_{k=0}^d \sum_{\substack{I\cup J\cup Q\\ =\{1,\ldots,N\}}} \langle D^I\partial^2\sigma^k(X_s), D^J(X_s), D^Q(X_s)\rangle dW^k_s\\
&= D^{2,\ldots,N}\sigma^{\omega_1}(X_{s_1}) + \sum_{\substack{I\cup J=\\ \{2,\ldots,N\}}} \langle D^I\partial\sigma^{\omega_1}(X_{s_1}), D^J L_{s_1}\rangle\\
&\quad - \sum_{\substack{I\cup J\cup Q\\ =\{2,\ldots,N\}}} \langle D^I\partial^2\sigma^{\omega_1}(X_{s_1}), D^J(X_{s_1}), D^Q(X_{s_1})\rangle\\
&\quad + \int_{s_1}^t \sum_{k=0}^d D^{1,\ldots,N}\sigma^k(X_s)dW^k_s + \int_{s_1}^t \sum_{k=0}^d \sum_{\substack{I\cup J=\{1,\ldots,N\};\\ I\neq\varnothing}} \langle D^I\partial\sigma^k(X_s), D^J L_s\rangle dW^k_s
\end{aligned}$$

$$- \int_{s_1}^t \sum_{k=0}^d \sum_{\substack{I\cup J\cup Q \\ =\{1,\dots,N\}}} \langle D^I \partial^2 \sigma^k(X_s), D^J(X_s), D^Q(X_s)\rangle dW_s^k$$
$$+ \int_{s_1}^t \sum_{k=0}^d \langle \partial\sigma^k(X_s), D^{1,2,\dots,N} L_s\rangle dW_s^k.$$

This fits the setting of Theorem 2.3.1 with:

$$\begin{aligned}
S_{s_1} &= D^{2,\dots,N}\sigma^{\omega_1}(X_{s_1}) + \sum_{\substack{I\cup J= \\ \{2,\dots,N\}}} \langle D^I \partial\sigma^{\omega_1}(X_{s_1}), D^J L_{s_1}\rangle \\
&\quad - \sum_{\substack{I\cup J\cup Q \\ =\{2,\dots,N\}}} \langle D^I \partial^2\sigma^{\omega_1}(X_{s_1}), D^J(X_{s_1}), D^Q(X_{s_1})\rangle, \\
F_s^k &= D^{1,\dots,N}\sigma^k(X_s) + \sum_{k=0}^d \sum_{\substack{I\cup J=\{1,\dots,N\} \\ I\neq\varnothing}} \langle D^I\partial\sigma^k(X_s), D^J L_s\rangle dW_s^k \\
&\quad - \sum_{\substack{I\cup J\cup Q \\ =\{1,\dots,N\}}} \langle D^I \partial^2\sigma^k(X_s), D^J(X_s), D^Q(X_s)\rangle, \\
E_s^k &= \partial\sigma^k(X_s).
\end{aligned}$$

Similarly to the proof of Theorem 3.2.1, the conditions of Theorem 2.3.1 are satisfied with suitable constants (polynomial in G and depending only on o, p, G), by the induction hypothesis and by the first inequality of Theorem 2.3.1.

The last two inequalities are derived similarly using the last part of Theorem 2.3.1. This concludes the proof. □

3.3. Bounds on the inverse Malliavin covariance matrix.

Here we show the invertibility of the Malliavin covariance matrix and prove upper bounds for the expected norms of the inverse. The main ingredient in the weak Hörmander case is our generalisation of the Norris Lemma 2.4.1.

We need the following classic lemma. The proof is inspired from [**49**] and from the explicit expansion on page 45 of [**2**], cf. also Lemma 4.2 page 23 in [**50**].

LEMMA 3.3.1. *Let γ be a symmetric non-negative definite $n\times n$ matrix. We assume that, for fixed $p\geq 2$, $\mathbb{E}(\|\gamma\|_{\mathrm{Fr}}^{p+1})<\infty$ where $\|\cdot\|_{\mathrm{Fr}}$ denotes the Frobenius norm, and that $\exists\, \epsilon_0, C_1, C_2 > 0$ such that $\forall \epsilon\leq\epsilon_0$,*

$$\sup_{\xi=1} \mathbb{P}\left(\langle\gamma\xi,\xi\rangle < C_1\epsilon\right) < C_2\epsilon^{p+1+2n}.$$

Then we have:

$$\mathbb{E}(\lambda_*(\gamma)^{-p}) \leq C_1^{-p}C_2\left(1+p(2\sqrt{n})^n + p\mathbb{E}(\|\gamma\|_{\mathrm{Fr}}^{p+1})\right)\epsilon_0^{-p}$$

where $\lambda_*(\gamma)$ denotes the smallest eigenvalue of γ. This immediately gives:

$$\mathbb{E}((\gamma^{-1})_{ij}^p) \leq C_1^{-p}C_2\left(1+p(2\sqrt{n})^n + p\mathbb{E}(\|\gamma\|_{\mathrm{Fr}}^{p+1})\right)\epsilon_0^{-p}.$$

PROOF. The proof is essentially the same as the proof in [**49**]. We assume that $C_1, C_2 = 1$, the more general case is a straightforward modification.

First, we will need the following elementary result:

LEMMA 3.3.2. *For any $\mu < 1/2$, a unit sphere in a space of dimension n can be covered by at most $\left(\frac{2\sqrt{n}}{\mu}\right)^n$ balls or radius μ with centres on the sphere.*

PROOF. Consider the grid of points whose coordinates are multiples of $\frac{\mu}{\sqrt{n}(1+\varepsilon)}$, for some small positive ε. For each grid point $x \in X$(where X is the set of points within $\frac{\mu}{\sqrt{n}}$ distance of the sphere), pick the point $s(x)$ on the sphere closest to x. The balls centered at the $s(x)$ and with radii $\frac{\mu}{2}$ cover the unit sphere. Indeed, by the triangle inequality, $B(s(x), \mu) \subset B(x, \frac{\mu}{2})$, and since the balls $B(x, \frac{\mu}{2})$ for x in the the grid cover the while space, those which intersect the sphere (precisely the ones with $x \in X$) cover the sphere. The cardinality of X is clearly less than $\left(\frac{2\sqrt{n}(1+\varepsilon)}{\mu}\right)^n$. Making $(1+\varepsilon)$ tend to zero gives the result. □

Now, we proceed as in [**49**] (for simplicity, we write λ for $\lambda_*(\gamma)$). Fix $\epsilon < 0$, let $v_1, v_2, \ldots, v_N$ be a finite set of points on the sphere satisfying the condition of the lemma above with $\mu = \frac{\epsilon^2}{2}$. Here $N = (4\sqrt{n})^n \epsilon^{-2n}$. Then we have:

$$\begin{aligned}\mathbb{P}(\lambda < \epsilon) &= \mathbb{P}\left(\inf_{|v|=1} v^T \gamma v < \epsilon\right) \\ &\leq \mathbb{P}\left(\inf_{|v|=1} v^T \gamma v < \epsilon, \|\gamma\|_{\text{Fr}} \leq \epsilon\right) + P\left(\|\gamma\|_{\text{Fr}} \geq \epsilon\right).\end{aligned}$$

Suppose that $\|\gamma\|_{\text{Fr}} \leq \epsilon$, and that $|v_k^T \gamma v_k| \geq 2\epsilon$ for all k, then for any unit vector v we have (choosing k such that v_k is closest to v, and in particular within $\frac{\epsilon^2}{2}$ of v)

$$\begin{aligned}v^T \gamma v &\geq v_k^T \gamma v_k - |v^T \gamma v - v_k^T \gamma v_k| \\ &\geq 2\epsilon - \left(|v^T \gamma v - v^T \gamma v_k| + |v^T \gamma v_k - v_k^T \gamma v_k|\right) \\ &\geq 2\epsilon - 2\|\gamma\|_{\text{Fr}} |v - v_k| \\ &\geq \epsilon.\end{aligned}$$

This means that if $\inf_{|v|=1} |v^T \gamma v| < \epsilon$, then $v_l^T \gamma v_k < \epsilon$ for all k. Therefore, we have:

$$\begin{aligned}\mathbb{P}(\lambda < \epsilon) &\leq \mathbb{P}(\exists k, \ |v_k^T \gamma v_k| < 2\epsilon) + \mathbb{P}(\|\gamma\|_{\text{Fr}} > \epsilon) \\ &\leq (4\sqrt{n})^n \epsilon^{-2n} (2\epsilon)^{p+1+2n} + \epsilon^{p+1} \mathbb{E}(\|\gamma\|_{\text{Fr}}^{p+1}) \\ &\leq \left((4\sqrt{n})^n 2^{p+1+2n} + \mathbb{E}(\|\gamma\|_{\text{Fr}}^{p+1})\right) \epsilon^{p+1}.\end{aligned}$$

Now,

$$\begin{aligned}\mathbb{E}(\lambda^{-p}) &= \int_0^\infty \mathbb{E}(\lambda^{-p} > y) dy \\ &= \int_0^{\epsilon_0^{-p}} 1 dy + \int_{\epsilon_0^{-p}}^\infty \mathbb{P}(\lambda^{-p} > y) dy \\ &\leq \epsilon_0^{-p} + \int_0^{\epsilon_0^{-p}} \left((4\sqrt{n})^n 2^{p+1+2n} + \mathbb{E}(\|\gamma\|_{\text{Fr}}^{p+1})\right) \epsilon^{p+1} \epsilon^{-p-1} (p-1) d\epsilon \\ &\leq \epsilon_0^p \left(p((4\sqrt{n})^n 2^{p+1+2n} + \mathbb{E}(\|\gamma\|_{\text{Fr}}^{p+1})) + 1\right).\end{aligned}$$

For the last part, observe that

$$|\gamma_{ij}^{-1}| = \langle \gamma_{ij}^{-1} e_j, e_i \rangle \leq |\gamma_{ij}^{-1} e_j| \leq \lambda_*(\gamma)^{-1}.$$

□

PROPOSITION 3.3.3. *Let (x, σ, F) be a $(1, g, G)$-tense ($g \geq 3$), uniformly H-elliptic system. Let $\Gamma_t^{i,j}$ denote the Malliavin covariance matrix (of $Y_t = F(X_t)$) at time t. For any $1 \leq p$ there exist strongly polynomial constants C_p and M_p such that for any $t \leq M_p$,*

$$\mathbb{E}\left(\lambda_*(\Gamma)^{-p}\right) \leq C_p t^{-p},$$

where for a symmetric matrix A, $\lambda_*(A)$ denotes the minimum eigenvalue of A.

PROOF. First note that by linearisation as in 3.1.1, we can suppose that F is linear. For any unit $v \in \mathbb{R}^n$, we have

$$v^T \gamma_t v = \int_0^t \sum_{k=1}^d (v^T * X_{0\to t} X_{0\to s}^{-1} \sigma^k(X_s))^2 ds.$$

From this we obtain further,

$$\begin{aligned}
\mathbb{E}(|v^T \Gamma_t v|^{p+1}) &\leq \mathbb{E}\left(\left(\int_0^t \sum_{k=1}^d |v^T * X_{0\to t} X_{0\to s}^{-1} \sigma^k(X_s)|^2 ds\right)^{p+1}\right) \\
&\leq G^{p+1} \mathbb{E}\left(\left(\int_0^t \ |v^T * X_{0\to t} X_{0\to s}^{-1}|^2 ds\right)^{p+1}\right) \\
&\leq G^{p+1} C_p e^{\beta_p t}
\end{aligned}$$

where C_p is a strongly polynomial constant and we have used Proposition 3.2.1. Therefore, for $t \leq M_p = \frac{1}{\beta_p}$ (M_p being a strongly polynomial), we do have that $\mathbb{E}(|\gamma_t|^{p+1})$ is bounded above by a proper constant.

Fix $\epsilon > 0$. Define for any ε the following event

$$\Omega = \Omega_{\epsilon,\varepsilon,v} = \{\exists s : t - \epsilon \leq s \leq t : (|v^T * (X_{t-\epsilon\to t} - \mathrm{Id})| \geq \varepsilon \vee |v^T * (X_{s\to t-\epsilon} - \mathrm{Id})| \geq \varepsilon)\}$$

where ε is some fixed constant to be fixed later. the matrix norm used is the operator norm (the constants would no longer be proper if we used the Frobenius norm). Outside Ω, we have for any $t - \epsilon \leq s \leq t$,

$$\begin{aligned}
&\sum_{k=1}^d |v^T * X_{s\to t} \sigma^k(X_s)|^2 \\
&= \sum_{k=1}^d |v^T * \sigma_s^k + v^T * (X_{s\to t} - \mathrm{Id})\sigma_s^k|^2 \\
&= \sum_{k=1}^d |v^T * \sigma_s^k|^2 + \sum_{k=1}^d (v^T * \sigma_s^k)(v^T * (X_{s\to t} - \mathrm{Id})\sigma_s^k) + \sum_{k=1}^d |v^T * (X_{s\to t} - \mathrm{Id})\sigma_s^k|^2 \\
&\geq \frac{H}{2} - \sqrt{|\sum_{k=1}^d v^T * \sigma_s^k|^2} \sqrt{|\sum_{k=1}^d v^T * (X_{s\to t} - \mathrm{Id})\sigma_s^k|^2} \\
&\quad \text{(Cauchy Schwarz)} \\
&\geq \frac{H}{2} - \sqrt{|v^T *|^2 G} \sqrt{(|v^T * (X_{s\to t} - \mathrm{Id})|^2 G)} \\
&\geq \frac{H}{2} - \sqrt{G^3} 3\varepsilon \sqrt{G^3}.
\end{aligned}$$

(Note that if $|v^T * (X_{t-\epsilon\to t} - \mathrm{Id})|, |v^T * (X_{s\to t-\epsilon} - \mathrm{Id})| \leq \varepsilon$, then $|v^T * (X_{s\to t} - \mathrm{Id})| \leq 3\varepsilon$). Therefore, we pick $\varepsilon = \frac{H}{15G^3}$ (ε is still a proper constant).

For this choice of ε, we now have, outside Ω, for any unit vector $v \in \mathbb{R}^n$:

$$\begin{aligned}
v^T \Gamma v &= \int_0^t \sum_{k=1}^d |v^T * \sigma_s^k + v^T * (X_{s\to t} - \mathrm{Id})\sigma_s^k|^2 ds \\
&\geq \int_{t-\epsilon}^t \sum_{k=1}^d |v^T * \sigma_s^k + v^T * (X_{s\to t} - \mathrm{Id})\sigma_s^k|^2 ds \\
&> \frac{H}{4}\epsilon.
\end{aligned}$$

This implies that $\mathbb{P}(\lambda_*(\Gamma) < \epsilon) \leq \mathbb{P}(\Omega)$.

Now, note that for any $1 \leq q \leq g$ and any $\epsilon \leq t$,

$$\begin{aligned}
\mathbb{P}(\Omega) &\leq \mathbb{P}\left(\sup_{t-\epsilon\leq s\leq t} |v^T*(X_{t-\epsilon\to s} - \mathrm{Id})| \geq \varepsilon\right) + \mathbb{P}\left(\sup_{t-\epsilon\leq s\leq t} |v^T*(X_{s\to t-\epsilon} - \mathrm{Id})| \geq \varepsilon\right) \\
&\leq \varepsilon^{-q}\mathbb{E}\left(\sup_{t-\epsilon\leq s\leq t} |v^T*(X_{t-\epsilon\to s} - \mathrm{Id})|^q\right) + \varepsilon^{-q}\mathbb{E}\left(\sup_{t-\epsilon\leq s\leq t} |v^T*(X_{s\to t-\epsilon} - \mathrm{Id})|^q\right) \\
&= \varepsilon^{-q}\int \mathbb{E}\left(\sup_{t-\epsilon\leq s\leq t} |v^T*(X_{t-\epsilon\to s} - \mathrm{Id})|^q | X_{t-\epsilon} = z\right) d\mu(z) \\
&+ \varepsilon^{-q}\int \mathbb{E}\left(\sup_{t-\epsilon\leq s\leq t} |v^T*(X_{s\to t-\epsilon} - \mathrm{Id})|^q | X_{t-\epsilon} = z\right) d\mu(z) \\
&\leq C_q \epsilon^{\frac{q}{2}} e^{\beta_q \epsilon} \leq C_q \epsilon^{\frac{q}{2}} e^{\beta_q t}
\end{aligned}$$

for some strongly polynomial constants β_q and C_q. Here $\mu(\cdot)$ denotes the probability measure associated to the random variable $X_{t-\epsilon}$, and we have used the fact that the results in Proposition 3.2.1 involve only proper constants, and are therefore independent of the starting point assuming uniformity.

Pick $q = 3 + p$ Set

$$M_p = \min\left(\frac{1}{\beta_q}, \frac{1}{2}\right)$$

(this is a strongly polynomial constant). Now we have that for any $t \leq M_p$,

$$\mathbb{P}(\Omega) \leq C_q e \epsilon^{\frac{q}{2}}.$$

To summarise, for any $\epsilon \leq \epsilon_0 := t$, we have that for any $v \in \mathbb{R}^n$,

$$\mathbb{P}(v^T \Gamma v \leq \epsilon) \leq \mathbb{P}(\Omega) \leq C_q e \epsilon^{\frac{q}{2}}.$$

Using Lemma 3.3.1 over all unit $v \in \mathbb{R}^n$, and using the fact that $\mathbb{E}(|\Gamma|^{p+1})$ is a strongly polynomial constant, we get that

$$\mathbb{E}(\lambda_*(\Gamma)^{-p}) \leq K_p t^{-p}$$

for some strongly polynomial constant K_p, as expected. □

We now proceed to the proof of our estimate on the Malliavin covariance matrix in the weak Hörmander case. This proof requires some of the notation and facts of Part 2.

PROPOSITION 3.3.4. *Let (x, σ, F) be a uniformly (L, g, G)-tense ($g \geq 3$), uniformly (L, H_L)-weak Hörmander system. Let $\Gamma_t^{i,j}$ denote the Malliavin covariance matrix (of $Y_t = F(X_t)$) at time t. For any $1 \leq p$, there exist polynomial constants M_p and C_p, such that for any $t \leq M_p$,*

$$\mathbb{E}\left(\lambda_*(\Gamma)^{-p}\right) \leq C_p t^{-p 2^{4(L-1)+1}},$$

where for a symmetric matrix A, $\lambda_*(A)$ denotes the minimum eigenvalue of A.

PROOF (with constants depending polynomially on m and faster than exponentially on d). We fix $0 < \rho < 2^{-4L-1}$, and fix $\ell > 2^{4(L+1)} > \frac{2}{\rho}$.

Claim 1: Let as usual τ denote the truncation operation on multi-indices. Let

$$s \in \operatorname{span}_{\substack{\#(\alpha)\leq\ell \\ \alpha_{\#\alpha}\neq 0}} e^{\alpha}$$

be such that

$$\sum_{\substack{\#(\alpha)\leq\ell \\ \alpha_{\#\alpha}\neq 0}} (s^{\alpha})^2 \geq K$$

for some polynomial K. For any $p \geq 1$, there are polynomial constants C_p and D_p such that for any $\epsilon \leq D_p$, we have

$$\mathbb{P}\left(\int_0^{\epsilon^\rho} (\sum_{\substack{\#(\alpha)\leq \ell \\ \alpha \neq (0)}} s^\alpha W^{\tau(\alpha)})^2 ds \leq \epsilon \right) \leq C_p \epsilon^p.$$

Proof of the Claim 1.

The proof of this is similar to classical proofs of the Hörmander theorem. Sketch of proof: Define the sets

$$\begin{aligned}
E_j &= \left\{ \int_{t-\epsilon^\rho}^t \left(\sum_{\substack{\#(\alpha)\leq \ell \\ \alpha \neq (0)}} s^\alpha W^{\tau^{j+1}(\alpha)} \right)^2 ds \leq \epsilon^{m(j)} \right\}, \\
m(j) &= 2^{-4(j-1)}, \\
F &= E_1 \cap E_2 \cap \ldots \cap E_\ell, \\
\Omega_{\text{free}} &= \left\{ \sup_{s \in [t-\epsilon^\rho, t]} \sup_{\#(\alpha)\leq \ell} |W^\alpha| \geq 1 \right\}.
\end{aligned}$$

On the event $F \cap \Omega_{\text{free}}^c$, we have

$$(L+1)\epsilon^{m(L)} > \bar{K}\epsilon^\rho$$

where $\bar{K}$ is a polynomial constant depending linearly on K coming from Stratonovich-Itô transformations.

In other words, for $\epsilon \leq \epsilon_0 = \min\left(\left(\frac{H}{4(L+1)} \right)^{\frac{1}{m(L)-\rho}}, t^{\frac{1}{\rho}} \right)$, we have that $\Omega_{\text{free}}^c \cap F = \varnothing$. An application of the Norris Lemma 2.4.1 with $\nu = \frac{1}{2}, r = \frac{1}{18}, q = 15$ shows

$$\mathbb{P}(E_i \cap E_{i+1}^c \cap \Omega^c) \leq C e^{-\epsilon^{-\nu}},$$

for some polynomial C, as long as

$$\epsilon \leq \min\left((d+1)^{-(L+1)240\times 2^{4(L-1)}}, t^{\frac{1}{\rho}}, (d+1)^{-(L+1)240\times 2^{4(L-1)}} \right).$$

An application of Theorem 2.1.1 shows, that there exist polynomial $D_{\text{free},p}$ an $C_{\text{free},p}$ such that for any $\epsilon \leq D_{\text{free},p}$, we have

$$\mathbb{P}(\Omega_{\text{free}}) \leq C_{\text{free},p}\epsilon^p. \tag{3.3.1}$$

Claim 1 follows.

Claim 2: Let

$$\Omega^J = \{ \sup_{t-\epsilon^\rho \leq s \leq t} |X_{s\to t-\epsilon^\rho}|, |X_{t-\epsilon^\rho \to s}| \geq 1/2 \}.$$

There exist polynomial constants $C_{J,p}$ and $D_{J,p}$ such that for any $\epsilon \leq D_{J,p}$,

$$\mathbb{P}(\Omega^J \geq 1/2) \leq C_{J,p}\epsilon^p. \tag{3.3.2}$$

Proof of Claim 2. For any fixed unit vectors $u, v \in \mathbb{R}^m$, Theorem 2.3.1 and Markov's inequality ensures that (for some proper constants D_p, C_p and for all $\epsilon \leq D_p$), we have

$$\mathbb{P}\left(\sup_{t-\epsilon^\rho \leq s \leq t} |X^{u,v}_{s\to t-\epsilon^\rho}|, |X^{u,v}_{t-\epsilon^\rho \to s}| \geq \varepsilon \right) \leq C_p \epsilon^p.$$

The claim follows upon bounding the operator norm by the Frobenius norm and taking $\varepsilon = \frac{1}{2m^2}$.

We now continue with the Proof of Proposition 3.3.4.

Let us write the following backwards stochastic Taylor expansion:

$$
\begin{aligned}
X_{s\to 0}V(X_s) =& V(X_0) + \sum_{k=1}^{d}\int_0^s X_{u\to 0}[\sigma^k, V](X_u)dW_u^k \\
&+ \int_0^s X_{u\to 0}\left([\sigma^0, V](X_u) + \frac{1}{2}\sum_{k=1}^{d}[\sigma^k,[\sigma^k, V]](X_u)\right)du \\
&= V(X_0) + \sum_{k=1}^{d}\int_0^s X_{u\to 0}[\sigma^k, V](X_u)\circ dW_u^k \\
&+ \int_0^s X_{u\to 0}\left([\sigma^0, V](X_u)\right)ds
\end{aligned}
$$

and similarly,

$$
\begin{aligned}
X_{t-s\to 0}V(X_{t-s}) =& V(X_t) + \sum_{k=1}^{d}\int_t^{t-s} X_{u\to t}[\sigma^k, V](X_u)dW_u^k \\
&+ \int_t^{t-s} X_{u\to t}\left([\sigma^0, V](X_u) + \frac{1}{2}\sum_{k=1}^{d}[\sigma^k,[\sigma^k, V]](X_u)\right)du \\
&= V(X_t) + \sum_{k=1}^{d}\int_0^s X_{t-u\to t}[\sigma^k, V](X_u)d\tilde{W}_u^k \\
&+ \int_0^s X_{t-u\to t}\left([-\sigma^0, V](X_{t-u}) + \frac{1}{2}\sum_{k=1}^{d}[\sigma^k,[\sigma^k, V]](X_{t-u})\right)du \\
&= V(X_t) + \sum_{k=1}^{d}\int_0^s X_{t-u\to t}[\sigma^k, V](X_u)\circ d\tilde{W}_u^k \\
&+ \int_0^s X_{t-u\to t}\left([-\sigma^0, V](X_{t-u})\right)du,
\end{aligned}
$$

where $\tilde{W}_u = W_{t-u} - W_t$ (which is distributed like a Brownian motion).

Now, using the above iteratively, writing $\tilde{\sigma}^0 = -\sigma^0$ and $\tilde{\sigma}^i = \sigma^i$ for each $i \neq 0$, we obtain for any unit $v \in \mathbb{R}^n$,

$$
\begin{aligned}
(3.3.3) \qquad |v^T\Gamma v| \geq& \int_0^{\epsilon^\rho}\sum_{k=1}^{d}(v^T{*}X_{t-s\to t}\sigma^k(X_{t-s}))^2 ds \\
=& \int_0^{\epsilon^\rho}\sum_{k=1}^{d}\Bigg(v^T{*}\sigma^k(X_t) + \sum_{i=1}^{d}\int_0^s v^T X_{t-u\to t}\sigma^{[i,k]}(X_{t-u})d\tilde{W}_u^i \\
&+ \int_0^s v^T X_{t-u\to t}\big(\tilde{\sigma}^{[0,k]}(X_{t-u}) + \frac{1}{2}\sum_{i=1}^{d}\tilde{\sigma}^{[i,[i,k]]}(X_{t-u})\big)du\Bigg)^2 ds \\
\cdots& \\
\geq& \int_0^{\epsilon^\rho}\Bigg(\sum_{\substack{\alpha_{\#\alpha}\neq 0\\ \#(\alpha)\leq \ell}}\tilde{W}_s^{[\alpha]}v^T{*}\tilde{\sigma}^{[\alpha]}(X_t) \\
&+ \sum_{\substack{\alpha_{\#\alpha}\neq 0\\ \#(\alpha)=\ell+1}}\int_0^s\cdots\int v^T{*}X_{t-\boldsymbol{\cdot}\to t}\tilde{\sigma}^{[\alpha]}(X_{t-\boldsymbol{\cdot}})\circ d\tilde{W}_{\boldsymbol{\cdot}}^{[\tau(\alpha)]}\Bigg)^2 ds.
\end{aligned}
$$

We can now write:

$$\begin{aligned}
&\int_0^{\epsilon^\rho} \Bigg(\sum_{\substack{\alpha_{\#\alpha}\neq 0 \\ \#(\alpha)\leq \ell}} \tilde{W}_s^{[\alpha]} v^T * \tilde{\sigma}^{[\alpha]}(X_t) \\
&\quad + \sum_{\substack{\alpha_{\#\alpha}\neq 0 \\ \#(\alpha)=\ell+1}} \int_0^s \cdots \int v^T * X_{t-\bullet\to t} \tilde{\sigma}^{[\alpha]}(X_{t-\bullet}) \circ d\tilde{W}_\bullet^{[\tau(\alpha)]} \Bigg)^2 ds \\
&= \int_0^{\epsilon^\rho} \Bigg(\sum_{\substack{\alpha_{\#\alpha}\neq 0 \\ \#(\alpha)\leq \ell}} \tilde{W}_s^{[\alpha]} v^T * \tilde{\sigma}^{[\alpha]}(X_t) \Bigg)^2 ds \\
&\quad + \sum_{\substack{\#(\alpha)\leq\ell, \#(\beta)=\ell+1 \\ \alpha_{\#(\alpha)}, \beta_{\#(\beta)}\neq 0}} \int_0^{\epsilon^\rho} \Big(v^T * \tilde{\sigma}^{[\alpha]}(X_t) \tilde{W}_s^{[\tau(\alpha)]} \Big) \\
&\qquad \times \left(\int_0^s \cdots \int v^T * X_{t-\bullet\to t} \tilde{\sigma}^{[\beta]}(X_{t-\bullet}) d\tilde{W}_\bullet^{[\tau(\beta)]} \right) ds \\
&\quad + \sum_{\substack{\#(\alpha), \#(\beta)=\ell+1 \\ \alpha_{\#(\alpha)}, \beta_{\#(\beta)}\neq 0}} \int_0^{\epsilon^\rho} \left(\int_0^s \cdots \int v^T * X_{t-\bullet\to t} \tilde{\sigma}^{[\alpha]}(X_{t-\bullet}) d\tilde{W}_\bullet^{[\tau(\alpha)]} \right) \\
&\qquad \times \left(\int_0^s \cdots \int v^T * X_{t-\bullet\to t} \tilde{\sigma}^{[\beta]}(X_{t-\bullet}) d\tilde{W}_\bullet^{[\tau(\beta)]} \right) ds.
\end{aligned}$$

We now define, for any multi-indices α, β with $\alpha_{\#(\alpha)}, \beta_{\#(\beta)} \neq 0$,

$$\begin{aligned}
A_s^\alpha &= \tilde{W}_s^{[\tau(\alpha)]}, \\
B_s^\beta &= \left(\int_0^s \cdots \int v^T * X_{t-\bullet\to t} \tilde{\sigma}^{[\beta])}(X_{t-\bullet}) \circ d\tilde{W}_\bullet^{[\tau(\beta)]} \right), \\
a^\alpha(X_t) &= v^T * \tilde{\sigma}^{[\alpha]}(X_t).
\end{aligned}$$

With this notation, Eq. (3.3.3) can be rewritten as

$$\begin{aligned}
|v^T \Gamma v| \geq &\int_0^{\epsilon^\rho} \Bigg(\sum_{\substack{\alpha_{\#\alpha}\neq 0 \\ \#(\alpha)\leq \ell}} \tilde{W}_s^{[\alpha]} v^T * \tilde{\sigma}^{[\alpha]}(X_t) \Bigg)^2 ds \\
&+ \sum_{\substack{\#(\alpha)\leq\ell, \#(\beta)=\ell+1 \\ \alpha_{\#(\alpha)}, \beta_{\#(\beta)}\neq 0}} \int_0^{\epsilon^\rho} A_s^\alpha a^\alpha(X_t) B_s^\beta ds \\
&+ \sum_{\substack{\#(\alpha), \#(\beta)=\ell+1 \\ \alpha_{\#(\alpha)}, \beta_{\#(\beta)}\neq 0}} \int_0^{\epsilon^\rho} B_s^\alpha B_s^\beta ds.
\end{aligned} \tag{3.3.4}$$

Now, by Claim 1, for each unit $v \in \operatorname{span}_{\#(\alpha)\leq\ell} e^{[\alpha]}$ and for any Brownian motion W, for $K = \#(\{\alpha : \#(\alpha) \leq \ell\})$, we can find $C_{v,p}$ such that

$$\mathbb{P}\left(\sum_{\#(\alpha)\leq\ell} v^\alpha \int_0^{\epsilon^\rho} (W^{[\tau(\alpha)]})^2 \leq \epsilon \right) \leq C_{v,p} \epsilon^{p+1+2K}.$$

Let

$$\Omega_{\text{absolute}} = \left\{ \exists \text{ unit } v \in \operatorname{span}_{\#(\alpha)\leq\ell} e^{[\alpha]} : \sum_{\#(\alpha)\leq\ell} v^\alpha \int_0^{\epsilon^\rho} (\tilde{W}^{[\tau(\alpha)]})^2 \leq \epsilon \right\}.$$

Using Lemma 3.3.1 and the above with p being replaced by $p+1+2K$ where in the second expression, p is the required value of p in the proposition, we obtain that there is a proper constant C_p, such that

$$\mathbb{P}\left(\Omega_{\text{absolute},\epsilon}\right) \le C_{\text{absolute},p}\epsilon^p. \tag{3.3.5}$$

Note that we have

$$\Omega^c_{\text{absolute},K_{\text{absolute}}} \cap \left\{ \int_0^{\epsilon^\rho} \Bigg(\sum_{\substack{\alpha_{\#\alpha}\neq 0\\ \#(\alpha)\le \ell}} \tilde{W}^{[\alpha]}_s v^T {*}\tilde{\sigma}^{[\alpha]}(X_t)\Bigg)^2 ds \le 2\epsilon \right\} = \varnothing. \tag{3.3.6}$$

We now use Eq. (3.3.4) to obtain, for any ϵ smaller than a polynomial quantity, and for any unit $v \in \mathbb{R}^n$

$$\begin{aligned}
\mathbb{P}(|v^T\Gamma v| \le \epsilon) &\le \mathbb{P}\Bigg(\int_0^{\epsilon^\rho} \Bigg(\sum_{\substack{\alpha_{\#\alpha}\neq 0\\ \#(\alpha)\le \ell}} \tilde{W}^{[\alpha]}_s v^T {*}\tilde{\sigma}^{[\alpha]}(X_t)\Bigg)^2 ds \le 2\epsilon \Bigg) \\
&\quad + \mathbb{P}\Bigg(\sum_{\substack{\#(\alpha)\le\ell,\#(\beta)=\ell+1\\ \alpha_{\#(\alpha)},\beta_{\#(\beta)}\neq 0}} \int_0^{\epsilon^\rho} A^\alpha_s a^\alpha(X_t) B^\beta_s ds \\
&\quad + \sum_{\substack{\#(\alpha),\#(\beta)=\ell+1\\ \alpha_{\#(\alpha)},\beta_{\#(\beta)}\neq 0}} \int_0^{\epsilon^\rho} B^\alpha_s B^\beta_s ds \ge \epsilon \Bigg) \\
&\le +\mathbb{P}(\Omega_{\text{absolute},K_{\text{absolute}}}) + \mathbb{P}(\Omega_{\text{free}}) + \mathbb{P}((\Omega^J)) \\
&\quad + \mathbb{P}(\Omega^c_{\text{absolute},K_{\text{absolute}}} \cap \int_0^{\epsilon^\rho} \Bigg(\sum_{\substack{\alpha_{\#\alpha}\neq 0\\ \#(\alpha)\le \ell}} \tilde{W}^{[\alpha]}_s v^T {*}\tilde{\sigma}^{[\alpha]}(X_t)\Bigg)^2 ds \le 2\epsilon) \\
&\quad + \sum_{\substack{\#(\alpha)\le\ell,\#(\beta)=\ell+1\\ \alpha_{\#(\alpha)},\beta_{\#(\beta)}\neq 0}} \mathbb{P}\Bigg(\Omega^c_{\text{free}} \cap (\Omega^J)^c \cap \{ \int_0^{\epsilon^\rho} A^\alpha_s a^\alpha(X_t) B^\beta_s ds \ge \frac{\epsilon}{(d+1)^{2\ell}} \} \Bigg).
\end{aligned} \tag{3.3.7}$$

Using Theorem 2.1.1 and the fact that $\ell > \frac{2}{\rho}$, we have that

$$\mathbb{P}\Bigg(\Omega^c_{\text{free}} \cap (\Omega^J)^c \cap \{ \int_0^{\epsilon^\rho} A^\alpha_s a^\alpha(X_t) B^\beta_s ds \ge \frac{\epsilon}{(d+1)^{2\ell}} \} \Bigg) \le C_{\text{rest},p}\epsilon^p \tag{3.3.8}$$

for some polynomial $C_{\text{rest},p}$.

Finally, combining all the estimates (3.3.2), (3.3.1), (3.3.8), (3.3.5) and (3.3.6), we obtain that for

$$\epsilon \le \frac{\min(D_{J,p}, D_{\text{free},p}, D_{\text{rest},p}, D_{\text{absolute}})}{2(d+1)^{2\ell+2}},$$

we can continue the calculation (3.3.7) as follows:

$$\mathbb{P}(|v^T\Gamma v| \le (C_{J,p} + C_{\text{free},p} + C_{\text{rest},p} + C_{\text{absolute}})((d+1)^{2(\ell+1)}\epsilon)^p.$$

Since the above is valid for any $p \ge 1$ and any $\in \mathbb{N}$, we can apply it for p being $p+2n+1$ where p is the p from our proposition, so that we can apply Lemma 3.3.1. This concludes the proof. □

3.4. Bounds on the derivatives of the inverse Malliavin covariance matrix

PROPOSITION 3.4.1. *Let (x,σ,F) be a $(1,g,G)$-tense ($g \ge 3$), uniformly H-elliptic system. Let $\Gamma^{i,j}_t$ denote the Malliavin covariance matrix (of $Y_t = F(X_t)$) at time t. For any $1 \le p$ and any $N \le g-3$, there exist strongly polynomial constants $C_{p,N}$ such that for any $t \le M_p$, any $1 \le p$, and any $s_1, s_2, \dots, s_N$,*

$$\sup_{s_1,s_2,\dots s_N \le t} \mathbb{E}\left(|D^N_{s_1,s_2,\dots,s_N}\Gamma^{-1}|^p\right) \le C_p t^{-p},$$

where $|\cdot|$ denotes the operator norm on the space of linear maps from $\mathbb{R}^{m\otimes N}$ to $\mathbb{R}^n$.

PROOF. A similar limiting argument to the proof of Lemma 2.1.6 in [**49**] shows that we can write symbolically calculations such as $D(\Gamma^{-1}\Gamma) = D(\Gamma^{-1})\Gamma + \Gamma^{-1}D(\Gamma)$ even though Γ is only almost surely invertible.

Now, note that by Proposition 3.2.1, we have that

$$\sup_{s_1,\ldots,s_N} \mathbb{E}(|D^N_{s_1,\ldots,s_N}\Gamma|^p) \leq Ct^p$$

for some proper constant C. Next we write

$$D^N_{s_1,\ldots,s_N}(\Gamma^{-1}) = -\sum_{a\cup b=\{1,2,\ldots,N\}} D^{\#a}_{s_a}(\Gamma^{-1})D^{\#b}_{s_b}(\Gamma)\Gamma^{-1}.$$

The result follows by induction using Proposition 3.3.3 as the initial case. $\square$

PROPOSITION 3.4.2. *Let (x,σ,F) be a uniformly (L,g,G)-tense, uniformly (L,H_L)-weak Hörmander system. Let $\Gamma^{i,j}_t$ denote the Malliavin covariance matrix (of $Y_t = F(X_t)$) at time t. For any $1 \leq p$ and any $N \leq g-3$, there exist polynomial constants $C_{p,N}$ such that for any $t \leq M_p$, any $1 \leq p$, and any $s_1, s_2, \ldots, s_N$*

$$\sup_{s_1,s_2,\ldots s_N\leq t} \mathbb{E}\left(|D^N_{s_1,s_2,\ldots,s_N}\Gamma^{-1}|^p\right) \leq C_p t^{-p2^{4(L-1)+1}},$$

where $|\cdot|$ denotes the operator norm on the space of linear maps from $\mathbb{R}^{m\otimes N}$ to $\mathbb{R}^n$.

PROOF. The proof is almost exactly the same as that of Proposition 3.4.1, except we take Proposition 3.4.2 as the initial case. $\square$

3.5. Alternative characterisation of uniform hypoellipticity and tenseness

LEMMA 3.5.1 (Alternative characterisation of uniform hypoellipticity and tenseness). *Let $v_1,\ldots,v_d \in \mathbb{R}^n$ satisfy* $\operatorname{span}(v_i) = \mathbb{R}^n$. *Let*

$$\beta_1 = \inf_{|x|=1}(\sum_i \langle v_i, x\rangle^2) > 0,$$

$$\beta_2 = \sup_{|v|=1}\left(\inf_{\lambda:\sum_i \lambda_i v_i = v}(|\lambda|^2)\right).$$

Then we have

$$\beta_1\beta_2 \geq 1.$$

Furthermore, if

$$\beta_3 = \sup_{|x|=1}(\sum_i \langle v_i, x\rangle^2),$$

we have:

$$\beta_1\beta_2 \leq \beta_3.$$

PROOF. Fix any unit $x \in \mathbb{R}^n$. We can write $x = \sum_{i=1}^d \lambda_i v_i$ for some λ with $\sum_{i=1}^d |\lambda_i|^2 \leq \beta_2$. This gives

$$\begin{aligned} 1 = \langle x, x\rangle &= \sum_{i=1}^d \lambda_i \langle v_i, x\rangle \\ &\leq \sqrt{\sum_{i=1}^d |\lambda_i|^2}\sqrt{\sum_i \langle v_i, x\rangle^2} \leq \sqrt{\beta_2}\sqrt{\sum_i \langle v_i, x\rangle^2}. \end{aligned}$$

Since x was arbitrary, the first part of the result follows.

For the second part, let $M = (v_1, v_2, v \ldots, v_d)$ be the matrix formed by the vectors $v_1,\ldots,v_d$. Consider the Moore-Penrose pseudo-inverse of M, defined as $N = M^T(MM^T)^{-1}$. For any unit

$v \in \mathbb{R}^n$, we can define $\lambda = Nv$, which ensures $M\lambda = v$ as expected. The problem therefore boils down to proving that the operator norm of $P : \mathbb{R}^n \to \mathbb{R}^d,\ v \mapsto M^t(MM^T)^{-1}$ is bounded above by $\frac{\beta_3}{\beta_2}$. To show this, note that MM^T is symmetric, and therefore diagonalisable. Then we have that the smallest eigenvalue of MM^T must be larger than β_1. It follows that the operator norm of $v \mapsto (MM^T)^{-1}v$ is bounded above by $\frac{1}{\beta_1}$. Furthermore, the operator norm of $w \to M^T w$ is bounded above by β_3. The result follows. $\square$

CHAPTER 4

Proof of Euclidean Bounds

Here we piece together the results from the previous chapters to arrive at a proof of our Euclidean bounds.

4.1. Integration by parts

The following is an adaptation of classics of Malliavin calculus (cf. [**2**], [**49**]).

THEOREM 4.1.1. *Let F_t be a random variable in $\mathbb{D}^{g,2^{3n}P}(\mathbb{R}^n)$ (P being a fixed integer) and let G be a random variable in $\mathbb{D}^{g,2^{3n}P}([0,1])$. Suppose that E is an event such that $1_{E^c}G = 0$. Let Γ be the Malliavin covariance matrix of F. Writing $\lambda_*(\cdot)$ for the lowest eigenvalue of a symmetric matrix, suppose that we have $\lambda_*(\Gamma)^{-1}1_E \in \mathbb{D}^{g,2^{3n}}(\mathbb{R})$. because Γ is invertible on E, we can write down expressions such as $\Gamma^{-1}G$ etc without problems of definition.*

Suppose that for some parameter $0 < t < 1$ and for some constants $C, K > 0$, we have the following bounds on the operator norms of the derivatives of $F1_E$, G, $\Gamma^{-1}1_E$ etc: For all $N \leq g$, for some given $a > 0$, $g \geq n$, and for all $p \leq 2^{2n+1}P$,

$$
\begin{aligned}
|D^N(F1_E)|_p &\leq Ct^{\frac{N}{2}},\\
|D^N(LF1_E)|_p &\leq Ct^{\frac{N}{2}},\\
|LF1_E|_p &\leq Ct^{\frac{1}{2}},\\
|D^N\Gamma^{-1}1_E|_p &\leq Ct^{\frac{N}{2}-a},\\
|D^N G|_p &\leq Ct^{\frac{N}{2}},
\end{aligned}
$$

where $|\cdot|_p$ denotes L^p norm in expectations. For $i = 0, 1, \ldots$ and for given vectors $v_1, v_2, \ldots \in \mathbb{R}^m$, define the Malliavin weights $H_i(F,G)$ inductively by:

$$
\begin{aligned}
H_0(F,G) &= G,\\
H_1(F,G) &= \delta(\Gamma^{-1}DF^{v_1}G),\\
H_{i+1}(F,G) &= \delta(\Gamma^{-1}DF^{v_{i+1}}H_i(F,G)),
\end{aligned}
$$

where δ denotes the Skorohod integral, adjoint of the Malliavin derivative operator.

We have that there exists a constant K', i, j, depending only on K, and a constant C' depending on n, C, p, and polynomial in C, such that for any i, j with $i + j \leq g$, and for any $p \leq P$,

$$\mathbb{E}(|D^j H_i(F,G)|_p) \leq Kt^{\frac{j-i(2a-1)}{2}}\sqrt{\mathbb{P}(E)}.$$

PROOF. We proceed by induction.

For $i = 0$, the result follows immediately from the assumptions:

$$\mathbb{E}(|D^jG|^p) = \mathbb{E}(|D^jG|^p1_E^p) \leq \mathbb{E}(|D^jG|^{2p})\mathbb{E}(1_E^{2p}) \leq C^pt^{\frac{pj}{2}}\sqrt{\mathbb{P}(E)}.$$

Supposing the result true for fixed i, we prove the result for $i + 1$: We have first

$$
\begin{aligned}
H_{i+1}(F,G) &= \delta(\Gamma^{-1}DF^{v_{i+1}}H_i(F,G))\\
&= \langle DF^{v_{i+1}}, D\Gamma^{-1}\rangle H_i(F,G) + \Gamma^{-1}\langle DH_i(F,G), DF^{v_{i+1}}\rangle\\
&\quad + \Gamma^{-1}H_i(F,G)LF.
\end{aligned}
$$

Then for the derivatives, for $j \leq g - i - 1$.

$$\begin{aligned}D^j H_{i+1}(F,G) &= D^j \delta(\Gamma^{-1} D F^{v_{i+1}} H_i(F,G)) \\ &= \sum_{k+l+m=j} \langle D^{k+1} F^{v_{i+1}}, D^{l+1}\Gamma^{-1}\rangle D^m H_i(F,G) \\ &\quad + \sum_{k+l+m=j} D^m \Gamma^{-1} \langle D^{k+1} H_i(F,G), D^{l+1} F^{v_{i+1}}\rangle \\ &\quad + \sum_{k+l+m=j} D^m \Gamma^{-1} D^k H_i(F,G) D^l LF.\end{aligned}$$

Now we can calculate, where $C_{p,g}$ denotes a constant depending only on p, g.

$$\begin{aligned}&\mathbb{E}(D^j H_{i+1}(F,G)^p) \\ &\leq \sqrt{\mathbb{E}(D^j H_{i+1}(F,G)^p)}\sqrt{\mathbb{E}(1_E^{2p})} \\ &\leq \sqrt{\mathbb{P}(E)} C_{p,g} \sum_{k+l+m=j} \mathbb{E}(|D^{k+1} F^{v_{i+1}}|_{8p})^p \mathbb{E}(|D^{l+1}\Gamma^{-1}|_{8p})^p \mathbb{E}(|D^m H_i(F,G)|_{8p})^p \\ &\quad + \sqrt{\mathbb{P}(E)} C_{p,g} \sum_{k+l+m=j} \mathbb{E}(|D^m \Gamma^{-1}|_{8p})^p \mathbb{E}(|D^{k+1} H_i(F,G)|_{8p})^p \mathbb{E}(|D^{l+1} F^{v_{i+1}}|_{8p})^p \\ &\quad + \sqrt{\mathbb{P}(E)} C_{p,g} \sum_{k+l+m=j} \mathbb{E}(|D^m \Gamma^{-1}|_{8p})^p \mathbb{E}(|D^k H_i(F,G)|_{8p})^p \mathbb{E}(|D^l LF|_{8p})^p \\ &\leq C_{p,g}\sqrt{\mathbb{P}(E)} \sum_{k+l+m=j} Ct^{p\frac{k+1}{2}} Ct^{p\frac{l+1-2a}{2}} t^{p\frac{m-i(2a-1)}{2}} \\ &\quad + C_{p,g}\sqrt{\mathbb{P}(E)} \sum_{k+l+m=j} Ct^{p\frac{m-2a}{2}} Ct^{p\frac{k+1-i(2a-1)}{2}} t^{p\frac{l+1}{2}} \\ &\quad + C_{p,g}\sqrt{\mathbb{P}(E)} \sum_{k+l+m=j} Ct^{p\frac{m-2a}{2}} Ct^{p\frac{k-i(2a-1)}{2}} t^{p\frac{l+1}{2}} \\ &\leq \sqrt{\mathbb{P}(E)} 3g^3 C_{p,g} C^3 t^{p\frac{j-i(2a-1)}{2}},\end{aligned}$$

for some K_1, K_2, K_3 depending only on K, as required. □

4.2. Concentration inequality for Euclidean bounds.

Let $v = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$ we write $1_{[v,\infty)}(.)$ for the function such that $1_{v\leq}(y) = 1$ if and only if for $i = 1, 2, \ldots, n$, $|v_i| \leq y_i$ and $\mathrm{sign}(v_i) = \mathrm{sign}(y_i)$.

We have the following result:

PROPOSITION 4.2.1. *Let $\mathcal{A} = (x, \sigma, F)$ be a $(1,1,G)$-tense system. We have, for $t \leq 1$,*

$$\mathbb{E}([1_{[v,\infty)}(Y_t)]) \leq 2e \exp\left(\frac{-|{*x} - v|^2}{8tG}\right).$$

PROOF. Without loss of generality, $*x_i \leq v_i$ for all i. Let $E = \{i : \frac{|v_i - *x_i|}{2\sqrt{G}} \leq t\}$.

Note that

$$Y_t = x + \int_0^t *\sigma^0(X_s)ds + \sum_{i=1}^d \int_0^t *\sigma^i(X_s)dW_s^i$$

and

$$\left|\int_0^t *\sigma^0(X_s)ds\right| \leq \sqrt{G}t.$$

Then we have, by the classic exponential martingale inequality,

$$\begin{aligned}\mathbb{E}([1_{[v,\infty)}(Y_t)]) &\le 2\exp\left(-\sum_{i\notin E}\frac{|v_i - *x_i|^2}{8Gt}\right)\\ &\le 2\exp\left(-\sum_{i\notin E}\frac{|v_i - *x_i|^2}{8Gt}\right)\exp\left(-\sum_{i\in E}\frac{|v_i - *x_i|^2}{8Gt}\right)\exp(t\#(E))\\ &\le 2\exp\left(\frac{-|*x - v|^2}{8Gt}\right)e^t \le 2e\exp\left(\frac{-|*x - v|^2}{8Gt}\right).\end{aligned}$$

□

4.3. Euclidean Integrable upper bound for an Elliptic system

THEOREM 4.3.1. *Let (x, σ, F) be a system that is $(1, g, G)$ tense ($g \ge n + 3$) and H-elliptic uniformly in a compact $\mathcal{K}$. Suppose that G is a random variable in $\mathbb{D}^{g,2^{3n+1}P}([0, 1])$. Suppose that E is an event and $\mathcal{K}' \subset \mathcal{K}$ a compact, such that $1_{E^c}G = 0$ and $x \in \mathcal{K}'$ whenever E holds. If the distance between the exterior of $\mathcal{K}$ and the interior of $\mathcal{K}'$ is a proper constant, there exists a proper constants M and strongly polynomial constants C, D such that for any $t \le D$, Y_t admits a density $p_t(x, \cdot)$ and, the density of Y_t perturbed by G,*

$$\mathbb{E}(\delta(Y_t(y))G),$$

satisfies the following for any $N \le g - 3 - n$ and any unit $v_1, v_2, \ldots, v_N \in \mathbb{R}^n$:

$$\left|\frac{\partial^N p_t^G(x, y)}{\partial v_1 \partial v_2 \ldots \partial v_N}\right| \le C\frac{\exp\left(\frac{-|*x-y|^2}{Mt}\right)}{t^{\frac{n+N}{2}}}\sqrt{\mathbb{P}(E)}.$$

PROOF. The existence of the density follows immediately from the invertibility 3.3.3 of the Malliavin covariance matrix. By the usual limiting arguments, we can perform integration by parts symbolically with a delta function inside the expectation. Then we have immediately, using the Malliavin IBP formula with weights defined as in Theorem 4.1.1 with $v_1, \ldots v_N$ taken from the assumptions of the present theorem, and $v_{N+i} = e^i$ for $i = 1, 2, \ldots n$, where e^i are the Euclidean basis vectors:

$$\begin{aligned}|\frac{\partial^N p_t^G(x, y)}{\partial v_1 \partial v_2 \ldots \partial v_N}| &= |\frac{\partial^N \mathbb{E}(\delta_y(Y_t)G)}{\partial v_1 \partial v_2 \ldots \partial v_N}| = |\mathbb{E}((\partial^{v_1,\ldots v_N}\delta_y)(Y_t)G)|\\ &= |\mathbb{E}(\delta_y(Y_t)H_N(F, G))| = |\mathbb{E}(1_{[y,\infty)}(Y_t)(-1)^{\xi}H_{N+n}(F, G))|\\ &\le \sqrt{\mathbb{E}(1_{[y,\infty)}(Y_t)^2)}\sqrt{\mathbb{E}(H_{N+n}(F, G)^2)}.\end{aligned}$$

Here $\xi = \#(i : y_i < *x_i)$.

Note that the conditions of Theorem 4.1.1 are satisfied by using Propositions 3.2.1, 3.3.3 and 3.4.1 to $X_t\phi(\sup_{0\le s\le t}|X_t|)$ where ϕ is a smooth bump function equal to 1 on $\mathcal{K}'$ and 0 outside $\mathcal{K}$. We have $a = 1$, and C is a strongly polynomial constant. It follows that we can use Theorem 4.1.1, and Proposition 4.2.1, to write further:

$$\begin{aligned}\left|\frac{\partial^N p_t^G(x, y)}{\partial v_1 \partial v_2 \ldots \partial v_N}\right| &\le \sqrt{\mathbb{E}(1_{[y,\infty)}(Y_t)^2)}\sqrt{\mathbb{E}(H_{N+n}(F, G)^2)}\\ &\le \sqrt{2e}\exp(\frac{-|*x - v|^2}{16tG})Kt^{-\frac{n+N}{2}}\sqrt{\mathbb{P}(E)}\end{aligned}$$

which is the required inequality with $C = \sqrt{2e}K$ (a proper constant) and $M = 16G$ (a proper constant). □

In particular, setting $G = 1$, and $\mathcal{K} = \mathcal{K}' = \mathbb{R}^m$, we have the following result, which is the easiest of the five main results of this thesis:

THEOREM 4.3.2 (Integrable Euclidean upper bounds with strongly polynomial constants for uniformly elliptic system). *Let (x, σ, F) be a uniformly $(1, g, G)$ tense ($g \geq n+3$), uniformly H-elliptic system. There exist strongly polynomial constants C, M, D such that for any $t \leq D$, Y_t admits a density $p_t(x, \cdot)$ satisfying the following for any $N \leq g-3-n$ and any unit $v_1, v_2, \ldots, v_N \in \mathbb{R}^n$:*

$$|\frac{\partial^N p_t(x,y)}{\partial v_1 \partial v_2 \ldots \partial v_N}| \leq C \frac{\exp\left(\frac{-|*x-y|^2}{Mt}\right)}{t^{\frac{n+N}{2}}}.$$

Work is ongoing in providing a more careful proof of the following extension:

THEOREM 4.3.3 (Integrable Euclidean upper bounds with proper constants for uniformly elliptic system). *Define the strong tension of a (linear) system as*

$$G = \sup_{v \in \mathbb{R}^m, |v|=1} \sup_{o \leq g} \sum_{\substack{w \in (\mathbb{R}^m)^{\otimes o} \\ w^i \in B}} \sum_{k=0}^{d} \langle \frac{\partial^o \sigma^k}{\Pi \partial w^i}, v \rangle^2 + |dF|^2,$$

where B is an orthonormal basis of $(\mathbb{R}^m)^{\otimes o}$ (This definition only uses one 'order parameter' g rather than two order parameters L and g.) Let (x, σ, F) be a strongly uniformly (g, G) tense ($g \geq n+3$), uniformly H-elliptic system. There exist proper constants C, M, D such that for any $t \leq D$, Y_t admits a density $p_t(x, \cdot)$ satisfying the following for any $N \leq g-3-n$ and any unit $v_1, v_2, \ldots, v_N \in \mathbb{R}^n$:

$$|\frac{\partial^N p_t(x,y)}{\partial v_1 \partial v_2 \ldots \partial v_N}| \leq C \frac{\exp\left(\frac{-|*x-y|^2}{Mt}\right)}{t^{\frac{n+N}{2}}}.$$

SKETCH OF PROOF. First, verify that there exists a proper constant $\bar{G}$ such that for all $i \leq g$, the operator norm of the map

$$\partial^i \sigma : (\mathbb{R}^d)^{\otimes i} \to \mathbb{R}^m$$

(viewing $(\mathbb{R}^d)^{\otimes i}$ as a free id-dimensional vector space) is bounded above by $\bar{G}$.

Then we can apply Lemma 2.3.1 to obtain a version of Theorem 3.2.1 that controls the operator norm of $\langle D^i X_t, v \rangle_{\mathbb{R}^m}$ for any fixed unit $v \in \mathbb{R}^m$ and for similar quantities such as LX_t etc. with proper constants. Here the value of m used in Theorem 2.3.1 to control $\langle D^i X_t, v \rangle_{\mathbb{R}^m}$ for any fixed unit $v \in \mathbb{R}^m$ is d^i. Then, as usual, further calculations (application of Theorem 4.1.1) are all done directly in the target space.

□

REMARK 4.3.1. The 'mixed tension' defined below (cf. 7.1.12) is properly controlled by the strong tension as defined above, which means only one definition is required.

4.4. Euclidean upper bound for systems satisfying the weak Hörmander condition

THEOREM 4.4.1. *Let (x, σ, F) be a system that is (L, g, G) tense ($g \geq n+3$) and (L, H_L) weak Hörmander, uniformly in a compact $\mathcal{K}$. Suppose that G is a random variable in the space $\mathbb{D}^{g, 2^{3n+1}P}([0,1])$. Suppose that E is an event and $\mathcal{K}' \subset \mathcal{K}$ a compact, such that $1_{E^c} DG = 0$ and $x \in \mathcal{K}'$ whenever E holds. If the distance between the exterior of $\mathcal{K}$ and the interior of $\mathcal{K}'$ is a proper constant, there exist polynomial constants D, M and C such that for any $t \leq D$, Y_t admits a density $p_t(x, \cdot)$ and the density of Y_t perturbed by G,*

$$\mathbb{E}(\delta(Y_t(y))G),$$

satisfies the following estimate for any $N \leq g-3-n$ and any unit $v_1, v_2, \ldots, v_N \in \mathbb{R}^n$:

$$\left| \frac{\partial^N p_t^G(x,y)}{\partial v_1 \partial v_2 \ldots \partial v_N} \right| \leq C \frac{\exp\left(\frac{-|*x-y|^2}{Mt}\right)}{t^{(n+N)2^{4L}}} \sqrt{\mathbb{P}(E)}.$$

PROOF. With all the same remarks as in the proof of 4.3.1, we can write, after killing off the diffusion outside $\mathcal{K}$, applying Propositions 3.2.1, 4.1.1 etc. (C is still a proper constant but this time $a = 2^{4(L-1)+1}$),

$$\begin{aligned}|\frac{\partial^N p_t^G(x,y)}{\partial v_1 \partial v_2 \dots \partial v_N}| &\leq \sqrt{\mathbb{E}[1_{[y,\infty)}(Y_t)]^2}\sqrt{\mathbb{E}(H_{N+n}(F,G)^2)} \\ &\leq \sqrt{2e}\exp\left(\frac{-|*x-v|^2}{16tG}\right)Kt^{-\frac{(n+N)(2.2^{4(L-1)+1}-1)}{2}}\sqrt{\mathbb{P}(E)} \\ &\leq C\frac{\exp\left(\frac{-|*x-y|^2}{Mt}\right)}{t^{(n+N)2^{4L}}}\sqrt{\mathbb{P}(E)},\end{aligned}$$

for some proper constants C and M, as expected. □

Again, setting $G = 1$, $\mathcal{K} = \mathcal{K}' = \mathbb{R}^m$, we get the following global result:

THEOREM 4.4.2 (Euclidean upper bounds for weak Hörmander systems, with proper constants). *Let (x, σ, F) be a uniformly (L, g, G) tense ($g \geq n+3$), uniformly (L, H_L)-weak Hörmander system. There exist polynomial constants D, M and C such that for any $t \leq D$, Y_t admits a density $p_t(x, \cdot)$ satisfying the following estimate for any $N \leq g-3-n$ and any unit $v_1, v_2, \dots, v_N \in \mathbb{R}^n$:*

$$\left|\frac{\partial^N p_t(x,y)}{\partial v_1 \partial v_2 \dots \partial v_N}\right| \leq C\frac{\exp\left(\frac{-|*x-y|^2}{Mt}\right)}{t^{(n+N)2^{4L}}}.$$

4.5. Localisation

The aim of this section is to show that systems whose vector fields are well-behaved inside a compact set can be proved to have densities (with associated bounds etc.) on a smaller compact set. The arguments are classic (see [**36**]).

THEOREM 4.5.1. *Let (x, σ, F) be a system that is $(1, g, G)$ tense ($g \geq n+3$) and H-elliptic uniformly for $y \in \mathcal{K} = B(*x, R) \subset \mathbb{R}^n$ ($R > 0$). There exist strongly polynomial constants D, M and C (dependent on R) such that for any $t \leq D$, Y_t admits a density $p_t(x, \cdot)$ inside $B(*x, R/4)$ and, the density of Y_t satisfies the following estimate for any $N \leq g-3-n$ and any unit $v_1, v_2, \dots, v_N \in \mathbb{R}^n$:*

$$p_t(x,y) \leq C\frac{\exp(\frac{-|*x-y|^2}{Mt})}{t^{\frac{n}{2}}}.$$

PROOF. Let $\phi : \mathcal{K} \to [0,1]$ be a smooth bump function with derivatives bounded by some K such that $\phi(u) = 0$ for $u \notin B(x, R)$ and $\phi(u) = 1$ for $u \in B(0, 3R/4)$. K can be taken chosen as a proper constant. Consider the random variable $G_t = \phi(\sup_{0\leq s\leq t}|Y_t|)$. By Proposition 3.2.1, G is in $\mathbb{D}^{g-3,\infty}$ and for fixed p, the $\mathbb{D}^{g-3,p}$ norm of G is a proper constant.

Let E denote the boundary of the ball of radius $R/2$ around x. Consider the following stopping times:

$$\begin{aligned}\tau_1 &= \inf(s \in [0,t] : \exists s_1 \leq s : |Y_{s_1}| > 3R/4 \wedge |X_s| \leq R/2) \\ \tau_2 &= \inf(s \in [\tau_1, t] : \exists s_1 \in [\tau_1, s] : |Y_{s_1}| > 3R/4 \wedge |X_s| \leq R/2) \\ &\dots \\ \forall N,\ \tau_N &= \inf(s \in [\tau_{N-1}, t] : \exists s_1 \in [\tau_1, s] : |Y_{s_1}| > 3R/4 \wedge |Y_s| \leq R/2).\end{aligned}$$

We define the following random variable:

$$N = \inf(n : \tau_i = \infty) - 1.$$

Define also the following localizing random variables:

$$G_t^{\tau_i} = \phi\left(\sup_{\tau_i \leq s \leq t}|Y_s|\right).$$

We have the following calculation, where $\mathbb{E}_{z,s}$ denotes expectation conditionally given that $X_s = z$.

$$\begin{aligned} p_t(x,y) &= \mathbb{E}(\delta(y)1_{N=0}) + \mathbb{E}(\delta(y)1_{N=1}) + \dots \\ &\leq \mathbb{E}(\delta(y)G) + \mathbb{P}(N \geq 1) \sup_{\substack{z\in E \\ s\in[0,t]}} \mathbb{E}_{z,s}(G^{\tau_1}) + \mathbb{P}(N \geq 2) \sup_{\substack{z\in E \\ s\in[0,t]}} \mathbb{E}_{z,s}(G^{\tau_2}) + \dots \\ &\leq C \frac{\exp\left(-\frac{|y-*x|^2}{Mt}\right)}{t^n} (\mathbb{P}(N \geq 0) + \mathbb{P}(N \geq 1) + P(N \geq 2) + \dots). \end{aligned}$$

For some strongly polynomial constants C, M and for all $t \leq D$ for some strongly polynomial constant D. Here we have used the fact that for any $z \in E$ and any $y \in B(*x, R/4)$, $|z-y| \geq |*x-y|$, and D is picked so that $\exp\left(-\frac{R^2}{16Mt}\right)/t^n$ is increasing for $t \leq D$.

Now we have for all i,

$$\mathbb{P}(N \geq i) \leq \exp\left(-\frac{iR^2}{16Mt}\right)$$

for some proper constant M. Thus $(\mathbb{P}(N \geq 0) + \mathbb{P}(N \geq 1) + P(N \geq 2) + \dots)$ is bounded by a proper constant, and the result follows. $\square$

Using exactly the same method of proof, the result for weak Hörmander systems follows:

THEOREM 4.5.2. *Let (x, σ, F) be a system that is (L, g, G) tense ($g \geq n+3$) and (L, H_L)-weak Hörmander uniformly for $y \in \mathcal{K} = B(*x, R) \subset \mathbb{R}^n$ ($R > 0$). There exist polynomial (dependent on R) constants D, M and C such that for any $t \leq D$, Y_t admits a density $p_t(x, \cdot)$ inside $B(*x, R/4)$ and, the density of Y_t satisfies the following estimate for any $N \leq g - 3 - n$ and any unit $v_1, v_2, \dots, v_N \in \mathbb{R}^n$:*

$$p_t(x,y) \leq C \frac{\exp\left(\frac{-|*x-y|^2}{Mt}\right)}{t^{n2^{4L}}}.$$

Part 2

From Euclidean bounds to integrable control bounds via auxiliary objects.

CHAPTER 5

Properties of 'distances' defined as the pushforwards of homogeneous norms

In this section, inspired by the ideas in [**37, 39**], we introduce a general framework for studying certain objects similar to control distances. This framework will be the main tool used to go from Gaussian bounds on the densities of auxiliary objects satisfying a form of scaling to local control bounds on the densities of the Kusuoka-Stroock Taylor approximation. The much more advanced and broad theory of regularity structures (cf. [**26**]) also uses graded spaces to model Taylor expansions but our basic approach is designed purely for proving control-type bounds on the densities of certain SDE locally, in particular, we usually only consider the Taylor expansion at one point.

It is worth noting that for application to systems satisfying the Progressive Hörmander condition (i.e. some of the most interesting aspects of this thesis), it is enough to use the versions of the theorems below where F is assumed linear, which are also much less technical.

REMARK 5.0.1. The set of notation of this chapter is distinct from that of the rest of the thesis.

5.1. Models

Our setting in this subsection is as follows: Let $V = \mathbb{R}^n$ and $U = \bigoplus_{i=1}^{I} U_i$ be a finite dimensional space and let b_k^i be an orthonormal basis for U_i for all i. Let $\nu_i = \dim(U_i)$ and $\nu = \sum_{i=1}^{I} \nu_i$. On U, define the'homogeneous norms' by

$$|u|_{U,t} = |u|_U = \sqrt{t + \sum_{i=1}^{I} (|\operatorname{pr}_{U_i}(u)|)^{\frac{2}{i}}} = \sqrt{t + \sum_{i=1}^{I} (|u^i|)^{\frac{2}{i}}}.$$

In general, we will write u^i for $\operatorname{pr}_{U_i}(u)$. Let $F_t : [0,T] \times U \to V$ be a time dependent function with derivatives of any order in any unit direction being uniformly bounded by K. We suppose that the quantity β, defined by

$$\beta = \beta(0) = \det(JF_t(0)JF_t^T(0)),$$

is strictly positive. We call the triple (U, V, F_t) a *model*. We define the time-dependent 'homogeneous distance'[1] on V by $|v|_t = \inf_{F_t(u)=v} |u|_U$. We may use the notation $o(U^i) = i$ or $o(v) = i$ for $v \in U_i$.

REMARK 5.1.1. In applications, U is some suitable finite-dimensional or proper-dimensional snapshot of path space.

The main theorem will be Theorem 5.1.1 below.

LEMMA 5.1.2. *$\forall u_1, u_2 \in U$, we have that*

$$\frac{1}{2}(|u_1|_U^2 + |u_2|_U^2) \le |u_1 + u_2|_U^2 \le 4(|u_1|_U^2 + |u_2|_U^2).$$

[1]Note that this 'distance' is a peculiar object: if $F_t(0) \neq 0$ (which is the case when the original problem to model has non-trivial drift), it is possible to have $0 \notin B_{||_t}(0, \sqrt{t})$ (the 'centre' of $B_{||_t}(0, \sqrt{t})$ is $F_t(0)$ which varies with t.)

PROOF. *Upper bound.* First note that for $a, b \in R$, we have $|a+b|^{2/i} \le 2^{2/i}(|a|^{2/i} + |b|^{2/i})$. Indeed, without loss of generality, $b = 1$ and $a \in [0,1]$, then clearly $(1+a)^{2/i} \le 2^{2/i} \le 2^{2/i}(1^{2/i} + a^{2/i})$. Now, for $i \le I$, we have $2^{2/i} \le 4$. This allows us to conclude that

$$\begin{aligned}
|u_1 + u_2|_U^2 &\le t + \sum_i (\mathrm{pr}_{U^i}(u_1 + u_2))^{2/i} \\
&= t + \sum_i (\mathrm{pr}_{U^i}(u_1) + \mathrm{pr}_{U^i}(u_2))^{2/i} \\
&\le t + 4\sum_i (\mathrm{pr}_{U^i}(u_1)^{2/i} + \mathrm{pr}_{U^i}(u_2)^{2/i}) \\
&\le 4(|u_1|_U^2 + |u_2|_U^2).
\end{aligned}$$

Lower bound. Similarly, we have $|a+b|^{2/i} \ge \frac{1}{2}(|a|^{2/i} + |b|^{2/i})$: without loss of generality, $b = 1$ and $a \in [0,1]$, then we have $(1+a)^{2/i} \ge 1^{2/i} \ge \frac{1}{2}(1^{2/i} + a^{2/i})$, which allows to conclude

$$\begin{aligned}
(|u_1|_U^2 + |u_2|_U^2) &\le t + \sum_i (\mathrm{pr}_{U_i}(u_1) + \mathrm{pr}_{U_i}(u_2))^{2/i}) \\
&\le t + 2\sum_i (\mathrm{pr}_{U_i}(u_1))^{2/i} + (\mathrm{pr}_{U_i}(u_2))^{2/i}) \\
&= 2|u_1 + u_2|_U^2.
\end{aligned}$$

□

Let $D = dF_t|_0$ be the Jacobian of F_t at 0 and $\bar{F}_t$ be the first order Taylor approximation of F at 0 ($\bar{F}_t(u) = F_t(0) + Du$). Let $\Lambda = \sum_{i=1}^{I} \dim(U_i)$, $\nu = \sum_{i=1}^{I} i.\dim(U_i)$. Furthermore, let $e_1, e_2, \ldots, e_\kappa \in \mathrm{Ker}(D)$ ($\kappa = \Lambda - \dim(V)$), and $e_{\kappa+1}, \ldots, e_\Lambda \in U$ be such that $e_1, \ldots, e_\Lambda$ is an orthonormal basis for U, and $JF_t^T(0)JF_t(0)$, expressed with respect to that basis, is diagonal. We denote this new matrix by $\mathcal{M} = \mathrm{diag}(0, 0, \ldots, \lambda_1, \lambda_2, \ldots, \lambda_{\Lambda-\kappa})$.

We pick $0 < \varepsilon < \frac{1}{2}$ and fix it for the rest of this section. Let $\alpha = \frac{1}{1-\varepsilon}$.

Let $r = \min(1/2, \frac{\varepsilon}{K}, \frac{\beta}{4n!K^{2n}})$, and $\bar{\rho} = \frac{I^{-I/2}r^I}{3}$ (for future reference).

LEMMA 5.1.3. *For $u \in B(0,r)$, we have*

$$\det(JF_t(u)JF_t^T(u)) > \frac{\beta}{2}.$$

PROOF. This is a routine computation. Let

$$N(u)_{i,j} = \sum_k \frac{\partial F_i}{\partial e_k}\frac{\partial F_j}{\partial e_k}.$$

We have, for $u \in B(0,r)$, that $|N(u)| \le K^2$. By the Leibniz differential rule, we also get $|\frac{\partial N}{\partial \bar{u}}(u)| \le 2K^2$ for any unit vector $\bar{u}$. It follows that for any permutation σ, we have

$$\Pi_{i\neq j} N_{i\sigma(i)} \frac{\partial N_{j\sigma(j)}}{\partial \bar{u}} \le 2K^{2n}.$$

Then we have

$$\frac{\partial \det(N)}{\partial \bar{u}} \le 2n!K^{2n}.$$

Integrating on the straight line to u yields the result, since $r \le \frac{\beta}{4n!K^{2n}}$. □

PROPOSITION 5.1.4. *Define $\underline{\beta}(u) = \sqrt{\lambda_*(JF_t(u)JF_t^T(u))}$. We have $\underline{\beta} \ge \beta^{1/2n}$.*

PROOF. The claim follows from the fact that β is the product of the eigenvalues of

$$JF_t(u)JF_t^T(u)).$$

□

PROPOSITION 5.1.5. *For $u \in B(0,r)$, we have*

$$\beta(u) \leq K^{2n},$$

where by definition, $\beta(u) = \det(JF_t(u)JF_t^T(u))$.

PROOF. Observe that by Cauchy-Schwarz,

$$|(JF_t(0)J^T F_t(0))_{ij}| \leq K^2.$$

Now, the matrix $(JF_t(0)J^T F_t(0))_{ij}$ is symmetric and therefore diagonalisable by an orthonormal matrix. We can deduce that if λ_i are the eigenvalues, $\forall i$,

$$|\lambda_i| \leq K^2.$$

Now, note that we also have

$$\beta = \det(\mathcal{M}) = \prod_{i=1}^{n} \lambda_i.$$

From the two equations above, we can deduce that

$$\beta \leq K^{2n}$$

as expected. □

LEMMA 5.1.6. *For all $\xi \in U$ with $|\xi| + \sqrt{t} < 1$, we have*

$$|\xi| \leq |\xi|_U \leq (I+1)^{(I-1)/2I} \sqrt{|\xi|^2 + t}^{1/I} \leq \sqrt{2(I+1)}(|\xi|^{\frac{1}{I}} + t^{\frac{1}{2I}}).$$

PROOF. For the left hand side, we have

$$|\xi|^2 = \sum_{i=1}^{I} |\xi_i|^2 \leq \sum_{i=1}^{I} |\xi_i|^{2/i} \leq |\xi|_U.$$

For the right hand side, we have, by Jensen's inequality,

$$|\xi|_U^{2I} = (t + \sum_{i=1}^{I} |\xi_i|^{2/i})^I \leq (I+1)^{I-1}(t^I + \sum_{i=1}^{I} |\xi_i|^{2I/i}) \leq (I+1)^{I-1}(t + \sum_{i=1}^{I} |\xi_i|^2).$$

□

LEMMA 5.1.7. *For $|u| \in B(0,r)$, we have the following control on the operator norm of $D^u - D$, where D^u is the Jacobian of F_t at u*

$$|D^u - D|_2 \leq \varepsilon.$$

PROOF. We have for any unit $v \in U$

$$(D^u - D)v = \sum_{i=1}^{\Lambda} \int_0^1 \frac{\partial^2 F_t(us)}{\partial v \partial e_i} \langle u, e_i \rangle ds \leq |u|K \leq rK.$$

Since this is uniform over v, we can write:

$$|D^u - D|^2 \leq r^2 K^2 \leq \varepsilon^2,$$

from where the lemma follows. □

DEFINITION. We say that the model (U, V, F_t) is *regular* if there exist constants

$$0 < \rho_1, \bar{\rho}_1, \rho_2 < \rho_3, \quad 0 < \gamma, \Gamma$$

and functions

$$R_1, R_2 : B(0, \rho_3) \to U$$

such that $R_2 : F_t^{-1}(\{F_t(u)\}) \cap B(0, \rho_3) \to U$ is differentiable with a differentiable inverse and for any $u \in B(0, \rho_3)$,

$$\begin{aligned}
&F_0(R_2(u)) = F_0(0),\\
&F_t(R_1(u)) = F_t(u),\\
&F_t(u_1) = F_t(u_2) \implies R_1(u_1) = R_1(u_2) \quad \forall u_1, u_2 \in B(0, \rho),\\
&\gamma^{-1}|F_t(u)|_V^2 \le |R_1(u)|_U^2 \le \gamma|F_t(u)|_V^2,\\
&\gamma^{-1}(|R_1(u)|_U^2 + |R_2(u)|_U^2) \le |u|_U^2 \le \gamma(|R_1(u)|_U^2 + |R_2(u)|_U^2),\\
&\Gamma^{-1} m_{F_0^{-1}(\{0\})}(R_2(u)) \le m_{F_t^{-1}(\{F_t(u)\})}(u) \le \Gamma m_{F_0^{-1}(\{0\})}(R_2(u)) \quad \text{and}\\
&|u| + \sqrt{t} \le \rho_1 \quad \implies \quad |R_1(u)|_U, |R_2(u)|_U \le \rho_2,\\
&|R_1(u)|_U, |R_2(u)|_U \le \rho_2 \implies |u| \le \bar{\rho}_1,
\end{aligned}$$

where for a manifold M and a point x on M, $m_M(u)$ denotes the volume element on M at x.

DEFINITION. Let (U, V, F_t) be a regular model, for any $0 < \rho < \min(\rho_2, \bar{\rho})$, we define the following subset of U

$$E_t(\rho) = \{u \in U : |R_1(u)|_U < \rho,\ |R_4(u)|_U < \rho\}.$$

We also define the quantity:

$$M_t(M, \rho) = \int_{F_0^{-1}(\{F_0(0)\}) \cap E} \frac{\exp\left(-\frac{|u|_U^2}{Mt}\right)}{t^{\frac{\nu}{2}}}\, dm_{F_0^{-1}(\{F_0(0)\}}(u).$$

PROPOSITION 5.1.8. *Let (U, V, F_t) ($t \in [0, 1]$) be a model such that ($\forall t \in [0, 1], \forall u \in U$) $F_t u = \phi(t) + Du$ for some matrix D. The model (U, V, F_t) is regular.*

PROOF. Pick $\rho_3 = \infty$ and R_1 such that $|R_1(u)|_U = |F_t(u)|_V$. There isn't necessarily a unique choice, but there is one by compactness. Alternatively, consider the alternative homogeneous norm $||_{\tilde{U}}$ on U defined by

$$|u|_{\tilde{U}} = \left(t^I + \sum_{i=1}^{I} |\operatorname{Pr}_i(u)|^{2I/i}\right)^{\frac{1}{2I}}$$

and define R_1 by $|R_1(u)|_{\tilde{U}} = \inf_{F_t(\bar{u}) = F_t(u)} |\bar{u}|_{\tilde{U}}$. In this case, R_1 is defined uniquely by convexity. Then, let $R_2(u) = u - R_1(u)$. This yields the result with $\Gamma = 1$, $\gamma = 4$.[2] Indeed for the last inequalities, note that if $|R_1(u)|_U, |R_2(u)|_U \le \rho$, then $|u|_U^2 \le 4(|R_1(u)|_U^2 + |R_2(u)|_U^2) \le 4\rho^2$, which implies by Lemma 5.1.6 that $|u| \le 2\rho$. On the other hand if $|u|, \sqrt{t} \le \rho_2$ for some ρ_2, then

$$|R_1(u)| \le |R_1(u)|_U \le |u|_U \le 2\sqrt{2(I+1)}\rho_2^{1/I},$$

and then by the triangle inequality, $|R_2(u)| \le \rho_2 + 2\sqrt{2(I+1)}\rho_2^{1/I}$. Then applying Lemma 5.1.6 again yields

$$\begin{aligned}
|R_2(u)|_U &\le \sqrt{2(I+1)}\left((\rho_2 + 2\sqrt{2(I+1)}\rho_2^{1/I})^{1/I} + \rho_2^{1/I}\right)\\
&\le 3 \times 2(I+1)\rho_2^{1/I^2} = 6(I+1)\rho_2^{1/I^2}.
\end{aligned}$$

So we can pick $\bar{\rho}_1 = 2\rho$ and $\rho_2 = (\frac{\rho}{6(I+1)})^{I^2}$. □

PROPOSITION 5.1.9. *Let X_t be a random variable in E_ρ for some $\rho \le \rho_2/2$ with density p_t satisfying*

$$p_t(u) \le \frac{\operatorname{Id}_E C_2 \exp\left(-\frac{|u|_U^2}{M_2 t}\right)}{t^{\nu/2}},$$

[2] If taking the alternative definition of R_1, γ is a constant depending on the ν_i's.

resp.,

$$\frac{\mathrm{Id}_E\, C_1 \exp\left(-\frac{|u|_U^2}{M_1 t}\right)}{t^\nu} \le p_t(u).$$

Let $Y_t = F_t(X_t)$, we have that Y_t has a density $\bar{p}_t$ satisfying

$$\bar{p}_t(v) \le \mathrm{Id}_{|v|_t<\rho}\, M_t(0, M_2', 2\rho) C_2' \exp\left(-\frac{|v|_t^2}{M_2' t}\right),$$

resp.

$$\mathrm{Id}_{|v|_t<\rho}\, C_1' \exp\left(-\frac{|v|_t^2}{M_1' t}\right) M_t(0, M_1', \rho) \le \bar{p}_t(v),$$

with $C_1' = \Gamma^{-1} K^{-2n} C_1$, $M_1' = M_2 \gamma^{-2}$, $C_2' = \Gamma\sqrt{\frac{2}{\beta}} C_2$, and $M_2' = M_2\gamma^2$.

PROOF. *Upper bound.* By formula 2.2.1, we have that

$$\begin{aligned}
\bar{p}_t(v) &= \int_{u\in E, F_t(u)=v} \det(JF^T(u)JF(u))^{-\frac{1}{2}} p(u)\, dm_{F_t^{-1}(v)}(u) \\
&\le \int_{u\in E, F_t(u)=v} \sqrt{\frac{2}{\beta}} p(u)\, dm_{F_t^{-1}(v)}(u) \\
&\le \int_{u\in E, F_t(u)=v} \sqrt{\frac{2}{\beta}} \frac{C_2 e^{-\frac{|u|_U^2}{M_2 t}}}{t^{\frac{\nu}{2}}}\, dm_{F_t^{-1}(v)}(u) \\
&\le \int_{\bar{u}\in E, F_0(\bar{u})=F_0(0)} \Gamma\sqrt{\frac{2}{\beta}} e^{-\frac{|R_1(u)|_U^2}{\gamma M_2 t}} \frac{C_2 e^{-\frac{|\bar{u}|_U^2}{\gamma M_2 t}}}{t^{\frac{\nu}{2}}}\, dm_{F_0^{-1}(F_0(0))}(u) \\
&\qquad \text{(By the change of variable } \bar{u} = R_2(u)) \\
&\le \int_{\bar{u}\in E, F_0(\bar{u})=F_0(0)} \Gamma\sqrt{\frac{2}{\beta}} e^{-\frac{|v|_V^2}{\gamma^2 M_2 t}} \frac{C_2 e^{-\frac{|\bar{u}|_U^2}{\gamma M_2 t}}}{t^{\frac{\nu}{2}}}\, dm_{F_0^{-1}(F_0(0))}(u) \\
&\le \Gamma\sqrt{\frac{2}{\beta}} C_2 e^{-\frac{|v|_V^2}{\gamma^2 M_2 t}} M(\gamma^2 M, \rho)
\end{aligned}$$

as expected.

Lower bound. Similarly, by the formula 2.2.1, we have that By formula 2.2.1, we have that

$$\begin{aligned}
\bar{p}_t(v) &= \int_{u\in E, F_t(u)=v} \det(JF^T(u)JF(u))^{-\frac{1}{2}} p(u)\, dm_{F_t^{-1}(v)}(u) \\
&\ge \int_{u\in E, F_t(u)=v} K^{-2n} p(u)\, dm_{F_t^{-1}(v)}(u) \\
&\ge \int_{u\in E, F_t(u)=v} \sqrt{\frac{2}{\beta}} \frac{C_1 e^{-\frac{|u|_U^2}{M_2 t}}}{t^{\frac{\nu}{2}}}\, dm_{F_t^{-1}(v)}(u) \\
&\ge \int_{\bar{u}\in E, F_0(\bar{u})=F_0(0)} \Gamma^{-1}\sqrt{\frac{2}{\beta}} e^{-\frac{\gamma|R_1(u)|_U^2}{M_2 t}} \frac{C_1 e^{-\frac{\gamma|\bar{u}|_U^2}{M_2 t}}}{t^{\frac{\nu}{2}}}\, dm_{F_0^{-1}(F_0(0))}(u) \\
&\qquad \text{(By the change of variable } \bar{u} = R_2(u)) \\
&\ge \int_{\bar{u}\in E, F_0(\bar{u})=F_0(0)} \Gamma^{-1} K^{-2n} e^{-\frac{\gamma^2|v|_V^2}{M_2 t}} \frac{C_1 e^{-\frac{\gamma|\bar{u}|_U^2}{M_2 t}}}{t^{\frac{\nu}{2}}}\, dm_{F_0^{-1}(F_0(0))}(u) \\
&\ge \Gamma^{-1} K^{-2n} C_2 e^{-\frac{\gamma^2|v|_V^2}{M_2 t}} M(\gamma^{-2} M, \rho),
\end{aligned}$$

as expected. $\square$

PROPOSITION 5.1.10. *Let $b \le \rho^2/t$, $\bar{\rho} \le \rho/2$, there exist constants Δ_1 and Δ_2, depending only on $o, K, \beta, \nu_1, \nu_2, \dots, \nu_I, I, \rho, b, M, \gamma, \Gamma$ (but not on the specific form of the function F), such that $\forall u \in E$, we have $\forall t \le T$,*

$$\frac{\Delta_1}{|B_t(\sqrt{bt})|} \le M_t(0, M, \rho) \le \frac{\Delta_2}{|B_t(\sqrt{bt})|}.$$

PROOF. *Upper bound.* Let ξ be a random variable supported on $E(\rho)$ with density

$$\frac{Q}{t^{\frac{\nu}{2}}} \exp\left(-\frac{|\xi|_U^2}{M\gamma^2 t}\right) \mathrm{Id}_{\xi \in \bar{E}}$$

for some appropriately chosen (time dependent) constant Q. Here

$$\bar{E} = E \cap \{\xi \in U : |\xi|_U^2 \le bt\}.$$

Note that

$$Q \ge Q_1 := \left(\int_{\delta_{\sqrt{\frac{1}{M\gamma^2 t}}}(\bar{E})} e^{-|\xi|_U^2} d\xi\right)^{-1} \sqrt{\frac{1}{M\gamma^2}}^{\nu}.$$

Let $F_t(\xi) = \zeta$. Let $q_t(v)$ be the density of ζ. We have, by Proposition 5.1.9, for $|v|_t \le tb$ (recall $bt \le \rho^2$),

$$\begin{aligned}
q_t(v) &\ge C_1' \exp\left(\frac{-|v|_t^2}{Mt}\right) M_t(M, \rho) \\
&\ge Q_1 \Gamma^{-1} K^{-2n} \exp\left(\frac{-|v|_t^2}{Mt}\right) M_t(M, \rho) \\
&\ge \left(\int_{\delta_{\sqrt{\frac{1}{M\gamma^2 t}}}(\bar{E})} e^{-|\xi|_U^2} d\xi\right)^{-1} \sqrt{\frac{1}{M\gamma^2}}^{\nu} \Gamma^{-1} K^{-2n} C e^{-b/M} M_t(M, \rho).
\end{aligned}$$

Integrating the above inequality over the ball $B_t(bt)$ yields:

$$\left(\int_{\delta_{\sqrt{\frac{1}{M\gamma^2 t}}}(\bar{E})} e^{-|\xi|_U^2} d\xi\right)^{-1} \sqrt{\frac{1}{M\gamma^2}}^{\nu} \Gamma^{-1} K^{-2n} e^{-b/M} M_t(M, \rho) |B_t(\sqrt{bt})| \le 1.$$

It follows that

$$M_t(M, \rho) \le \frac{M^{\nu/2} \gamma^{\nu} \Gamma K^{2n} e^{b/M} \left(\int_{\delta_{\sqrt{\frac{1}{M\gamma^2 t}}}(\bar{E})} e^{-|\xi|_U^2} d\xi\right)}{|B_t(\sqrt{bt})|} = \frac{\Delta_2}{|B_t(\sqrt{bt})|}.$$

So we have the upper bound with $\Delta_2 = M^{\nu/2} \gamma^{\nu} \Gamma K^{2n} e^{b/M} \left(\int_{\delta_{\sqrt{\frac{1}{M\gamma^2 t}}}(\bar{E})} e^{-|\xi|_U^2} d\xi\right)$.

Lower bound. Let ξ be a random variable supported on E with density

$$\frac{Q}{t^{\frac{\nu}{2}}} \exp\left(-\frac{\gamma^2 |\xi|_U^2}{Mt}\right) \mathrm{Id}_{\xi \in \bar{E}}$$

for some appropriately chosen constant Q (Q clearly only depends on M, b, ν_i, Λ, E). Here

$$\bar{E} = E \cap \{\xi \in U : |\xi|_U^2 \le bt\}.$$

Let $F_t(\xi) = \zeta$. Let $q_t(v)$ be the density of ζ. Note that similarly to the upper bound,

$$Q \le Q_1 := \left(\int_{\delta_{\sqrt{\frac{\gamma^2}{Mt}}}(\bar{E})} e^{-|\xi|_U^2} d\xi\right)^{-1} \sqrt{\frac{\gamma^2}{M}}^{\nu}.$$

We have, by Proposition 5.1.9, for $|v|_t \le tb$,

$$\begin{aligned} q_t(v) &\le M_t(M,\rho)\, Q\sqrt{\frac{2}{\beta}}\Gamma \exp\left(-\frac{|v|_t^2}{Mt}\right) \\ &\le M_t(M,\rho)\, Q\sqrt{\frac{2}{\beta}}\Gamma. \end{aligned}$$

Now, we integrate the above inequality over the ball $B_t(0,\sqrt{bt})$ (note that by construction of $\bar{E}$, this is the whole support of $q_t(v)$, law of ζ). This yields:

$$1 \le M_t(M,\rho) Q\sqrt{\frac{2}{\beta}}\Gamma|B_t(\sqrt{bt})|.$$

From this, it follows that

$$M_t(M,\rho) \ge \frac{\Gamma^{-1}\sqrt{\frac{\beta}{2}}\frac{1}{Q}}{|B_t(\sqrt{bt})|} \ge \Gamma^{-1}\sqrt{\frac{\beta}{2}}\,\frac{\left(\int_{\delta_{\sqrt{\frac{\gamma^2}{Mt}}}(\bar{E})} e^{-|\xi|_U^2} d\xi\right)\sqrt{\frac{M}{\gamma^2}}^{\nu}}{|B_t(\sqrt{bt})|}.$$

We therefore have the lower bound with

$$\Delta_1 = \Gamma^{-1}\sqrt{\frac{\beta}{2}}\left(\int_{\delta_{\sqrt{\frac{\gamma^2}{Mt}}}(\bar{E})} e^{-|\xi|_U^2} d\xi\right)\sqrt{\frac{M}{\gamma^2}}^{\nu}. \qquad \square$$

We are now in a position to prove the following key theorem:

THEOREM 5.1.1. *Let (U,V,F_t) be a regular model. Let X_t^1 a stochastic process in U, and let $G_t \in [0,1]$ be a localising random variable.*

Assume that, for some constants C, M, β, ϵ (not dependent on t), and for any $t \le 1$,

(1) $|G_t|_{n+3,2^{3n+1}}, |X_t^1|_{n+3,2^{3n+1}} \le C$*;*
(2) X_t^1*, localised by G_t, has a density that satisfies*

$$p_{X_t^1}^{\phi(G_t)}(u) = \mathbb{E}(\delta_u(X_t^1)\phi(G_t)) \le \frac{Ce^{-\frac{|u|_U^2}{Mt}}}{t^{\frac{\nu}{2}}},$$

where ν is the homogeneous dimension of U;

(3) $\mathbb{P}\left(\sup_{s\le t}(|X_t^1|_U) \ge r\right) \le e^{-\frac{Mr}{t}} \quad (\forall r, t > 0)$*;*
(4) $\mathbb{P}(|G_t| \ge 1) \le Ce^{-\frac{1}{Mt}}, \quad \forall t \le T$*;*
(5) $\det(|JF_t(0)J^T F_t(0)|) \ge \beta$*, $\forall t \le T$, where T in the exponent denotes matrix transposition.*

Let $X_t^2 = F(X_t^1)$. There exist constants $C', C_1', C_2', \omega, M', \rho$, depending only on $o, K, C, M, \beta, \nu_i, \gamma, \Gamma$, and such that C_2' doesn't depend on M, and a time dependent localising random variable $0 \le \bar{G}_t \le 1$, such that that X_t^2, localised by $\bar{G}_t$, admits a density, and

(1) $\mathbb{E}(1-\bar{G}) \le C'e^{-\frac{\omega}{t}}$ *and* $\forall |v|_t \le \rho$,
(2) $\mathbb{E}(\delta(X_t^2 = v)\bar{G}_t) \le \frac{C_1' \exp\left(-\frac{|v|_t^2}{M't}\right)}{|B_{||_t}(0,\sqrt{t})|}$,
(3) $\mathbb{E}(\delta(X_t^2 = v)\bar{G}_t) \le \frac{C_2' \exp\left(-\frac{|v|_t^2}{M't}\right)}{|B_{||_t}(0,\sqrt{tM'})|}$.

If M is an absolute constant, we can choose the second formula. If the dependence of M on other parameters is difficult to control, we can chose the last formula for a bound that is properly locally integrable in space-time even if M is not proper.

PROOF. First, let $\bar{G}_t = G_t\phi(\frac{4X_t^1}{\rho_1})$ with E_ρ defined as above. Note that $\bar{G}_t = 0$ whenever $X_t^1 \notin E_\rho$ By by the fourth and fifth assumptions, as well as the regularity of the model, we have that

$$\mathbb{P}(\bar{G}_t = 0) \leq Ce^{-\frac{1}{Mt}} + Ce^{-\frac{\rho_1}{Mt}} \leq 2Ce^{-\frac{\rho_1}{Mt}} = C'e^{-\frac{\omega}{t}}$$

with $C' = 2C$ and $\omega = \rho_1/M$, as required. Now, recall that by assumption

$$\mathbb{E}(\bar{G}_t\delta(X_t^1 = \tilde{u})) \leq \mathbb{E}(\phi(G)\delta(X_t^1 = \tilde{u})) \leq C\frac{e^{-\frac{|\tilde{u}|_U^2}{Mt}}}{t^{\nu/2}}.$$

Now, by Proposition 5.1.9, we have that

$$\begin{aligned}\mathbb{E}(\bar{G}\delta(X_t^2 = v)) &\leq M_t(M_2', \rho)C_2' \exp\left(-\frac{|\tilde{u}|_U^2}{M_2't}\right)\\ &= M_t(M\gamma^{-2}, \rho)C\Gamma\sqrt{\frac{2}{\beta}} \exp\left(-\frac{\gamma^2|\tilde{u}|_U^2}{Mt}\right).\end{aligned}$$

Then, by Proposition 5.1.10, we have

$$\begin{aligned}\mathbb{E}(\bar{G}\delta(X_t^2 = v)) &\leq M_t(M\gamma^{-2}, \rho)C\Gamma\sqrt{\frac{2}{\beta}}e^{-\frac{\gamma^2|\tilde{u}|_U^2}{Mt}}\\ &\leq C\Gamma\sqrt{\frac{2}{\beta}}e^{-\frac{\gamma^2|\tilde{u}|_U^2}{Mt}}\frac{\Delta_2(M\gamma^{-2})}{|B_t(\sqrt{bt})|}.\end{aligned}$$

Hence we have on the one hand (for $b = 1$)

$$\mathbb{E}(\bar{G}\delta(X_t^2 = v)) \leq C\Gamma\sqrt{\frac{2}{\beta}}e^{-\frac{\gamma^2|\tilde{u}|_U^2}{Mt}}\frac{M^{\nu/2}\Gamma K^{2n}e^{\gamma^2/M}\left(\int_{\delta_{\sqrt{\frac{\gamma^2}{Mt}}}(\bar{E})} e^{-|\xi|_U^2}d\xi\right)}{|B_t(\sqrt{t})|},$$

and on the other hand ($b = M/\gamma^2$)

$$\mathbb{E}(\bar{G}\delta(X_t^2 = v)) \leq C\Gamma\sqrt{\frac{2}{\beta}}e^{-\frac{\gamma^2|\tilde{u}|_U^2}{Mt}}\frac{M^{\nu/2}\Gamma K^{2n}e\left(\int_{\delta_{\sqrt{\frac{\gamma^2}{Mt}}}(\bar{E})} e^{-|\xi|_U^2}d\xi\right)}{|B_t(\sqrt{Mt/\gamma^2})|}$$

as expected. □

Furthermore, note that we also have a similar results for lower bounds (we don't worry about dependence of the constants on anything here):

THEOREM 5.1.2. *Let (U, V, F_t) be a regular model. Let X_t^1 and G_t be two stochastic processes in U. Let δ_s be the family of dilations defined by*

$$\forall u \in U, \quad \mathrm{pr}_{U_i}(\delta_s(u)) = s^i\,\mathrm{pr}_{U_i}(u),\ \delta_s(t, u) = (ts, \delta_s(u)).$$

Let $\phi : U \to \mathbb{R}^+$ be a localising function such that $\phi(u) = 0$ whenever $|u| \geq 2$ and $\phi(u) = 1$ whenever $|u| \leq 1$ Assume that, for some constants C, M, β, ϵ (not dependent on t), and for any $t \leq 1$, we have:

(1) $|\phi(G)|_{n+3,2^{3n+1}}, |X_t^1|_{n+3,2^{3n+1}} \leq C$;
(2) X_t^1*, localised by $\phi(G_t)$ has a density that satisfies* $p_{X_t^1}^{\phi(G_t)}(u) \geq Ce^{-\frac{M|u|_U^2}{t}}$;
(3) $\mathbb{P}(\sup_{s\leq t}(|X_t^1|_U) \geq r\forall i \leq C)e^{-\frac{Mr}{t}}$ $(\forall r, t > 0)$;
(4) $\mathbb{P}(|G| \geq 1) \leq Ce^{-\frac{M}{t}}, \forall t \leq T$;
(5) $\det(|JF_t(0)J^TF_t(0)|) \geq \beta, \forall t \leq T$*, where T in the exponent denotes matrix transposition.*

Let $X_t^2 = F(X_t^1)$. There exist constants $C', C_1', \omega, M', \rho$ depending only on o, K, C, M, β, ν_i, Γ, γ, and a time dependent random variable $\bar{G}_t$, such that that $X_t^2\bar{G}_t$ admits a density and

(1) $\mathbb{E}(1-\phi(G)) \leq C'e^{-\frac{\omega}{t}}$ *and* $\forall |v|_t \leq \rho$

(2) $\mathbb{E}(\delta(X_t^2 = v)\bar{G}_t) \geq \frac{C_1' e^{-\frac{M'|v|_t^2}{t}}}{|B_{||_t}(0,\sqrt{t})|}$.

PROOF. First, like in the proof of the upper bound, let $\bar{G}_t = \phi(G_t)\phi(\frac{4X_t^1}{\bar{\rho}_1})$. By the fourth and fifth assumptions, we have that

$$\mathbb{P}(\bar{G}) \leq Ce^{-\frac{1}{Mt}} + Ce^{-\frac{\bar{\rho}_1}{Mt}} \leq C'e^{-\frac{\rho_1}{Mt}}.$$

Then, still similarly to the proof of the upper bound, recall that our density satisfies

$$\mathbb{E}(\bar{G}\delta(X_t^1 = \tilde{u})) \geq C\frac{e^{-\frac{|\tilde{u}|_U^2}{Mt}}}{t^{\nu/2}}.$$

Now, by Proposition 5.1.9, we have that

$$\begin{aligned}
\mathbb{E}(\bar{G}\delta(X_t^2 = v)) &\geq M_t(0, M_1', \rho)C_1'\frac{e^{-\frac{|\tilde{u}|_U^2}{M_1' t}}}{t^{\nu}} \\
&= M_t(0, M_1\gamma^{-2}, \rho)C\Gamma^{-1}K^{-2n}e^{-\frac{\gamma^2|\tilde{u}|_U^2}{M_1 t}} \\
&\geq \frac{\Delta_1(M_1')}{|B_t(\sqrt{bt})|}C\Gamma^{-1}K^{-2n}e^{-\frac{\gamma^2|\tilde{u}|_U^2}{M_1 t}}, \qquad \text{by Proposition 5.1.10.}
\end{aligned}$$

Plugging in $b = 1$, we obtain the required result. □

We finish this subsection with the following doubling conditions. One important remark is that the constant appearing in the theorem below only depends on other constants involved in the definition of the problem. In particular, when we apply this to the situation where our point models are modeling control distances, we will immediately get uniformity with respect to the centre.

PROPOSITION 5.1.11. *Let (U, V, F) be a regular model satisfying the usual conditions. There exist constants $D, \Delta \geq 0$, depending only on $n, \beta, T, K, I, \nu_i, n, \Lambda, \gamma, \Gamma$ such that $\forall d \leq \Delta, t \leq T$.*

$$|B_t(2d)| \leq D|B_t(d)|$$

PROOF. From Proposition 5.1.10, we set $\Delta = \frac{\rho}{2}$. Apply Proposition 5.1.10 with, on the right, $b = \frac{4d^2}{t}$, and on the left, $b = \frac{d^2}{t}$, we obtain that for $d \leq \rho/2$ and for any M,

$$|B_t(2d)| \leq |B_t(d)|\frac{\Delta_2^{b=\frac{4d^2}{t}}}{\Delta_1^{b=\frac{d^2}{t}}} = D|B_t(d)|.$$

Replacing the values of $\Delta_2^{b=\frac{4d^2}{t}}$ and $\Delta_1^{b=\frac{d^2}{t}}$, we obtain:

$$D = \frac{M^{\nu/2}\gamma^{\nu}\Gamma K^{2n}e^{4d^2/tM}\left(\int_{\delta_{\sqrt{\frac{1}{M\gamma^2 t}}}(\bar{E})} e^{-|\xi|_U^2}d\xi\right)}{\Gamma^{-1}\sqrt{\frac{\beta}{2}}\left(\int_{\delta_{\sqrt{\frac{\gamma^2}{Mt}}}(\bar{E})} e^{-|\xi|_U^2}d\xi\right)\sqrt{\frac{M}{\gamma^2}}^{\nu}}.$$

We see that a good choice for M would be $M = \frac{d^2}{t}$. This leaves us with, for b smaller than a constant only depending on ρ,

$$\begin{aligned}D =& \frac{\gamma^{2\nu}\Gamma^2 K^{2n} e^4 \left(\int_{\delta_{1/d\gamma}(\bar{E})} e^{-|\xi|_U^2} d\xi\right)}{\sqrt{\frac{\beta}{2}}\left(\int_{\delta_{\gamma/d}(\bar{E})} e^{-|\xi|_U^2} d\xi\right)} \\ &\leq \gamma^{2\nu}\Gamma^2 K^{2n} e^4 \sqrt{\frac{2}{\beta}} \frac{\left(\int_{|\xi|_U \leq 2/\gamma} e^{-|\xi|_U^2} d\xi\right)}{\left(\int_{|\xi|_U \leq \gamma} e^{-|\xi|_U^2} d\xi\right)}.\end{aligned}$$

□

We also have the following result, also providing an alternative proof of doubling conditions.

LEMMA 5.1.12. *Let (U, V, F) be a model satisfying the usual conditions. Let $(U, V, \tilde{F}) = (U, V, JF)$ be the model obtained by replacing F by its first order Taylor approximation JF. Denote by $||_{V,t}$ and $||_{\tilde{V},t}$ the corresponding 'distances' on V. There exist constants κ and δ, depending only on $I, n, \beta, K, \nu_i, \gamma, \Gamma$, such that for any $s, t \leq \delta$,*

$$\kappa^{-1}|B_{\tilde{V}}(s)| \leq |B_V(s)| \leq \kappa |B_{\tilde{V}}(s)|,$$

where $|B_{\tilde{V}}(s)|$ denotes the volume (in V) of the ball of radius s with respect to the 'metric' $||_{V,t}$.

PROOF. Suppose $x, y \in B_U(0, R)$ for some $R < \rho/4$. Let $z = \mathrm{Pr}_{\mathrm{Ker}\,\tilde{F}}(y) + \mathrm{Pr}_{\mathrm{Ker}\,\tilde{F}^\perp}(x)$. We have first, by Lemma 5.1.2, $R^2 \geq |x|_U^2 \geq \frac{1}{2}(|\mathrm{Pr}_{\mathrm{Ker}\,\tilde{F}^\perp}(x)|_U^2 + |\mathrm{Pr}_{\mathrm{Ker}\,\tilde{F}}(x)|_U^2)$, and therefore $|\mathrm{Pr}_{\mathrm{Ker}\,\tilde{F}^\perp}(x)|_U^2 \leq 2R^2$. Similarly $|\mathrm{Pr}_{\mathrm{Ker}\,\tilde{F}}(y)|_U^2 \leq 2R^2$. and finally, another application of Lemma 5.1.2 ensures that $|z|_U^2 \leq 4(2R^2 + 2R^2) = 8R^2$.

It follows that

$$\mathrm{Pr}_{\mathrm{Pr}_{\ker \tilde{F}^\perp}(x) + \ker \tilde{F}}(F^{-1}(\{F(x)\}) \cap B_U(0, R)) \subset \tilde{F}^{-1}(\{\tilde{F}(x)\}) \cap B_U(0, 4R)$$

Then we have

$$\mathrm{Vol}(F^{-1}(\{F(x)\}) \cap B_U(0, R)) \leq \Gamma\,\mathrm{Vol}(\tilde{F}^{-1}(\{\tilde{F}(x)\}) \cap B_U(0, 4R))$$

Now, the disintegration formula 2.2.1 ensures:

$$\begin{aligned}\mathrm{Vol}(B_V(0, R)) &= \int_{x \in B_U(0,R)} \frac{\sqrt{JF_x JF_x^T}}{\mathrm{Vol}(F^{-1}(\{F(x)\}) \cap B_U(0, R))} dx \\ &\geq \frac{\sqrt{\beta/2}}{\Gamma} \int_{x \in B_U(0,R)} \frac{1}{\mathrm{Vol}(\tilde{F}^{-1}(\{\tilde{F}(x)\}) \cap B_U(0, 4R))} dx \\ &\geq \frac{\sqrt{\beta/2}}{K^n \Gamma} \int_{x \in B_U(0,R)} \frac{\sqrt{\tilde{F}\tilde{F}^T}}{\mathrm{Vol}(\tilde{F}^{-1}(\{\tilde{F}(x)\}) \cap B_U(0, 4R))} dx \\ &= \frac{4^\nu \sqrt{\beta/2}}{K^n \Gamma} \mathrm{Vol}(B_{\tilde{V}}(0, R)).\end{aligned}$$

The other inequality is proved analogously.

An alternative proof is to use Proposition 5.1.10 and the fact that the constants only depend on the constants $\beta, \gamma, \ldots$ from the assumptions, which can be left unchanged after switching to the distance $||_V$. □

5.2. Equivalences between models

The next lemma begins to hint at the link between the quantity β involved in Lemma 2.2.1 and the hypoellipticity constant of a system, usually denoted β, but in this section, denoted $\underline{\beta}$ to distinguish it from the former.

LEMMA 5.2.1. *Let $F : \mathbb{R}^m \to \mathbb{R}^n$ be a function with C^o constant bounded by K. For $x \in \mathbb{R}^m$, let $\beta_F(x) = \det(JF(x)J^T F(x))$ and let*

$$\underline{\beta} = \sqrt{\inf_{|w|=1,\, w\in\mathbb{R}^n} |wJF|^2}$$

($\underline{\beta}$ corresponds to some sort of 'hypoellipticity constant'). We have the following inequality:

$$K \geq \underline{\beta} \geq \frac{\sqrt{\beta}}{K^{(n-1)}}$$

PROOF. The left hand side is trivial.

For the right hand side, begin by observing that the matrix $JF(x)J^T F(x)$ is symmetric, and therefore diagonalisable by an orthonormal matrix. Now, let us write the eigenvalues $\lambda_1, \dots, \lambda_n$. Since the diagonalising matrix is orthogonal, we have that $|\lambda_i| \leq K^2$ for each i. Since $\beta = \Pi_i \lambda_i$ and $\underline{\beta}^2 = \inf_i |\lambda_i|$, the lemma follows. □

Recall also that $\frac{1}{\sqrt{\underline{\beta}}}$ is an upper bound on the operator norm of any inverse function to the first order expansion of F.

The following proposition is a simpler but more explicit version of the implicit functions theorem. We include it both because it is for our problems to know on what quantities the radii involved depend (we eventually need a parametrised version of it), and because we want to highlight the beautiful similarities between this theorem the fact that Hörmander's condition implies the existence of a path between two points.

PROPOSITION 5.2.2 (Inverse functions theorem.). *Let $F : \mathbb{R}^m \to \mathbb{R}^n$ be a function with all derivatives bounded in operator norm by K and $\beta_F(0) = \det(JF(0)JF^T(o)) > 0$. Let $0 < \varepsilon < (\beta/2)^{1/2n}/2$, where as usual, $\underline{\beta}$ is the square root of the smallest eigenvalue of $JF(0)JF^T(0)$. Let*

$$\rho_1 = \frac{1}{2}\min\left(\frac{\varepsilon}{K}\,\frac{\beta}{4n!K^{2n}}\right).$$

Let $\rho_2 = (\beta/2)^{1/2n}\rho_1$. For any $y \in \mathbb{R}^m$ such that $|y - F(0)| \leq \rho_2$, there exists an $x \in \mathbb{R}^m$ such that $F(x) = y$.

Furthermore, $|x| \leq \frac{4|y|}{(\beta/2)^{1/2n}}$, and if $m = n$, f has a uniquely defined and differentiable inverse $f^{-1} : B(0, \rho_2) \to B(0, \rho_1)$ with derivative satisfying

$$Df^{-1}(y) = (Df(f^{-1}(y))^{-1}.$$

In particular, iteratively differentiating the above equation shows that F^{-1} is smooth.

PROOF. Similarly to the above subsections, we observe that $\forall x \in B(0, 2\rho_1)$, we have:

$$|DF_x - DF_0| \leq \varepsilon$$

and

$$\beta(x) \geq \beta(0) - \sum_{\sigma \text{ permutation of } \{1,\dots,n\}} 2(K^2)^{n-1}(K^2) \geq \frac{\beta(0)}{2}.$$

Now, we fix $D = DF_0$ (an $n \times m$ matrix)[3] we can find $x_1 \in \mathbb{R}^m$ with $Dx_1 = y$ and $|x_1| \leq \frac{|y|}{\underline{\beta}(0)}$. We define the x_i's iteratively by

$$\begin{aligned}
&x_0 = 0,\\
&D(x_i - x_{i-1}) = y - F(x_{i-1}), \quad \text{and}\\
&|x_i - x_{i-1}| \leq \frac{|F(y) - F(x_{i-1})|}{(\beta/2)^{1/2n}}.
\end{aligned}$$

[3] it it perhaps slightly unnatural to fix the Jacobian at zero but we do it this way to simplify the calculations and to stay in line with the analogous theorems below

The sequence (x_i) is a Cauchy sequence. Indeed, using first the qualitative part of the induction relation, we have

$$\begin{aligned}|y - F(x_i)| &= |y - F(x_{i-1}) + F(x_{i-1}) - F(x_i)| \\ &= |D(x_i - x_{i-1}) + F(x_{i-1}) - F(x_i)| \\ &\leq |\int_0^1 JF_{x_{i-1}+s\frac{x_i - x_{i-1}}{|x_i - x_{i-1}|}} \frac{x_i - x_{i-1}}{|x_i - x_{i-1}|} ds| \\ &\leq \varepsilon |x_i - x_{i-1}|.\end{aligned}$$

Now, using the quantitative part of the induction relation, we have:

$$|x_{i+1} - x_i| \leq \frac{|y - F(x_i)|}{(\beta/2)^{1/2n}} \leq \frac{\varepsilon |x_i - x_{i-1}|}{(\beta/2)^{1/2n}} \leq \frac{|x_i - x_{i-1}|}{2}.$$

Now, since (x_i) is a Cauchy sequence, it has a limit, which we call x. Now, observe that

$$|y - F(x_i)| < \varepsilon |x_i - x_{i-1}| \leq \varepsilon \left(\frac{\varepsilon}{(\beta/2)^{1/2n}}\right)^{i-1} < \varepsilon \left(\frac{1}{2}\right)^{i-1}.$$

Letting i tend to infinity, we get that $F(x) = y$ as expected. Finally, observe that

$$\begin{aligned}|x_i| &= |x_1 + \sum_{k=2}^{i} (x_k - x_{k-1})| \\ &< |x_1| \left(1 + \sum_{k=0}^{\infty} (\frac{\varepsilon}{(\beta/2)^{1/2n}})^k\right) \\ &< 2|x_1| < \frac{2|y|}{(\beta/2)^{1/2n}}\end{aligned}$$

as expected.

For uniqueness, suppose that $x_1, x_2 \in B(0, \rho_1)$ and $f(x_1) = f(x_2) \in B(0, \rho_2)$. We have immediately, for $\phi : [0,1] \to B(0, \rho_1)$, $t \mapsto (1-t)x_1 + tx_2$,

$$\begin{aligned}|f(x_2) - f(x_1)| &= |Df.(x_2 - x_1) + \int_0^1 \phi''(t)dt| \\ &\geq (\beta/2)^{1/2n}|x_2 - x_1| - K|x_2 - x_1| \geq ((\beta/2)^{1/2n} - \epsilon/2)|x_2 - x_1|.\end{aligned} \tag{5.2.1}$$

For the differentiability, we have for any $y_1, y_2 \in B(0\rho_2)$, using Eq. (5.2.1),

$$\begin{aligned}&\lim_{y_2 \to y_1} \frac{|y_2 - y_1 - (Df(f^{-1}(y_1)))^{-1}(y_2 - y_2)|}{|y_2 - y_1|} \\ &\leq \lim_{y_2 \to y_1} \frac{K}{((\beta/2)^{1/2n} - \epsilon/2)} \frac{|Df(f^{-1}(y_1))(f^{-1}(y_2) - f^{-1}(y_1)) - (y_2 - y_2)|}{|f^{-1}(y_2) - f^{-1}(y_1)|} \\ &= 0.\end{aligned}$$

□

We have the following simple but occasionally useful result (similar results have been proved directly for control distances in [**37**], [**13**], [**47**] etc.):

LEMMA 5.2.3. *Let (U, V, F_t) be a model such that the first derivative of F_t is uniformly bounded by some K, and with $\beta > 0$ There exists a constant C, depending only on I and β, such that for any $v \in V$ with $|v| \leq \frac{1}{K}$, we have*

$$C^{-1}|v|_t^I \leq |v| \leq C|v|_t.$$

PROOF. For the upper bound, note that for any $u \in U$, we immediately have $|u| \leq |u|_U$, the upper bound follows immediately with $C = K$.

For the lower bound, note that by Theorem 5.2.2, there exists a $u \in U$ with $|u| \leq 4\frac{|v|}{\underline{\beta}}$. Then we have

$$|v|_t \leq |u|_U \leq \sqrt{I}|u|^{\frac{1}{I}} \leq \left(4\frac{|v|}{(\beta/2)^{1/2n}}\right)^{\frac{1}{I}}.$$

□

We now introduce a few technical lemmas that are useful to show local equivalences between 'distances' defined by models. The final result, Proposition 5.2.6, is already in a general form amenable to applications in Part 3: the particular case where $\phi_k^j = 0$ for all j, k is enough to prove the upper bounds in Part 2.

DEFINITION 5.2.4. Let $U = U_1 \oplus U_2 \oplus \ldots \oplus U_I$ be a graded space with as usual $\dim(U_i) = \nu_i$, $\nu = \sum_{i=1}^{I} \nu_i$. Let $f : U \to \mathbb{R}^n$ be a function. We say that f is k-scaled, for some $k \in \mathbb{N}$, if for any $u \in U, s \in \mathbb{R}^+$, we have $f(\delta_s(u)) = s^i f(u)$.

LEMMA 5.2.5. *Let $U = U_1 \oplus U_2 \oplus \ldots \oplus U_I$ be a graded space and let f be a k- scaled function for some k. Suppose that there exist K, M (with $M \geq 1$) such that for any $v \in U$ with $|v|^2_{\text{eucl}} = 1$, and for any x with $|x|_{\text{eucl}} \leq M$, we have*

$$|\frac{\partial f}{\partial v}(x)| \leq K.$$

There exists a constant C, depending only on K, M, I such that, for any $i \in \{1, \ldots, I\}$, $x \in U$ with $|x|_U \leq 1$, and for any $y_i \in U_i$ with $|y_i|_{\text{eucl}} = 1$,

$$|\frac{\partial f}{\partial y_i}(x)| \leq C|x|_U^{k-i}$$

In particular, if $z \in U$ and for all i, $|z_i| \leq \bar{C}|x|_U^{m+i}$ and $|z_i| \leq \tilde{C}|x|_U^i$ for some $m \geq 0$ and some $\bar{C}, \tilde{C} \geq 1$, then for any x with $|x| \leq 1$, we have

$$|f(x+z) - f(x)| \leq C\bar{C}I(4(1+\tilde{C}\sqrt{I}))^I|x|_U^{k+m}$$

PROOF. For any $x \in U$, let $\omega(x) := \frac{1}{\inf(s:|\delta_s(x)|_{\text{eucl}} \geq 1)}$ There exists a constant C depending only on I such that for any $x \in U$, $C^{-1}|x|_U \leq \omega(x) \leq C|x|_U$.

We also write $\phi(x)$ for the point of the sphere $\{\xi : \xi \in U, |\xi|_{\text{eucl}} = |x|_{\text{eucl}}\}$ that can be obtained as $\delta_s(x)$ for $s = \omega(x)^{-1}$. We express the derivative $\frac{\partial f}{\partial y_i}(x)$ in a curvilinear coordinate system with components $(\omega(x), \phi(x))$.

Using the chain rule, we see that (using the notation x_i for $\text{Pr}_{U_i}(x)$, and writing C for a constant dependent only on K, M, I that changes from line to line)

$$\begin{aligned}
\left|\frac{\partial f}{\partial y_i}(x)\right| &= \left|\frac{\partial f}{\partial \omega(x)}\frac{\partial \omega(x)}{\partial y_i} + \frac{\partial f}{\partial \phi(x)}\frac{\partial \phi(x)}{\partial y_i}\right| \\
&= \left|k\omega(x)^{k-1}\frac{\frac{2}{i}|x_i|^{\frac{2-i}{i}}\frac{\partial |x_i|}{\partial y_i}}{2\omega(x)} + \frac{\partial f}{\partial \phi(x)}\frac{\partial \phi(x)}{\partial y_i}\right| \\
&\leq C\omega(x)^{k-1}\omega(x)^{\frac{2-i}{i}.i}\omega(x)^{-1} + K\omega(x)^k\omega(x)^{-i} \\
&\leq C\omega(x)^{k-i} \leq C|x|_U^{k-i},
\end{aligned}$$

which concludes the proof of the first part.

For the second part, we integrate the first part along the line $x + tz$ ($t \in [0,1]$). Note first that by Lemma 5.1.2, we have $|x + tz|_U \leq 4(|x|_U + \tilde{C}\sqrt{I}|x|_U)$. Then we have

$$|f(x+z) - f(x)| = |\int_0^1 \frac{\partial f}{\partial y_i}(x+zt)zdt|$$

$$\leq C(4(1+\tilde{C}\sqrt{I}))^I \sum_{i=1}^{I} |x|_U^{k-i} \bar{C} |x|_U^{m+i}$$
$$= C\bar{C}(4(1+\tilde{C}\sqrt{I}))^I \sum_{i=1}^{I} |x|_U^{k+m}$$
$$= C\bar{C}I(4(1+\tilde{C}\sqrt{I}))^I |x|_U^{k+m},$$

as expected. □

The following proposition is the main tool for proving results about progressive Hörmander and separated progressive Hörmander systems:

PROPOSITION 5.2.6 (**Main tool for deterministic study of 'progressive' structures**). *Let $U = U_1 \oplus U_2 \oplus \ldots \oplus U_I$ and $V = V_1 \oplus V_2 \oplus \ldots \oplus V_J$ ($I \geq J$) be two graded spaces with dimensions $\nu = \sum_{i=1}^{I} \nu_i$ and $\bar{\nu} = \sum_{j=1}^{J} \bar{\nu}_j$, and let $F : U \to V$ be a function which can be expressed for some N in the form*

$$F(u) = \sum_{j=1}^{J} M_j u_j + \sum_{j=1}^{J} \psi^j(u) + \sum_{j=1}^{J} \sum_{k=j+1}^{N} \phi_k^j(u),$$

where for each j and k, $\phi_k^j : U \to V_j \simeq \mathbb{R}^{\bar{\nu}_j}$ is a k-scaled function, M_j is a $\bar{\nu}_j \times \nu_j$ matrix, and (for each $j \leq J$) ψ^j is a j-scaled function of $u_1, u_2, \ldots, u_{i-1}$.

Suppose also that, uniformly over j, k, u with $|u|_{\text{eucl}} = 1$, we have

$$|\frac{\partial \phi_k^j(u)}{\partial u}| \leq K,$$
$$|\frac{\partial \psi^j(u)}{\partial u}| \leq K \quad \textit{and}$$
$$|v^T M_j M_j^T v| \geq H^{-2} \quad \forall v \in \mathbb{R}^{\bar{\nu}_j} \quad \textit{with } |v| = 1,$$

for some K, H (here $|\cdot|$ denotes the operator norm, and $\frac{\partial \phi_k^j(u)}{\partial u}$ should be seen as a Jacobian matrix).

There exist constants M and G, depending only on I, J, K, N such that for any $v \in V$ with $|v|_V \leq M$, there exists a $u \in U$ such that

$$F(u) = v \quad \textit{and} \quad |u|_U \leq G|v|_V.$$

REMARK 5.2.7. Because there clearly exist constants M, G depending only on I, J, K, N such that for any $|u|_U \leq M$, $|F(u)|_V \leq G|u|_U$, the above theorem is an 'equivalence theorem'.

PROOF. In the calculations below, unless otherwise stated, $|\cdot|$ denotes the Euclidean norm. Note that we can work equivalently with either $|\cdot|_U$ and $|\cdot|_V$ or the homogeneous norm $\omega(y) = (\inf_{|\delta_s(y)|=1}(s))^{-1}$ (which can be defined in either U or V). For any element u of U (resp. v of V), we will write u_i (resp. v_i) for $Pr_{U_i}(u)$(resp. $Pr_{V_i}(v)$). We will use suffixes to denote steps in our iterative construction procedure.

For a non square matrix $M \in \mathbb{R}^n \otimes \mathbb{R}^m$ ($m \geq n$), such that the smallest eigenvalue of MM^T is strictly positive, by abuse of notation, we define, for any $x \in \mathbb{R}^n$, $y = M^{-1}(x)$ to be the unique element of $\mathbb{R}^m$ such that $My = x$ and y is orthogonal to the kernel of the linear map represented by M, i.e. M^{-1} is used to denote the Moore-Penrose inverse.

We will use the following notation: $\xi_i^n = (u_1^n, \ldots, u_{i-1}^n, u_i^{n-1}, u_{i+1}^{n-1}, \ldots)$.

Using this notation, we now define the iterative procedure that will converge to our inverse element u:

We set $u^{-1}, u^0 = 0$ and then $\forall n \geq 1, \forall 1 \leq j \leq J, \forall k \geq j+1$, we set:

$$u_j^n = u_j^{n-1} + M_j^{-1}(v_j - \mathrm{Pr}_{V_j}(F(\xi_j^n))) \quad \text{and} \quad u_k^n = 0$$

Set $\Gamma = NCI(4(1+\sqrt{I}))^I + 1$. For $|v|_V \leq M = \frac{\Gamma^{-J}}{2}$, we will show by induction the following three identities:

$$
\begin{aligned}
|u_j^n - u_j^{n-1}| &\leq \Gamma^{(n-1)J+j-1}|v|_V^{n+j-1} \leq |v|_V^j \\
|u_j^n| &\leq 2\Gamma^j|v|_V^j \\
|v^j - \mathrm{Pr}_{V_i}(F(u^n))| &\leq H^{-1}\Gamma^{nJ}|v|_V^{n+j}.
\end{aligned}
$$

First, note that the second inequality follows immediately from the first one:

$$
|u_j^n| \leq \sum_{k=1}^{n} |u_j^k - u_j^{k-1}| \leq \sum_{k=1}^{n} \Gamma^{(k-1)J+j-1}|v|_V^{k+j} \leq 2|v|_V^j\Gamma^{j-1} \leq 2|v|_V^j\Gamma^j.
$$

This means we only need to prove the first and last inequalities. For this, we proceed by induction over n.

First note that for $n=0$, the result is clear. We now suppose it holds for n, and prove it for $n+1$. In all the calculations below, we will make much use of the fact that the operator norm of M_j^{-1} is bounded above by H.

We prove the first inequality by induction over j. For $j=1$, we have

$$
\begin{aligned}
|u_1^{n+1} - u_1^n| &\leq H|v_1 - \mathrm{Pr}_{V_1}(F(u^n))| \leq HH^{-1}\Gamma^{nJ}|v|_V^{n+j} \\
&= \Gamma^{nJ}|v|_V^{n+j} = \Gamma^{nJ+1-1}|v|_V^{n+1+1-1}
\end{aligned}
$$

as required. Then for the induction step (over j), we have:

$$
\begin{aligned}
|u_j^{n+1} - u_{j+1}^n| &\leq H|v^{j+1} - \mathrm{Pr}_{V_{j+1}}(F(\xi_{j+1}^{n+1}))| \\
&\leq H\left(|v^{j+1} - \mathrm{Pr}_{V_{j+1}}(F(u^n))| + |\mathrm{Pr}_{V_{j+1}}(F(u^n)) - \mathrm{Pr}_{V_{j+1}}(F(\xi_{j+1}^{n+1}))|\right) \\
&\leq H\left(H^{-1}\Gamma^{nJ}|v|_V^{n+j+1} + \sum_{k=1}^{j} |\mathrm{Pr}_{V_{j+1}}(F(\xi_{k+1}^{n+1})) - \mathrm{Pr}_{V_{j+1}}(F(\xi_k^{n+1}))|\right)
\end{aligned}
$$

Now, we see that

$$
|\xi_{k+1}^{n+1} - \xi_k^{n+1}| = |u_{k+1}^{n+1} - u_k^{n+1}| \leq \Gamma^{nJ+k}|v|_V^{n+k} \leq 1|v|_V^k\Gamma^k.
$$

Note also that for $k \leq j$, trivially $\Gamma^{nJ+k}|v|_V^{n+k} \leq \Gamma^{nJ+k}|v|_V^{n+j}$. This allows one to apply Lemma 5.2.5 on ψ^{j+1}, with $\bar{C} = \Gamma^{nJ+j}$ and $\tilde{C} = 1$, giving, where C is the constant from Lemma 5.2.5 (relative to the K appearing in this theorem),

$$
\begin{aligned}
|u_j^{n+1} - u_{j+1}^n| &\leq H\left(H^{-1}\Gamma^{nJ}|v|_V^{n+j+1} + \sum_{k=1}^{j} |\mathrm{Pr}_{V_{j+1}}(F(\xi_{k+1}^{n+1})) - \mathrm{Pr}_{V_{j+1}}(F(\xi_k^{n+1}))|\right) \\
&\leq H\left(H^{-1}\Gamma^{nJ}|v|_V^{n+j+1} + C\bar{C}I(4(1+\tilde{C}\sqrt{I}))^I|v|_U^{j+1+n}\right) \\
&= H\left(H^{-1}\Gamma^{nJ}|v|_V^{n+j+1} + C\Gamma^{nJ+j}I(4(1+\sqrt{I}))^I|v|_U^{j+1+n}\right) \\
&\leq H\left(H^{-1}\Gamma^{nJ+j}|v|_V^{n+j+1} + H^{-1}C\Gamma^{nJ+j}I(4(1+\sqrt{I}))^I|v|_U^{j+1+n}\right) \\
&\leq |v|_U^{j+1+n}\Gamma^{nJ+j+1},
\end{aligned}
$$

as expected. (Recall that $\Gamma \geq CI(4(1+\sqrt{I}))^I + 1$, by definition.) This concludes the induction step for the first inequality.

Now we only have to complete the induction step for the third inequality. We have immediately, using the first inequality and Lemma 5.2.5 on the ϕ_k^j for $k = j+1, \ldots, N$ (recall $\psi^j(u^{n+1})$ only depends on $u_1^{n+1}, u_2^{n+1}, \ldots u_{j-1}^{n+1}$), for any $1 \leq j \leq J$:

$$
|v_i - \mathrm{Pr}_{V_j}(F(u^{n+1}))| = |v_j - \mathrm{Pr}_{V_i}(F(\xi_{j+1}^{n+1})) + \mathrm{Pr}_{V_j}(F(\xi_{j+1}^{n+1})) - \mathrm{Pr}_{V_j}(F(u^n))|
$$

$$\begin{aligned}
&= 0 + |\operatorname{Pr}_{V_j}(F(\xi_{j+1}^{n+1})) - \operatorname{Pr}_{V_j}(F(u^{n+1}))| \\
&\leq \sum_{k=j+1}^{N} \Gamma^{nJ+J-1} CI(4(1+\sqrt{I}))^I |v|_V^{k+n} \\
&\leq NCI(4(1+\sqrt{I}))^I \Gamma^{nJ+J-1} |v|_V^{n+j+1} \\
&\leq |v|_V^{n+j+1} \Gamma^{(n+1)J},
\end{aligned}$$

as expected. (Recall that by definition, $\Gamma \geq NCI(4(1+\sqrt{I}))^I$.) This concludes the proof of the three inequalities.

To show the theorem, observe that the first inequality shows that for any j, $(u_j^n)_{n\in\mathbb{N}}$ is a Cauchy sequence. Define its limit as u_j. Now taking the limit as n tends to infinity in the third and last inequality yields $v_j - \operatorname{Pr}_{V_j}(F(u)) = 0$. Since this is valid for each j, this implies $v = F(u)$, as required. Finally, taking the limit as n tends to infinity in the second inequality yields: $|u_j| \leq 2\Gamma^j |v|_V^j$, which implies $|u|_U \leq \sqrt{I} 2^I \Gamma |v|_V$, which is our required second inequality with $G = \sqrt{I} 2^I \Gamma$. □

To finish this section, we note the following trivial fact about families of non centred seminorms:

PROPOSITION 5.2.8. *Let $V = \mathbb{R}^n$, and suppose that we are given two families of non centered seminorms $|\cdot|_{1,t}$ and $|\cdot|_{2,t}$ (i.e. $|v|_t = |v - v_t|_{*t}$ where $v_t \in V$ is a smooth curve starting at 0 and $|\cdot|_{*t}$ is a family of seminorms.) Suppose that $|\cdot|_{1,t}$ satisfies a doubling condition in the sense that there exist constants Δ_1, T and D_1 such that*

$$\forall d \leq \Delta_1, \quad |B_{1,t}(2d)| \leq D_1 |B_{1,t}(d)|.$$

Suppose also that the families of non centred seminorms are locally equivalent in the sense that there exist constants $\bar{\Delta}$ and $\bar{D}$ such that

$$\forall v \in V, \forall t \leq T \text{ such that } |v|_{1,t} \leq \bar{\Delta}, \text{ we have } |v|_{2,t} \leq \bar{D} |v|_{1,t}.$$

Then we have that there exist constants Δ_2 and D_2 such that

$$\forall d \leq \Delta_2, \text{ we have } |B_{2,t}(2d)| \leq D_2 |B_{2,t}(d)|.$$

Explicitly, $D_2 = D_1^{1+2\lceil \frac{\log(\bar{D})}{\log(2)} \rceil}$, and $\Delta_2 = \frac{\min(\bar{\Delta}, \Delta_1)}{2^{\lceil \frac{\log(2\bar{D}^2)}{\log(2)} \rceil}}$.

PROOF. Fix $d \leq \frac{\min(\bar{\Delta}, \Delta_1)}{2^{\lceil \frac{\log(2\bar{D}^2)}{\log(2)} \rceil}}$. Set $\bar{d} = \frac{d}{\bar{D}^2}$ and $K = 2\bar{D}^2$

$$\begin{aligned}
|B_{2,t}(2d)| = |B_{2,t}(K\bar{d})| &\leq |B_{1,t}(K\bar{D}\bar{d})| \\
&\leq |B_{1,t}(2^{\lceil \frac{\log(K)}{\log(2)} \rceil} \bar{D}\bar{d})| \leq D_1^{\lceil \frac{\log(K)}{\log(2)} \rceil} |B_{1,t}(\bar{D}\bar{d})| \\
&= D_1^{1+2\lceil \frac{\log(\bar{D})}{\log(2)} \rceil} |B_{1,t}(\bar{D}\bar{d})|.
\end{aligned}$$

□

CHAPTER 6

Auxiliary systems and objects

Here we introduce some of the auxiliary objects that will be used to prove our bounds.

6.1. Truncated signatures

Here we recall one of the basic objects of rough path theory, the truncated signature, and some of its basic properties. We begin by a recalling the theory for a general d-dimensional path. We refer to [**22**], [**44**], [**43**], [**40**], [**31**], [**11**] or [**21**] for more details.

REMARK 6.1.1. Note that it is possible to prove upper bounds whose space-time integrals are polynomial constants without any understanding of the algebraic structure of the objects below. Indeed, controlling the equivalence constant between homogeneous and control norms could be avoided by using the homogeneous norm both in the ball (in the denominator) and inside the exponential, which would not change the space-time integral of our upper bound. It does no harm to simply see the log-signature as the projection of the signature onto the tangent space of its intrinsic space at 0 and see the exponential map as the inverse of that transformation. The logarithm is just an orthogonal projection from a manifold in a Euclidean vector space to a subspace of that Euclidean vector space (for a proof see [**52**]).

Recall from rough path theory the tensor space

$$T^l(\mathbb{R}^d) = \bigoplus_{k=0}^{l} (\mathbb{R}^d)^{\otimes k},$$

which we will call the *truncated signature space of order* l *and generator of size* d. The dimension of that space is $\sum_{i=0}^{l} d^i$.

We will always work with the following basis for $T^l(\mathbb{R}^{d+1})$:

Let $e_1, \ldots, e_d$ be the canonical basis for $\mathbb{R}^d$. An orthonormal basis for the vector space $T^l(\mathbb{R}^{d+1})$ is given by:

$$e^1, \ldots, e^d, \ldots, e^i \otimes e^j, \ldots, e^{i_1} \otimes e^{i_2} \otimes \ldots \otimes^{i_l}, \ldots$$

We also use the short-cut $e^{i,j}$ for $e^i \otimes e^j$ and $e^{[i,j]}$ for $e^i \otimes e^j - e^j \otimes e^i$, and more generally, we define iteratively $e^{i,\alpha} = e^i \otimes e^\alpha$ and $e^{[i,\alpha]} = [e^i, e^\alpha] = e^i \otimes e^\alpha - e^\alpha \otimes e^i$.

DEFINITION 6.1.2. If $v = (v^i)_{i=0,1,2,\ldots,d}$ is collection of $d+1$ vector fields, for any multi-index α, we define higher order derivatives (v^α) and brackets $(v^{[\alpha]})$ by the induction relations $v^{(i)} = v^i$, $v^{i,\alpha} = \frac{\partial v^\alpha}{\partial v^i}$ and $v^{[i,\alpha]} = [v^i, v^\alpha]$. In fact, it will sometimes be useful to view $v^\cdot$ as a linear map from the tensor space $T^l(\mathbb{R}^{d+1})$ to $\mathbb{R}^{d+1}$ which makes the operation of directional derivation correspond to the tensor product and the two operations of taking brackets correspond to each other (brackets on the tensor space are defined via the formula $[e_1, e_2] = e_1 \oplus e_2 - e_2 \oplus e_1$ and linear extension).

From this perspective, the above definition of v^α corresponds to using the shorthand v^α for $v^{\bigoplus_{k=1}^{\#\alpha} e_{\alpha_k}}$.

It is a non-trivial fact that the iterative definition of $v^{[\alpha]}$ coincides with the definition of $v^{e^{[\alpha]}}$ (this result goes back to Kusuoka-Stroock [**37**]):

LEMMA 6.1.3. *For any $r+1$ differentiable family of vector fields $\sigma^0, \sigma^1, \dots, \sigma^d$ on $\mathbb{R}^m$ and any multi-index α with $|\alpha| \leq r$, we have*

$$\sigma^{[\alpha]} = \sigma^{e^{[\alpha]}}.$$

PROOF. *Claim:* for any multi-index α, if we write $\alpha = \sum_\beta \sigma^{(\beta)}$ (uniquely), then we have for any $i \in \{0, 1, \dots, d\}$,

$$\sigma^{[\alpha]}(\sigma^i) = \sum_\beta \lambda_\beta \sigma^{(\beta,i)}.$$

Proof of Claim:

We proceed by induction on the length of α: If $\#(\alpha) = 1$ and $\alpha = (j)$, then indeed $\sigma^\alpha(\sigma^i) = \sigma^j(\sigma^i) = \sigma(j,i)$.

For the induction step, if $\alpha = (j, \gamma)$, then if we write $\sigma^{[\gamma]} = \sum_\beta \lambda_\beta^\gamma \sigma^\beta$, we have

$$\begin{aligned}
\sigma^{[\alpha]}(\sigma^i) &= (\sigma^j(\sigma^{[\gamma]}))(\sigma^i) \\
&= \sum_\beta \lambda_\beta^\gamma (\sigma^j(\sigma^\beta))(\sigma^i) - \sum_\beta \lambda_\beta^\gamma (\sigma^\beta(\sigma^j))(\sigma^i) \\
&= \sum_\beta (\sigma^{(j,\beta,i)}) - \frac{\partial^2 \sigma^i}{\partial \sigma^j \partial \sigma^\beta}) - \sum_\beta (\sigma^{(\beta,j,i)} - \frac{\partial^2 \sigma^i}{\partial \sigma^j \partial \sigma^\beta}) \\
&= \sum_\beta \lambda_\beta^\gamma (\sigma^{(j,\beta,i)} - \sigma^{(\beta,j,i)}).
\end{aligned}$$

Since $\sum_\beta \lambda_\beta^\gamma (\sigma^{(j,\beta)} - \sigma^{(\beta,j)})$ is the expansion of $\sigma^{(j,\gamma)}$ in our orthogonal basis of $T^l(\mathbb{R}d+1)$, this concludes the proof of the claim.

Proof of the theorem: Again, we proceed by induction over the length of α. If $\alpha = (i)$ then the result is clear.

For the induction step, suppose $\alpha = (j, \gamma)$, we have immediately:

$$\begin{aligned}
\sigma^{[j,\gamma]} &= \sigma^j(\sigma^{[\gamma]}) - \sigma^{[\gamma]}(\sigma^j) \\
&= \sigma^j(\sigma^{e^{[\gamma]}}) - \sigma^{e^{[\gamma]}}(\sigma^j) \\
&= \sigma^{e^{(j,[\gamma])}} - \sigma^{e^{[\gamma]}}(\sigma^j) \\
&= \sigma^{e^{(j,[\gamma])}} - \sigma^{e^{([\gamma],j)}} = \sigma^{e^{[(j,\gamma)]}},
\end{aligned}$$

where at the last next to last line, we have used the definition of iterated derivatives only, and at the last line, we have used the claim proved above. This concludes the proof. $\square$

REMARK 6.1.4. The order of bracketing is not relevant: for instance, if α is a Lyndon word and $[[\alpha]]$ is the standard bracketing of α, it can be shown similarly that $\sigma^{[\alpha]} = \sigma^{e^{[[\alpha]]}}$.

REMARK 6.1.5. It can be shown by induction using the Jacobi identity that the space of arbitrary bracketings of words is the space

$$\operatorname{span}_{|\alpha| \leq L} e^{[\alpha]} = \operatorname{span}_{\substack{\alpha \in \mathcal{L}(\{0,1,\dots,d\} \\ |\alpha| \leq L)}} e^{[[\alpha]]}.$$

We can now use expressions such as $\sigma^{[\alpha]}$ or $\sigma^{e^{[\alpha]}}$ and σ^α or σ^{e^α} interchangeably.

Recall that for a sufficiently regular[1] path $\gamma : [0,T] \to \mathbb{R}^d$ we define iteratively the *iterated integrals* of γ by

$$\gamma_t^{\alpha,i} = \int_0^t \gamma_s^\alpha d\gamma_s^i.$$

[1]For instance, a deterministic path of bounded variation, or an Itô process.

If γ is a deterministic path of bounded variation, the integral above is standard Lebesgue integration. If γ is an Itô process, the integral above is to be interpreted as a *Stratonovich integral*. Formally, this means we suppose we are already given a geometric rough path lift of the path γ (otherwise we would formally need to change the probability space to make sure both γ and its integrals are well defined random variables on that same probability space).

DEFINITION 6.1.6. The *truncated signature (of order l)* of the path γ at time t is the the highly degenerate object

$$(s(\gamma)_t^{\alpha})_{\#(\alpha)\leq l} \in T^l(\mathbb{R}^d).$$

Let $G^l(\mathbb{R}^d)$ be the free nilpotent group of degree l, that is, the group defined by the exponential of Lie polynomials up to degree l in $T^l(\mathbb{R}^d)$.

It can be shown that the truncated signature s, as a function of time, of a path of bounded variation, as well as the rough path lift of Brownian motion (which is the same as considering the the process together with its Stratonovich integrals), lie in the space of geometric rough paths, that is $s_t \in G^l(\mathbb{R}^d)$ and for any u, t, $s_t = s_u \otimes s_{t-u}$.

Note that $G^l(\mathbb{R}^d)$ is a special manifold in the tensor space $T^l(\mathbb{R}^d)$, endowed with a Carnot group structure.

The signature is a very exciting tool that can be used in a variety of contexts including geometry ([**16**], where the tool was defined for the first time), rough path theory, control theory, character recognition ([**23**]), improved Monte Carlo simulations ([**45**], [**41**]), filtering [**19**], time series analysis/machine learning ([**18**],[**25**]) and of course (like the present thesis) applications to SDE ([**37**], [**32**]).

We note from [**22**] the following full characterisation of the algebraic relations that an element of the tensor space $T^l(\mathbb{R}^d)$ must satisfy to be the signature of a path:

THEOREM 6.1.1 (Shuffle product formula). *A tensor element*

$$a = (1, a^1, \cdots, a^l) \in T^l(\mathbb{R}^d)$$

belongs to $G(\mathbb{R}^d)$ if and only if

$$a^m \otimes a^n = \sum_{\sigma\in\mathcal{S}(m,n)} \mathcal{P}(a^{m+n}), \quad \forall m, n \geq 1,$$

where $\mathcal{S}(m,n)$ denotes the set of (m,n) shuffles in the permutation group of order $m+n$ and $\mathcal{P}^\sigma : V^{\otimes(m+n)} \to V^{\otimes(m+n)}$ is the permutation operator given by

$$\mathcal{P}^\sigma(v_1 \otimes \ldots \otimes v_{m+n}) = v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(m+n)}.$$

We define also the *log-signature space* $\mathcal{L}^l(\mathbb{R}^d)$ to be the Lie algebra associated with the Lie group $G^l(\mathbb{R}^d)$. More explicitly:

$$\mathcal{L}^l(\mathbb{R}^d) = \bigoplus_{\#(\alpha)\leq l} e^{[\alpha]}.$$

The *log-signature* of the path γ is the orthogonal projection $\log(S(\gamma))$ of the signature on the log-signature space.

We note (cf. [**44**]) that the log signature determines the signature uniquely and the map that associates the signature to the log signature and its inverse have the series expansions of the exponential and logarithm functions respectively:

We write $S = 1 + {}^1S + {}^2S\ldots$ for the expression of the signature in the space $T^l(\mathbb{R}^d)$. ${}^iS$ is an element of $(\mathbb{R}^d)^{\otimes i}$ and ${}^1S + {}^2S\ldots + {}^lS$ is an element of $\bigoplus_{k=1}^l(\mathbb{R}^d)^{\otimes k}$

$$\log(S) = ({}^1S + {}^2S\ldots + {}^lS) - \frac{({}^1S + {}^2S\ldots + {}^lS)^2}{2} + \ldots$$

(in this formula, multiplication must be interpreted as the operation $\otimes$ described above), and for the exponential, if $R \in \mathcal{L}^l(\mathbb{R}^d)$,

$$\exp(R) = \sum_{k=0}^{l} \frac{R^k}{k!}.$$

REMARK 6.1.7. In both of the formulae above, because we are looking at truncated signatures rather than the full signature, we must use the convention that $e^\alpha \otimes e^\beta = 0$ if $\#\alpha + \#\beta > l$.

6.2. Truncating at order rather than length.

The above objects can also be defined for the tensor space

$$T^\infty(\mathbb{R}^{d+1}) = \bigoplus_{k=0}^{l} (\mathbb{R}^{d+1})^{\otimes k},$$

where we use the notation

$$e^0, e^1, \dots, e^d, \dots, e^i \otimes e^j, \dots, e^{i_1} \otimes e^{i_2} \otimes \dots \otimes e^{i_l}, \dots$$

for the basis vectors (for $i, j, i_1, i_2, \dots, i_l \in \{0, 1, \dots, d\}$).

Recall that for a multi-index $\alpha \in \mathrm{Multi}(\{0, 1, \dots, d\})$, the order $|\alpha|$ of α is defined as $2o_0(\alpha) + o_1(\alpha)$ where $o_0(\alpha)$ and $o_1(\alpha)$ denote the number of zero and non-zero indices in α respectively.

We now define the following subspaces of $T^\infty(\mathbb{R}^{d+1})$ for any fixed l:

The truncated signature space

$$\mathcal{T}^l(\mathbb{R}^d, \mathbb{R}) = \mathrm{span}_{|\alpha| \le l}\, e^\alpha,$$

and the truncated log-signature space

$$\mathcal{L}^l(\mathbb{R}^d, \mathbb{R}) = \mathrm{span}_{|\alpha| \le l}\, e^{[\alpha]}.$$

Warning Those notations may vary from the standard rough path literature.

One key observation as that the Formulae 6.1 and 6.1 show that for $s_1 \in T^\infty(\mathbb{R}^{d+1})$, $s_2 \in \mathcal{L}^\infty(\mathbb{R}^d)$, $\mathrm{Pr}_{\mathcal{L}^l(\mathbb{R}^d,\mathbb{R})}(\log(s_1))$ and $\mathrm{Pr}_{\mathcal{T}^l(\mathbb{R}^d,\mathbb{R})}(\exp(s_2))$ are uniquely determined by $\mathrm{Pr}_{\mathcal{T}^l(\mathbb{R}^d,\mathbb{R})}(s_2)$ and $\mathrm{Pr}_{\mathcal{L}^l(\mathbb{R}^d,\mathbb{R})}(s_1)$ respectively.

For a path $\gamma \in \mathcal{P}_T^{d+1}$, we will denote by $s(\gamma) \in \mathcal{T}^l(\mathbb{R}^d, \mathbb{R})$ its signature, truncated at order l. When considering a system, we will use the notation $S = s(*W) \in \mathcal{T}^l(\mathbb{R}^d, \mathbb{R})$ for the truncated signature of the path $(s, W_s^1, W_s^2, \dots, W_s^d)$ ($s \in [0, \infty]$). As usual, the integrals are interpreted in the Stratonovich sense, which is equivalent to supposing that we are already given a (length 2) geometric rough path lift of the driving Brownian motions.

6.3. The standard KST approximation for a general system

Let $\mathcal{A}$ be a system, the Kusuoka-Stroock-Taylor (KST) approximation of Y_t of order l is

$$T_t^l = y + \sum_{|\alpha| \le l} (*\sigma)^\alpha(0) W_t^\alpha.$$

REMARK 6.3.1. Note that $(*\sigma)^\alpha$ is not the same thing as $*\sigma^\alpha$ (unless F is linear). The latter is used in the definition the weak Hörmander condition, whilst the former is used in the KST approximation. If F is not linear, the KSTA written with $*\sigma^\alpha$ would not be close enough to the actual solution.* However, any system can be turned into a system where F is linear by just adding n components to the background space equal to the solution $Y_t = F(X_t)$, and replacing F by the projection onto that space. So this is not a particularly interesting point.

Similarly to the definition of $Y_t(\gamma)$ for a deterministic driving path, we also have that for any driving deterministic smooth driving path Γ in $\mathbb{R}^{d+1}$,

$$T_t^l(\Gamma) = y + \sum_{|\alpha|\leq l} (*\sigma)^\alpha(0)(*\gamma)^\alpha.$$

For a path $\gamma \in \mathbb{R}^d$, we denote by $*\gamma$ the path in $\mathbb{R}^{d+1}$ such that, $\forall s \in [0,1]$, $*\gamma_s^0 = s$ and $*\gamma_s^i = \gamma_s^i$, $\forall i = \{1, 2, .., d\}$.

We need the following definition of homogeneous length for a path in $\mathbb{R}^{d+1}$ whose first component must be given different scaling:

DEFINITION 6.3.2. For a curve $(\tau, \gamma)_s \in \mathbb{R} \oplus \mathbb{R}^d$ parametrised over $0 \leq s \leq T$, we define the following homogeneous metric:

$$|(\tau,\gamma)|_{L^2,1} = \sqrt{\int_0^T T|\frac{\partial\gamma}{\partial s}|^2 + |\frac{\partial\tau}{\partial s}|ds}$$
$$|(\tau,\gamma)|_{L^2} = \sqrt{\int_0^T |\frac{\partial\gamma}{\partial s}|^2 + |\frac{\partial\tau}{\partial s}|ds}.$$

PROPOSITION 6.3.3. *The definition of* $|(\tau,\gamma)|_{L^2,1}$ *is independent of parametrisation.*

PROOF. Trivial. □

PROPOSITION 6.3.4. *If* $({}^1\tau, {}^1\gamma)$ *and* $({}^2\tau, {}^2\gamma)$ *are two smooth curves in* $\mathbb{R}^{d+1}$*, we have*

$$|({}^1\tau, {}^1\gamma) \otimes ({}^2\tau, {}^2\gamma)|_{L^2,1} \leq |({}^1\tau, {}^1\gamma)|_{L^2,1} + |({}^2\tau, {}^2\gamma)|_{L^2,1} \quad \textit{and}$$
$$|({}^1\tau, {}^1\gamma) \otimes ({}^2\tau, {}^2\gamma)|_{L^2} \leq |({}^1\tau, {}^1\gamma)|_{L^2} + |({}^2\tau, {}^2\gamma)|_{L^2}.$$

PROOF. Follows immediately from the fact that $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ for any positive real numbers a, b. □

LEMMA 6.3.5. *Let* $\Gamma : [0,1] \to \mathbb{R}^{d+1}$ *be a smooth driving path with* $\Gamma = (\tau, \gamma)$ *with* $\tau_s \in \mathbb{R}$ *and* $\gamma_s \in \mathbb{R}^d$*. Let* A_s *be a functional from* $[0,1]$ *to* $\mathbb{R}^n$*. Define the iterated integrals* $(A\Gamma)^\alpha = \int^{\circ\alpha} A(d\Gamma)^\alpha$ *of* A *with respect to* Γ *by* $(A\Gamma)^i = \int_0^1 Ad(\Gamma_s)^i$ *and*

$$(A\Gamma)^{\alpha,i} = \int_0^1 ((A\Gamma)^\alpha)_s d(\Gamma_s)^i.$$

Suppose that $A_s \leq K, \forall s \in [0,1]$ *(without loss of generality* $K \geq 1$*), we have that:*

$$\sup_{0\leq s\leq 1} |(A\Gamma)_s^\alpha|_{L^2} \leq |(A\Gamma)^{\bar{\alpha}}|_{L^2}|\gamma|_{L^2} \quad \textit{if} \quad b \neq 0$$
$$\leq |(A\Gamma)^{\bar{\alpha}}|_{L^2}|\tau|_{L^2} \quad \textit{if} \quad b = 0$$

where as usual, $\bar{\alpha}$ *is the multi-index composed of the first* $\#(\alpha) - 1$ *indices of* α*, such that* $\alpha = (\bar{\alpha}, b)$ *for some* $b \in \{0, 1, \ldots d\}$*.*

This means in particular that

$$\sup_{0\leq s\leq t} |(A\Gamma)_s^\alpha| \leq K(|\gamma|_{L^2})^{o_1(\alpha)}|\tau|_{L^2}^{o_0(\alpha)} \leq K(|\Gamma|_{L^2})^{|\alpha|},$$

where $|\Gamma|_{L^2}$ *denotes the homogeneous metric defined above.*

More precisely, if we have real processes A_s^α *for each* α*, then we have:*

$$\sup_{0\leq s\leq 1} |\sum_{|\alpha|=L} (A\Gamma)_s^\alpha| \leq C\sqrt{\sum_{|\alpha|=L}\int_0^1 (A_s^\alpha)^2 ds}|\Gamma|_{L^2}^L$$

for some constant C *dependent only on* L*.*

PROOF. We prove the first result by induction on $\#\alpha$ (not $|\alpha|$).

For $\#\alpha = 1$,

(1) if $\alpha = (0)$ then

$$\sup_{0\le s\le 1} |(A\Gamma)^{\alpha}_{s}| = \sup_{0\le s\le 1} |\int_0^s A_{s_1} d\tau_{s_1}| \le (\int_0^1 A_s^2 ds)^{\frac{1}{2}} (\int_0^1 |\frac{\partial \tau_s}{\partial s}|^2 ds)^{\frac{1}{2}} \le K|\tau|_{L^2}.$$

(2) if $\alpha = (j)$ for some $j \le d$ then

$$\sup_{0\le s_1\le 1} |(A\Gamma)^{\alpha}| = \sup_{0\le s_1\le 1} |\int_0^{s_1} A_s d\gamma^j_s| \le K|\int_0^1 1 d\gamma^j_s| \le K|\gamma|_{L^2}.$$

For the induction step, we have:

(1) If $\alpha = (\bar\alpha, 0)$ for some $\bar\alpha$, then

$$\begin{aligned}\sup_{0\le s_1\le t} |(A\Gamma)^{\alpha})_{s_1}| &\le t \sup_{0\le s_1\le t} |(A\Gamma)^{\bar\alpha}_{s_1}| \\ &\le Kt|\gamma|^{o_1(\bar\alpha)} t^{o_0(\bar\alpha)} \le K t^{o_0(\alpha)} |\gamma|^{o_1(\alpha)} \\ &\le K(|\Gamma|)^{|\alpha|},\end{aligned}$$

as expected.

(2) If $\alpha = (\bar\alpha, j)$, then

$$\begin{aligned}\sup_{0\le s_1\le t} |(A\Gamma)^{\alpha}_{s_1}| &= \sup_{0\le s_1\le t} |\int_0^{s_1} (A\Gamma)^{\bar\alpha} d\gamma^j_s| \le |\gamma^j| \sup_{0\le s\le t} |((A\Gamma)^{\bar\alpha})_s| \\ &\le |\gamma| K |\gamma|^{o_1(\bar\alpha)} t^{o_0(\bar\alpha)} \\ &\le K|\gamma|^{o_1(\alpha)} t^{o_0(\alpha)} \le K(|\Gamma|)^{|\alpha|},\end{aligned}$$

as expected.

For the last, more precise part, we write $r(\alpha)$ for the multi-index obtained by replacing every non zero component of α by 1. We also write $\tau(\alpha)$ for the multi-index obtained by deleting the last component of α.

We first perform the following inductive calculation, for any $\delta \in \mathrm{Multi}(\{0,1\}, L)$

$$\begin{aligned}&\sum_{r(\alpha)=\delta} \sup_{0\le s\le 1} |(A^{\alpha}\Gamma)^{\alpha}_s| \\ &\le \sum_{\substack{i\in\{1,2,\dots,d\}, r(\alpha)=\delta \\ \alpha=(\beta,i)}} \mathrm{Id}_{r(\alpha)_{\#\alpha}=1} \sup_{0\le s\le 1} |\int_0^1 (A^{\alpha}\Gamma)^{\beta}_s \frac{\partial \Gamma^i_s}{\partial s} ds| \\ &\quad + \mathrm{Id}_{r(\alpha)_{\#\alpha}=0} \sup_{0\le s\le 1} |\int_0^1 (A^{\alpha}\Gamma)^{\beta}_s \frac{\partial \Gamma^0_s}{\partial s} ds| \\ &\le \sqrt{\sum_{\substack{i\in\{1,2,\dots,d\}, r(\alpha)=\delta \\ \alpha=(\beta,i)}} \int_0^1 ((A^{\alpha}\Gamma)^{\beta}_s)^2 ds} |\Gamma|_{L^2} \mathrm{Id}_{r(\alpha)_{\#\alpha}=1} \\ &\quad + \sqrt{\sum_{\substack{r(\alpha)=\delta \\ \alpha=(\beta,0)}} \int_0^1 ((A^{\alpha}\Gamma)^{\beta}_s)^2 ds} |\Gamma|^2_{L^2} \mathrm{Id}_{r(\alpha)_{\#\alpha}=0} \\ &\le (\sum_{\substack{i\in\{1,2,\dots,d\}, r(\alpha)=\delta \\ \alpha=(\beta,i)}} \sup_{0\le s\le 1} |(A^{\alpha}\Gamma)^{\beta}_s|) |\Gamma|_{L^2} \mathrm{Id}_{r(\alpha)_{\#\alpha}=1} \\ &\quad + \sup_{0\le s\le 1} (| \sum_{\substack{i\in\{1,2,\dots,d\}, r(\alpha)=\delta \\ \alpha=(\beta,i)}} \int_0^1 ((A^{\alpha}\Gamma)^{\beta}_s|)|\Gamma|^2_{L^2} \mathrm{Id}_{r(\alpha)_{\#\alpha}=0} \\ &\le \dots\end{aligned}$$

$$\leq \sqrt{\sum_{r(\alpha)=\delta} \int_0^1 (A_s^\alpha)^2 ds |\Gamma|^{|\alpha|}}.$$

Using this, we finally obtain:

$$\begin{aligned}
\sup_{0\leq s\leq 1} |\sum_{|\alpha|=L} (A\Gamma)_s^\alpha| &\leq \sum_{|\delta|=L} |\sum_{r(\alpha)=\delta} (A\Gamma)_s^\alpha| \\
&\leq \sum_{|\delta|=L} \sqrt{\sum_{r(\alpha)=\delta} \int_0^1 (A_s^\alpha)^2 ds |\Gamma|_{L^2}^L} \\
&\leq 2^L |\Gamma|_{L^2}^L \sqrt{\sum_{|\alpha|=L} \int_0^1 (A_s^\alpha)^2 ds}.
\end{aligned}$$

This concludes the proof. □

REMARK 6.3.6. We cannot make l tend to ∞ in the theorem below: the aim, as in [**37**], is to use the expansion for a given sufficiently large l.

THEOREM 6.3.1. *Let $\mathcal{A} = (x, \sigma, F)$ be an (L, H_L)-weak Hörmander, (L', g, G)-tense system, with $g, L' \geq 2 + l$ with $l \geq L$*

(1) *If Γ is a smooth driving path in $\mathbb{R}^{d+1}$, we have $\forall i = 1, 2, \dots, n$*
$$|Y_t(\Gamma) - T^l(\Gamma)| \leq C|\Gamma|_{L^2,1}^{l+1}$$
for some C dependent only on G and l, and polynomial in G.

(2) *Writing as usual Y_t for the solution to the system, for $R \geq 1$ we have that there exist constants C, C_2, dependent only on G, l, R, d, and polynomial in G, d, such that*
$$\mathbb{P}(|Y_t - T_t^l| \geq R) \leq C_1 e^{\frac{-(R/C_2)^{\frac{2}{l+1}}}{8t}}$$

PROOF. For the deterministic part, using Lemma 6.3.5 yields the result immediately with

$$\begin{aligned}
C &\leq 2^{l+3} (\sup_{|v|=1; v\in\mathbb{R}^n} \sum_{l+1\leq|\alpha|\leq l+2} \langle \sigma^\alpha, v\rangle^2)^{1/2} \\
&\leq 2^{l+3} G^{1/2}.
\end{aligned}$$

For the random part, observe that

$$|Y_t - T_t^l| = |\sum_{\alpha\in A} \int^{\circ\alpha} (*\sigma)^\alpha (\circ dW)^\alpha| \tag{6.3.1}$$

where the iterated integrals are understood as Stratonovich integrals and the set A is defined by $A = \cup_{|\alpha|=l} \cup_{i=0}^d (i, \alpha)$.

We convert the above integrals into Itô integrals so that we can apply Theorem 2.1.1.

We will use the following notation:

$$A = \cup_{|\alpha|=l} \cup_{i=0}^d \{(i, \alpha)\}$$
$$A_2 = \cup_{|\alpha|=l} \cup_{i=1}^d \{(i, i, \alpha)\}.$$

We will denote by $c : A \to \mathcal{P}(\mathrm{Multi}(\{0, 1, \dots, d\})),\ \alpha \mapsto c(\alpha)$ the function such that where $c(\alpha)$ is the set of all multi-indices that can be obtained from α by successively replacing consecutive repeated indices by 0 any number of times. We will denote by

$$\bar{c} : A_2 \to \mathcal{P}(\mathrm{Multi}(\{0, 1, \dots, d\})),\ \alpha \mapsto \bar{c}(\alpha),$$

the function such that $\bar{c}(\alpha)$ is the set of all multi-indices that can be obtained from α by successively replacing consecutive repeated indices by 0 any number of times, with the constraint that the first two indices (which are equal and non -zero since $\alpha \in A_2$), must be replaced. For $\beta \in c(\alpha)$

or $\beta \in \bar{c}(\alpha)$, we will write $\varepsilon(\alpha,\beta)$ for the number of such replacements required to go from α to β. here are some examples to clarify the definitions:

$$\begin{aligned} c((1,2,2,5,0,0,9,9)) &= \{(1,2,2,5,0,0,9,9),(1,2,2,5,0,0,0),\\ &\qquad (1,0,5,0,0,9,9),(1,0,5,0,0,0)\}\\ c((1,1,1,1)) &= \{(0,1,1),(1,1,0),(0,0),(1,0,1)\}\\ \bar{c}((5,5,4,3,0,1,1)) &= \{(0,4,3,0,1,1),(0,4,3,0,0)\}\\ \varepsilon((5,5,4,3,0,1,1),(0,4,3,0,1,1)) &= 1\\ \varepsilon((5,5,4,3,0,1,1),(0,4,3,0,0)) &= 2\\ \varepsilon((1,2,2,5,0,0,9,9),(1,2,2,5,0,0,0)) &= 1\\ \varepsilon((1,2,2,5,0,0,9,9),(1,0,5,0,0,0)) &= 2. \end{aligned}$$

We now have the required notation to convert Eq. (6.3.1) into Itô form:

$$\begin{aligned} |Y_t - T_t^l| &= \sum_{\alpha\in A}\int^{\circ\alpha}(*\sigma)^\alpha(\circ dW)^\alpha \\ &= \sum_{\beta\in A}\sum_{\alpha\in c(\beta)}\int^{\alpha}\sigma^\beta(dW)^\alpha 2^{-\varepsilon(\alpha,\beta)} \\ &+ \sum_{\beta\in A_2}\sum_{\alpha\in\bar{c}(\beta)}\int^{\alpha}\sigma^\beta(dW)^\alpha 2^{-\varepsilon(\alpha,\beta)}. \end{aligned} \tag{6.3.2}$$

Now note that

$$\begin{aligned} &\sup_{|v|=1;v\in\mathbb{R}^n}\sum_{l+1\le|\alpha|\le l+2}|\sum_{\substack{\beta\in A:\\ \alpha\in c(\beta)}}\langle\sigma^\beta,v\rangle 2^{-\varepsilon(\alpha,\beta)} + \sum_{\substack{\beta\in A_2:\\ \alpha\in\bar{c}(\beta)}}\langle\sigma^\beta,v\rangle 2^{-\varepsilon(\alpha,\beta)}|^2 \\ &\le \sum_{l+1\le|\alpha|\le l+2} 2^{l+2}(\sum_{\substack{\beta\in A_2:\\ \alpha\in\bar{c}(\beta)}}|\langle\sigma^\beta,v\rangle|^2 + \sum_{\substack{\beta\in A:\\ \alpha\in c(\beta)}}|\langle\sigma^\beta,v\rangle|^2) \\ &\le \sum_{l+1\le|\beta|\le l+2} 2^{2(l+2)}|\langle\sigma^\beta,v\rangle|^2 \\ &\le 2^{2(l+2)}G. \end{aligned} \tag{6.3.3}$$

Using Eqs. (6.3.2) and (6.3.3) and applying Theorem 2.1.1, we immediately obtain the result with $C_2 = 2^{2(l+2)}G$ and $C_1 = 2^{2l+4}$. □

6.4. Comparing the densities of the KSTA and the original process

There are two possible related ways to obtain a fixed given sufficiently large order of approximation l such that the density of T^l is close enough to the density of Y^l for the purposes of the proof of our upper bounds.

Informally, we have two random variables which are both unlikely to be far away from each other (cf. 6.3.1) and reasonably smooth, and we show that their densities must be close to each other.

We could either (re-) use Malliavin smoothness directly, or use the smoothness of the density and derivatives in the form of Gaussian bounds on the densities and their derivatives (in terms of the Euclidean distance, obtained with Malliavin calculus), coupled with Fourier Analysis.

The first approach is the approach used in the modern Bally-Pigato-Caramellino literature (see Theorem 2.4, Lemma 2.7 etc in [**50**], Theorem 2.1 in [**7**] then its use in [**4**] etc.), the second approach is the one used in the original Kusuoka-Stroock article [**37**]. We adopt the second approach.

6.4.1. Using Fourier Analysis to qualify the precision of the KSTA. The following is inspired from the main arguments in the proof of Proposition 4.6 on page 423 of [**37**]:

PROPOSITION 6.4.1. *Let*

$$p\colon [0,1]\times\mathbb{R}^n \to \mathbb{R}^+,\ (t,y)\mapsto p_t(y) \quad \textit{and} \quad q\colon \mathbb{R}^+\times\mathbb{R}^n \to \mathbb{R}^+,\ (t,y)\mapsto q_t(y)$$

be the perturbed densities, perturbed by the localising function $\phi(U)$, (ϕ smooth, $U \in \mathbb{D}^\infty$, $\phi(U)\leq 1$ a.s.) of two processes P_t and Q_t on $\mathbb{R}^n$ satisfying, uniformly over $(t,y)\in[0,T]\times\mathbb{R}^n$:

$$\left|\frac{\partial^k p_t(y)}{\pi_{i=1}^k \partial z_i}\right| \leq C\frac{\exp\left(-\frac{M|y-x|^2}{t}\right)}{t^\mu},$$

$$\left|\frac{\partial^k q_t(y)}{\pi_{i=1}^k \partial z_i}\right| \leq C\frac{\exp\left(-\frac{M|y-x|^2}{t}\right)}{t^\mu}$$

$$\mathbb{P}(|P_t-Q_t|\geq R)\leq C\exp\left(-\left(\frac{R}{Mt}\right)^{\frac{2}{m}}\right) \quad \textit{and}$$

$$\mathbb{P}(\phi(U)<1)\leq Ce^{-\frac{c}{t}},$$

for some fixed constants C,c,M,μ,m (with $m\geq 2(n+1)(\mu-n)$), for some fixed $x\in\mathbb{R}^n$, and for all $|z_i|=1$, $i\in\{0,1,\dots k\}$, $k\leq n+1$, and for all $R>0$.

There exists a constant C_1, depending only on C,M,n, such that

$$|p_t(y)-q_t(y)|\leq C_1 t^{\frac{m}{2(n+1)}-\mu+n}.$$

PROOF. Let us define the following Fourier transforms of p and q:

$$\hat{p}_t(\xi)=\int_{\mathbb{R}^n} e^{-i\langle\xi,y\rangle}p_t(y)dy$$

$$\hat{q}_t(\xi)=\int_{\mathbb{R}^n} e^{-i\langle\xi,y\rangle}q_t(y)dy.$$

First note that by using the first two conditions, we obtain:

$$\begin{aligned}
|\hat{p}_t(\xi)-\hat{q}_t(\xi)| &\leq |\hat{p}_t(\xi)|+|\hat{q}_t(\xi)| \\
&\leq \frac{1}{|\xi|^{n+1}}(|\int_{\mathbb{R}^n} e^{-i\langle\xi,y\rangle}\frac{\partial^{n+1}p_t(y)}{\pi_{i=1}^{n+1}\partial z_i}dy| \\
&+|\int_{\mathbb{R}^n} e^{-i\langle\xi,y\rangle}\frac{\partial^{n+1}q_t(y)}{\pi_{i=1}^{n+1}\partial z_i}dy|) \\
&\leq \frac{2\bar{C}}{|\xi|^{n+1}t^{\mu-n}},
\end{aligned}$$

for some constant $\bar{C}$ dependent only on C and n.

Furthermore,

$$\begin{aligned}
|\hat{p}_t(\xi)-\hat{q}_t(\xi)| &= |\mathbb{E}\left((e^{i\langle\xi,P_t\rangle}-e^{i\langle\xi,Q_t\rangle})\phi(U)\right)| \\
&\leq \mathbb{E}(|1-\phi(U)|)+\mathbb{E}(|e^{i\langle\xi,P_t\rangle}-e^{i\langle\xi,Q_t\rangle}|) \\
&\leq \mathbb{E}(|1-\phi(U)|)+\mathbb{E}(|e^{i\langle\xi,P_t-Q_t\rangle}-1|) \\
&\leq e^{-\frac{c}{t}}+\mathbb{E}(|e^{i\langle\xi,P_t-Q_t\rangle}-1|) \\
&\leq Kt^{m/2}+\mathbb{E}(|e^{i\langle\xi,P_t-Q_t\rangle}-1|) \\
&\leq Kt^{m/2}+\mathbb{E}(|\xi||P_t-Q_t|) \\
&\leq K(1+|\xi|)t^{m/2},
\end{aligned}$$

where K is some constant changing at each line that only depends on C, n, μ, m, M. Here at the last line we have used the condition $\mathbb{P}(|P_t - Q_t| \geq R) \leq Ce^{-(\frac{R}{Mt})^{\frac{2}{m}}}$, integrated over R.

Next we write, for any $R > 0$,

$$\begin{aligned}|p_t(y) - q_t(y)| &\leq \frac{1}{2\pi}\|\hat{p}_t(\xi) - \hat{q}_t(\xi)\|_{L^1} \\ &\leq \int_{|\xi|\leq R} K(1+|\xi|)t^{m/2}d\xi + \int_{|\xi|\geq R} \frac{2\bar{C}}{|\xi|^{n+1}t^{\mu-n}}d\xi \\ &\leq K_1 R^{n+1}t^{m/2} + K_2 t^{-\mu+n}R^{-1}\end{aligned}$$

for some constants K_1, K_2, dependent only on C, n, μ, m, M.

Now set $R = t^{-\frac{m+2(\mu-n)}{2(n+2)}}$. This gives

$$\begin{aligned}|p_t(y) - q_t(y)| &\leq (K_1 + K_2)t^{\frac{m-2(\mu-n)(n+1)}{2(n+2)}} \\ &\leq (K_1 + K_2)t^{\frac{m}{2(n+1)}-(\mu-n)}.\end{aligned}$$

□

As a consequence, we have the following very useful result:

THEOREM 6.4.1. *Let $\mathcal{A} = (x, \sigma, F)$ be a uniformly (L, g, G)-tense, (L, H_L)-weak Hörmander system. Let p_t^l denote the density of the lth order Kusuoka-Stroock-Taylor approximation $\bar{Y}_t$ of Y_t, localised by the localising function $U \in [0, 1]$ such that $U = 0$ when $|\operatorname{logsig}(*W)| \geq 1/2$, and write p_t for the density of the actual solution Y_t, also localised by U. Suppose that $g \geq (2n+2)^2 2^{4L} + 3$. For any $l \geq (2n+2)^2 2^{4L}$, if $g \geq l + n + 3$, we have for some constant C*[2]*:*

$$|p_t(y) - p_t^l(y)| < Ct^{\frac{l+1}{2(n+1)}-(2n+1)2^{4L}+n}.$$

In particular, for $l = (2n+2)^2 2^{4L}$ we have

$$|p_t(y) - p_t^l(y)| < Ct.$$

PROOF. Let as usual w be the free vector fields in $T^l(\mathbb{R}^{d+1})$, the system $(0, w, F^{ST})$, restricted to $|\operatorname{logsig}(*W)| \geq 1/2$, is uniformly $(L, H_L/2)$ weak Hörmander and $(L, g, \min(G, C_{\text{free}}))$-tense, where we have used the fact that $g \geq l+n+3$[3], and where C_{free} is the tension of the system $(0, w, Id)$ which we know to be an absolute constant from Lemma 8.1.1. Therefore, we can apply Theorem 4.4.1 to obtain

$$\left|\frac{\partial^k p_t(y)}{\pi_{i=1}^k \partial z_i}\right| \leq C\frac{e^{-\frac{M|y-x|^2}{t}}}{t^{(n+k)2^4 L}}$$

for any $k \leq n+1$. A very important observation is that the quantity $\mu = (n+k)2^4L \leq (2n+1)2^4L$ does not depend on l. In our proof this is because the weak Hörmander constant of the system $(0, w, F^{ST})$ is L rather than l(even though the weak Hörmander order of the system $(0, w, Id)$ is of course l). In [**37**], the proof of this fact is done more specifically and directly by controlling the effect of the rest on the density.

Of course, because $g \geq n + 3$, we can also use Theorem 4.4.1 to immediately obtain the equivalent bound for the density of the actual process Y_t.

Furthermore, Theorem 6.3.1 now ensures that the last condition of Proposition 6.4.1 is also satisfied.

The result now follows immediately by applying Proposition 6.4.1 with $\mu = (2n+1)2^4L$. □

---

[2] The constant is proper but we shall not use that fact

[3] We need $g \geq l$ for the KSTA to be even well defined. We need $g \geq l + 2$ to be able to control the remainder. In fact as long as $g \geq l + 3$, the system $(0, w, F^{ST})$ is (Lg, G_g)-tense for any g, with the precise value of G_g coming from the exponential map.

CHAPTER 7

Application of models to an auxiliary object: deterministic results

In this section, we show the meaning of the abstract setting described above in terms of application to systems.

Let $\mathcal{A} = (X, \sigma, F)$ be a (L, g, G)-tense (L, H_L)-Weak Hörmander system. We define the following generalisations of the control distance:

DEFINITION 7.0.1. Let $\mathcal{P}^d_{[0,t]}$ be the set of smooth paths in $\mathbb{R}^d$ parametrised on the interval $[0, t]$ such that $\gamma_0 = 0$. For $y \in B$, we define the following distance, which (for $F = \mathrm{Id}$) is the same as the distance d_t defined in [**39**]:

$$\tilde{d}_t(y) = \inf_{\gamma \in \mathcal{P}_t, Y_t(*\gamma)=y} |\gamma|_{L^2,1} = \inf_{\gamma \in \mathcal{P}_t, Y_t(*\gamma)=y} |\gamma|.$$

Here $|\gamma|$ denotes the length of γ, and $Y_t(*\gamma)$ denotes the solution to the ODE corresponding to the system driven by γ instead of Brownian motions. The equivalence between the two definitions comes from the fact that the condition $Y_t(*\gamma) = y$ is invariant by reparametrisation.

We call it the Léandre distance.

DEFINITION 7.0.2. Let $\mathcal{P}^{d+1}_{[0,t]}$ be the set of smooth paths in $\mathbb{R}^{d+1}$ parametrised on the interval $[0, t]$ such that $\gamma_0 = 0$. For $y \in B$, we define the following distance, which is the one we shall be working with for our upper bounds.

$$d_t(y) = \inf_{\substack{\Gamma \in \mathcal{P}_t, Y_t(\Gamma)=y \\ \Gamma^0_t = t}} |\Gamma|_{L^2,1} = \inf_{\substack{\Gamma \in \mathcal{P}_1, Y_t(\Gamma)=y \\ \Gamma^0_1 = t}} |\Gamma|_{L^2,1} \overset{\forall s}{=} \inf_{\substack{\Gamma \in \mathcal{P}_s, Y_s(\Gamma)=y \\ \Gamma^0_s = t}} |\Gamma|.$$

Here $|\Gamma|$ denotes the homogeneous length of Γ:

$$|\Gamma| = |\Gamma|_{\mathbb{R}^d,\mathbb{R}} = \int_0^T \sqrt{|\dot{\Gamma}^0_s| + \sum_{i=1}^d (\dot{\Gamma}^i_s)^2},$$

and as usual $Y_t(\Gamma)$ denotes the solution to the ODE corresponding to the system driven by Γ. Again, the last identity comes from the possibility of reparametrizing the curve optimally.

DEFINITION 7.0.3. Let us define the following functions:

$$F^{ST}_l : T^l(\mathbb{R}^d) \to \mathbb{R}^n,\ S \mapsto F(x) + \sum_{|\alpha| \le l} S^\alpha (*\sigma)^\alpha(x)$$

and

$$F^{\log(S)T}_l : \mathcal{L}^l(\mathbb{R}^d, \mathbb{R}) \to \mathbb{R}^n,\ S \mapsto F(x) + \sum_{|\alpha| \le l} \exp(S)^\alpha (*\sigma)^\alpha(x).$$

Let $\mathcal{S}$ be the space obtained by ignoring the component in $\alpha = (0)$ of elements in $\mathcal{L}^l(\mathbb{R}^d, \mathbb{R})$:

$$\mathcal{S} := \operatorname{span}_{\substack{|\alpha| \le l \\ \alpha \neq (0)}} e^{[\alpha]}.$$

We have $\mathbb{R} \oplus \mathcal{S} = \mathcal{L}^l(\mathbb{R}^d, \mathbb{R})$. We use the following notational shorthand:

$$F^{\log(S)T}_{t,l}(\cdot) = F^{\log(S)T}_l(t, \cdot).$$

For any bounded variation path in $\mathbb{R}^d$ parametrised over $[0, t]$, we write $\operatorname{logsig}^-(*\gamma)$ for the projection of the log-signature of $*\gamma$ on $\mathcal{S}$. We can give $(\mathcal{S}, F^{\log(S)T}_{t,l}, \mathbb{R}^n)$ the structure of a model by assigning to any linear combination of $\sigma^{[\alpha]}$ with $|\alpha|$ fixed the homogeneous degree components $|\alpha|$. We call the homogeneous 'distance' associated to this model d_t^l.

For $[0,t]$, we write $\mathrm{logsig}^-(\gamma)$ for the projection of the log-signature of γ on $\mathcal{S}$.

The following is immediate:

PROPOSITION 7.0.4. *We have the following equivalent description of d_t: Let $\mathcal{P}_1^{d+1}$ be the set of paths $\Gamma \in \mathbb{R}^{d+1}$ indexed over $[0,1]$, with $\Gamma_0 = 0$. For $\Gamma \in P_1^{d+1}$, we write $\mathrm{Sol}_{x,t}(\Gamma)$ or simply $\mathrm{Sol}_x(\Gamma)$ for the solution to the following ODE*

$$X_0 = x, \tag{7.0.1}$$
$$dX_s = \sigma^{\{1,2,\dots d\}}(X_s)d\Gamma_s^{\{1,2,\dots d\}} + t\sigma^0(X)d\Gamma_s^0.$$

Then we have:

$$d_t(y) = \inf_{\substack{\Gamma \in \mathcal{P}_1^{d+1} \\ F(\mathrm{Sol}_x(\Gamma))=y;\Gamma_1^{(0)}=1}} |\Gamma|_{L^2},$$

and

$$\tilde{d}_t(y) = \inf_{\substack{\gamma \in \mathcal{P}_1^{d} \\ F(\mathrm{Sol}_x(*\gamma))=y}} |\gamma|_{L^2}.$$

REMARK 7.0.5. If the drift is null, and F_t is the identity (we have an SDE instead of a system), then both $\tilde{d}_t(y)$ and $d_t(y)$ reduce to the Carnot-Carathéodory distance.

We note the following trivial properties of d_t:

PROPOSITION 7.0.6. *Let $\mathcal{A} = (x, \sigma, F)$ be a uniformly (L, g, G)-tense, uniformly (L, H_L)-weak Hörmander system, for any $0 = s_0 \le s_1 \le s_2 \le \ldots \le s_N = t$ and any $x_1, x_2, \ldots, x_N$, we have*

$$d_t(x, *x_N) \le \sum_{i=0}^{N} d_{s_{i+1}-s_i}(x_i, *x_{i+1}).$$

PROPOSITION 7.0.7. *Let $\mathcal{A} = (x, \sigma, F)$ be a uniformly (L, g, G)-tense, uniformly (L, H_L)-weak Hörmander system, we have, for any $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$ and any t,*

$$d_t(x, y) \ge |*x - y|\, G^{-\frac{1}{2}}(1-t).$$

Note that d_t and $\tilde{d}_t$ depend on the complete local behaviour of the vector fields, whilst d_t^l only depends on the iterated derivatives of σ up to order l, evaluated at 0.

Aim. In this section, we aim to show some local equivalence between the distances d_t^l and d_t.

REMARK 7.0.8. Locally, we can *morally* view the trio (solution map, space of driving paths, target space) as a model whose associated homogeneous distance is the control distance. The equivalence of d_t^l and d_t is therefore locally an infinite dimensional version of Proposition 5.2.6 But since the space is infinite dimensional, we must be more careful. However, while the infinite dimensionality of the model forces us to prove things independently, the idea of the proof is still guided by the above remark and remains similar to another declension of the proof of the inverse functions theorem.

We will prove the theorems step by step with a few preliminary propositions.

The following proposition shows the link between the quantity $\underline{\beta}$ defined for models, and the hypoellipticity constant H_L defined for systems.

PROPOSITION 7.0.9. *Let $\mathcal{A} = (x, \sigma, F)$ be a system that is (L', g, G)-tense and (L, H_L)-weak Hörmander at x (with $L' \ge L$). For any $l \le L'$, there exist constants B_l^1 and B_l^2, depending only on l, such that*

$$\sqrt{\frac{H_l}{B_l^1}} \le \underline{\beta}_{FST} \le \sqrt{\frac{H_l}{B_l^2}},$$

where $\underline{\beta}_{FST}$ is the square root of the smallest eigenvalue of $JF^{ST}(JF^{ST})^T$.

In particular, since $H_L \leq H_l \leq G$, we have that $\underline{\beta}_{FST}$ is a proper constant.

PROOF. We prove the theorem for the case of zero drift (i.e. multi-indices are over the set $\{1, 2, \dots, d\}$), the general case is very similar. The key is to observe that if α and $\bar{\alpha}$ are two multi-indices which are not exactly a reordering of each other, then e^{α} and $e^{\bar{\alpha}}$ are orthogonal. Clearly this means the same is true of $e^{[\alpha]}$ and $e^{[\bar{\alpha}]}$.

Let $J(l_0, l_1, l_2, \dots, l_\kappa)$ where $2l_0 + \sum_{k=1}^{\kappa} = \bar{l} \leq l$ be the multi-set of cardinality $\bar{l} - l_0$ with elements in $\{0, 1, 2, 3, \dots, \kappa\}$ containing k exactly l_k times for all $0 \leq k \leq \kappa$. Note that there are only as many such multi-sets as there are choices of numbers $1 \leq l_1, l_2 ... l_\kappa$ and l_0 such that $2l_0 + \sum_{k=0}^{\kappa} = \bar{l} \leq l$. The exact number is not relevant: what matters is getting rid of the dependence on d. That being said, since

$$\sum_{i=0}^{\lfloor l/2 \rfloor} 2^{l-2i} = \lfloor \frac{2^{l+2} - 1}{3} \rfloor,$$

that there are $\lfloor \frac{2^{l+2}-1}{3} \rfloor$ follows from the fact that there are 2^l choices with $l_0 = 0$, and that fact can be proved in two ways:

Proof 1: For fixed κ and fixed $\bar{l}$, we can view the choice as picking the $l_i - 1$ such that $\sum_i l_i - 1 = \bar{l} - \kappa$, which is a choice of a way of dividing a set of $\bar{l} - \kappa$ elements into κ types. There are

$$\binom{\bar{l} - \kappa + \kappa - 1}{\kappa - 1} = \binom{\bar{l} - 1}{\kappa - 1}$$

such choices. Summing over κ, we get $2^{\bar{l}-1}$ choices if $\bar{l} \geq 1$ (otherwise we get one choice). Summing over $\bar{l}$, we get 2^l choices.

Proof 2: There is a bijection between the choices of J and the set of subsets of $\{1, 2, \dots, l\}$: Let H be a subset of $\{1, 2, \dots, l\}$. Set $k_0 = \min(a : a \in H)$, $k_1 = \min(b : b \notin H, b \geq k_0)$, then similarly,

$$k_i = \min(b : b \in H, b \geq k_{i-1}) \quad \text{if} \quad k_{i-1} \notin H \quad \text{and}$$
$$k_i = \min(b : b \notin H, b \geq k_{i-1}) \quad \text{if} \quad k_{i-1} \in H.$$

Then set $l_i = k_i - k_{i-1}$ for all i.

For each such J we now consider an orthonormal basis $\{h_k^J\}_{k \in \{1,2,\dots,K\}}$ (for some $K \leq \#(J)!$) of the space[1]

$$\text{span}_{\alpha \in P_J} e^{[\alpha]},$$

where P_J denotes the set of multi-indices which contain the same indices, and the same number of times, as J. Let $\lambda_{J,k}^{\alpha}$ be numbers such that $\forall J, k, \sum_{\alpha \in P_J} \lambda_{J,k}^{\alpha} e^{\alpha} = h_k^J$, minimising $\sum_{\alpha} (\lambda_{J,k}^{\alpha})^2$ (i.e. take the $|\lambda_{J,k}|$ orthogonal to the kernel of the linear map that sends a vector $\mu \in \text{span}_{\alpha \in P_J} e^{[\alpha]}$ to $\sum_{\alpha \in P_J} \mu^{\alpha} e^{[\alpha]}$).

Let I be a multi-set of elements of $\{0, 1, 2, \dots, d\}$. Let ι be the function that renames '1' the smallest non-zero element of I, '2' for the second smallest, etc. We take as an orthonormal basis of $\text{span}_{\alpha \in P_I}(e^{[\alpha]})$ the vectors h_k^I defined by:

$$h_k^I = \sum_{\alpha \in P_I} \lambda_{\iota(I),k}^{\iota(\alpha)} e^{[\alpha]}.$$

This is the fixed orthonormal basis of the log-signature space that we will always assume we are working with.

Let $\tilde{P}$ be the matrix whose columns are the expressions of the e^{α}'s, where α runs over multi-indices with elements in $\{1, 2, \dots, l\}$, in the basis formed by the $h_k^{J(l_1, l_2, \dots, l_k)}$ (for all values of

[1] Note that for some choices of J, the space is $\{0\}$, and the basis is empty. For instance $J()$, $J(j)$ and $J(0, i)$ (for any $i, 2j \leq l$) each satisfy $\text{span}_{\alpha \in P_J} e^{\alpha} = \{0\}$

$(l_1, l_2, \ldots, l_k))$ Let P be the matrix whose columns are the expressions of the $e^{[\alpha]}$'s in the basis formed by the h_k^I for all multi-sets I.

P is a block diagonal matrix (elements are in the same block if they correspond to multi-indices that are exact reorderings of each other), and the block that corresponds to I has the same entries as the block that corresponds to ιI. So there are only 2^l different blocks. We can also group all the blocks into bigger blocks each of which is identical to $\tilde{P}$. Let B_l^1 and B_l^2 respectively be inverses of the largest and smallest eigenvalues of $\tilde{P}\tilde{P}^T$. Those are in fact also the smallest and largest eigenvalues of PP^T.

The theorem follows upon observing that H_l is the minimum eigenvalue of the matrix

$$JF^{ST}PP^T(JF^{ST})^T$$

whilst $\underline{\beta}_{F^{ST}}$ is the square root of the minimum eigenvalue of $JF^{ST}(JF^{ST})^T$. This concludes the proof. $\square$

REMARK 7.0.10. If polynomial dependence inside the exponential is not required, we do not need polynomial dependence of B_l and we can simply observe that it is a positive constant depending only (at most) on l, $\underline{\beta}$ and d. In this case we could simply pick any orthonormal basis of $\mathcal{S}$.

LEMMA 7.0.11. *Let $l \geq 1$. There exists a polynomial constant M_l such that for any $S \in \mathcal{L}^l(\mathbb{R}^d)$, there exists a $\Gamma \in \mathcal{P}^d$ such that*

$$\mathrm{logsig}_l(\Gamma) = S$$
$$|\Gamma|_{L,1} \leq C|S|_{\mathcal{L}^l(\mathbb{R}^d)},$$

where $||_{\mathcal{L}^l(\mathbb{R}^d)}$ denotes the homogeneous norm in $\mathcal{L}^l(\mathbb{R}^d)$ assigning the same degree for all indices.

PROOF. This is a well known and standard application of the BCH formula: By reparametrising, we can work with lengths rather than the norm $||_{L,1}$.

Claim: There exists a polynomial C_1 such that for any $\lambda \in \mathbb{R}$ and for any multi-index α of length less than l, there exists a path Γ of length $|\Gamma| \leq \lambda^{\frac{1}{l}} C_l$ and $\mathrm{logsig}_l(\Gamma) = \lambda e^{[\alpha]}$.

Proof of the Claim: The claim is obvious for $l = 1$ by taking a linear path. For the induction case, suppose that the result is true for $l-1$, we prove that it is true for l. Suppose that $\alpha = (a, \beta)$ for some $a \in \{0, 1, \ldots, d\}$. By the induction hypothesis, there is a path $\gamma_{e^{[\beta]}}$ of length less than C_{l-1} such that $\mathrm{logsig}(\gamma) = e^{[\beta]}$, and rescaling, there is a path of $\gamma_{(\lambda/2)^{\frac{l-1}{l}} e^{[\beta]}}$ of length less than $C_{l-1}(\frac{\lambda}{2})^{\frac{1}{l}}$ such that $\mathrm{logsig}(\gamma) = (\lambda/2)^{\frac{l-1}{l}} e^{[\beta]}$. We then consider the following concatenation:

$$\tilde{\gamma} = \gamma_{(\lambda/2)^{\frac{1}{l}} e^a} \otimes \gamma_{(\lambda/2)^{\frac{l-1}{l}} e^{[\beta]}} \otimes \gamma_{-(\lambda/2)^{\frac{1}{l}} e^a} \otimes \gamma_{-(\lambda/2)^{\frac{l-1}{l}} e^{[\beta]}}$$

By the BCH formula, $\mathrm{logsig}_l(\tilde{\gamma}) = \lambda e^{[\alpha]}$. By the induction hypothesis,

$$|\tilde{\gamma}| \leq 2(\frac{\lambda}{2})^{\frac{1}{l}}(1 + C_{l-1}).$$

This proves the claim.

Proof of the theorem: We also prove this theorem by induction. Similarly to the claim, the case $l = 1$ is trivial. Writing $\mathcal{M}$ for the constant involved.

For the induction step, suppose that $S = \zeta + \xi$ with $\zeta \in \mathcal{L}^{l-1}(\mathbb{R}^d)$ and $\xi \in \mathrm{span}_{|\bar{\alpha}|=l}\, e^{[\bar{\alpha}]}$.

By the induction hypothesis, we have a path γ_ζ with length less than $\mathcal{M}_{l-1}|\zeta|_{\mathcal{L}^{l-1}(\mathbb{R}^d)}$ whose log signature coincides with ζ on the space $\mathcal{L}^{l-1}(\mathbb{R}^d)$. Being projections of multiple integrals of γ_ζ, the components of $\mathrm{logsig}(\gamma_\zeta)$ in $\mathrm{span}_{|\bar{\alpha}|=l}\, e^{[\alpha]}$ are all bounded above by $(M_l|\zeta|_{\mathcal{L}^{l-1}(\mathbb{R}^d)})^l$. So if we fix a way of writing $\mathrm{logsig}(\gamma_\xi) = \sum_{\bar{\alpha}} \mu_{\bar{\alpha}} e^{[\bar{\alpha}]}$, we have $|\mu_{[\bar{\alpha}]}| \leq 2^l(M_l|\zeta|_{\mathcal{L}^{l-1}(\mathbb{R}^d)})^l$.

We also fix a way of writing $\xi = \sum_{|\bar{\alpha}|=l} e^{[\bar{\alpha}]}\xi_{\bar{\alpha}}$.

Now let $\{a_1, a_2, \ldots, a_{d^l}\}$ be any ordering of $\{\bar{\alpha} : |\bar{\alpha}| = l\}$, consider the following concatenation

$$\Gamma = \gamma_\zeta \otimes_{i=1}^{d^l} \gamma_{\xi_{a_i} - \mu_{a_i}}$$

We indeed have that

$$\mathrm{logsig}_l(\Gamma) = \xi + \zeta.$$

Furthermore,

$$|\Gamma| \leq M_{l-1}|\zeta|_{\mathcal{L}^{l-1}(\mathbb{R}^d)} + \sum_{|\bar{\alpha}|=l} \left((2M_{l-1}|\zeta|_{\mathcal{L}^{l-1}(\mathbb{R}^d)})^l + C_l^l \xi_{[\bar{\alpha}]}| \right)^{\frac{1}{l}}.$$

The result follows. □

We now formulate the following easy generalisation of Lemma 7.0.11:

LEMMA 7.0.12. *Let $l \geq 1$. There exists a polynomial constant M_l such that for any $S \in \mathcal{L}^l(\mathbb{R}^d, \mathbb{R})$, there exists a $(\tau, \gamma) \in \mathcal{P}^{d+1}$ such that*

$$\mathrm{logsig}_l((\tau, \gamma)) = S$$
$$|(\tau, \gamma)| \leq M_l |S|_{\mathcal{L}^l(\mathbb{R}^{d+1})},$$

where $|\cdot|_{\mathcal{L}^l(\mathbb{R}^d,\mathbb{R})}$ denotes the homogeneous norm in $\mathcal{L}^l(\mathbb{R}^d, \mathbb{R})$ assigning double degree to zero indices, and similarly $||$ is the homogeneous length.

PROOF. The proof is almost identical to that of Lemma 7.0.11:

Claim: There exists a polynomial C_l such that for any $\lambda \in \mathbb{R}$ and for any multi-index α of order less than l, there exists a path (τ, γ) of homogeneous length $|(\tau, \gamma)| \leq \lambda^{\frac{1}{|\alpha|}} C_l$ and $\mathrm{logsig}_l((\tau, \gamma)) = \lambda e^{[\alpha]}$

Proof of claim: The claim is obvious for $l = 1$ and for $l = 2$ and $\gamma = 0$, by taking a linear path. For the induction case, suppose that $\alpha = (a, \beta)$ for some $a \in \{0, 1, \ldots, d\}$. By the induction hypothesis, there is a path $\gamma_{e^{[\beta]}}$ of homogeneous length less than C_{l-1} such that $\mathrm{logsig}(\gamma) = e^{[\beta]}$. We then consider the following concatenation:

If $a \neq 0$:

$$\tilde{\gamma} = \gamma_{(\lambda/2)^{\frac{1}{|\alpha|}} e^a} \otimes \gamma_{(\lambda/2)^{\frac{|\alpha|-1}{|\alpha|}} e^{[\beta]}} \otimes \gamma_{-(\lambda/2)^{\frac{1}{|\alpha|}} e^a} \otimes \gamma_{-(\lambda/2)^{\frac{|\alpha|-1}{|\alpha|}} e^{[\beta]}}.$$

and if $a = 0$:

$$\tilde{\gamma} = \gamma_{(\lambda/2)^{\frac{2}{|\alpha|}} e^a} \otimes \gamma_{(\lambda/2)^{\frac{|\alpha|-2}{|\alpha|}} e^{[\beta]}} \otimes \gamma_{-(\lambda/2)^{\frac{2}{|\alpha|}} e^a} \otimes \gamma_{-(\lambda/2)^{\frac{|\alpha|-2}{|\alpha|}} e^{[\beta]}}.$$

By the BCH formula, $\mathrm{logsig}_l(\tilde{\gamma}) = \lambda e^{[\alpha]}$. By the induction hypothesis, and by Lemma 6.3.4, its homogeneous length is less than $2(\frac{\lambda}{2})^{\frac{1}{|\alpha|}}(1 + C_{l-1})$. This proves the claim.

Proof of the theorem: We also prove this theorem by induction over the order, which we write l here, not the length. Similarly to the claim, the case $l = 1$ is trivial.

For the induction step, suppose that $S = \zeta + \xi$ with $\zeta \in \mathrm{span}_{|\bar{\alpha}| \leq l-1} e^{[\bar{\alpha}]}$ and $\xi \in \mathrm{span}_{|\bar{\alpha}|=l} e^{[\bar{\alpha}]}$. By induction hypothesis, we have a path γ_ζ with homogeneous length less than $M_{l-1}|\zeta|_{\mathcal{L}^{l-1}(\mathbb{R}^d,\mathbb{R})}$ whose log signature coincides with ζ on the space $\mathrm{span}_{|\bar{\alpha}|=l} e^{[\bar{\alpha}]}$. Being projections of multiple integrals of γ_ζ, the components of $\mathrm{logsig}(\gamma_\zeta)$ in $\mathrm{span}_{|\bar{\alpha}|=l} e^{[\alpha]}$ are all bounded above by $(M_{l-1}|\zeta|_{\mathcal{L}^{l-1}(\mathbb{R}^d,\mathbb{R})})^l$. This means that if we fix a way of writing

$$\mathrm{Pr}_{\mathrm{span}_{|\bar{\alpha}|=l}(e^{[\bar{\alpha}]})}(\gamma_\zeta) = \sum_{|\bar{\alpha}|=l} e^{[\bar{\alpha}]} \mu_{\bar{\alpha}},$$

we have $|\mu_{\bar{\alpha}}| \leq 2^l (M_{l-1}|\zeta|_{\mathcal{L}^{l-1}(\mathbb{R}^d,\mathbb{R})})^l$.

We fix a way of writing $\xi = \sum_{|\bar{\alpha}|=l} e^{[\bar{\alpha}]} \xi_{[\bar{\alpha}]}$.

Choose any ordering $a_1, a_2, \ldots, a_K$ of the set $\{\bar{\alpha} : |\bar{\alpha}| = l\}$. (Here $K = \#(\{\bar{\alpha} : |\bar{\alpha}| = l\})$.) We can consider the following concatenation:

$$\Gamma = \gamma_\zeta \otimes_{i=1}^{K} \gamma_{\xi_{a_i} - \mu_{a_1}}.$$

Now we indeed have $\operatorname{logsig}_l(\Gamma) = \xi + \zeta$. Furthermore, by Lemma 6.3.4,

$$|\Gamma| \leq M_{l-1}|\zeta|_{\mathcal{L}^{l-1}(\mathbb{R}^d,\mathbb{R})} + \left(\sum_{|\bar{\alpha}|=l} 2^l (M_{l-1}|\zeta|_{\mathcal{L}^{l-1}(\mathbb{R}^d,\mathbb{R})})^l + C_l^l |\xi_{[\alpha]}|\right)^{\frac{1}{l}}.$$

The result follows. $\square$

We will need the following lemma:

THEOREM 7.0.1. *Let $\mathcal{A} = (x, \sigma, F)$ be an (L, g, G)- tense, (L, H_l)-weak Hörmander system. Fix $l \geq L$. Suppose that $g \geq l + 3$. There exist polynomial constants C_1, C_2 and C_3 such that for all $y \in \mathbb{R}^n$ such that $\max(d_t(y), \sqrt{t}) < C_1$, we have*

$$d_t^l(y) + \sqrt{t} \leq C_2 d_t(y),$$

and for all $y \in \mathbb{R}^n$ such that $\max(d_t^l(y), \sqrt{t}) < C_1$,

$$d_t(y) \leq C_3(d_t^l(y) + \sqrt{t}).$$

PROOF. *First inequality* Let $\Gamma \in \mathcal{P}_t^{d+1}$ be a path such that $Y_t(\Gamma) = y$, $\Gamma_t^{(0)} = t$ and $|\Gamma| \leq d_t(y)(1 + \varepsilon)$ where as usual ε is a fixed quantity greater than $\frac{1}{2}$. Let $S = \operatorname{logsig}^-(\Gamma)$.

From the usual Taylor expansion, and Theorem 6.3.1, we have the following, where M_1 is a constant depending only on G, n, H_L:

$$\begin{aligned} |Y_t(\Gamma) - T_t^l(\Gamma)| &\leq M_1|\Gamma|_{L^2,1}^{l+1} \\ &\leq M_1(d_t(y)(1+\varepsilon))^{l+1}. \end{aligned} \tag{7.0.2}$$

We also have

$$|S| \leq |(t, S)|_{\mathcal{L}^l(\mathbb{R}^d,\mathbb{R})} \leq M_2(d_t(y)(1+\varepsilon) + \sqrt{t})$$

for some polynomial constant M_2. Let ρ be the ρ obtained from our development on Models, relative to the model $(\mathcal{S}, F_t^{\log(S)T}, \mathbb{R}^n)$. As long as $t \leq 1/2$, this is clearly a polynomial quantity (in fact, it is proper, as can be seen from Proposition 7.0.9 and the definition of the weak Hörmander constant H_L) We know that for any $S \in \mathcal{S}$ such that $|s| \leq \rho$, we have

$$\beta_{F_t^{\log(S)T}(S+\boldsymbol{\cdot})} \geq \frac{\beta_{F_t^{\log(S)T}}}{2}.$$

Also, $\rho_S = \rho(\mathcal{L}^l(\mathbb{R}^{d+1}), F_t^{\log(S)T}(S + \boldsymbol{\cdot}), \mathbb{R}^n) \geq \Theta$ for any S with $|S| \leq \rho$ and for some polynomial quantity Θ.

Now, by Theorem 5.2.2, we know that for any $v \in \mathbb{R}^n$ such that $|v| \leq \underline{\beta}(S)\rho_S$, there exists an $s \in \mathcal{S}$ with $|s| \leq \frac{4|v|}{\underline{\beta}(S)}$ such that

$$F_t^{\log(S)T}(S + s) = T^l(\Gamma) + v. \tag{7.0.3}$$

Therefore, we set $C_1 = \min(\frac{\rho_0}{2(1+\varepsilon)M_2}, \frac{\Theta\underline{\beta}(0)}{2M_1}, 1/\sqrt{2})$. This implies $M_2(d_t(y)(1+\varepsilon) + \sqrt{t}) < \rho_0$ and $M_1 d_t(y)(1+\varepsilon) < \frac{\underline{\beta}_0\Theta}{2} < \underline{\beta}_S\Theta$. By Eq. (7.0.2), this ensures that we can apply the above with $v = Y_t(\Gamma) - T^l(\Gamma)$. With

$$|s| \leq \frac{4|Y_t(\Gamma) - T_t^l(\Gamma)|}{\underline{\beta}(S)} \leq M_3 d_t(y)^{l+1}$$

for some polynomial constant M_3.

By Lemma 5.2.3, this implies that $|s|_{\mathcal{L}^l(\mathbb{R}^{d+1})} \leq M_4 d_t(y)^{\frac{l+1}{l}}$ (for some polynomial constant M_4). Then by Lemma 5.1.2, we have that $|S+s|_{\mathcal{L}^l(\mathbb{R}^{d+1})} \leq 2d_t(1+\varepsilon+M_4)$. Again from (7.0.3), this allows us to conclude that

$$d_t^l(y) \leq (2M_4+2)d_t(y).$$

Noting that trivially, $\sqrt{t} \leq d_t(y)$, we get our inequality with $C_2 = (2M_4+2)$. Since our constant is polynomial in $\beta, \underline{\beta}$ etc., it is polynomial in G,d, H_l.

Second inequality.

The spirit of the proof is the same, but the infinite dimensionality of path space means that we must actually repeat and adapt the steps of the proof of our inverse functions Theorem 5.2.2. We use the representation given in Eq. (7.0.1) from Proposition 7.0.4. The requirement $\Gamma_1 = 1$ when Γ is parametrised over $[0,1]$ will be achieved by making sure that the first iteration has ${}^1\Gamma_1^0 = 1$ and all the other iterations have 0 as $e^{(0)}$ components of their log signatures.

First note that there clearly exists a polynomial R such that for any $x' \in \mathbb{R}^m$ and $t_1 \in [0,t]$ such that $|x'-x| \leq R$, we have $H_L(x') > \frac{H_L(x)}{2}$, which will imply that $\rho(x'), \underline{\beta}(x'), \beta(x')$ (i.e. the $\rho, \underline{\beta}$ and β relative to the model $(\mathcal{S}, \mathbb{R}^n, F_{x',t_1}^{\log(S)T})$) are controlled by polynomial quantities. There is also a $\bar{R}$, uniform over all x' such that $|x'-x| \leq R$, such that for any $y \in \mathbb{R}^n$ such that $|y - F(x')| \leq \bar{R}$, we can apply Lemma 5.2.2 to obtain an s such that $F_{x',t_1}^{\log(S),T}(s) = y$.

Let C_{PLR} be the constant from Theorem 5.2.3, involved in the application of the implicit functions theorem.

Let D be such that the original system $\mathcal{A}$ is (g,g,D)-tense. This clearly a polynomial constant.

Let C_{constr} be the constant from Lemma 7.0.11.

Let $y \in \mathbb{R}^n$ with $\min(t, d_{x,t}^l(y)) \leq C_1$ for

$$C_1 = \min(\frac{R}{2DC_{\text{constr}}}, \frac{\bar{R}}{2D(1+\varepsilon)^{l+1}C_{\text{constr}}}, C_{\text{constr}}^{-1}, \frac{D^{-1+l}}{2C_{\text{PLR}}^l}(1+\varepsilon)^{-1-l}C_{\text{constr}}^{-1-2l}).$$

Here, $K = \sup_{|s|\leq 1}\sup_{N\in\mathbb{N}} |\partial^N F_{x,t}^{\log(S)T}|$, where $||$ denotes the operator norm. This is clearly a polynomial quantity by the assumptions on the system, and the (polynomial) form of the exponential function restricted to the truncated signature space.

Let ${}^1S \in \operatorname{span}_{\alpha\neq(0),|\alpha|\leq l}(e^{[\alpha]}) \subset \mathcal{S}_l$ be such that $F_{x,t}^{\log(S),T}(S) = y$ and $|{}^1S|_{\mathcal{S}} \leq (1+\varepsilon)d_t^l(y)$. (${}^1S$ exists by the finiteness of C_1). By Lemma 7.0.11, there exists a path ${}^1\Gamma = ({}^1\tau, {}^1\gamma) \in \mathcal{P}_1^{d+1}$ such that

$$\operatorname{logsig}({}^1\Gamma) = (1, \delta_{\frac{1}{t},1}({}^1S)),$$

where $\delta_{\frac{1}{t},1}$ denotes the (linear) dilation such that $\delta_{\frac{1}{t},1}(e^\alpha) = \frac{e^\alpha}{t^{o_0(\alpha)}}$ for each α, and

$$|{}^1\Gamma|_{L^2,1} \leq C_{\text{constr}}\max(t, d_t(y)).$$

Now we have, again by the usual Taylor expansion and the results of the previous section:

$$\begin{aligned}
|F(\mathrm{Sol}_{x,t}({}^1\Gamma)) - y| &\leq D|{}^1\Gamma|_{L^2,1}^{l+1} \\
&\leq D(1+\varepsilon)^{l+1}C_{\text{constr}}^{l+1}\max(\sqrt{t}, d_t^l(y))^{l+1} \\
|\mathrm{Sol}_{x,t}({}^1\Gamma) - x| &\leq D|{}^1\Gamma|_{L^2,1} \\
&\leq DC_{\text{constr}}\max(\sqrt{t}, d_t^l(y))
\end{aligned} \tag{7.0.4}$$

for some polynomial M, and where C_{constr} is the constant from Proposition 7.0.12.

We define $x_1 = x$ and $x_2 = \mathrm{Sol}_{x,t}({}^1\Gamma)$.

Now, the second equation in (7.0.4) ensures that

$$|x_1 - x| \leq DC_1C_{\text{constr}} \leq \frac{R}{2}$$

and the first part ensures that $|F(x_1) - y| \leq D(1+\varepsilon)^{l+1} C_{\text{constr}} C_1 \leq \frac{\bar{R}}{2}$ and there exists an ${}^2S \in \mathcal{S}_l$ such that

$$F^{\log(S),T}_{x_1,t}({}^2S) = y,$$

and then a path ${}^2\Gamma \in \mathbb{P}^{d+1}_1$ with

$$\operatorname{logsig}({}^2\Gamma) = (0, \delta_{\frac{1}{t},1}({}^2S)).$$

Continuing in this way, we define and obtain, using Lemma 5.2.3:

$$\begin{aligned}
x_n - x_{n-1} &= \operatorname{Sol}_{x_{n-1},t}({}^n\Gamma) \\
\operatorname{logsig}({}^n\Gamma) &= (0, \delta_{\frac{1}{t},1}({}^nS)) \quad \text{for} \quad n > 1 \\
y &= F^{\log(S),T}_{x_n,t}({}^nS) \\
|{}^nS|_{\mathcal{S}_l} &\leq C_{\text{PLR}} D^{\frac{1}{l}} (1+\varepsilon)^{\frac{l+1}{l}} C_{\text{constr}}^{\frac{1+l}{l}} |{}^{n-1}S|_{\mathcal{S}}^{\frac{l+1}{l}} \\
&\leq \frac{1}{2} |{}^{n-1}S|_{\mathcal{S}} \quad \text{for} \quad n > 1 \\
|{}^n\Gamma|_{L^2,1} &\leq C_{\text{constr}} D C_{\text{PLR}} D^{\frac{1}{l}} (1+\varepsilon)^{\frac{l+1}{l}} C_{\text{constr}}^{\frac{1+l}{l}} |{}^{n-1}\Gamma|_{L^2,1} \\
&\leq \frac{1}{2} |{}^{n-1}\Gamma|_{L^2,1} \quad \text{for} \quad n > 1
\end{aligned}$$

where at the last lines, we have used the facts that $C_1^{\frac{1}{l}} C_{\text{PLR}} D^{\frac{1}{l}} (1+\varepsilon)^{\frac{l+1}{l}} C_{\text{constr}}^{\frac{l+1}{l}} < \frac{1}{2}$ and $C_1^{\frac{1}{l}} C_{\text{PLR}} D^{\frac{1}{l}} (1+\varepsilon)^{\frac{l+1}{l}} C_{\text{constr}}^{\frac{l+1}{l}} D C_{\text{constr}} < \frac{1}{2}$.

Note that we have

$$\begin{aligned}
\sum_n |x_n - x_{n-1}| \leq D \sum_n |{}^n\Gamma|_{L^2,1} \leq 2 D C_{\text{constr}} C_1 \leq R \quad \text{and} \\
|y - F(x_n)| \leq |y - F(x)| \leq \bar{R},
\end{aligned}$$

which means we can indeed apply our implicit functions Theorem 5.2.2 at each step.

Now, we set:

$$\Gamma = \otimes_{n=1}^{\infty} {}^n\Gamma.$$

We have $|\Gamma|_{L^2,1} \leq 2C_{\text{constr}} \max(\sqrt{t}, d^l_t(y))$ and if we reparametrise Γ so that it is parametrised over $[0,1]$, we get that $\operatorname{Sol}_{x,t}(\Gamma) = y$, and $\Gamma^0_1 = 1$ this concludes the proof.

REMARK 7.0.13. In fact, it is easy to convince oneself that the constants involved in the proof of the first inequality can be made proper at the cost of assuming that the original system is (g, g, G)-tense. The constants in the second inequality are only polynomial because of the construction in the proof of Proposition 7.0.11.

□

COROLLARY 7.0.14. *Fix $l \geq L$. Let $\mathcal{A} = (x, \sigma, F)$ be a (g, g, G)-tense, (L, H_L)-weak Hörmander system. Suppose $g \geq l+3$. There exist proper constants $D, \bar{D}, K$ and a polynomial constant $\bar{K}$ such that for any $t \leq D$ and any y such that $|{*x} - y| \leq \bar{D}$, we have:*

$$\begin{aligned}
|d^l_t(x,y)| &\leq K|*\bar{x}_t - y|^{\frac{1}{l}} \qquad \textit{and} \\
|d_t(x,y)| &\leq \bar{K}(|*\bar{x}_t - y|^{\frac{1}{l}} + \sqrt{t}).
\end{aligned}$$

PROOF. This is a consequence of Theorems 7.0.1 and 5.2.2. Indeed, the inverse operator norm of the Jacobian of $F^{\log(S),T}$ is a proper constant (uniformly in t), we also know from the Taylor series expansion of the exponential function on $\mathcal{L}(\mathbb{R}^{d+1})$ and the algebraic structure (similarly to Proposition 7.0.9)that the derivatives in any direction are bounded by a proper constant:

For $v \in \mathcal{L}^l(\mathbb{R}^d, \mathbb{R})$ and $x \in \mathcal{L}^l(\mathbb{R}^d, \mathbb{R})$ with $|x| \leq 1$, we have

$$\frac{\partial F^{\log(S),T}}{\partial v}(x) = JF^{ST}\frac{\partial \exp(x)}{\partial v}(x).$$

We will show that the RHS is bounded by a proper constant for $|x|, |v| \leq 1$. By abuse of notation, we will write x for (t, x) and $x^{(0)} = t$. This means that in articular we are assuming that $t \leq 1$, which is not a problem as 1 is a proper constant. It suffices to prove this for $v \in \mathcal{H}_L$ our basis of $\mathcal{L}^l(\mathbb{R}^d, \mathbb{R})$. Since any $b \in \mathcal{H}_L$ can be expressed in $T^l(\mathbb{R}^{d+1})$ as a linear combination of elements of the form e^α with the number of terms being less than a proper constant, it is enough to show the bound where the exponential function is replaced by its formal extension (via the Taylor series) on the whole of $T^l(\mathbb{R}^{d+1})$, and for $v = e^\alpha$ for some $\alpha \in \text{Multi}(\{0, 1, \ldots, d\})$. Since JF^{ST} is bounded by proper constants, it is enough to bound $\frac{\partial \exp(x)}{\partial v}(x)$

We now have the following calculation (from the series expansion of the exponential):

$$\begin{aligned}
&\frac{\partial \exp}{\partial e^\alpha}(x) \\
&= \sum_{n=1}^{l} \frac{1}{n!} \Pr_{\mathcal{T}^l(\mathbb{R}^d,\mathbb{R})} \left(\sum_{i=1}^{n} \sum_{\substack{\beta_1,\ldots,\beta_{i-1},\beta_{i+1},\ldots,\beta_n \\ \in \text{Multi}(\{0,1,\ldots,d\})}} \Pi_{j=1}^{i-1} x^{\beta_j} \Pi_{j=i+1}^{n} x^{\beta_j} e^{(\beta_1,\ldots,\beta_{i-1},\alpha,\beta_{i+1},\ldots,\beta_n)} \right).
\end{aligned}$$

For multi-indices α, γ and for any natural number n, let $K(n, \alpha, \gamma)$ be the set of $(n-1)$-tuples of natural numbers that can be written $(\kappa_1, \ldots, \kappa_{i-1}, \kappa_{i+1}, \ldots \kappa_n)$ in such a way that if we set $\kappa_i = \#(\alpha)$, then $\sum_{j=1}^{n} \kappa_j = \#(\gamma)$ and

$$(\gamma_{\sum_{j=1}^{i-1}\kappa_j+1}, \gamma_{\sum_{j=1}^{i-1}\kappa_j+2}, \ldots, \gamma_{\sum_{j=1}^{i-1}\kappa_j+\#(\alpha)}) = \alpha$$

The maximal cardinality of $K(n, \alpha, \gamma)$ when $|\alpha|, |\gamma| \leq L$ is a combinatorial function of L only, which we write K_L.

Write also $\beta(\kappa, \gamma, j)$ for the multi-index composed of the components of γ with positions $\sum_{k=1}^{j-1} \kappa_k + 1, \ldots, \sum_{k=1}^{j} \kappa_k$. We have $e^{(\beta(\kappa,\gamma,1),\ldots,\beta(\kappa,\gamma,n))} = e^\gamma$.

Now, we have, for any γ,

$$\begin{aligned}
&\langle \frac{\partial \exp(x)}{\partial e^\alpha}(x), e^\gamma \rangle \\
&= \sum_{\kappa \in K(n,\alpha,\gamma)} (\Pi_{j \neq i(\kappa)} x^{\beta(\kappa,\gamma,j)}).
\end{aligned}$$

Finally, we have

$$\begin{aligned}
&|\frac{\partial \exp(x)}{\partial e^\alpha}(x)|^2 \\
&= \sum_\gamma (\sum_{\kappa \in K(n,\alpha,\gamma)} (\Pi_{j \neq i(\kappa)} x^{\beta(\kappa,\gamma,j)}))^2 \\
&\leq K_L^2 \sum_\gamma \sum_{\kappa \in K(n,\alpha,\gamma)} (\Pi_{j \neq i} x^{\beta(\kappa,\gamma,j)})^2 \\
&\leq K_L^2 L \sum_{n=1}^{L} \sum_{\sum_{j=1}^{n-1} k_j + \#(\alpha) \leq L} \Pi_{j=1}^{n-1} (\sum_{\#(\beta)=k_j} (x^\beta)^2) \\
&\leq L K_L^4 |x|^2 \leq L K_L^4
\end{aligned}$$

where at the last lines, we have used the fact that

$$\sum_{\#(\beta)=k_j} (x^\beta)^2 \leq |x|^2 \leq 1.$$

It follows that the model $(0, \mathrm{Pr}_{\mathcal{L}^l(\mathbb{R}^d,\mathbb{R})}(v_f), F^{\log(S),T})$ (where v_f denote the vector fields defined by $v_f^{(\alpha,i)}(x) = x^\alpha$) satisfies the standard boundedness and degeneracy assumptions with proper constants. We can therefore use Theorem 5.2.2 to deduce that there exist proper constants $\bar{D}$ and A such that for any y such that $|{*x} - y| \leq \bar{D}$, there exists an element $s \in \mathcal{L}^l(\mathbb{R}^d, \mathbb{R})$ such that $F^{\log(S),T}(s) = y$ and $|s| \leq A|{*x} - y|$. Since $|s|_{\mathcal{L}^l(\mathbb{R}^d,\mathbb{R})} \leq |s|^{\frac{1}{l}}$, we immediately deduce the first inequality.

The second inequality now follows from Theorem 7.0.1 (the constant D is comes from the application of that theorem). □

There is an alternative way of constructing a continuously defined version of d_t^l:

DEFINITION. We define the $d_{t,l,\infty}$ distance as follows:

$$d_{t,l,\infty}(y) = \inf(|\operatorname{logsig}(\Gamma)|_{\mathcal{L}^l(\mathbb{R}^d,\mathbb{R})}|Y_t(\Gamma) = y).$$

We note the following important consequence of the above corollary:

PROPOSITION 7.0.15. *Let $\mathcal{A} = (x, \sigma, F)$ be a uniformly (L', g, G)-tense, uniformly (L, H_L)-weak Hörmander system. Fix $l \geq L$ and suppose that $L' \geq l + 2$, $g \geq l + 3$. There exists a polynomial constant C and a proper constant M such that we have, for any $y \in \mathbb{R}^n$:*

$$d_t(x, y) \leq C(|{*x} - y| + 1)$$
$$d_t^\infty(x, y) \leq M(|{*x} - y| + 1).$$

PROOF. Let $N = \lceil |{*x} - y| K \rceil$, where K is the constant from Corollary 7.0.14. Then for $i = 0, 1, 2, \ldots, N$, define $y_i = {*x} + \frac{i|{*x}-y|}{N}$. By Corollary 7.0.14, we can iteratively construct x_i such that $d_t(x_i, y_{i+1}) \leq C|x_i, x_{i+1}|^{\frac{1}{L}} \leq C$ for some polynomial constant C. Then we have $d_t(x, y) \leq CN \leq C(|{*x} - y| + 1)$. The second inequality is proved similarly. (The only real source of non-proper dependence in the first inequality comes from the construction of a control with given log signature(and the calculation of its homogeneous length), which is no longer required when defining the distance directly in terms of homogeneous norms in the log signature space.) □

We finish with the following doubling condition:

PROPOSITION 7.0.16. *Let $\mathcal{A} = (x, \sigma, F)$, F linear, be a uniformly (L, g, G)-mixed tense, uniformly (L, H_L) weak Hörmander system with $g \geq L + 3$. There exist constants D and M such that for any $t \leq D$ any $r > 0$ and any $\bar{x} \in \mathbb{R}^m$,*

$$|B_{d_t}(x, 2r)| \leq M|B_{d_t}(x, r)|.$$

PROOF. This follows from, Theorem 7.0.1, and Propositions 5.1.11, 5.2.8, and 7.0.15. □

Using a similar method of proof to Theorem 7.0.1, and using Proposition 7.1.6, we can show the following:

THEOREM 7.0.2. *Let $\mathcal{A} = (x, \sigma, F)$ be an (L', g, G)- tense, (L, H_l)-weak Hörmander system. Fix $l \geq L$. Suppose $L' \geq l + 2$, $g \geq l + 3$. There exist proper constants C_1, C_2 and C_3 such that $\forall y \in \mathbb{R}^n$ s.t. $\max(d_t(y), \sqrt{t}) < C_1$, we have*

$$d_{t,\infty,l}(y) \leq C_2 d_t^l(y)$$

and $\forall y \in \mathbb{R}^n$ s.t. $\max(d_t^l(y), \sqrt{t}) < C_1$

$$d_{t,\infty,l}(y) \leq C_3(d_t^\infty(y))$$

Because all the d_t^l for $L \leq l \leq L_2$ are properly equivalent (for fixed L_2, this follows from our inverse functions Theorem 5.2.2), and the only choices for l that we are interested in are $l = L$ and $l = (2n + 2)^2 2^{4L}$, we can pick $l = L$ and write $d_{t,\infty}$ for $d_{t,\infty,l}$.

7.1. Systems satisfying the Progressive Hörmander condition

DEFINITION. The system $\mathcal{A} = (x, F, \sigma)$ satisfies the Progressive Hörmander condition (at x) with constants (L, H_L) if for any unit $u \in \mathcal{L}^L(\mathbb{R}^d, \mathbb{R})$ and for any unit $v \in \mathbb{R}^n$, there exist $\lambda_{\alpha,\beta} \in \mathbb{R}$ and $\lambda_{v,\beta}$ such that

$$\sum_{\substack{|\beta|\leq|\alpha|\\ \beta\neq[0]}} \lambda_{\alpha,\beta} * \sigma^{[\beta]} = *\sigma^u \quad \text{and}$$

$$\sum_{\substack{|\beta|\leq L\\ \beta\neq[0]}} \lambda_{v,\beta} * \sigma^{[\beta]} = v$$

and for any unit $u \in \mathcal{L}^L(\mathbb{R}^d, \mathbb{R})$ and for any unit $v \in \mathbb{R}^n$

$$\sum_{|\beta|\leq|\alpha|} \lambda_{v,\beta}^2 \leq H_L \quad \text{and}$$

$$\sum_{|\alpha|\leq L} (\sum_{|\beta|\leq L} \lambda_{\alpha,\beta} u_\beta)^2 \leq H_L^{-1}$$

REMARK 7.1.1. Note that by Proposition 7.0.9, the requirement

$$\sum_{|\alpha|\leq L} (\sum_{|\beta|\leq L} \lambda_{\alpha,\beta} u_\beta)^2 \leq H_L^{-1}$$

is, up to a proper constant, equivalent to requiring that

$$\sum_{\substack{|\alpha|\leq L\\ \beta\in H}} \langle \lambda_{\alpha,\cdot}, u\rangle^2 \leq H_L^{-1},$$

where H denotes the basis of $\mathcal{S}$ defined in Proposition 7.0.9.

REMARK 7.1.2. Up to a strongly polynomial constant, this definition is equivalent to the following:

- The Hörmander condition holds with constant H, *and*
- for any multi-index α containing non zero indices and with $|\alpha| \leq L$ there exist real numbers $\lambda_{\alpha,\beta}$ such that

$$\sum_{\substack{|\beta|\leq|\alpha|\\ \beta\neq[0]}} \lambda_{\alpha,\beta} * \sigma^{[\beta]} = *\sigma^\alpha \quad \text{and} \sum_\beta (\lambda_{\alpha,\beta})^2 \leq H^{-1}$$

REMARK 7.1.3. Proposition 3.5.1 shows that the progressive Hörmander condition implies the Hörmander condition.

For the whole of this section, we will use the following notation:

$$\bar{x}^t = x + \sum_{\substack{\alpha\in\mathrm{Multi}(\{0\})\\ |\alpha|\leq L}} \sigma^\alpha(x) \frac{t^{|\alpha|/2}}{(|\alpha|/2)!}.$$

The following 'log-homogeneous distance' can be defined for any system, but is meaningful (i.e. equivalent to the some control distance) only under the progressive Hörmander condition.

DEFINITION 7.1.4. Let $\mathcal{A} = (x, F, \sigma)$ be a system, the log-homogeneous distance $d(\cdot, \cdot)$: $\mathbb{R}^m \otimes \mathbb{R}^n \to \mathbb{R}$ is defined by

$$d_{t,\log}(x, y) = \inf\Big(\sqrt{\sum_{i\leq L} \big(\sum_{\substack{\beta\in H\\ \beta\neq[0];|\beta|=i}} \lambda_\beta^2\big)^{\frac{1}{i}}} : \sum_{\beta\in\mathcal{L}_L} \lambda_\beta \sigma^\beta = y - \bar{x}^t\Big).$$

REMARK 7.1.5. By Proposition 7.0.9, the above is equivalent, up to a proper constant, to the following definition:

$$d_{t,\log,\perp}(x,y) = \inf\Big(\sqrt{\sum_{i\le L}(\mathrm{Pr}_{U_i}(\lambda))^{\frac{1}{i}}} : \sum_{\beta\in\mathcal{L}_L} \lambda_\beta \sigma^{[[\beta]]} = y - \bar{x}^t\Big),$$

where $U_i = \mathrm{span}_{|\alpha|=1;\alpha\neq(0)}\, e^{\alpha}$. Here $\mathcal{L}_L$ is the set of Lyndon words[2] of *order* less than L, and for a Lyndon word α, $[[\alpha]]$ denotes the standard bracketing of α.

7.1.1. The compensated signature. Let $\mathcal{A} = (x,\sigma,F)$ be a (L,H_L)-progressive Hörmander system. Let

$$F^{SR_r} : T^L(\mathbb{R}^d,\mathbb{R}) \to (\mathbb{R}^n)^{\otimes L},\ s \mapsto \otimes_{i=1}^{L}\Big(\sum_{\substack{|\alpha|=i\\ \alpha\neq(0)}} \sigma^{\alpha} s^{\alpha}\Big)$$

For each $\alpha \in T^L(\mathbb{R}^{d+1})$ with $|\alpha| \le L$ and $\alpha \neq (0)$, pick $\lambda_{\alpha,\cdot} \in \mathcal{S}$ such that

$$\begin{aligned}
\langle \lambda_{\alpha,\cdot}, (*\sigma)^{\cdot}\rangle &= *\sigma^{\alpha},\\
\mathrm{Pr}_{\mathrm{span}_{|\beta|\ge|\alpha|+1}\, e^{[\beta]}}(\lambda_{\alpha,\cdot}) &= 0 \quad \text{and}\\
\mathrm{Pr}_{\mathrm{Ker}(F^{SR_r})}(\lambda_{\alpha,\cdot}) &= 0.
\end{aligned}$$

Here $\lambda_{\alpha,\cdot}$ is the Moore-Penrose pseudo-inverse of $*\sigma^{\alpha}$ by the map

$$F^{SR_r} : T^L(\mathbb{R}^d,\mathbb{R}) \to (\mathbb{R}^n)^{\otimes L},\ s \mapsto \sum_{\substack{|\beta|\le|\alpha|\\ \beta\neq(0)}} \sigma^{\beta} s^{\beta}.$$

We can then extend linearly to obtain a function

$$\Psi : T^L(\mathbb{R}^{d+1})/\,\mathrm{span}(e^{(0)}) \to \mathcal{S}/\,\mathrm{Ker}(F^{SR_r}),$$

(where $\mathcal{S}/\,\mathrm{Ker}(F^{SR_r})$ denotes the orthogonal complement of $/\,\mathrm{Ker}(F^{SR_r})$ in $\mathcal{S}$) such that, writing H for the orthonormal basis of $\mathcal{S}$ from Proposition 7.0.9, for any $s \in T^L(\mathbb{R}^d,\mathbb{R})/\,\mathrm{span}(e^{(0)})$, we have

$$\begin{aligned}
\sum_{h\in H} \Psi^h(s)\sigma^h &= \sum_{|\alpha|\le L} \sigma^{\alpha} s^{\alpha}\\
|\Psi(s)|^2 \le H_L^{-1}|s|^2 &\quad \text{and} \forall i \le L\\
s \in \mathrm{span}_{\substack{|\alpha|\le i\\ \alpha\neq(0)}}\, e^{\alpha} &\implies \Psi(s) \in \mathrm{span}_{\substack{|\alpha|\le i\\ \alpha\neq(0)}}\, e^{[\alpha]}.
\end{aligned}$$

We now define the following function:

$$\Psi_1 : T^L(\mathbb{R}^d,\mathbb{R})/\,\mathrm{span}(e^{(0)}) \to \mathcal{S}/\,\mathrm{Ker}(F^{SR_r}),\ s \mapsto \mathrm{Pr}_{\mathcal{S}/\,\mathrm{Ker}(F^{SR_r})}(s) + \Psi(s - \mathrm{Pr}_{\mathcal{S}}(s))$$

Note that Ψ_1 and Ψ depend on the initial point x and on the tensor $\sigma^{\cdot}$.

Now, by the progressive Hörmander condition, we can also define the following function:

$$\Psi_2 : \mathbb{R}^n \to \mathcal{S}/\,\mathrm{Ker}(F^{SR_r}),\ v \mapsto \Psi_2(v),$$

such that

$$\sum_{h\in H} \Psi_2^h(v)\sigma^h = y - \bar{x}_t,$$

and Ψ_2 is bounded in operator norm by H_L.

Writing S for the l^{th} order signature, we can now define the l^{th} order *(finite order) compensated signature* R_l by

$$\begin{aligned}
\tilde{S} &:= \mathrm{Pr}_{T^L(\mathbb{R}^{d+1})/\,\mathrm{span}(e^{(0)})}(S)\\
R_l &:= \Psi_1(\tilde{S}) + \Psi_2(F^{ST}(S-\tilde{S}) - \bar{x}_t),
\end{aligned}$$

[2]For an excellent explanation of the relevance of this concept to log-signatures, see **[51]**

and the *compensated signature* R_∞ by

$$R_\infty := \Phi_1(\tilde{S}) + \Psi_2(X_t - F^{ST}(\tilde{S})). \tag{7.1.1}$$

We write $\mathcal{F}$ for the reduced signature space $\operatorname{span}_{\substack{\alpha\neq(0)\\|\alpha|\leq L}} e^{[\alpha]}/\operatorname{Ker}(F^{SR_r})$, and F^{RT} for the restriction of F^{ST} to $\mathcal{F}$.

Since the quantity R_∞ only depends on X and $\log(S)$, we can define the function

$$F^{\log(S)X,R_\infty} : (\log(S), X) \to R_\infty$$

as the function giving the reduced signature as a function of the log signature and the solution process evaluated at time t. We can define $F^{S,X,R_\infty} : (S, X) \to R_\infty$ similarly.

The above definitions of R_∞ and R_l ensure that

$$F^{RT}(R_\infty) = Y_t.$$

and

$$F^{RT}(R_l) = \bar{Y}_t = \bar{x}_t + \sum_{|\alpha|\leq l} \sigma^\alpha \sigma^\alpha.$$

7.1.2. A continuously defined version of the log homogeneous distance. In this section, we assume that the progressive Hörmander condition holds uniformly.

We begin with a few notational simplifications.

We write, throughout this Part of the manuscript, $B = \cup_{i,k}\{h_k^i\}$ for the orthornormal basis of $\mathcal{L}^l(\mathbb{R}^d, \mathbb{R})$ that we have defined in Proposition 7.0.9, and H for $B \setminus \{e^{(1)}\}$.

In this section we will write $a_1, a_2, \mathit{\Delta}_1, \mathit{\Delta}_2 ...$ for elements of $\mathcal{L}^l(\mathbb{R}^d, \mathbb{R})$. Then $\mathit{\Delta} = (\mathit{\Delta}_1, ...\mathit{\Delta}_o)$ is a vector in $(\mathcal{L}^l(\mathbb{R}^d, \mathbb{R}))^{\otimes o}$.

We write $|a| = i$ if $a = h_j^i$, $||$ is the homogeneous degree. We will also write $|\mathit{\Delta}| = \sum_{i=1}^o |\mathit{\Delta}_o|$.

PROPOSITION 7.1.6. *Let $U = \oplus_{i=1}^I U_i$, where $U_i = \mathbb{R}^{\nu_i}$ be a graded space and let us define $\nu = \sum_{i=1}^I \nu_i$.*

Suppose we are given the natural orthonormal basis $B = \{b_1, b_2 \dots, b_\nu\}$ of U such that $\{b_1, b_2, \dots, b_{\nu_1}\}$ is an orthonormal basis of U_1, etc. Furthermore, for any smooth path $\gamma \in \mathcal{P}^U_{[0,1]}$, we define the following homogeneous energy:

$$|\gamma_{L^2,U}| = (\sum_{i=1}^I (\int_0^1 |\frac{\partial \gamma_t^{U_i}}{\partial t}|^2)^{\frac{1}{i}})^{\frac{1}{2}}.$$

Write $\operatorname{Multi}(B)$ for the set of multi-indices on B. For a multi index $\Delta = (\Delta_1, \Delta_2, \dots, \Delta_o) \in \operatorname{Multi}(B)$, we define the homogeneous degree $|\Delta| = \sum_{k=1}^o |\Delta_k|$ where by definition $|\Delta_k| = i$ for $\Delta_k \in U_i$. Write $\operatorname{Multi}(B, l)$ for the set of multi-indices in $\operatorname{Multi}(B)$ of homogeneous degree less than l.

Let $A : [0, 1]^o \otimes \operatorname{Multi}(B, l) \to \mathbb{R}^m,\ (t, u) \mapsto A_t^\Delta$ be a smooth process. Suppose also that

$$G := \max\left(\sup_{\tilde{t}\in[0,1]^o} \sup_{\substack{|v|=1\\ v\in\mathbb{R}^m}} \sum_{\Delta\in\operatorname{Multi}(B,l)} \langle A_{\tilde{t}}^\Delta, v\rangle^2_{\mathbb{R}^m}, 1 \right) < \infty.$$

As usual, for any process $B \in U$ parametrised over $[0, 1]^o$ and any $\Delta \in \operatorname{Multi}_o(B)$, we define the iterated integrals $\int^\Delta B(d\gamma)^\Delta$ iteratively by

$$(\int^b B(d\gamma)^b)_{t_2,\dots,t_o} = \int_0^1 B_{t_1,\dots,t_o} d\gamma_{t_1}^b$$

$$\int^{(\Delta,b)} (d\gamma)^{(\Delta,b)} = \int_0^1 (\int^\Delta (d\gamma)^\Delta)_{t_o} d\gamma_{t_o}^b.$$

$$\int^{(\Delta,b),t}(d\gamma)^{(\Delta,b)} = \int_0^t (\int^{\Delta}(d\gamma)^{\Delta})_{t_o} d\gamma^b_{t_o}.$$

Then there exists a constant C, dependent only on I,l, but not on the ν_i's or m, such that

$$\left| \sum_{\substack{\Delta\in\mathrm{Multi}(B)\\|\Delta|=l}} \int^{\Delta} A^{\Delta}(d\gamma)^{\Delta} \right|_{\mathbb{R}^m} \leq CG|\gamma|^l_{L^2,U}$$

(explicitly, $C = I^{\frac{l}{2}}$).

PROOF. For any $\Delta \in \mathrm{Multi}(B)$, we define $r(\Delta) \in \mathrm{Multi}(\{1,2,\ldots,I\})$ to be the multi-index in $\mathrm{Multi}(\{1,2,\ldots,I\})$ obtained by replacing each index Δ_k of Δ by $|\Delta_k|$. Fix $i_1 \in \{1,2,\ldots,I\}$ and $v \in U_{i_1}$ and $\delta \in \mathrm{Multi}(\{1,2,\ldots,I\},l)$.

Claim: If E^{Δ} is a process in $\mathbb{R}^m$ defined for each Δ, parametrised over $[0,1]$, and such that $r(\Delta)=\delta$ and

$$\sup_{\tilde{t}\in[0,1]^{|\delta|}} \sum_{r(\Delta)=\delta} \langle E^{\Delta}_{\tilde{t}}, v\rangle^2 \leq K$$

for some $K \geq 1$, for any δ of length less than o,

$$\left| \sum_{r(\Delta)=\delta} \int^{\Delta} \langle E^{\Delta}, v\rangle (d\gamma)^{\Delta} \right| \leq K|\gamma|^{|\Delta|}_{L^2,U}$$

(here $\langle\cdot,\cdot\rangle$ denotes the Euclidean inner product).

Proof of claim: We proceed by induction over the length o of $\delta = (\delta_1,\ldots,\delta_o)$ (not the order). For $o=1$ we have, by Cauchy-Schwarz,

$$\begin{aligned}
\left| \sum_{\Delta\in B\cap U_\delta} \int_0^1 \langle E^{\Delta}_t, v\rangle d\gamma^{\Delta}_t \right| &\leq \left(\sum_{\Delta\in B\cap U_\delta} \int_0^1 \langle E^{\Delta}_t, v\rangle^2 dt \right)^{\frac{1}{2}} \left(\int_0^1 \sum_{\Delta\in B\cap U_\delta} |\frac{\partial\gamma^{\Delta}_t}{\partial t}|^2 dt \right)^{\frac{1}{2}} \\
&= \left(\sum_{r(\Delta)=\delta} \int_0^1 \langle E^{\Delta}_t, v\rangle^2 dt \right)^{\frac{1}{2}} \left(\int_0^1 |\frac{\partial\gamma^{U_\delta}_t}{\partial t}|^2 dt \right)^{\frac{1}{2}} \\
&\leq \left(\sum_{r(\Delta)=\delta} \int_0^1 \langle E^{\Delta}_t, v\rangle^2 dt \right)^{\frac{1}{2}} |\gamma|^{|\Delta|}_{L^2,U} \\
&= \left(\sum_{r(\Delta)=\delta} \int_0^1 \langle E^{\Delta}_t, v\rangle^2 dt \right)^{\frac{1}{2}} |\gamma|^{|\delta|}_{L^2,U} \\
&\leq \sqrt{K}|\gamma|^{|\delta|}_{L^2,U} \leq K|\gamma|^{|\delta|}_{L^2,U}.
\end{aligned}$$

Now for the induction step, suppose that consider the multi-indices $\Delta = (\Delta_1,\Delta_2,\ldots,\Delta_o) = (\bar{\Delta},\Delta_o)$ in $\mathrm{Multi}(B,l)$ such that $r(\Delta)=\delta$.

We have:

$$\begin{aligned}
&\left| \sum_{r(\Delta)=\delta} \int^{\Delta} \langle E^{\Delta}, v\rangle (d\gamma)^{\Delta} \right| \\
&= \sum_{\Delta_o\in B\cap U_{\delta_o}} \int_0^1 \sum_{r(\bar{\Delta})=\bar{\delta}} \int^{\bar{\Delta},t_o} \langle E^{\bar{\Delta},\Delta_o}_{\tilde{t}}, v\rangle (d\gamma)^{\bar{\Delta}} d\gamma^{\Delta_o}_{t_o} \\
&\leq \sum_{\Delta_o\in B\cap U_{\delta_o}} \left(\left(\int_0^1 \left(\sum_{r(\bar{\Delta})=\bar{\delta}} \int^{\bar{\Delta},t_o} \langle E^{\bar{\Delta},\Delta_o}_{\tilde{t}}, v\rangle (d\gamma)^{\bar{\Delta}} \right)^2 dt_o \right)^{\frac{1}{2}} \left(\int_0^1 |\frac{\partial\gamma^{U_{\delta_o}}_t}{\partial t}|^2 dt \right)^{\frac{1}{2}} \right)
\end{aligned}$$

(by Cauchy-Schwarz)

$$\leq \sum_{\Delta_o\in B\cap U_{\delta_o}} \left(\sup_{t_o\in[0,1]} \left| \sum_{r(\bar\Delta)=\bar\delta} \int^{\bar\Delta,t_o} \langle E^{\bar\Delta,\Delta_o}_{\tilde t}, v\rangle (d\gamma)^{\bar\Delta} \right| |\gamma|^{|\Delta_o|}_{L^2,U} \right)$$

$$= |\gamma|^{|\delta_o|}_{L^2,U} \sum_{\Delta_o\in B\cap U_{\delta_o}} \left(\sup_{t_o\in[0,1]} \left| \sum_{r(\bar\Delta)=\bar\delta} \int^{\bar\Delta,t_o} \langle E^{\bar\Delta,\Delta_o}_{\tilde t}, v\rangle (d\gamma)^{\bar\Delta} \right| \right) |\gamma|^{|\delta_o|}_{L^2,U}$$

$$\leq |\gamma|^{|\delta_o|}_{L^2,U} \sum_{\Delta_o\in B\cap U_{\delta_o}} |\gamma|^{|\bar\delta|}_{L^2,U} \max \left(\sum_{r(\bar\Delta)=\bar\delta} \sup_{\tilde t\in[0,1]^o} \langle E^{\bar\Delta,\Delta_o}_{\tilde t}, v\rangle^2, 1 \right)$$

(by the induction hypothesis)

$$\leq |\gamma|^{|\delta_o|}_{L^2,U} |\gamma|^{|\bar\delta|}_{L^2,U} \max \left(\sum_{r(\Delta)=\delta} \sup_{\tilde t\in[0,1]^o} \langle E^{\Delta}_{\tilde t}, v\rangle^2 dt, 1 \right) \leq K|\gamma|^{|\delta|}_{L^2,U},$$

as expected. This proves the claim.

Proof of the theorem:

We have immediately:

$$\left| \sum_{\substack{\Delta\in\mathrm{Multi}(B)\\ |\Delta|=l}} \int^{\Delta} A^{\Delta}(d\gamma)^{\Delta} \right|^2_{\mathbb{R}^m} = \left| \sum_{|\Delta|=l} \int^{\Delta} A^{\Delta}(d\gamma)^{\Delta} \right|^2$$

$$\leq \sum_{|\delta|=l} \left| \sum_{r(\Delta)=\delta} \int^{\Delta} A^{\Delta}(d\gamma)^{\Delta} \right|^2$$

$$\leq \sum_{|\delta|=l} |\gamma|^{2|\delta|}_{L^2,U} (\sup_{\substack{\tilde t\in[0,1]^o, |\delta|=l\\ v\in\mathbb{R}^m, |v|=1}} \sum_{r(\Delta)=\delta} \langle A^{\Delta}_t, v\rangle^2)^2$$

$$\leq \sum_{|\delta|=l} |\gamma|^{2l}_{L^2,U} G^2 = I^l |\gamma|^{2l}_{L^2,U} G^2.$$

This is the required inequality with $C = I^{l/2}$. □

We define the following continuous global version of the log homogeneous distance:

DEFINITION. Let $\mathcal{P}_{x,y}$ be the set of smooth paths

$$\theta : [0,1] \to \mathcal{L}^l(\mathbb{R}^{d+1}),\ t \mapsto \theta_t,$$

such that the solution $x_t \in \mathbb{R}^n$ of the following ODE:

$$\begin{aligned} x_0 &= x, \\ dx_t &= \sum_{a\in B} (*\sigma)^a(x_t) \frac{\partial \theta^a_t}{\partial t} dt \\ &= \sum_{a\in B} *\sigma^a(x_t) \frac{\partial \theta^a_t}{\partial t} dt, \end{aligned} \tag{7.1.2}$$

satisfies $F(x_1) = y$.

Since $\mathcal{L}^l(\mathbb{R}^d, \mathbb{R})$ is a Euclidean space with a homogeneity structure, we can define, as in Lemma 7.1.6, the homogeneous energy of θ:

$$|\theta|^2_{L^2,\mathcal{L}^l(\mathbb{R}^d,\mathbb{R})} = \sum_{i=1}^{L} (\int_0^1 | \frac{\partial \theta^{U_i}_t}{\partial t} |^2 dt)^{\frac{1}{i}}$$

where $U_i = \operatorname{span}_k h^i_k = \operatorname{span}_{|\alpha|=i} \sigma^{[\alpha]}$.

We also write $\mathcal{P}_{x,y}^{[0]=ts}$ for the set of paths $\theta \in \mathcal{P}_{x,y}$ such that

$$\theta_s^{[0]} = ts \quad \forall \quad 0 \leq s \leq 1.$$

Then the continuously defined log-homogeneous distance $d_{t,\log,\infty}$ is:

$$d_{t,\log,\infty}(x,y) = \inf_{\theta \in \mathcal{P}_{x,y}^{[0]=ts}} (|\theta|_{L^2, \mathcal{L}^l(\mathbb{R}^{d+1})}),$$

where as usual $||_{\mathcal{L}^l(\mathbb{R}^{d+1})}$ denotes the homogeneous norm on the space $\mathcal{L}^l(\mathbb{R}^{d+1})$.

REMARK 7.1.7. If the system $\bar{\mathcal{A}} = (x, \sigma, Id)$ also satisfies the progressive weak Hörmander condition, then

$$d_{t,\log,\infty}(x,y) = \inf_{z \in \mathbb{R}^m, F(z)=y} (d_{t,\log,\infty}(x,z)).$$

We note the following trivial properties of d_t, similar to Propositions 7.0.6 and 7.0.7:

PROPOSITION 7.1.8. *Let $\mathcal{A} = (x, \sigma, F)$ be a uniformly (L, g, G)-tense, uniformly (L, H_L) weak progressive Hörmander system, for any $0 = s_0 \leq s_1 \leq s_2 \leq \ldots \leq s_N = t$ and any $x_1, x_2, \ldots, x_N$, we have*

$$d_{t,\log,\infty}(x, *x_N) \leq \sum_{i=0}^{N} d_{s_{i+1}-s_i,\log,\infty}(x_i, *x_{i+1}).$$

PROPOSITION 7.1.9. *Let $\mathcal{A} = (x, \sigma, F)$ be a uniformly (L, g, G)-tense, uniformly (L, H_L) weak progressive Hörmander system, we have, for any $x \in \mathbb{R}^m, y \in \mathbb{R}^n$ and any t,*

$$d_{t,\log,\infty}(x,y) \geq |*x - y| G^{-\frac{1}{2}}(1-t).$$

And the next important consequence:

PROPOSITION 7.1.10. *Let $\mathcal{A} = (x, \sigma, F)$ be a uniformly (L, g, G)-tense, uniformly (L, H_L)-weak progressive Hörmander system, there exists a proper constant C such that we have, for any $y \in \mathbb{R}^n$:*

$$d_t(x,y) \leq C(|*x - y| + 1).$$

PROOF. The proof is the same as the proof of Proposition 7.0.15 except that the constant is proper because the constant from Proposition 7.1.1 is proper. □

The following intuitive definition gives the idea of the detailed-Progressive Hörmander condition. The precise definition will come in 7.1.2.

DEFINITION (Intuitive definition). We say that the system $\mathcal{A} = (x, \sigma, F)$ satisfies the detailed-Progressive Hörmander condition, with constants (L, H) is for any vector v obtained as the image by F of a composition of the vector fields σ with any order of differentiation and no more than k terms can be expressed as

$$v = \sum_{|\alpha| \leq k} \lambda_{[\alpha]} \sigma^{[\alpha]} \quad \text{with}$$

$$\sum_{\substack{|\alpha| \leq k \\ \alpha \neq (0)}} \lambda_{[\alpha]}^2 \leq H^{-1}.$$

REMARK 7.1.11. This is slightly stronger than the Progressive Hörmander condition, but reasonable examples that satisfy the Progressive Hörmander condition will also satisfy the detailed-Progressive Hörmander condition.

For the detailed-Progressive Hörmander condition, we must also include elements such as

$$v = (\sigma^1 \sigma^2)(\sigma^3) = \frac{\partial \sigma^3}{\partial \frac{\partial \sigma^2}{\partial \sigma^1}} = \sum_{i,j} \sigma_i^1 \frac{\partial \sigma_j^2}{\partial x_i} \frac{\partial \sigma^3}{\partial x_j} \quad \text{and}$$

$$w = \frac{\partial^2 \sigma^3}{\partial \sigma^2 \partial \sigma^1} = \sum_{i,j} \sigma_i^1 \sigma_j^2 \frac{\partial^2 \sigma^3}{\partial x_i \partial x_j},$$

not just elements of the form

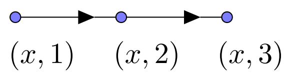
$$\begin{aligned} V = \sigma^{(1,2,3)} = \sigma^1(\sigma^2(\sigma^3)) = \frac{\partial \frac{\partial \sigma^3}{\partial \sigma^2}}{\partial \sigma^1} &= \sum_{i,j} \sigma_i^1 \frac{\partial \sigma_j^2}{\partial x_i} \frac{\partial \sigma^3}{\partial x_j} + \sigma_i^1 \sigma_j^2 \frac{\partial^2 \sigma^3}{\partial x_i \partial x_j} \\ &= \sigma_i^1 \sigma_j^2 \frac{\partial^2 \sigma^3}{\partial x_i \partial x_j} + \frac{\partial^2 \sigma^3}{\partial \sigma^2 \partial \sigma^1} = v + w. \end{aligned}$$

The natural indexing set over which to consider such derivatives is the set of rooted trees, as has been described before in slightly different contexts (cf. [**48, 24, 14**] etc.)

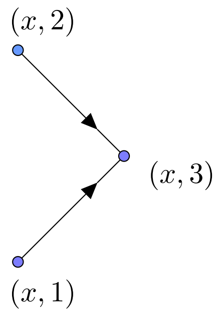
Let $\mathcal{T}(\mathbb{R}^m \otimes \{0, \ldots, d\}, k)$ be the set of rooted trees with less than k vertices in $\mathbb{R}^m \otimes \{0, 1, \ldots, d\}$ (possibly repeated), there is a natural way of associating a mixed multiple derivative of order less than k to each element $\alpha \in \mathcal{T}(\mathbb{R}^m \otimes \{0, \ldots, d\}, k)$. For instance, v and w from above correspond to the figures A and B below (respectively), and the following is what corresponds to the tree from figure C:

$$\frac{\partial^2 \sigma^1(x)}{\partial \sigma^3(z) \partial \frac{\partial \sigma^2(y)}{\partial \sigma^4(u)}} = \sum_{i,j} \frac{\partial \sigma^1}{\partial x_i \partial x_j}(x) \sigma_i^3(z) \sum_k \frac{\partial \sigma_j^2}{\partial x_k}(y) \sigma_k^4(u).$$

For instance, $v = (\sigma^1 \sigma^2)(\sigma^3) = \frac{\partial \sigma^3}{\partial \frac{\partial \sigma^2}{\partial \sigma^1}} = \sum_{i,j} \sigma_i^1 \frac{\partial \sigma_j^2}{\partial x_i} \frac{\partial \sigma^3}{\partial x_j}$ corresponds to the following:

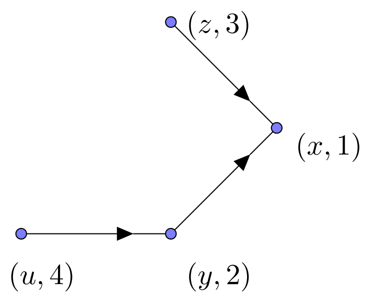


$w = \frac{\partial^2 \sigma^3}{\partial \sigma^2 \partial \sigma^1} = \sum_{i,j} \sigma_i^1 \sigma_j^2 \frac{\partial^2 \sigma^3}{\partial x_i \partial x_j}$ corresponds to the following:



And this formula corresponds to the picture below it:

$$\frac{\partial^2 \sigma^1(x)}{\partial \sigma^3(z) \partial \frac{\partial \sigma^2(y)}{\partial \sigma^4(u)}} = \sum_{i,j} \frac{\partial \sigma^1}{\partial x_i \partial x_j}(x) \sigma_i^3(z) \sum_k \frac{\partial \sigma_j^2}{\partial x_k}(y) \sigma_k^4(u)$$



For any multi-index $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_{\#(\alpha)}) \in \mathrm{Multi}(\mathbb{R}^m \otimes \{0, 1, \ldots, d\})$, we define $\mathcal{T}(\alpha)$ to be the set of rooted trees whose vertices are the indices in α, and such that for any i, j with $i \leq j$, α_j is not a leaf of α_i.

It is easy to convince oneself that

$$\sigma^\alpha = \sum_{\tau\in\mathcal{T}(\alpha)} \sigma^\tau.$$

Since the cardinality of $\mathcal{T}(\alpha)$ is a proper constant, we see that the detailed-Progressive Hörmander condition implies the progressive Hörmander condition with the same constant up to a properly constant multiplicative factor.

It is possible to define analogously quantities such as σ^τ for $\tau \in \mathcal{T}(\Delta)$ for $\Delta \in \mathrm{Multi}(\mathbb{R}^m \otimes \mathcal{L}^l(\mathbb{R}^d, \mathbb{R}))$.

DEFINITION 7.1.12. The mixed tension $\mathcal{G}$, defined when F is linear, of order (L, g) and localisation parameter r of the system $\mathcal{A} = (x, \sigma, F)$ (F is assumed to be linear as usual in this part of the thesis) is defined as

$$\mathcal{G} = \sup_{x\in\mathbb{R}^m} \sup_{o\leq g} \sup_{\substack{v\in\mathbb{R}^m \\ |v|=1}} \sup_{\substack{w\in(\mathbb{R}^m)^{\otimes o} \\ \forall i, |w^i|=1}} \sum_{\substack{\alpha\in\mathrm{Multi}(B_{Eucl}(x,r)\otimes\{0,1,\ldots,d\}); \\ |\alpha|\leq L}} \sum_{\tau\in\mathcal{T}(\alpha)} \langle \frac{\partial^o\sigma^\tau}{\Pi\partial w^i}, v\rangle^2 + |F|.$$

DEFINITION. Let $\mathcal{E}_L$ be the free vector space over rooted trees of order less than L and $\mathcal{L}$ be the set of Lyndon words of order less than L. We say that the system $\mathcal{A} = (x, \sigma, F)$ is (H_L, L)-detailed weak progressive Hörmander if it is (L, H_L)-progressive Hörmander and, in addition, there exists a function $\Psi : \mathcal{E}_L \to \mathcal{S}$ such that for any $\tau \in \mathcal{E}_L$ and any $u \in \mathcal{S}$,

$$*\sigma^\tau(x) = \sum_{\substack{|\beta|\leq|\alpha| \\ \beta\neq(0),\beta\in\mathcal{L}_L}} *\sigma^{\Psi^{[[\beta]]}(\tau)}(x) \quad \text{and}$$

$$\sum_{\substack{|\beta|\leq|\alpha| \\ \beta\neq(0),\beta\in\mathcal{L}_L}} \langle \Psi^{[[\beta]]}, u\rangle^2 \leq H_L^{-1},$$

and for any $|\beta| \leq L$, $\Psi(e^{[[\beta]]}) = e^{[[\beta]]}$. By Proposition 7.0.9, this definition is equivalent, up to a proper constant, to the definition obtained by replacing $e^{[\beta]}$ with $b \in \mathcal{L}_L$ by e^b with $b \in H$.

REMARK 7.1.13. Up to a strongly polynomial constant, this definition is equivalent to the following:

- The Hörmander condition holds with constant H, *and*
- for any rooted tree τ containing non zero indices and with $|\tau| \leq L$ there exist real numbers $\lambda_{\tau,\beta}$ such that

$$\sum_{\substack{|\beta|\leq|\tau| \\ \beta\neq[0]}} \lambda_{\tau,\beta} *\sigma^{[\beta]} = *\sigma^\tau \quad \text{and} \sum_\beta (\lambda_{\tau,\beta})^2 \leq H^{-1}$$

DEFINITION. We say that the vector fields $\sigma^0, \sigma^1, \ldots, \sigma^d$ [3] on $\mathbb{R}^m$ are (L, H_L)-uniformly progressively finitely generated (UPFG)[4] in the set $U \in \mathbb{R}^m$ if there exists a function $\Psi : \mathcal{E}_L \to \mathcal{S}$ such that for any $\tau \in \mathcal{E}_L$ and any $u \in \mathcal{S}$,

$$\sigma^\tau(x) = \sum_{\substack{|\beta|\leq|\alpha| \\ \beta\neq(0),\beta\in\mathcal{L}_L}} \sigma^{\Psi^{[\beta]}(\tau)}(x) \quad \text{and}$$

$$\sum_{\substack{|\beta|\leq|\alpha| \\ \beta\neq(0),\beta\in\mathcal{L}_L}} \langle \Psi^{[\beta]}, u\rangle^2 \leq H_L^{-1}$$

and for any $|\beta| \leq L$, $\Psi(e^{[\beta]}) = e^{[\beta]}$.

[3]The definition depends on which vector field is called 'v^0'

[4]The non progressive equivalent of this condition, namely the "Uniformly Finitely Generated" (UFG) condition is the standard condition to use when proving gradient bounds. See [**35**]

THEOREM 7.1.1. *Let $\mathcal{A} = (x, \sigma, F)$ (with F linear), be a $(r, 2L, g, \mathcal{G})$-mixed tense, (H_L, L)-detailed weak progressive Hörmander system. Suppose that $g \geq L + 1$. There exists a proper constant C (i.e. depending only on $\mathcal{G}$, H,g,L,n and not on m, d, polynomial in $\mathcal{G}$, H) such that for all $x \in \mathbb{R}^m, y \in \mathbb{R}^n$ and $t \leq 1$ such that*

$$d_{t,\log,\infty}(x, y), d_{t,\log}(x, y) \leq 1/2,$$

we have:

$$\max(d_{t,\log}(x, y), \sqrt{t}) \leq C d_{t,\log,\infty}(x, y). \tag{7.1.3}$$

Furthermore, if the σ's are (L, H_L)-uniformly detailed-Progressively finitely generated on a ball of radius r around x, then there are proper constants C_2, C_3 (i.e. depending only on $\mathcal{G}$, H,g,L,n,r, and not on m,d, polynomial in $\mathcal{G}$, H) such that for all $y \in \mathbb{R}^m$ such that

$$d_{t,\log,\infty}(x, y), d_{t,\log}(x, y) \leq C_3,$$

we have

$$d_{\mathcal{A},t,\log,\infty}(x, y) \leq C_2 \max(d_{\bar{\mathcal{A}},t,\log}(x, y), \sqrt{t}). \tag{7.1.4}$$

In particular, for all $x \in \mathbb{R}^m, y \in \mathbb{R}^n$, (with $d_{\mathcal{A},t,\log,\infty}(x, y), d_{\mathcal{A},t,\log}(x, y) \leq C_3$),

$$d_{t,\log,\infty}(x, y) \leq C_2 \max(d_{t,\log}(x, y), \sqrt{t}).$$

PROOF. *Proof of inequality* (7.1.4)*:*

Note first that we have the following expansion where Δ is a multi-index with indices in B:

$$y = *x + \sum_{\substack{|\Delta| \leq L \\ \Delta = (\Delta_1, \Delta_2, \ldots, \Delta_o)}} *\sigma^{\Delta} \int_0^1 \int_0^{t_{o-1}} \cdots \int_0^{t_2} \int_0^{t_1} d\theta_{t_1}^{\Delta_1} d\theta_{t_2}^{\Delta_2} \ldots d\theta_{t_o}^{\Delta_o}$$
$$+ \sum_{\Delta \in \epsilon} \int_0^1 \int_0^{t_{o-1}} \cdots \int_0^{t_2} \int_0^{t_1} *\sigma^{\Delta}(x_{t_o}) d\theta_{t_1}^{\Delta_1} d\theta_{t_2}^{\Delta_2} \ldots d\theta_{t_o}^{\Delta_o}$$

for some set ϵ such that $\Delta \in \epsilon \implies L + 1 \leq |\Delta| \leq 2L$.

Using the progressive Hörmander condition, some simplifying notation, and projecting on the space $U_i = \operatorname{span}_{\substack{|\alpha| = i \\ \alpha \notin \mathrm{Multi}(\{0\})}} e^{[\alpha]}$, we see that if we define the following $\mu \in \mathcal{S}$:

$$\mu^{U_i} = \theta_1^{U_i} + \sum_{\substack{|\Delta| \leq L \\ \Delta = (\Delta_1, \Delta_2, \ldots, \Delta_o), o \geq 2}} \Psi_{x_0}^{U_i}(\Delta) \int^{\Delta} (d\theta)^{\Delta}$$
$$+ \Psi_{x_0,L}^{U_i}(\sum_{\Delta \in \epsilon} \int^{\Delta} *\sigma^{\Delta}(x_{\bullet})(d\theta)^{\Delta}),$$

we will certainly have that:

$$y = *x + \sum_{\substack{\beta \in \mathcal{L}_L \\ |\beta| \leq L}} \mu^{[\beta]} * \sigma_x^{[\beta]} + \sum_{\substack{\alpha \in \mathrm{Multi}(\{0\}) \\ \#(\alpha) \leq \lfloor \frac{L}{2} \rfloor}} \sigma_x^{\alpha} \frac{t^{\#(\alpha)}}{(\#(\alpha))!}.$$

Now, Proposition 7.1.6 applied to $\sum_{\Delta \in \epsilon} \int^{\Delta} *\sigma^{\Delta}(x_{\bullet})(d\theta)^{\Delta}$ ensures that

$$|\sum_{\Delta \in \epsilon} \int^{\Delta} *\sigma^{\Delta}(x_{\bullet})(d\theta)^{\Delta}| \leq C\mathcal{F}(L)\mathcal{G}|\theta|_{L^2,\mathcal{S}}^{L+1},$$

for some proper constant C, and a combinatorial function $\mathcal{F}(L)$. Then note that Lemma 3.5.1 and the assumption on the tension guarantee that there exists a function $\Psi_{x_0,L} : \mathbb{R}^n \to \mathcal{S}$ such that for any unit $u \in \mathcal{S}$ and $v \in \mathbb{R}^n$, $\langle \Psi(v), u \rangle^2 \leq H_L^{-1}\mathcal{G}$. Therefore the above implies further that

$$|\Psi_{x_0,L}^{U_i}(\sum_{\Delta \in \epsilon} \int^{\Delta} *\sigma^{\Delta}(x_{\bullet})(d\theta)^{\Delta})| \leq (C\mathcal{G}^2\mathcal{F}(L)H_L^{-1})^{1/2}|\theta|_{L^2,\mathcal{S}}^{L+1}. \tag{7.1.5}$$

Proposition 7.1.6, applied to

$$\sum_{\substack{|\Delta|\leq L \\ \Delta=(\Delta_1,\Delta_2,\dots,\Delta_o),o\geq 2}} \Psi_{x_0}^{U_i}(\Delta)\int^{\Delta}(d\theta)^{\Delta},$$

together with the progressive Hörmander condition, ensures that

$$\left|\sum_{\substack{|\Delta|\leq L \\ \Delta=(\Delta_1,\Delta_2,\dots,\Delta_o),o\geq 2}} \Psi_{x_0}^{U_i}(\Delta)\int^{\Delta}(d\theta)^{\Delta}\right|^2 \leq C|\theta|^i_{L^2,\mathcal{S}}, \tag{7.1.6}$$

for some proper constant C. Here we have identified Δ with the element of $\mathcal{E}_{2L}$ obtained by decomposing Δ into rooted trees. As usual, by the symmetry of the scalar product in the tensor space with respect to relabelling of the indices (cf. Proposition 7.0.9), the decomposition of Δ contains a proper number of terms in $\mathcal{E}_{2L}$ each multiplied by proper constants, and conversely each term appears only in the decomposition of a proper number of compound multi-indices Δ.

Combining Eqs. (7.1.5) and (7.1.6) yields that

$$\begin{aligned}
|\mu|_{\mathcal{S}} &= \left(\sum_{i=1}^{L}|\mu^{U_i}|_i^{\frac{2}{i}}\right)^{\frac{1}{2}} \\
&\leq (\sum_{i=1}^{L}()C|\theta|^i_{L^2,\mathcal{L}^L(\mathbb{R}^d,\mathbb{R})})^{\frac{2}{i}})^{\frac{1}{2}} \\
&\leq \sqrt{L}C|\theta|_{L^2,\mathcal{L}^L(\mathbb{R}^d,\mathbb{R})}
\end{aligned}$$

for some proper constant C.

Proof of inequality (7.1.3)*:*

We will write, as above, H for our basis of $\mathcal{L}^L(\mathbb{R}^d,\mathbb{R})$ defined in Proposition 7.0.9. We can define a function $\bar{\Psi}_{x_0,L}:\mathbb{R}^n\to\mathcal{S}$ such that for any $v\in\mathbb{R}^n$, $F^{ST}(t,\bar{\Psi}_{x_0,L}(v))=v$ and $|\bar{\Psi}_{x_0,L}(v)|^2_{\mathcal{S}}=\inf(|u|^2_{\mathcal{S}}:F^{ST}(t,u)=v)$.

Then we set

$$\begin{aligned}
\mu_1 &= \bar{\Psi}_{x_0,L}(y-F(\bar{x}^t))+\sum_{i=1}^{\lfloor\frac{L}{2}\rfloor}e^{\otimes^i(0)}\frac{t^i}{i!}, \\
\mu_s &= \delta_s(\mu_1) \quad \forall s\leq 1, \\
x_s &= F^{ST}_{\bar{\mathcal{A}}}(\mu_s) \quad \forall s\leq 1,
\end{aligned}$$

where δ_t is the homogeneous dilation on $\mathcal{L}^L(\mathbb{R}^d,\mathbb{R})$ and $\bar{\mathcal{A}}=(x,\sigma,Id)$.

Set C_3 small enough that $|x-x_s|\leq\frac{r}{2}$ for any $s\in[0,1]$. This can clearly be done whilst keeping C_3 a proper constant.

Note that, since $\mu_s^{U_i}=s^{\frac{i}{2}}\mu_1^{U_i}$, writing $U=\bar{\mathcal{S}}=\mathcal{L}^L(\mathbb{R}^d,\mathbb{R})\otimes(\mathrm{span}_{i\leq\lfloor\frac{L}{2}\rfloor}e^{\otimes^i(0)})$, we have

$$\begin{aligned}
|\mu|^2_{L^2,\bar{\mathcal{S}}} &= \sum_{i=1}^{L}(\int_0^1|\frac{\partial\mu_s^{U_i}}{\partial s}|^2ds)^{\frac{1}{i}} \\
&= \sum_{i=1}^{L}(|\mu_1^{U_i}|^2\int_0^1\frac{i}{2}s^{\frac{i-2}{2}}dt)^{\frac{1}{i}} \\
&= \sum_{i=1}^{L}(|\mu_1^{U_i}|^2)^{\frac{1}{i}} = |\mu_1|^2_U.
\end{aligned}$$

Now, for any $0 \leq s \leq 1$, we have the following Taylor expansion for σ_0:

$$
\begin{aligned}
\sigma_0^{a_1} &= \sigma_s^{a_1} + \sum_{a_2} \int_s^0 \sigma^{((a_2,0),(a_1,v))} d\mu_v^{a_2} \\
&= \sigma_s^{a_1} + \sum_{a_2} \sigma^{((a_2,s),(a_1,s))} \int_s^0 d\mu_v^{a_2} + \sum_{a_3,a_2} \int_s^0 \int_s^v \frac{\partial^2 \sigma^{(a_1,u)}}{\partial\sigma^{(a_2,s)} \partial\sigma^{(a_3,0)}} d\mu_u^{a_3} d\mu_v^{a_2} \\
&\quad + \sum_{a_3,a_2} \int_1^0 \int_1^s \frac{\partial \sigma^{(a_1,v)}}{\partial \frac{\partial \sigma^{(a_2,u)}}{\partial\sigma^{(a_3,0)}}} d\mu_u^{a_3} d\mu_v^{a_2} \\
&= \dots \\
&= \sum_{\substack{\Delta=((\Delta_1,s),\dots,(\Delta_o,s))\in \mathrm{Multi}(H\otimes\{s\}) \\ \Delta_o=a_1;|\Delta|\leq L}} \sigma^\Delta \int_{t_1,\dots,t_{o-1}\in[0,s]^{\#(\bar\Delta)}}^{\bar\Delta} \xi_{t_1,\dots,t_{o-1}} d\mu_{t_1,\dots,t_{o-1}}^{\bar\Delta} \\
&\quad + \int_{[s,0]^{\#(\bar\Delta)}}^{\bar\Delta} \sum_{\substack{\Delta=((\Delta_1,t_1),\dots,(\Delta_o,t_o))\in \mathrm{Multi}(H\otimes[0,s]) \\ \Delta_o=a_1;t_1=0;\Delta\in\epsilon\otimes[0,s]^{\#(\Delta)}}} \sum_{\tau\in\mathcal{T}(\Delta)} \sigma^\tau \phi_{t_1,\dots,t_{o-1}} d\mu_{t_1,\dots,t_{o-1}}^{\bar\Delta}
\end{aligned}
$$

for some $\phi_{t_2,\dots,t_o}, \xi_{t_2,\dots,t_o}$ taking values in $\{-1,0,1\}$. Here $\bar\Delta$ means the multi-index obtained by deleting the last index of Δ.

This motivates the following choice of θ:

$$
\begin{aligned}
\theta_s^{a_1} &= \sum_{\substack{\Delta\in \mathrm{Multi}(H) \\ |\Delta|\leq L}} \Psi_{x_s}^{a_1}(\Delta) \int_{t_1,\dots,t_o\in[0,s]^{\#(\Delta)}}^{\Delta} \tilde\xi_{t_1,\dots,t_o} d\mu_{t_1,\dots,t_o}^{\Delta} \\
&\quad + \Psi_{x_s,L}^{a_1}\Big(\int_{[0,s]^{\#(\Delta)}}^{\Delta} \sum_{\substack{\Delta=((\Delta_1,t_1),\dots,(\Delta_o,t_o))\in \mathrm{Multi}(B\otimes[0,t]) \\ t_1=0;\Delta\in\epsilon\otimes[0,s]^{\#(\Delta)}}} \sum_{\tau\in\mathcal{T}(\Delta)} \sigma_{t_1,\dots,t_o}^\tau \tilde\phi_{t_1,\dots,t_o} d\mu_{t_1,\dots,t_o}^{\Delta}\Big) \\
&= \mu_s^{a_1} + \sum_{\substack{\Delta\in \mathrm{Multi}(H) \\ |\Delta|\leq L;\#(\Delta)\geq 2}} \Psi_{x_s}^{a_1}(\Delta) \int_{t_1,\dots,t_o\in[0,t]^{\#(\Delta)}}^{\Delta} \tilde\xi_{t_1,\dots,t_o} d\mu_{t_1,\dots,t_o}^{\Delta} \\
&\quad + \Psi_{x_s,L}^{a_1}\Big(\int_{[0,s]^{\#(\Delta)}}^{\Delta} \sum_{\substack{\Delta=((\Delta_1,t_1),\dots,(\Delta_o,t_o))\in \mathrm{Multi}(B\otimes[0,s]) \\ t_1=0;\Delta\in\epsilon\otimes[0,s]^{\#(\Delta)}}} \sum_{\tau\in\mathcal{T}(\Delta)} \sigma^\tau \tilde\phi_{t_1,\dots,t_o} d\mu_{t_1,\dots,t_o}^{\Delta}\Big)
\end{aligned}
$$

and $\theta_s^{(0)} = \mu_s^{(0)} = st$. For some $\tilde\xi, \tilde\phi$ taking values only in $\{-1,0,1\}$

Rewriting this over the whole space U_i, we get:

$$
\begin{aligned}
\theta_s^{U_i} &= \mu_s^{U_i} + \sum_{\substack{\Delta\in \mathrm{Multi}(B) \\ |\Delta|\leq L;\#(\Delta)\geq 2}} \Psi_{x_s}^{U_i}(\Delta) \int_{t_1,\dots,t_o\in[0,s]^{\#(\Delta)}}^{\Delta} \tilde\xi_{t_1,\dots,t_o} d\mu_{t_1,\dots,t_o}^{\Delta} \\
&\quad + \Psi_{x_s}^{U_i}\Big(\int_{[0,s]^{\#(\Delta)}}^{\Delta} \sum_{\substack{\Delta=((\Delta_1,t_1),\dots,(\Delta_o,t_o))\in \mathrm{Multi}(H\otimes[0,s]) \\ t_1=0;\Delta\in\epsilon\otimes[0,s]^{\#(\Delta)}}} \sum_{\tau\in\mathcal{T}(\Delta)} \sigma^\tau \tilde\phi_{t_1,\dots,t_o} d\mu_{t_1,\dots,t_o}^{\Delta}\Big).
\end{aligned}
$$

This still ensures that θ satisfies the ODE (7.1.2):

$$
\begin{aligned}
x_0 &= x, \\
dx_t &= \sum_{a\in B} \sigma^a(x_t) \frac{\partial \theta_t^a}{\partial t} dt.
\end{aligned}
$$

Define a linear operation δ on the free vector space over rooted trees with indices in

$$\mathcal{T}(\{0,1,\dots,d\})$$

by requiring that for τ a rooted tree with vertices in $\mathcal{T}(\{0,1\dots,d\})$,

$$\delta(\tau)=\sum_{\tilde\tau\in G(\tau)} e^{\tilde\tau}$$

where $G(\tau)$ is the set of rooted trees with vertices in $\{0,1,\dots,d\}$ obtained by gluing the root of each vertex τ_i of τ to one of the vertices of τ_j where τ_i is a leaf of τ_j. Then we have

$$\sigma^{\Delta}=\sum_{\tau\in\mathcal{T}(\Delta)}\sigma^{\delta(\tau)}.$$

We can extend the above operation to rooted trees with vertices in $B\otimes[0,1]$ similarly.

Each vertex $\Delta_i\in B_{2L}\otimes[0,1]$ can be written as a sum of rooted trees $\sum_{\iota(\tau)=O(\Delta)}\lambda_{\tau,\Delta}e^{\tau}$, where $\iota(\tau)$ denotes the multi-set composed of the vertices of τ, $O(\Delta_i)$ is a fixed multi-set depending only on Δ_i, and $|\lambda_{\tau,\Delta}|\le(2L)!2^{2L}$.

It follows that for any unit $v\in\mathbb{R}^n$, writing $\mathcal{T}^2(\Delta)$ for the set of rooted trees belonging to $G(\tau)$ for some τ being a rooted tree with root in $\mathcal{T}(\Delta_{\#(\Delta)})$ and other vertices in $\mathcal{T}(\Delta_i)$ for $i=1,2,\dots,\#(\Delta)-1$,

$$\begin{aligned}&\sum_{\substack{\Delta=((\Delta_1,t_1),\dots,(\Delta_o,t_o))\in\mathrm{Multi}(H\otimes[0,s])\\ t_1=0;\Delta\in\epsilon\otimes[0,s]^{\#(\Delta)}}}\sum_{\tau\in\mathcal{T}(\Delta)}\langle\sigma^{\tau},v\rangle^2\\
&=\sum_{\substack{\tau\in\mathrm{Multi}(\{0,1,\dots,d\}\otimes[0,1])\\|\tau|\le 2L}}\langle\sigma^{\tau}\bar\lambda_{\tau},v\rangle^2\\
&\le\sup_{\tau}(|\bar\lambda_{\tau}|^2)\mathcal{G},\end{aligned}$$

for some $\bar\lambda_\tau$ less than some combinatorial function of L in absolute value. In other words, the expression above is a proper constant.

Next, applying Theorem 7.1.6 to

$$\int_{[0,s]^{\#(\Delta)}}^{\Delta}\sum_{\substack{\Delta=((\Delta_1,t_1),\dots,(\Delta_o,t_o))\in\mathrm{Multi}(H\otimes[0,s])\\ t_1=0;\Delta\in\epsilon\otimes[0,s]^{\#(\Delta)}}}\sum_{\tau\in\mathcal{T}(\Delta)}\sigma^{\tau}\tilde\phi_{t_1,\dots,t_o}d\mu^{\Delta}_{t_1,\dots,t_o}$$

ensures that

$$\begin{aligned}&\left|\int_{[0,s]^{\#(\Delta)}}^{\Delta}\sum_{\substack{\Delta=((\Delta_1,t_1),\dots,(\Delta_o,t_o))\in\mathrm{Multi}(H\otimes[0,s])\\ t_1=0;\Delta\in\epsilon\otimes[0,s]^{\#(\Delta)}}}\sum_{\tau\in\mathcal{T}(\Delta)}\sigma^{\tau}\tilde\phi_{t_1,\dots,t_o}d\mu^{\Delta}_{t_1,\dots,t_o}\right|^2\\
&\le K|\theta|^{2(L+1)}_{L^2,\mathcal{L}^L(\mathbb{R}^d,\mathbb{R})}\end{aligned}$$

for some proper constant K. By the weak Hörmander condition, it then follows that

$$\begin{aligned}&\left|\Psi^{U_i}_{x_s}(\int_{[0,s]^{\#(\Delta)}}^{\Delta}\sum_{\substack{\Delta=((\Delta_1,t_1),\dots,(\Delta_o,t_o))\in\mathrm{Multi}(H\otimes[0,s])\\ t_1=0;\Delta\in\epsilon\otimes[0,s]^{\#(\Delta)}}}\sum_{\tau\in\mathcal{T}(\Delta)}\sigma^{\tau}\tilde\phi_{t_1,\dots,t_o}d\mu^{\Delta}_{t_1,\dots,t_o})\right|^2\\
&\le K|\theta|^{2(L+1)}_{L^2,\mathcal{L}^L(\mathbb{R}^d,\mathbb{R})}\end{aligned}$$

for some proper constant K.

As in the proof of the other side of the inequality, we also have, using the progressive Hörmander condition, the symmetry of the inner product on the tensor space with respect to relabelling, and Theorem 7.1.6,

$$\left| \sum_{\substack{\Delta \in \mathrm{Multi}(B) \\ |\Delta| \leq L; \#(\Delta) \geq 2}} \Psi_{x_s}^{U_i}(\Delta) \int_{t_1,\ldots,t_o \in [0,s]^{\#(\Delta)}}^{\Delta} \tilde{\xi}_{t_1,\ldots,t_o} d\mu_{t_1,\ldots,t_o}^{\Delta} \right|^2 \leq K|\theta|_{L^2,\mathcal{L}^L(\mathbb{R}^d,\mathbb{R})}^{2i}$$

for some proper constant K.

Similarly to the proof of the inequality (7.1.4):

Note first that we have the following expansion where $\varDelta$ is a multi-index with indices in B:

$$\begin{aligned} y =& *x + \sum_{\substack{|\varDelta| \leq L \\ \varDelta = (\varDelta_1, \varDelta_2, \ldots, \varDelta_o)}} *\sigma^{\varDelta} \int_0^1 \int_0^{t_{o-1}} \ldots \int_0^{t_2} \int_0^{t_1} d\theta_{t_1}^{\varDelta_1} d\theta_{t_2}^{\varDelta_2} \ldots d\theta_{t_o}^{\varDelta_o} \\ &+ \sum_{\varDelta \in \epsilon} \int_0^1 \int_0^{t_{o-1}} \ldots \int_0^{t_2} \int_0^{t_1} *\sigma^{\varDelta}(x_{t_o}) d\theta_{t_1}^{\varDelta_1} d\theta_{t_2}^{\varDelta_2} \ldots d\theta_{t_o}^{\varDelta_o} \end{aligned}$$

for some set ϵ such that $\varDelta \in \epsilon \implies L+1 \leq |\varDelta| \leq 2L$.

Using the progressive Hörmander condition, some simplifying notation, and projecting on the space $U_i = \operatorname{span}_{\substack{|\alpha|=i \\ \alpha \notin \mathrm{Multi}(\{0\})}} e^{[\alpha]}$, we see that if we define the following $\mu \in \mathcal{S}$:

$$\begin{aligned} \mu^{U_i} = \theta_1^{U_i} &+ \sum_{\substack{|\varDelta| \leq L \\ \varDelta = (\varDelta_1, \varDelta_2, \ldots, \varDelta_o), o \geq 2}} \Psi_{x_0}^{U_i}(\varDelta) \int^{\varDelta} (d\theta)^{\varDelta} \\ &+ \Psi_{x_0,L}^{U_i}\left(\sum_{\varDelta \in \epsilon} \int^{\Delta} *\sigma^{\varDelta}(x_{\bullet})(d\theta)^{\Delta}\right), \end{aligned}$$

we will certainly have that:

$$y = *x + \sum_{\substack{\beta \in \mathcal{L}_L \\ |\beta| \leq L}} \mu^{[\beta]} * \sigma_x^{[\beta]} + \sum_{\substack{\alpha \in \mathrm{Multi}(\{0\}) \\ \#(\alpha) \leq \lfloor \frac{L}{2} \rfloor}} \sigma_x^{\alpha} \frac{t^{\#(\alpha)}}{(\#(\alpha))!}.$$

Now, Proposition 7.1.6 applied to $\sum_{\varDelta \in \epsilon} \int^{\Delta} *\sigma^{\varDelta}(x_{\bullet})(d\theta)^{\Delta}$ ensures that

$$\left| \sum_{\varDelta \in \epsilon} \int^{\Delta} *\sigma^{\varDelta}(x_{\bullet})(d\theta)^{\Delta} \right| \leq C\mathcal{F}(L)\mathcal{G}|\theta|_{L^2,\mathcal{S}}^{L+1},$$

for some proper constant C, and a combinatorial function $\mathcal{F}(L)$. Then note that Lemma 3.5.1 and the assumption on the tension guarantee that there exists a function $\Psi_{x_0,L} : \mathbb{R}^n \to \mathcal{S}$ such that for any unit $u \in \mathcal{S}$ and $v \in \mathbb{R}^n$, $\langle \Psi(v), u \rangle^2 \leq H_L^{-1}\mathcal{G}$. Therefore the above implies further that

$$\left| \Psi_{x_0,L}^{U_i}\Big(\sum_{\varDelta \in \epsilon} \int^{\Delta} *\sigma^{\varDelta}(x_{\bullet})(d\theta)^{\Delta} \Big) \right| \leq (C\mathcal{G}^2\mathcal{F}(L)H_L^{-1})^{1/2}|\theta|_{L^2,\mathcal{S}}^{L+1}. \tag{7.1.7}$$

Proposition 7.1.6, applied to

$$\sum_{\substack{|\varDelta| \leq L \\ \varDelta = (\varDelta_1, \varDelta_2, \ldots, \varDelta_o), o \geq 2}} \Psi_{x_0}^{U_i}(\varDelta) \int^{\varDelta} (d\theta)^{\varDelta},$$

together with the progressive Hörmander condition, ensures that

$$\left| \sum_{\substack{|\varDelta| \leq L \\ \varDelta = (\varDelta_1, \varDelta_2, \ldots, \varDelta_o), o \geq 2}} \Psi_{x_0}^{U_i}(\varDelta) \int^{\varDelta} (d\theta)^{\varDelta} \right|^2 \leq C|\theta|_{L^2,\mathcal{S}}^{i}, \tag{7.1.8}$$

for some proper constant C. Here we have identified $\varDelta$ with the element of $\mathcal{E}_{2L}$ obtained by decomposing $\varDelta$ into rooted trees. As usual, by the symmetry of the scalar product in the tensor space with respect to relabelling of the indices (cf. Proposition 7.0.9), the decomposition of Δ contains a proper number of terms in $\mathcal{E}_{2L}$ each multiplied by proper constants, and conversely each term appears only in the decomposition of a proper number of compound multi-indices Δ.

Combining Eqs. (7.1.7) and (7.1.8) yields that

$$\begin{aligned}|\mu|_{\mathcal{S}} &= (\sum_{i=1}^{L} |\mu^{U_i}|^{\frac{2}{i}})^{\frac{1}{2}} \\ &\leq (\sum_{i=1}^{L} C|\theta|^{i}_{L^2,\mathcal{L}^L(\mathbb{R}^d,\mathbb{R})})^{\frac{2}{i}})^{\frac{1}{2}} \\ &\leq \sqrt{L}C|\theta|_{L^2,\mathcal{L}^L(\mathbb{R}^d,\mathbb{R})}.\end{aligned}$$

for some proper constant C.

Proof of inequality (7.1.3)*:*

We will write, as above, H for our basis of $\mathcal{L}^L(\mathbb{R}^d,\mathbb{R})$ defined in Proposition 7.0.9. We can define a function $\bar{\Psi}_{x_0,L} : \mathbb{R}^n \to \mathcal{S}$ such that for any $v \in \mathbb{R}^n$, $F^{ST}(t, \bar{\Psi}_{x_0,L}(v)) = v$ and $|\bar{\Psi}_{x_0,L}(v)|^2_{\mathcal{S}} = \inf(|u|^2_{\mathcal{S}} : F^{ST}(t,u) = v)$.

Then we set

$$\begin{aligned}\mu_1 &= \bar{\Psi}_{x_0,L}(y - F(\bar{x}^t)) + \sum_{i=1}^{\lfloor \frac{L}{2} \rfloor} e^{\otimes^i(0)} \frac{t^i}{i!} \\ \mu_s &= \delta_s(\mu_1) \quad \forall s \leq 1 \\ x_s &= F^{ST}_{\bar{\mathcal{A}}}(\mu_s) \quad \forall s \leq 1\end{aligned}$$

where δ_t is the homogeneous dilation on $\mathcal{L}^L(\mathbb{R}^d,\mathbb{R})$ and $\bar{\mathcal{A}} = (x, \sigma, Id)$.

Set C_3 small enough that $|x - x_s| \leq \frac{r}{2}$ for any $s \in [0,1]$. This can clearly be done whilst keeping C_3 a proper constant.

Note that, since $\mu^{U_i}_s = s^{\frac{i}{2}} \mu^{U_i}_1$, writing $U = \bar{\mathcal{S}} = \mathcal{L}^L(\mathbb{R}^d,\mathbb{R}) \otimes (\operatorname{span}_{i \leq \lfloor \frac{L}{2} \rfloor} e^{\otimes^i(0)})$, we have

$$\begin{aligned}|\mu|^2_{L^2,\bar{\mathcal{S}}} &= \sum_{i=1}^{L} (\int_0^1 |\frac{\partial \mu^{U_i}_s}{\partial s}|^2 ds)^{\frac{1}{i}} \\ &= \sum_{i=1}^{L} (|\mu^{U_i}_1|^2 \int_0^1 \frac{i}{2} s^{\frac{i-2}{2}} dt)^{\frac{1}{i}} \\ &= \sum_{i=1}^{L} (|\mu^{U_i}_1|^2)^{\frac{1}{i}} = |\mu_1|^2_U.\end{aligned}$$

Now, for any $0 \leq s \leq 1$, we have the following Taylor expansion for σ_0:

$$\begin{aligned}\sigma^{a_1}_0 &= \sigma^{a_1}_s + \sum_{a_2} \int_s^0 \sigma^{((a_2,0),(a_1,v))} d\mu^{a_2}_v \\ &= \sigma^{a_1}_s + \sum_{a_2} \sigma^{((a_2,s),(a_1,s))} \int_s^0 d\mu^{a_2}_v + \sum_{a_3,a_2} \int_s^0 \int_s^v \frac{\partial^2 \sigma^{(a_1,u)}}{\partial \sigma^{(a_2,s)} \partial \sigma^{(a_3,0)}} d\mu^{a_3}_u d\mu^{a_2}_v \\ &+ \sum_{a_3,a_2} \int_1^0 \int_1^s \frac{\partial \sigma^{(a_1,v)}}{\partial \frac{\partial \sigma^{(a_2,u)}}{\partial \sigma^{(a_3,0)}}} d\mu^{a_3}_u d\mu^{a_2}_v \\ &= \dots \\ &= \sum_{\substack{\Delta=((\Delta_1,s),\dots,(\Delta_o,s)) \in \mathrm{Multi}(H \otimes \{s\}) \\ \Delta_o = a_1; |\Delta| \leq L}} \sigma^{\Delta} \int_{t_1,\dots,t_{o-1} \in [0,s]^{\#(\bar{\Delta})}}^{\bar{\Delta}} \xi_{t_1,\dots,t_{o-1}} d\mu^{\bar{\Delta}}_{t_1,\dots,t_{o-1}}\end{aligned}$$

$$+\int_{[s,0]^{\#(\bar{\Delta})}}^{\bar{\Delta}} \sum_{\substack{\Delta=((\Delta_1,t_1),\ldots,(\Delta_o,t_o))\in\mathrm{Multi}(H\otimes[0,s])\\ \Delta_o=a_1;t_1=0;\Delta\in\epsilon\otimes[0,s]^{\#(\Delta)}}} \sum_{\tau\in\mathcal{T}(\Delta)} \sigma^\tau \phi_{t_1,\ldots,t_{o-1}} d\mu^{\bar{\Delta}}_{t_1,\ldots,t_{o-1}}$$

for some $\phi_{t_2,\ldots,t_o}, \xi_{t_2,\ldots,t_o}$ taking values in $\{-1,0,1\}$. Here $\bar{\Delta}$ means the multi-index obtained by deleting the last index of Δ.

This motivates the following choice of θ:

$$\begin{aligned}
\theta^{a_1}_s = &\sum_{\substack{\Delta\in\mathrm{Multi}(H)\\ |\Delta|\le L}} \Psi^{a_1}_{x_s}(\Delta)\int^{\Delta}_{t_1,\ldots,t_o\in[0,s]^{\#(\Delta)}} \tilde{\xi}_{t_1,\ldots,t_o} d\mu^{\Delta}_{t_1,\ldots,t_o}\\
&+\Psi^{a_1}_{x_s,L}\Big(\int^{\Delta}_{[0,s]^{\#(\Delta)}} \sum_{\substack{\Delta=((\Delta_1,t_1),\ldots,(\Delta_o,t_o))\in\mathrm{Multi}(B\otimes[0,t])\\ t_1=0;\Delta\in\epsilon\otimes[0,s]^{\#(\Delta)}}} \sum_{\tau\in\mathcal{T}(\Delta)} \sigma^\tau_{t_1,\ldots,t_o}\tilde{\phi}_{t_1,\ldots,t_o} d\mu^{\Delta}_{t_1,\ldots,t_o}\Big)\\
= \mu^{a_1}_s + &\sum_{\substack{\Delta\in\mathrm{Multi}(H)\\ |\Delta|\le L;\#(\Delta)\ge 2}} \Psi^{a_1}_{x_s}(\Delta)\int^{\Delta}_{t_1,\ldots,t_o\in[0,t]^{\#(\Delta)}} \tilde{\xi}_{t_1,\ldots,t_o} d\mu^{\Delta}_{t_1,\ldots,t_o}\\
&+\Psi^{a_1}_{x_s,L}\Big(\int^{\Delta}_{[0,s]^{\#(\Delta)}} \sum_{\substack{\Delta=((\Delta_1,t_1),\ldots,(\Delta_o,t_o))\in\mathrm{Multi}(B\otimes[0,s])\\ t_1=0;\Delta\in\epsilon\otimes[0,s]^{\#(\Delta)}}} \sum_{\tau\in\mathcal{T}(\Delta)} \sigma^\tau\tilde{\phi}_{t_1,\ldots,t_o} d\mu^{\Delta}_{t_1,\ldots,t_o}\Big)
\end{aligned}$$

and $\theta^{(0)}_s = \mu^{(0)}_s = st$. For some $\tilde{\xi}, \tilde{\phi}$ taking values only in $\{-1,0,1\}$.

Rewriting this over the whole space U_i, we get:

$$\begin{aligned}
\theta^{U_i}_s = \mu^{U_i}_s + &\sum_{\substack{\Delta\in\mathrm{Multi}(B)\\ |\Delta|\le L;\#(\Delta)\ge 2}} \Psi^{U_i}_{x_s}(\Delta)\int^{\Delta}_{t_1,\ldots,t_o\in[0,s]^{\#(\Delta)}} \tilde{\xi}_{t_1,\ldots,t_o} d\mu^{\Delta}_{t_1,\ldots,t_o}\\
&+\Psi^{U_i}_{x_s}\big(\int^{\Delta}_{[0,s]^{\#(\Delta)}} \sum_{\substack{\Delta=((\Delta_1,t_1),\ldots,(\Delta_o,t_o))\in\mathrm{Multi}(H\otimes[0,s])\\ t_1=0;\Delta\in\epsilon\otimes[0,s]^{\#(\Delta)}}} \sum_{\tau\in\mathcal{T}(\Delta)} \sigma^\tau\tilde{\phi}_{t_1,\ldots,t_o} d\mu^{\Delta}_{t_1,\ldots,t_o}\big).
\end{aligned}$$

This still ensures that θ satisfies the ODE (7.1.2):

$$\begin{aligned}
x_0 &= x,\\
dx_t &= \sum_{a\in B}\sigma^a(x_t)\frac{\partial\theta^a_t}{\partial t}dt.
\end{aligned}$$

Define a linear operation δ on the free vector space over rooted trees with indices in

$$\mathcal{T}(\{0,1,\ldots,d\})$$

by requiring that for τ a rooted tree with vertices in $\mathcal{T}(\{0,1\ldots,d\})$,

$$\delta(\tau) = \sum_{\tilde{\tau}\in G(\tau)} e^{\tilde{\tau}}$$

where $G(\tau)$ is the set of rooted trees with vertices in $\{0,1,\ldots,d\}$ obtained by gluing the root of each vertex τ_i of τ to one of the vertices of τ_j where τ_i is a leaf of τ_j. Then we have

$$\sigma^{\varDelta} = \sum_{\tau\in\mathcal{T}(\varDelta)} \sigma^{\delta(\tau)}.$$

We can extend the above operation to rooted trees with vertices in $B\otimes[0,1]$ similarly.

Each vertex $\varDelta_i \in B_{2L}\otimes[0,1]$ can be written as a sum of rooted trees $\sum_{\iota(\tau)=O(\varDelta)}\lambda_{\tau,\varDelta}e^\tau$, where $\iota(\tau)$ denotes the multi-set composed of the vertices of τ, $O(\varDelta_i)$ is a fixed multi-set depending only on $\varDelta_i$, and $|\lambda_{\tau,\varDelta}| \le (2L)!2^{2L}$.

It follows that for any unit $v \in \mathbb{R}^n$, writing $\mathcal{T}^2(\Delta)$ for the set of rooted trees belonging to $G(\tau)$ for some τ being a rooted tree with root in $\mathcal{T}(\Delta_{\#(\Delta)})$ and other vertices in $\mathcal{T}(\Delta_i)$ for $i = 1, 2, \ldots, \#(\Delta) - 1$,

$$\begin{aligned}
&\sum_{\substack{\Delta=((\Delta_1,t_1),\ldots,(\Delta_o,t_o))\in\mathrm{Multi}(H\otimes[0,s])\\ t_1=0;\Delta\in\epsilon\otimes[0,s]^{\#(\Delta)}}} \sum_{\tau\in\mathcal{T}(\Delta)} \langle\sigma^\tau, v\rangle^2 \\
&= \sum_{\substack{\tau\in\mathrm{Multi}(\{0,1,\ldots,d\}\otimes[0,1])\\ |\tau|\le 2L}} \langle\sigma^\tau\bar{\lambda}_\tau, v\rangle^2 \\
&\le \sup_\tau(|\bar{\lambda}_\tau|^2)\mathcal{G},
\end{aligned}$$

for some $\bar{\lambda}_\tau$ less than some combinatorial function of L in absolute value. In other words, the expression above is a proper constant.

Next, applying Theorem 7.1.6 to

$$\int_{[0,s]^{\#(\Delta)}}^{\Delta} \sum_{\substack{\Delta=((\Delta_1,t_1),\ldots,(\Delta_o,t_o))\in\mathrm{Multi}(H\otimes[0,s])\\ t_1=0;\Delta\in\epsilon\otimes[0,s]^{\#(\Delta)}}} \sum_{\tau\in\mathcal{T}(\Delta)} \sigma^\tau\tilde{\phi}_{t_1,\ldots,t_o} d\mu^{\Delta}_{t_1,\ldots,t_o}$$

ensures that

$$\begin{aligned}
&\left|\int_{[0,s]^{\#(\Delta)}}^{\Delta} \sum_{\substack{\Delta=((\Delta_1,t_1),\ldots,(\Delta_o,t_o))\in\mathrm{Multi}(H\otimes[0,s])\\ t_1=0;\Delta\in\epsilon\otimes[0,s]^{\#(\Delta)}}} \sum_{\tau\in\mathcal{T}(\Delta)} \sigma^\tau\tilde{\phi}_{t_1,\ldots,t_o} d\mu^{\Delta}_{t_1,\ldots,t_o}\right|^2 \\
&\le K|\theta|^{2(L+1)}_{L^2,\mathcal{L}^L(\mathbb{R}^d,\mathbb{R})}
\end{aligned}$$

for some proper constant K. By the weak Hörmander condition, it then follows that

$$\begin{aligned}
&\left|\Psi^{U_i}_{x_s}\Big(\int_{[0,s]^{\#(\Delta)}}^{\Delta} \sum_{\substack{\Delta=((\Delta_1,t_1),\ldots,(\Delta_o,t_o))\in\mathrm{Multi}(H\otimes[0,s])\\ t_1=0;\Delta\in\epsilon\otimes[0,s]^{\#(\Delta)}}} \sum_{\tau\in\mathcal{T}(\Delta)} \sigma^\tau\tilde{\phi}_{t_1,\ldots,t_o} d\mu^{\Delta}_{t_1,\ldots,t_o}\Big)\right|^2 \\
&\le K|\theta|^{2(L+1)}_{L^2,\mathcal{L}^L(\mathbb{R}^d,\mathbb{R})}
\end{aligned} \tag{7.1.9}$$

for some proper constant K.

As in the proof of the other side of the inequality, we also have, using the progressive Hörmander condition, the symmetry of the inner product on the tensor space with respect to relabelling, and Theorem 7.1.6,

$$\left|\sum_{\substack{\Delta\in\mathrm{Multi}(B)\\ |\Delta|\le L;\#(\Delta)\ge 2}} \Psi^{U_i}_{x_s}(\Delta)\int_{t_1,\ldots,t_o\in[0,s]^{\#(\Delta)}}^{\Delta} \tilde{\xi}_{t_1,\ldots,t_o} d\mu^{\Delta}_{t_1,\ldots,t_o}\right|^2 \le K|\theta|^{2i}_{L^2,\mathcal{L}^L(\mathbb{R}^d,\mathbb{R})} \tag{7.1.10}$$

for some proper constant K.

Similarly to the proof of inequality (7.1.3), using inequalities (7.1.9) and (7.1.10), we obtain

$$\begin{aligned}
d_{\mathcal{A},t,\log,\infty}(x,y)^2 &\le |\theta|^2_{L^2,\mathcal{L}^L(\mathbb{R}^d,\mathbb{R})} \le K|\mu|^2_{L^2,\bar{\mathcal{S}}} = K(|\mu_1|^2_U + t) \\
&\le K\max(d_{\mathcal{A},\tilde{t},\log}(x,y)^2, t),
\end{aligned}$$

for some proper constant K changing from line to line. □

We finish with the following doubling condition:

PROPOSITION 7.1.14. *Let $\mathcal{A} = (x, \sigma, F)$, F linear, be a uniformly $(r, L, g, \mathcal{G})$-mixed tense, uniformly (H, L) detailed weak progressive Hörmander system. There exist proper constants D and M such that for any $t \le D$ any $r > 0$ and any $\bar{x} \in \mathbb{R}^m$,*

$$|B_{d_{t,\log,\infty}}(x, 2r)| \le M|B_{d_{t,\log,\infty}}(x, r)|.$$

PROOF. This follows from Proposition 5.1.11, Theorem 7.1.1 and the Propositions 5.2.8 and 7.1.10. $\square$

7.2. Comparisons between distances

Here we compare the Léandre, homogeneous, and Carnot-Carathéodory distances in certain cases. This type of results are well known in both the classical situation (cf. [**1**]) and in geometric settings.

DEFINITION. Let $\mathcal{A} = (x, \sigma, F)$ be a (L, g, G)-tense, (L, H_L)-weak Hörmander system. The Carnot-Carathéodory distance is defined by

$$d_c(y) = \inf\Big(|\gamma|_{L^2} : Y_t((\tau, \gamma)) = y,\ \gamma \in \mathcal{P}_1^d,\ \tau_s = 0\,\forall s\Big).$$

First, we note the following trivial facts:

PROPOSITION 7.2.1. *Let $\mathcal{A} = (x, \sigma, F)$ be a (L, g, G)-tense, (L, H_L)-weak Hörmander system. For any y, we have*

$$d_t(y) \le \tilde{d}_t(y).$$

PROPOSITION 7.2.2. *Let $\mathcal{A} = (x, \sigma, F)$ be a (L, g, G)-tense, (L, H_L)-weak Hörmander system with $\sigma^{(0)} = 0$ uniformly. For any y, we have*

$$d_t(y) = \tilde{d}_t(y) = \sqrt{d_c(y)^2 + t}.$$

Now we consider the link between the distances in slightly less trivial cases.

PROPOSITION 7.2.3. *Let $\mathcal{A} = (x, \sigma, F)$ be a uniformly (L, g, G)-tense, (L, H_L)-weak Hörmander system such that there exist functions λ_i (for $i = \{1, \ldots, d\}$) and a fixed constant $\bar{H}$ such that we have, uniformly in x:*

$$\sigma^{(0)}(x) = \sum_{i=1}^{d} \lambda_i(x)\sigma^i(x) \quad \textit{and} \quad \sum_i |\lambda_i(x)|^2 \le \bar{H}^2.$$

(This implies the strong Hörmander condition.)

Then there exists a polynomial constant C such that for any $y \in \mathbb{R}^n, t \in \mathbb{R}+$ such that $d_t(y), d_c(y), t \le 1/2$,

$$C^{-1} d_c(y) \le d_t(y) \le C(d_c(y) + \sqrt{t}).$$

PROOF. For the left hand side, following the characterisation from Proposition 7.0.4, let $\Gamma \in \mathcal{P}_1^{d+1}$ be a control such that $\mathrm{Sol}_{x,t}(\Gamma) = y$. Write X_s ($s \in [0,1]$) for the solution curve, so that $X_1 = \mathrm{Sol}(\Gamma)$. We begin by re-parametrising Γ so that for all s,

$$\sqrt{|\dot{\Gamma}^0| + \sum_{i=1}^{d} |\dot{\Gamma}_s^i|^2} = d_t(y).$$

Pick $\gamma \in \mathcal{P}_1^d$ such that for all $s \in [0,1]$, $i = 1, 2, \ldots, d$,

$$d\gamma_s^i = d\Gamma_s^i + \lambda^i(X_s)d\Gamma_s^0.$$

This ensures that the solution to the SDE

$$X_0 = x,$$

$$dX_s = \sum_{i=1}^{d} \sigma^i(X_s)d\gamma_s^i$$

satisfies $F(X_1) = y$. Since $d_t(y) \le 1/2$, we have that $|\dot{\Gamma}_s^0| < 1$. Therefore,

$$d_c(y)^2 \le |\gamma|_{L^2}^2 \le \sum_{i=1}^{d} \int_0^1 (\dot{\Gamma}_s^i + \lambda_s^i \dot{\Gamma}_s^0)^2 ds$$

$$\leq 2\sum_{i=1}^{d}\int_0^1 (\dot{\Gamma}_s^i)^2 + (\lambda_s^i\dot{\Gamma}_s^0)^2 ds$$

$$\leq 2\int_0^1 \bar{H}^2(\dot{\Gamma}_s^0)^2 + \sum_{i=1}^{d}(\dot{\Gamma}_s^i)^2 ds$$

$$\leq 2\bar{H}^2\int_0^1 \sum_{i=0}^{d}(\dot{\Gamma}_s^i)^2 ds$$

$$\leq 2\bar{H}^2\int_0^1 |\dot{\Gamma}_s^0| + \sum_{i=1}^{d}(\dot{\Gamma}_s^i)^2 ds$$

$$\leq 2\bar{H}^2 d_t(y)^2.$$

The proof of the right hand side is very similar. □

PROPOSITION 7.2.4. *Let $\mathcal{A} = (x, \sigma, F)$ be a uniformly (L, g, G)-tense system (with $g \geq 2$) such that the vector fields σ are uniformly (L, H_L)-weakly progressively finitely generated, and such that there exist functions λ_α and a fixed constant $\bar{H}$ such that we have, uniformly in x:*

$$\sigma^0(x) = \sum_{|\alpha|=\#(\alpha)\leq 2} \lambda_\alpha(x)\sigma^{[\alpha]}(x) \tag{7.2.1}$$

$$\sum |\lambda_\alpha|^2 \leq \frac{1}{\bar{H}}.$$

There exist a constants D_1, D_2, C_1, C_2, polynomial in $G, H_L, \bar{H}, d$, such that or any y, $t \in \mathbb{R}^+$ such that $\max(\sqrt{t}, d_{t,\log,\infty}(y)) \leq D_1$,

$$d_c(y) \leq C_1(d_{t,\log,\infty}(y) + \sqrt{t}),$$

and for any y, $t \in \mathbb{R}+$ such that $\max(\sqrt{t}, d_c(y)) \leq D_2$,

$$d_{t,\log,\infty}(y) \leq C_1(d_c(y) + \sqrt{t}).$$

PROOF. By applying Theorem 7.1.1 to the system $\mathcal{B}$ obtained by replacing $\sigma^{(0)}$ by 0 uniformly in $\mathcal{A}$, we know that $d_c(y)$ is polynomially locally equivalent to the 'distance' $d_{c,\log,\perp,L}(y)$, defined as the homogeneous 'distance' associated to the system $\mathcal{B}$.

This means we only need to show

$$d_{c,\log,\perp,L}(y) \leq C_2(d_{t,\log,\perp,L}(y) + \sqrt{t})$$

and

$$d^l_{t,\log,\perp,L}(y) \leq C_3(d_{c,\log,\perp,L}(y) + \sqrt{t}).$$

Proof of the second inequality

We use as usual the notation H for the basis of

$$\mathcal{S}_1 = \operatorname{span}_{\substack{\alpha\in\mathrm{Multi}(\{0,1,\ldots,d\})\\ |\alpha|\leq L, \alpha\neq(0)}} e^{[\alpha]}$$

constructed in Proposition 7.0.9. We will use the notation H_1 for the set obtained similarly for

$$\mathcal{S}_2 = \operatorname{span}_{\substack{\alpha\in\mathrm{Multi}(\{1,\ldots,d\})\\ |\alpha|\leq L,}} e^{[\alpha]}.$$

Let $u \in H_1$ be such that

$$\sum_{h\in H_1} u^h\sigma^h = y - x$$

$$|u|_{\mathcal{L}^L(\mathbb{R}^d)} \leq d_{c,\log,\perp,L}(y)(1+\epsilon),$$

for some small $\epsilon > 0$.

By differentiating the condition (7.2.1), we get that there exists a polynomial constant $\tilde{H}$ and some λ^i_ξ ($i \in \{1, \dots \lfloor L/2 \rfloor\}$, $\xi \in H_1$) such that

$$\sigma^{\otimes^i(0)} = \sum_{\substack{\xi \in H_1 \\ |\xi| \le 2i}} \lambda^i_\xi \sigma^\xi,$$

$$\sum_{\substack{\xi \in H_1 \\ |\xi| \le 2i}} |\lambda^i_\xi|^2 \le \tilde{H}.$$

Then we set, for $h \in H_1$,

$$\bar{u}^h = u^h - \sum_{i \le \lfloor L/2 \rfloor} \lambda^i_h \frac{t^i}{i!}.$$

This ensures that $F^{ST}(\bar{u}^h) = y$. Now note that for $h \in H_1 \subset H$,

$$\begin{aligned} |\bar{u}^h|^2 &\le 2 \left(|u^h|^2 + \left(\sum_{\substack{i \le L, h \in H_1 \\ |h| \le 2i}} |\lambda^i_h| \frac{t^i}{i!} \right)^2 \right) \\ &\le 2 \left(|u^h|^2 + \tilde{H} \sqrt{t}^{2|h|} \right) \\ &\le 4\tilde{H} \max(|u|_{\mathcal{S}_2}, \sqrt{t})^{2|h|} \end{aligned}$$

for some polynomial constant M. Summing the contributions of the constant degree spaces to calculate the homogeneous norm, we get

$$\begin{aligned} |\bar{u}|_{\mathcal{S}_2} &\le M \max(|u|_{\mathcal{S}_2}, \sqrt{t}) \\ &\le M(1+\epsilon) \max(d_{c,\log,\perp,L}(y), \sqrt{t}), \end{aligned}$$

for some polynomial constant M. Taking for instance $\epsilon = 1$ gives the right hand side.

Proof of the first inequality

The proof is very similar. We write H_0 for the set $\cup_{i \le \lfloor L/2 \rfloor} \{e^{\otimes^i(0)}\}$. By differentiating condition 7.2.1, we obtain a polynomial constant $\tilde{H}$ and some λ^ζ_ξ (with $\zeta \in H \setminus H_1, \xi \in H_1$) such that

$$\sigma^\zeta = \sum_{\xi \in H_1, |\xi| \le |\zeta|} \lambda^\zeta_\xi \sigma^\xi$$

$$\sum_{\xi \in H_1, |\xi| \le |\zeta|} |\lambda^\zeta_\xi|^2 \le \tilde{H}.$$

That allows one to define $\tilde{u} \in \mathcal{S}_2$ by

$$\tilde{u}^h = u^h + \sum_{\substack{\zeta \in H \setminus \cup(H_1, H_0) \\ |\zeta| \ge |h|}} \lambda^\zeta_h u^\zeta + \sum_{\substack{\zeta \in H_0 \\ |\zeta| \ge |h|}} \lambda^\zeta_h \frac{\sqrt{t}^{|\zeta|}}{(\frac{|\zeta|}{2})!}.$$

This ensures that $F^{\log(S),T}_{\mathcal{B}}(\tilde{u}) = y$.

Then we note

$$\begin{aligned} |\tilde{u}^h| &\le |u^h| + \sum_{\substack{\zeta \in H \setminus \cup(H_1, H_0) \\ |\zeta| \ge |h|}} |\lambda^\zeta_h||u^\zeta| + \sum_{\substack{\zeta \in H_0 \\ |\zeta| \ge |h|}} |\lambda^\zeta_h| \frac{\sqrt{t}^{|\zeta|}}{(\frac{|\zeta|}{2})!} \\ &\le |u^h| + \sum_{\substack{\zeta \in H \setminus \cup(H_1, H_0) \\ |\zeta| \ge |h|}} |\lambda^\zeta_h||u|^{|h|}_{\mathcal{S}_1} + \sum_{\substack{\zeta \in H_0 \\ |\zeta| \ge |h|}} |\lambda^\zeta_h| \sqrt{t}^{|h|} \\ &\le M_2 \max(|u|_{\mathcal{S}_1}, \sqrt{t})^{|h|}, \end{aligned}$$

for some polynomial constant M_2. Again, summing contributions from each constant degree spaces, we get

$$\begin{aligned} d_{c,\log,\perp,L}(y) &\leq |\tilde{u}|_{\mathcal{S}_2} \\ &\leq M_3 \max(|u|_{\mathcal{S}_1}, \sqrt{t}) \\ &\leq M_3(1+\epsilon)\max(d_{t,\log,L}, \sqrt{t}) \end{aligned}$$

for some polynomial M_3. □

7.3. Symmetrisation and integrability of upper bounds

PROPOSITION 7.3.1. *Let $d_t : \mathbb{R}^m \to \mathbb{R}^+$ be a continuous time-dependent function with $t \in [0,T]$ for some fixed $T \leq 1$, such that there exist $\tilde{x} \in \mathbb{R}^m$ and a constant K such that for any $y \in \mathbb{R}^m$,*

$$d_t(y) + 1 \geq K_2|\tilde{x} - y|,$$

and satisfying a doubling condition with constant D. For any C, M, the function

$$E_t(x,y) = C\frac{e^{-\frac{M d_t(y)^2}{t}}}{|B_{d_t}(\sqrt{t})|}$$

is space-time integrable and we have

$$\int_{t\in[0,T]}\int_{y\in\mathbb{R}^m} E_t(x,y) \leq U$$

for some constant U that depends only on K, C, M, D, but not $\tilde{x}$ or m

PROOF. This is simple calculation. □

We note the following classical result from [**13**] (or [**47**], [**37**] etc.):

LEMMA 7.3.2. *Let $\sigma^1, \sigma^2, \ldots, \sigma^d$ be some smooth vector fields on $\mathbb{R}^n$, and let d be the control distance associated to those vector fields. For every $\beta', \beta > 0$ there exists $c(\beta, \beta')$ such that*

$$\frac{e^{-\frac{\beta d(x,y)^2}{t}}}{|B_d(x,\sqrt{t})|} < c(\beta,\beta')\frac{e^{-\frac{\beta' d(x,y)^2}{t}}}{|B_d(y,\sqrt{t})|}.$$

The following corollary follows immediately:

COROLLARY 7.3.3. *Let $\sigma^1, \sigma^2, \ldots, \sigma^d$ be some smooth vector fields on $\mathbb{R}^n$, and let d be the control distance associated to those vector fields. We have for any $\beta > 0$,*

$$\int_{t\in[0,T]}\int_{y\in\mathbb{R}^n} \frac{e^{-\frac{\beta d(x,y)^2}{t}}}{|B_d(y,\sqrt{t})|}dy < \infty.$$

This is shows that in the case of *trivial drift*, the upper bounds are integrable in both y and x. It is integrability in y that is relevant to some applications such as generalisations of Löcherbach's theorem.

We cannot do this in the case of non trivial drift because our 'distance' is not symmetric. However, we can do the following:

LEMMA 7.3.4. *Suppose that we are given a sufficiently smooth function $F : \mathbb{R}^m \to \mathbb{R}^n$, and a time dependent function $d_t : \mathbb{R}^m \otimes \mathbb{R}^m \to \mathbb{R}^+$, continuous in t. For $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$, we write $d_t(x,y)$ for $\min_{\substack{x_2\in\mathbb{R}^m \\ F(x_2)=y}} d(x,x_2)$. Balls $B_t(x,r)$ are always taken as sets in $\mathbb{R}^n$. Suppose that for some constants D and M, we have*

$$\begin{aligned} \forall t, r \leq D, \forall x, \quad & |B_t(x,r)| \leq M|B_t(x,2r)| \\ \forall s \leq t,\ \forall x_1, x_2, x_3 \in \mathbb{R}^m, \quad & d_t(x_1,x_3) = d_s(x_1,x_2) + d_{t-s}(x_2,x_3) \\ \forall s, r, t \leq D, \forall x \in \mathbb{R}^m, \quad & |B_s(x,r)| \leq M|B_t(x,r)|. \end{aligned}$$

Then the following holds: For any $x_1, x_2 \in \mathbb{R}^m$ and $t, r \leq D$ such that $d_t(x_1, x_2) < r$,

$$M^{-2}|B_t(x_1, r)| \leq |B_t(x_2, r)| \leq M^2|B_t(x_1, r)|.$$

PROOF. Note that by continuity of d in t, there exists an $\epsilon > 0$ such that $d_{t-\epsilon}(x_1, x_2) < r$. Clearly if $\mathbb{R}^n \ni y \in B_\epsilon(x_2)$. Then there exists an $x_3 \in \mathbb{R}^m$ such that $d_\epsilon(x_2, x_3) \leq r$ and $F(x_3) = y$. Then we have

$$d_t(x_1, y) = d_t(x_1, x_3) \leq d_{t-\epsilon}(x_1, x_2) + d_\epsilon(x_2, x_3) < r + r = 2r,$$

from which it follows that $y \in B_t(x_1, 2r)$.

This means that $B_\epsilon(x_2, r) \subset B_t(x_1, 2r)$. Then we have

$$|B_t(x_2)| \leq M|B_\epsilon(x_2, r)| \leq M|B_t(x_1, 2r)| \leq M^2|B_t(x_1, r)|.$$

The other inequality is proved similarly. □

DEFINITION. For such a time dependent 'distance' as in Lemma 7.3.4, we define the backwards balls around x, by

$$\overleftarrow{B}^{\mathbb{R}^m}_{t,x}(r) = \{y \in \mathbb{R}^m : d_t(y, x) < r\},$$
$$B_{-t,x}(r) = \overleftarrow{B}_{t,x}(r) = \overleftarrow{B}^{\mathbb{R}^n}_{t,x}(r) = F(\overleftarrow{B}^{\mathbb{R}^m}_{t,x}(r)).$$

PROPOSITION 7.3.5. *Let $\mathcal{A} = (x, \sigma, F)$ be a uniformly detailed weak progressive Hörmander system such that the σ's are uniformly mixed finitely progressively generated in the background space. There exist proper constants D and M such that for any $t \leq D$ and for any $x_1, x_2 \in \mathbb{R}^m$ such that $d_t(x_1, x_2) \leq D\sqrt{t}$, we have*

$$|B_t(x_2, \sqrt{t})| \leq M|B_t(x_1, \sqrt{t})|,$$
$$|B_{-t}(x_2, \sqrt{t})| \leq M|B_t(x_1, \sqrt{t})|.$$

PROOF. First, note that by simple integration theory, the coarse log-homogeneous distance $d^l_{t,\log}$ satisfies all the requirements of Lemma 7.3.4 with proper constants.

It follows that for $t, d^l_{t,\log}(x_1, x_2) \leq D_1$ (D_1 proper), we have for some proper M_1,

$$|B_{d^l_{t,\log}}(x_2, \sqrt{t})| \leq M_1|B_{d^l_{t,\log}}(x_1, \sqrt{t})|.$$

Now, using Theorem 7.1.1, which has proper constants, we can write, for C being a proper constant from Theorem 7.1.1 and $\bar{D}$ being the doubling constant for the $d_{t,\log,\infty}$ distance, and for $t, d^l_{t,\log}(x_1, x_2) \leq \frac{D_1}{C}$:

$$\begin{aligned}
|B_{d_{t,\log,\infty}}(x_2, \sqrt{t})| &\leq |B_{d^l_{t,\log}}(x_2, \sqrt{t}C)| \leq |B_{d^l_{t,\log}}(x_2, \sqrt{t})|\bar{D}^{\log_2(C)} \\
&\leq M_1\bar{D}^{\log_2(C)}|B_{d^l_{t,\log}}(x_1, \sqrt{t})| \\
&\leq M_1\bar{D}^{\log_2(C)}|B_{d_{t,\log,\infty}}(x_1, C\sqrt{t})| \\
&\leq M_1\bar{D}^{2\log_2(C)}|B_{d_{t,\log,\infty}}(x_1, \sqrt{t})|.
\end{aligned}$$

All constants involved are proper. This proves the first inequality.

For the second one, note that the reverse distance d_{-t} is just the 'distance' associated to the vector fields $-\sigma^i$ ($i \neq 0$) and σ^0, which doesn't affect the progressive Hörmander constants. In fact even for negative values of s, t, we have:

$$\forall s, r, t \text{ with}, r, |s|, |t| \leq D,\ \forall x \in \mathbb{R}^m, \quad |B_{s,\log,\infty}(x, r)| \leq M|B_{t,\log,\infty}(x, r)|,$$

so the proof works similarly. □

PROPOSITION 7.3.6. *Let $\mathcal{A} = (x, \sigma, F)$ be a uniformly (L, H_L) weak Hörmander, uniformly (L, g, G)-tense system ($g \geq L + 1$). There exist constants D, M, C, such that for any x_1, x_2 such that $\sqrt{t} \leq D$, $d_t(x_1, x_2) \leq \sqrt{t}D$,*

$$|B^l_t(x_2, \sqrt{t})| \leq C|B^l_t(x_1, \sqrt{t})|$$

$$\begin{aligned}
|B^l_{-t}(x_2,\sqrt{t})| &\leq C|B^l_t(x_1,\sqrt{t})| \\
|B^\infty_t(x_2,\sqrt{t})| &\leq C|B^l_t(x_1,\sqrt{t})| \\
|B^\infty_{-t}(x_2,\sqrt{t})| &\leq C|B^l_t(x_1,\sqrt{t})| \\
|B_t(x_2,\sqrt{t})| &\leq C|B_t(x_1,\sqrt{t})| \\
|B_{-t}(x_2,\sqrt{t})| &\leq C|B_t(x_1,\sqrt{t})|.
\end{aligned}$$

PROOF. Again, by observation of the Jacobian of the function $F_t^{\log(S),T}$ for different values of t (including negative ones, note that the reverse distance d_{-t} is just a forward distance associated to the reversed vector fields $\sigma^i(-1)^{1_{i\neq 0}}$ etc.), we have that the distance d^l_t satisfies all the requirements of Lemma 7.3.4 (with bad constants from doubling).[5] The result follows. □

THEOREM 7.3.1. *Let $\mathcal{A} = (x,\sigma,F)$ be a uniformly (L,H_L) weak Hörmander, uniformly (L,g,G)-tense system. For any $\beta > 0$, there exist constants $c(\beta)$, D and $m(\beta)$ such that*

$$\frac{e^{-\frac{\beta d_t(y,x)^2}{t}}}{|B_{d_t}(y,\sqrt{t})|} < c(\beta)\frac{e^{-\frac{m(\beta)d_t(y,x)^2}{t}}}{|\overleftarrow{B}_{d_t}(x,\sqrt{t})|}.$$

PROOF. For the case $d(y,x) \leq D\sqrt{t}$ (D being the constant relative to $d_{t,\infty}$ from Proposition 7.3.6), we have by Proposition 7.3.6:

$$\frac{e^{-\frac{\beta d_t(y,x)^2}{t}}}{|B_{d,t}(y,\sqrt{t})|} \leq C\frac{e^{-\frac{m(\beta)d_{t,\infty}(y,x)^2}{t}}}{|\overleftarrow{B}_{d_t}(x,\sqrt{t})|}$$

where C is the constant from Proposition 7.3.6.

For the case $d(y,x) \geq D\sqrt{t}$, we have first:

$$B_t(x,\sqrt{t}) \subset B_{2t}(y,\sqrt{t}+d(y,x))$$

which gives, using Proposition 5.1.11, for some constants M and ν,

$$\begin{aligned}
|B_t(x,\sqrt{t})| \leq |B_{2t}(y,\sqrt{t}+d(y,x))| &\leq M|B_t(y,\sqrt{t}+d(y,x))| \\
&\leq M(\frac{\sqrt{t}+d(y,x)}{\sqrt{t}})^\nu |B_t(y,\sqrt{t})|.
\end{aligned}$$

This allows us to conclude, when $d(y,x) \geq D\sqrt{t}$,

$$\begin{aligned}
\frac{e^{-\frac{\beta d_t(y,x)^2}{t}}}{|B_{d,t}(y,\sqrt{t})|} &\leq M\left(\frac{\sqrt{t}+d(y,x)}{\sqrt{t}}\right)^\nu \frac{e^{-\frac{\beta d_t(y,x)^2}{t}}}{|B_{d,t}(x,\sqrt{t})|} \\
&\leq M^2\left(\frac{\sqrt{t}+d(y,x)}{\sqrt{t}}\right)^\nu \frac{e^{-\frac{\beta d_t(y,x)^2}{t}}}{|\overleftarrow{B}_{d,t}(x,\sqrt{t})|} \\
&\leq M^2\left(1+\frac{1}{D}\right)\left(\frac{d(y,x)}{\sqrt{t}}\right)^\nu \frac{e^{-\frac{\beta d_t(y,x)^2}{t}}}{|\overleftarrow{B}_{d,t}(x,\sqrt{t})|} \\
&\leq c(\beta)\frac{e^{-\frac{m(\beta)d_t(y,x)^2}{t}}}{|\overleftarrow{B}_{d_t}(x,\sqrt{t})|}.
\end{aligned}$$

□

[5] The same observation is related to the following similar result from [**39**]: up to a constant C uniform in x, balls of radius $\sqrt{t}$ around x behave like $\sum_{\substack{|A|\leq 2L,\\ A_1,A_2,\ldots\in\mathrm{Multi}(\{0,1,\ldots,d\})}} \sqrt{t}^{|A|}\det(\sigma^{[A_1]}\sigma^{[A_2]}\ldots\sigma^{[A_{\#(A)}]})$

Now, for systems satisfying the progressive Hörmander condition, we have the following analogous result (the proof is the same, replacing the distances, constants become proper because doubling condition constants are proper):

THEOREM 7.3.2. *Let $\mathcal{A} = (x, \sigma, F)$ be an $(L, g, \mathcal{G})$-mixed tense, (L, H_L) weak detailed-Progressive Hörmander system ($g \geq L+1$). For any $\beta > 0$, there exist proper constants $c(\beta)$, D and $m(\beta)$ such that*

$$\frac{e^{-\frac{\beta d_{t,\log,\infty}(y,x)^2}{t}}}{|B_{d_{t,\log,\infty}}(y, \sqrt{t})|} < c(\beta) \frac{e^{-\frac{m(\beta) d_{t,\log,\infty}(y,x)^2}{t}}}{|\overleftarrow{B}_{d_{t,\log,\infty}}(x, \sqrt{t})|}.$$

REMARK 7.3.7. Note that the coarse log-homogeneous distance has the following nice symmetry property, which follows directly from the linearity of the model function F^{RT}:

$$\overleftarrow{d}^{\,l}_{x,t,\log}(\bar{x}_t + v) = d^l_{x,t,\log}(\bar{x}_{-t} + v)$$

where $\overleftarrow{d}^{\,l}_{x,t,\log}$ denotes the coarse log homogeneous distance associated to the time reversed equation, $v \in \mathbb{R}^m$, and as usual, for any $s \in \mathbb{R}$ (here $s = t$ or $-t$),

$$\bar{x}_t = x + \sum_{\substack{\alpha \in \mathrm{Multi}_0(L) \\ |\alpha| \leq L}} \sigma^\alpha(x) \frac{s^{\#(\alpha)}}{\#(\alpha)!}.$$

In fact, we even have, more generally, for any s, t such that $|s|, |t| \leq D$ for some proper constant,

$$d^l_{x_s,\log}(\bar{x}_s + v) = d^l_{x_t,\log}(\bar{x}_t + v).$$

CHAPTER 8

Application of models to an auxiliary object: probabilistic results

8.1. Some Lemmas and preliminary estimates

To prove our results, we need a more flexible and general version of the scaling argument in the proof of Theorem (3.12) on page 411 of [**37**].

THEOREM 8.1.1 (Scaling). *Let $U = \bigoplus_{i=1}^I U_i$ be a Euclidean space endowed with the following homogeneous norm:*

$$|u|_U = |u|_{U,t} = \left(\sum_i (|u^i|)^{2/i}\right)^{1/2}.$$

where for $u \in U$, u^i is the projection of u onto the space U_i. Let $\nu = \sum_i (i \dim(U_i))$. We introduce the family of dilations δ_s defined by:

$$\mathrm{pr}_{U_i}(\delta_s(u)) = u^i s^{i/2}.$$

Suppose that $\forall t \le T$, ξ_t is a random variable in U, such that there exist $T \le 1$, $K \le 1$, $k \in \mathbb{N}$, $C \ge 1$ and $M \le 1$ such that

$$\forall s, t \in \mathbb{R}, u \in U \text{ such that } s, t \le T, s \ge 1, |u|_{\text{eucl}} \le K,$$

$$\mathbb{P}(\delta_s(\xi_t) = u) \le \frac{C}{(st)^k} \exp\left(-\frac{|u|^2_{\text{eucl}}}{st}\right).$$

Then, there exist $C_1 \ge 1$, $M_1 \le 1$ such that $\forall t \le T$, $\forall u \in U$ with $|u|_{\text{eucl}} \le 1$,

$$\mathbb{P}(\xi_t = u) \le \frac{C_1}{t^{\frac{\nu}{2}}} \exp\left(-\frac{M_1 |u|^2_U}{t}\right).$$

Furthermore, M_1 depends only on M and I, and the constant C_1 is polynomial in C, M and only depends on C, M, ν, T.

PROOF. For the sake of this proof, we also introduce the following alternative homogeneous norm:

$$\omega(u) = (\sup(s : |\delta_s(u)| \le 1))^{-\frac{1}{2}}.$$

For I a natural number, we write

$$D_I = \left(\max\left(\sup_{a \in \mathbb{R}^i, a \neq 0} \frac{\sum_i a_i^2}{\sum_i |a_i|^{2/i}}, \sup_{a \in \mathbb{R}^i, a \neq 0} \frac{\sum_i |a_i|^{2/i}}{\sum_i a_i^2}\right)\right)^{1/2}.$$

We first note the following crucial fact: $\forall u \in U$, we have:

$$D_I^{-1} \omega(u) \le |u|_U \le D_I \omega(u).$$

We can now begin the proof per se:

Case 1 $\frac{t}{\omega(u)^2} \le T$. We can set $s = \frac{1}{\omega(u)^2}$, then $st \le T$ and $s \ge 1$. Furthermore, $|\delta_s(u)| \le 1$, so we can use our first assumption to obtain:

$$\begin{aligned}\mathbb{P}(\xi_t = u) &= \frac{1}{s^{\frac{\nu}{2}}} \mathbb{P}(\delta_s(\xi_t) = \delta_s(u)) \\ &\le \frac{1}{s^{\frac{\nu}{2}}} C \frac{e^{-\frac{M|\delta_s(u)|^2}{st}}}{(st)^k}\end{aligned}$$

$$
\begin{aligned}
&= \omega(u)^{\nu} C \frac{e^{-\frac{\frac{M}{t}}{\omega(u)^2}}}{(\frac{t}{\omega(u)^2})^k} \\
&\leq \frac{C\omega(u)^{2\nu}}{t^{\frac{\nu}{2}}} \frac{e^{-\frac{\frac{M}{t}}{\omega(u)^2}}}{(\frac{t}{\omega(u)^2})^{k-\frac{\nu}{2}}} \\
&\leq \frac{CK_k \omega(u)^{2\nu}}{M^{k-\frac{\nu}{2}} t^{\frac{\nu}{2}}} e^{-\frac{M_1\omega(u)^2}{t}} \\
&\leq \frac{CK_k}{M^{k-\frac{\nu}{2}} t^{\frac{\nu}{2}}} e^{-\frac{M\omega(u)^2}{2t}} \\
&\leq \frac{CK_k}{M^{k-\frac{\nu}{2}} t^{\frac{\nu}{2}}} e^{-\frac{MD_I^{-1}|u|_U^2}{2t}} \\
&\leq \frac{CK_k}{M^{k-\frac{\nu}{2}} t^{\frac{\nu}{2}}} e^{-\frac{M_1|u|_U^2}{2t}}.
\end{aligned}
$$

We have the following justifications:

1. at the first line we have used the change of variables $v = \delta_s(u)$;
2. at the second line, we have used the first assumption;
3. at the third line, we have used the expression of our choice of s and the fact that by definition of $\omega(u)$, we have $|\delta_s(u)| = 1$;
4. at the fifth line, we have defined a constant K_k, dependent on k only, as follows: $K_k = \max_{x\leq 1}(\frac{e^{-\frac{1}{2x}}}{x^k})$;
5. at the sixth line, we have used the fact that $\omega(u) \leq 1$;
6. at the last line, we have defined $M_1 = MD_I^{-1.}$

Case 2 $\frac{t}{\omega(u)^2} \geq T$. We set $s = \frac{T}{t}$ Then we have:

$$
\begin{aligned}
\mathbb{P}(\xi_t = u) &= \frac{1}{s^{\frac{\nu}{2}}} \mathbb{P}(\delta_s(\xi_t) = \delta_s(u)) \\
&\leq \frac{1}{s^{\frac{\nu}{2}}} C \frac{e^{-\frac{M|\delta_s(u)|^2}{st}}}{(st)^k} \\
&\leq t^{\frac{\nu}{2}} \frac{e^{-\frac{M|\delta_s(u)|^2}{T}}}{T^k} \\
&\leq \frac{1}{t^{\frac{\nu}{2}} T^k} \\
&\leq \frac{e^{MT}}{t^{\frac{\nu}{2}} T^k} e^{-\frac{M\omega(u)^2}{t}} \\
&\leq \frac{e^{MT}}{t^{\frac{\nu}{2}} T^k} e^{-\frac{M_1|u|_U^2}{t}}
\end{aligned}
$$

where we have the following justifications:

1. at the first line we have used the change of variables $v = \delta_s(u)$;
2. at the second line, we have used the first assumption;
3. at the fourth line, we have used the fact that $t \leq T \leq 1$;
4. at the fifth line, we have used the fact that $\frac{t}{\omega(u)^2} \geq T$;
5. at the last line, we have used the fact that $M_1 = \frac{MD_I^{-1}}{2} \leq MD_I^{-1}$

Conclusion Combining both of the above cases, we see that the result is true with

$$C_1 = \max(\frac{e^{MT}}{T^k}, \frac{CK_k}{M^{k-\frac{\nu}{2}}}) \quad \text{and} \quad M_1 = \frac{MD_I^{-1}}{2}.$$

□

THEOREM 8.1.2. *Let (x,σ,F) be a uniformly (L,g,G)-tense with $g \geq n+3$, uniformly (L,H_L) weak Hörmander system, set $l = L$, there exist polynomial constants C,D,M, and ν, such that for any $y \in \mathbb{R}^n$ with $d_{t,\infty}(x,y) \leq D$ and any $t \leq D$,*

$$p_t(x,y) \leq C\frac{e^{-\frac{Md_t(x,y)^2}{t}}}{t^\nu}.$$

PROOF. Define $q(x,y,\xi)$ for $\xi \in \mathbb{R}^n$ to be the following variation of the Fourier transform of $p_t(x,y)$:

$$\hat{p}_t(\xi) = \int_{\mathbb{R}^n} e^{-i\langle\xi,y\rangle} p_t(x,y)\phi((y-\xi)t^{-L})dy$$

where ϕ is a fixed localizing function with arbitrary derivatives bounded by a proper constant, such that for any $y \in \mathbb{R}^n$, $\phi(y) \leq 1$ and for any $y \in \mathbb{R}^n$ with $|y| \leq 1$, $\phi(y) = 0$.

Note that we have already proved in the previous part of the thesis that for any $N \geq 0$ there exists a polynomial constants $D > 0, M, C, \mu$ such that for any $t \leq D$ and

$$p_t(x,y) \leq C\frac{e^{-M\frac{|*x-y|^2}{t}}}{t^\mu}.$$

Using that, we can conclude, similarly to the proof of Theorem 6.4.1, that; (for any $t \leq D$),

$$\begin{aligned}\hat{p}_t(\xi) &\leq \frac{1}{|\xi|^{n+1}}|\int_{\mathbb{R}^n} e^{-i\langle\xi,y\rangle}\frac{\partial^{n+1}p_t(y)}{\pi_{i=1}^{n+1}\partial z_i}dy| \\ &\leq \frac{K}{|\xi|^{n+1}t^{\mu-n}}\end{aligned}$$

for some polynomial constant K.

Next, we have for any $R > 0$

$$\begin{aligned}p_t(x,y) &\leq \frac{1}{2\pi}|\hat{p}_t(x,\cdot)|_{L^2} \\ &\leq \int_{B_0(R)\subset\mathbb{R}^n}\hat{p}_t(x,\xi)dx + \int_{B_0(R)^C}\hat{p}_t(x,\xi)dx \\ &\leq K(R^n\mathbb{P}(|Y_t-y| \leq t^L) + R^{-1}\frac{1}{t^{\mu-n}})\end{aligned}$$

for some (other) polynomial constant K.

Setting $R = t^{-\frac{\mu-n}{n+1}}\mathbb{P}(|Y_t-y| \leq t^L)^{-\frac{1}{n+1}}$, we get

$$p_t(x,y) \leq Kt^{-\frac{(\mu-n)n}{n+1}}\mathbb{P}(|Y_t-y| \leq t^L)^{\frac{1}{n+1}}.$$

Hence, we only need to check that there exist polynomial constants C, M such that for any $t \leq D$,

$$\mathbb{P}(|Y_t-y| \leq t^L) \leq Ce^{-\frac{Md_t(x,y)^2}{t}}.$$

To do that, we use Corollary 7.0.14 as follows:

Let $\bar{Y}_t$ denote the KST approximation ($\bar{X}_t$ its background space counterpart), and let K denote the minimum of the constant from Corollary 7.0.14.

First note that if $d^l_{x,t}(y) \leq 3Kt^{\frac{1}{2}}$, we have

$$\mathbb{P}(|Y_t-y| \leq t^L) \leq 1 \leq 1.e^{-\frac{d_t(x,y)^2}{9K^2t}}.$$

Therefore, forcing $C \geq 1$ and $M \leq \frac{1}{9K^2}$ ensures we only need to worry about the case $d_{x,t}(y) \geq 3Kt^{\frac{1}{2}}$. In that case we have first

$$|Y_t-y|^{\frac{1}{L}} \geq |Y_t-y|^{\frac{1}{l}} \geq \frac{1}{K}d_{\frac{t-t^{l+1}}{2}}(X_t,y)$$

$$\geq \frac{1}{K}(d_t(x,y) - d_{\frac{t-t^{l+1}}{2}}(x,\bar{Y}_t) - d_{t^{l+1}}(\bar{X}_t,Y_t)).$$

Set D small enough that $t \leq D$ implies $4G^{-1}Kt^{\frac{1}{2}} \leq \bar{D}$ where $\bar{D}$ is the constant from Corollary 7.0.14. Then

$$d_{\frac{t-t^{l+1}}{2}}(x,z) \leq K \implies |{*x} - z| < 2G^{-1}Kt^{\frac{1}{2}} \implies |z-y| < 4G^{-1}Kt^{\frac{1}{2}} \leq \bar{D}.$$

For $z \in \mathbb{R}^m$, write $\mathcal{R}^l_s(z) = z + \sum_{\substack{\#(\alpha)\leq\frac{l}{2}\\ \nexists i:\alpha_i\neq 0}} \sigma^\alpha_z \frac{s^{\#(\alpha)}}{\#(\alpha)!}$. Then (because $t \leq 1$), writing S for the signature, and using Theorem 2.1.1, as well as Corollary 7.0.14 and Theorem 7.0.1,

$$\begin{aligned}
&\mathbb{P}(|Y_t - y| \leq t^L)\\
&\quad\leq \mathbb{P}(d_{\frac{t-t^{l+1}}{2}}(x,\bar{Y}_t) > d_t(x,y)/3)\\
&\qquad + \mathbb{P}(d_{t^{t+1}}(\bar{X}_t,Y_t) > d_t(x,y)/3 \vee d_{\frac{t-t^{l+1}}{2}}(x,\bar{Y}_t) \leq Kt^{\frac{1}{2}})\\
&\quad\leq \mathbb{P}(d_{\frac{t-t^{l+1}}{2}}(x,\bar{Y}_t) > Bd_t(x,y)/3)\\
&\qquad + \mathbb{P}(d_{t^{t+1}}(\bar{X}_t,Y_t) > d_t(x,y)/3 \vee d_{\frac{t-t^{l+1}}{2}}(x,\bar{Y}_t) \leq Kt^{\frac{1}{2}})\\
&\quad\leq \mathbb{P}(|S^{\log}_{\frac{t-t^{l+1}}{2}}|_{\mathcal{S}} > Bd_t(x,y)/3) + \mathbb{P}(|{*\mathcal{R}^l_{t^{l+1}}}(\bar{X}_t) - Y_t| \geq (d_t(x,y)/(3K))^l)\\
&\quad\leq K_L \exp\left(-\frac{(Bd_t(x,y))^2}{72tL}\right) + \mathbb{P}(|{*\mathcal{R}^l_{t^{l+1}}}(\bar{X}_t) - Y_t| \geq (d_t(x,y)/(3K))^l)\\
&\qquad\qquad \text{(by Theorem 2.1.1).}
\end{aligned}$$

Now note that

$$|{*\mathcal{R}^l_{t^{l+1}}}(\bar{X}_t) - {*\bar{X}_t}| < LGt^{l+1} \leq \frac{LG}{(3K)^{l+1}}(d^l_t(x,y))^{2(l+1)}.$$

Then we can continue the above calculation to obtain, using Theorem 6.3.1 (consequence of Theorem 2.1.1):

$$\begin{aligned}
&\mathbb{P}(|Y_t - y| \leq t^L)\\
&\quad\leq K_L \exp\left(-\frac{(Bd_t(x,y))^2}{72tL}\right)\\
&\qquad + \mathbb{P}\left(|{*\bar{X}_t} - Y_t| \geq (d_t(x,y)/(3K))^l - \frac{LG}{(3K)^{l+1}}(d_t(x,y))^{2(l+1)}\right)\\
&\quad\leq K_L \exp\left(-\frac{(Bd_t(x,y))^2}{72tL}\right) + E\exp(-\frac{\bar{E}d_t(x,y)^{\frac{2l}{l+1}}}{t})
\end{aligned}$$

for some polynomial constants E and $\bar{E}$. The result follows. □

THEOREM 8.1.3. *Let (x,σ,F) be a uniformly $(L,g,\mathcal{G})$ mixed-tense ($g \geq n+3$), uniformly (L,H_L) detailed weak progressive Hörmander system, set $l = L$, there exists polynomial constants C,D,M and ν such that for any $y \in \mathbb{R}^n$ with $d_{t,\infty,\log}(x,y) \leq D$ and any $t \leq D$,*

$$p_t(x,y) \leq C\frac{e^{-\frac{Md_{t,\log,\infty}(x,y)^2}{t}}}{t^\nu}.$$

PROOF. The proof is exactly the same as the proof of Theorem 8.1.2. □

We note the following, which is interesting in itself:

LEMMA 8.1.1. *Let $d \in \mathbb{N}$. Consider the following vector fields in $\mathcal{T}^l(\mathbb{R}^d,\mathbb{R})$: for $i = 0,1,2,\ldots d$, for $z \in \mathcal{T}^l(\mathbb{R}^d,\mathbb{R})$, $w^i_z = \sum_{|(\alpha,i)|\leq l} z^\alpha \frac{\partial}{\partial z^{(\alpha,i)}}$ (with the convention that $z^\varnothing = 1$) The tension G of order $(\bar{l},\infty)$ of the SDE driven by w, restricted to a Euclidean ball of radius R ($R > 0$), is bounded above by a constant that depends on $\bar{l}$ and R only.*

PROOF. We have, by Cauchy-Schwarz:

$$G \leq \sum_{N\geq 0} \sup_{\substack{v\in T^l(\mathbb{R}^d,\mathbb{R})\\ |v|=1}} \sup_{x\in B(0,R)} \sup_{\substack{v^1,v^2\ldots v^N\in T^l(\mathbb{R}^d,\mathbb{R})\\ |v^1|,|v^2|,\ldots|v^N|=1}} \sum_{\substack{\gamma\in\mathrm{Multi}(\{0,1,\ldots,d\})\\ |\gamma|\leq\bar{l}}} \langle \frac{\partial w_x^\gamma}{\partial v^1\partial v^2\ldots\partial v^N}, v\rangle^2$$

$$\leq \sum_{N\geq 0} \sup_{\substack{v\in T^l(\mathbb{R}^d,\mathbb{R})\\ |v|=1}} \sup_{x\in B(0,R)} \sum_{\substack{|\alpha^k|\leq l\forall k\\ \alpha\in\mathrm{Multi}(\{0,1,\ldots,d\})}} \sum_{\substack{\gamma\in\mathrm{Multi}(\{0,1,\ldots,d\})\\ |\gamma|\leq\bar{l}}} \langle \frac{\partial w_x^\gamma}{\partial e^{\alpha^1}\partial e^{\alpha^2}\ldots\partial e^{\alpha^N}}, v\rangle^2.$$

Now, observe the following crucial fact:

$$\frac{\partial w_z^\gamma}{\partial e^\beta} = \frac{\partial}{\partial z^{(\beta,\gamma)}},$$

so we can continue the above calculation as follows:

$$G \leq \sum_{N\geq 0} \sup_{\substack{v\in T^l(\mathbb{R}^d,\mathbb{R})\\ |v|=1}} \sup_{x\in B(0,R)} \sum_{\forall k,|\alpha^k|\leq l} \sum_{\substack{\gamma\in\mathrm{Multi}(\{0,1,\ldots,d\})\\ |\gamma|\leq\bar{l}}} \langle \frac{\partial w_x^\gamma}{\partial e^{\alpha^1}\partial e^{\alpha^2}\ldots\partial e^{\alpha^N}}, v\rangle^2$$

$$\leq \sum_{N\geq 0} \sup_{\substack{v\in T^l(\mathbb{R}^d,\mathbb{R})\\ |v|=1}} \sup_{x\in B(0,R)} \left(\sum_{\substack{\gamma\in\mathrm{Multi}(\{0,1,\ldots,d\})\\ |\gamma|\leq\bar{l}}} \langle \sum_\beta \frac{x^\beta\partial}{\partial x^{(\beta,\gamma)}}, v\rangle^2 + \sum_{\substack{\gamma\in\mathrm{Multi}(\{0,1,\ldots,d\})\\ |\gamma|\leq\bar{l}}} \sum_\beta \langle \frac{\partial}{\partial x^{(\beta,\gamma)}}, v\rangle^2 \right)$$

$$\leq \bar{l}|v|^2|x|^2 + \bar{l}|v|^2 = \bar{l} + \bar{l}|x|^2 \leq \bar{l}(1+R^2).$$

□

LEMMA 8.1.2. *Let $d\in\mathbb{N}$. Consider the following vector fields in $\mathcal{L}^l(\mathbb{R}^d,\mathbb{R})$: for $i = 0,1,2,\ldots d$, for $z\in\mathcal{T}^l(\mathbb{R}^d,\mathbb{R})$, $w_z^i = \sum_{\#(\alpha)\leq l-1} z^\alpha \frac{\partial}{\partial z^{(\alpha,i)}}$ (with the convention that $z^\varnothing = 1$). The weak Hörmander constant H_l, restricted to a Euclidean ball of radius R ($R>0$) in $\mathcal{L}^l(\mathbb{R}^d,\mathbb{R})$, is bounded below by a proper constant (depending on l and R only, not on d).*

PROOF. By Lemmas 3.5.1 and 8.1.1, is equivalent to prove that any vector v in $\mathcal{L}^l(\mathbb{R}^d,\mathbb{R})$ there exist some λ_α such that $\sum_{|\alpha|\leq l}\lambda_\alpha w^{[\alpha]} = v$. It is equivalent to prove the result over $\mathcal{L}^l(\mathbb{R}^d)$ rather than $\mathcal{L}^l(\mathbb{R}^d,\mathbb{R})$, indeed, it follows from the orthogonality of multi-indices that are not a reordering of each other that if

$$v\in\mathcal{L}^l(\mathbb{R}^d,\mathbb{R})\subset\mathcal{L}^l(\mathbb{R}^{d+1})$$

and

$$\sum_{\substack{\#(\alpha)\leq l\\ \alpha\in\mathrm{Multi}(\{0,1,\ldots,d\})}} \lambda_\alpha w^{[\alpha]} = v,$$

then

$$\sum_{\substack{|\alpha|\leq l\\ \alpha\in\mathrm{Multi}(\{0,1,\ldots,d\})}} \lambda_\alpha w^{[\alpha]} = v$$

(identifying $\mathcal{L}^l(\mathbb{R}^d,\mathbb{R})$ and its corresponding subset of $\mathcal{L}^l(\mathbb{R}^{d+1})$)

We prove this result by induction over l. The result is clearly true for $l=1$. Now suppose the result is true for l. There is a constant C_l such that for any unit $v\in\mathcal{L}^l(\mathbb{R}^d)$ and any $z\in\mathcal{L}^l(\mathbb{R}^d)$ with $|z|\leq R$, there exist λ such that

$$\sum_{\#(\alpha)\leq l} \lambda_\alpha w^{[\alpha]} = v,$$

and $\sum_\alpha \lambda_\alpha^2 \leq C_l^{-1}$ Now, for fixed $z \in \mathcal{L}^{l+1}(\mathbb{R}^d)$, let $u \in \mathcal{L}^{l+1}(\mathbb{R}^d)$ be a unit vector. Set $v = \Pr_{\mathcal{L}^l(\mathbb{R}^d)}(u)$ and apply the induction hypothesis. We now have a $\lambda \in \mathcal{L}^l(\mathbb{R}^d)$ such that $\sum_{\#(\alpha)\leq l} \lambda_\alpha^2 \leq C_l^{-1}$ and

$$\Pr_{\mathcal{L}^l(\mathbb{R}^d)}(u) = \Pr_{\mathcal{L}^l(\mathbb{R}^d)}(\sum_{\#(\alpha)\leq l} \lambda_\alpha w^{[\alpha]}).$$

Recall the explicit formula for $w^{[\alpha]}$: $w^{[\alpha]} = \sum_\beta z_\beta \frac{\partial}{\partial z^{(\beta,[\alpha])}}$. In particular, $|w^{[\alpha]}| \leq R$. It follows that

$$|\Pr_{\mathrm{span}_{\#(\alpha)=l+1}(e^{[\alpha]})}(\sum_{\#(\alpha)\leq l} \lambda_\alpha w^{[\alpha]})|^2 \leq R^2 C_l^{-1}.$$

Now note that for $\#\alpha = l+1$, $w^{[\alpha]} = \frac{\partial}{\partial z^{[\alpha]}}$. Therefore, using also Proposition 7.0.9, we see that there exist some λ_α (with $\#(\alpha) = l+1$) such that

$$\sum_{\#(\alpha)=l+1} w^{[\alpha]} = \Pr_{\mathrm{span}_{\#(\alpha)=l+1} e^{[\alpha]}}(u) - \Pr_{\mathrm{span}_{\#(\alpha)=l+1} e^{[\alpha]}}(\sum_{\#(\alpha)\leq l} \lambda_\alpha w^{[\alpha]}),$$

and

$$\sum_{\#(\alpha)=l+1} |\lambda_\alpha|^2 \leq 2(R^2 C_l^{-1} + R^2) B_l,$$

where B_l is the constant from Proposition 7.0.9. Clearly this implies

$$\sum_{\#(\alpha)\leq l+1} \lambda_\alpha w^{[\alpha]} = u \qquad \text{and}$$

$$\sum_{\#(\alpha)\leq l+1} |\lambda_\alpha|^2 \leq 2(R^2 C_l^{-1} + R^2) B_l + C_l^{-1}.$$

We can therefore set $C_{l+1} = (2(R^2 C_l^{-1} + R^2) B_l + C_l^{-1})^{-1}$, which concludes the proof. □

8.2. Global upper bound for a weak Hörmander system with drift

Here we prove the analogue the result in [**37**] for systems with drift, with constants which are still independent of m, but exhibit fast dependence in d. The key, local part of the estimate was already achieved in the SDE case in [**39**]. The difference is that our estimate is global (we use Theorem 8.1.2), but that is at the cost of using the distance d_t instead of $\tilde{d}_t$ (as in [**39**]). Of course, no lower bound could hold with the distance d_t in general[1]. Still, our global bound is space-time integrable.

LEMMA 8.2.1. *Let (x, σ, F) be a uniformly (L, g, G)-tense, uniformly (L, H_L) weak Hörmander system. For $l \geq L$, if $g \geq l+3$ the model $(\mathcal{S}, \mathbb{R}^n, F_t^{\log(S)T})$ is regular, and the corresponding constants $\gamma, \Gamma, \rho_1, \rho_2, \rho_3, \bar{\rho}_1$ only depend on the constants from the assumptions, and in particular, do not depend on the initial point x.*

PROOF. Similar results have been proved, in a different formulation, in the drift-free case in [**37**] and in the general case in [**39**]. Our proof is similar but we work on (projections of) log-signatures rather than signatures. Our notation is closer to [**37**] (cf. Lemmas (3.15) and (3.23)) than [**39**], but our result, like [**39**] but unlike [**37**], is general enough to apply to systems with non trivial drifts.

Let $r_1 = \frac{1}{2}\min\left(\frac{\varepsilon}{K}\ \frac{\beta}{4n!K^{2n}}\right)$. For $u \in B(0, r_1) \subset \mathcal{S}_l$, consider the map

$$f_u \colon B(0, r_1) \to R^n,\ \xi \mapsto F_0^{\log(S),T}(u + (JF_0^{\log(S),T}|_u)^T \xi).$$

We have $\beta_{f_u} \geq \beta/2$, and the derivatives of f are uniformly bounded by a (proper) constant independent of u and x. In particular, applying Lemma 5.2.2, we see that there are proper constants $r_2, r_3 < r_1$ and C (not dependent on u) such that there exists an open subset W_u of $B(u, r_2)$

[1]For specific cases where it is possible, see the separated weak Hörmander condition defined later

such that $f_u : W_u \to B(f_u(u), r_3)$ is diffeomorphic, and for any $y \in \mathbb{R}^n$, setting $g_u(y) = u + (JF_0^{\log(S),T}|_u)^T f_u^{-1}(y)$, we have

$$|y - f_u(u)| \leq r_3 \implies |g_u(y)| \leq C|y - f_u(0)|. \tag{8.2.1}$$

For $u \in B(0, r_2)$, similarly to the proof of the regularity of linear models (cf. 5.1.8) we define $R_1(u)$ such that

$$|R_1(u)|_U = \min(|\bar{u}|_U : F_t(\bar{u}) = F_t(u)).$$

For a continuously differentiable curve h in $\mathbb{R}^{d+1}$ we will write $\mathrm{sol}(x, h)$ for the solution to the differential equation

$$X_0 = x$$

$$dX_s = \sum_{i=0}^{d} \sigma^i(X_s) dh_s^i.$$

Note that for any $x_1, x_2 \in \mathbb{R}^m$ and any smooth curve h,

$$|\,\mathrm{sol}(x_1, h) - \mathrm{sol}(x_2, h)| \leq e^G |x_1 - x_2|. \tag{8.2.2}$$

Now, let h^1 be a smooth curve in $\mathbb{R}^{d+1}$, parametrised over $[0, 1]$ such that $\mathrm{logsig}(h^1) = (t, u)$. Let also h^2 be a curve in $\mathbb{R}^{d+1}$ (also parametrised over $[0, 1]$) such that $\mathrm{logsig}(h^2) = (t, R_1(u))^{-1}$, where $(t, R_1(u))^{-1}$ denotes the group inverse of $(t, R_1(u))$. We can ensure $|h^2|_{L^2}^2 \leq C(t + |u|_U^2)$ and $|h^1|_{L^2}^2 \leq C(|u|_U^2 + t)$ for some new constant C still not dependent on x or u. Note that if we write $\bar{h}^2$ for the curve that satisfies $\bar{h}_s^2 = h_{1-s}^2$ for all s, then $\mathrm{logsig}(\bar{h}^2) = (t, R_1(u))$.

Let

$$v = \mathrm{Pr}_{\mathcal{S}_l}\left((t, u) \otimes (t, R_1(u))^{-1}\right).$$

In the calculations that follow, C, r_1, r_2 is are constants that can change in each line but only depends on the constants in the original assumptions, and in particular does not depend on u or x.

We have on the one hand:

$$|\,\mathrm{sol}(x, \bar{h}^2) - F_{x,t}^{\log(S),T}(u)| = |\,\mathrm{sol}(x, \bar{h}^2) - F_{x,t}^{\log(S),T}(R_1(u))| \leq C|R_1(u)|_U^{l+1}.$$

and on the other hand

$$|\,\mathrm{sol}(x, h^1) - F_{x,t}^{\log(S),T}(u)| \leq C|u|_U^{l+1}$$

Combining both of the above, we have

$$|\,\mathrm{sol}(x, h^1) - \mathrm{sol}(x, \bar{h}^2)| \leq C(|u|_U^{l+1} + |R_1(u)|_U^{l+1}),$$

and therefore, using Eq. (8.2.2),

$$\begin{aligned} |x - \mathrm{sol}(x, h^1 \otimes h^2)| &= |\,\mathrm{sol}(x, \bar{h}^2 \otimes h^2) - \mathrm{sol}(x, h^1 \otimes h^2)| \\ &\leq C(|u|_U^{l+1} + |R_1(u)|_U^{l+1}). \end{aligned}$$

But we also have

$$|F_{x,0}^{\log(S),T}(v) - \mathrm{sol}(x, h^1 \otimes h^2)| \leq C(|u|_U^{l+1} + |R_1(u)|_U^{l+1}).$$

Therefore, combining the above, we obtain

$$|F_{x,0}^{\log(S),T}(v) - x| \leq C(|u|_U^{l+1} + |R_1(u)|_U^{l+1}) \leq C|u|_U^{l+1}.$$

Using (8.2.1), it follows (after making r_1 smaller, but still not dependent on x or u) that for $|u|_U \leq r_1$, we have

$$|g_v(F_{x,0}^{\log(S),T}(v) - x) - v| \leq C|u|_U^{l+1}. \tag{8.2.3}$$

We therefore define

$$R_2(u) = g_v(F_{x,0}^{\log(S),T}(v) - x).$$

Equation (8.2.3), and Lemmas 5.1.2, Lemma 14 and Proposition 6.3.4 ensure that

$$|R_2(u)|_U \leq C|v|_U + C|u|_U^{(l+1)/l} \leq C(|R_1(u)|_U + |u|_U + |u|_U^{(l+1)/l}) \leq C|u|_U,$$

and we can set $\gamma = C$ for the C from the above equation. Note that the use of Proposition 6.3.4 is a crucial step of the proof.

The smoothness and differentiability of $R_2|_{F_t^{-1}(\{u\})}$ follow from the corresponding properties of g_u, and the comparison between the volume forms on the fibers $F_t^{-1}(\{F_t(u)\})$ and $F_0^{-1}(\{F_0(0)\})$ follows from the bounds on the determinant of g_u and its inverse.

Finally, similarly to the the proof of the regularity of linear models 5.1.8, if

$$|u|,\ \sqrt{t} \leq (2(I+1))^{-I/2}(r_1/2)^I,$$

then $|u|_U \leq r_1$, so we can set $\rho_3 = (2(I+1))^{-I/2}(r_1/2)^I$. Then setting $\rho_2 = \rho_3/2$, we certainly have again that $|u|, \sqrt{t} \leq \rho_1 := (\frac{\rho_2}{2C\sqrt{2(I+1)}})^I$ implies $|R_1(u)|_U \leq |u|_U \leq \rho_2/C$ and $|R_2(u)| \leq \rho_2$, as required. □

REMARK 8.2.2. It is interesting to see how simple addition of vectors in the proof of 5.1.8 replaces approximate Lie group multiplication in the proof of 8.2.1. This is related to the intuitive idea, explained in the introduction, that the definition of $d_{t,\log,\infty}$ obtained from a definition of the control distance through getting rid of the exponential function in the RDE

$$dz = \partial f.d\boldsymbol{\gamma} = \partial f.d\exp(\log(\boldsymbol{\gamma}))$$

for the control.

THEOREM 8.2.1. *Let (x, σ, F) be a uniformly (L, g, G)-tense, uniformly (L, H_L) weak Hörmander system, with $g \geq (2n+2)^2 2^{4L} + n + 3$. There exist constants D, C and M, depending only on d, n, H_L, L, G, such that for any x, y and any $t \leq D$,*

$$p_t(x, y) \leq C\frac{e^{-\frac{d_t(x,y)^2}{Mt}}}{|B_{d_t}(x, \sqrt{t})|}.$$

PROOF. Let $l = (2n+2)^2 2^{4L}$. Let S denote the projection of the log signature of W on $\mathcal{S}$ as a random variable. Consider the system $(0, \mathrm{Pr}_{\mathcal{S}}(w^i), \mathrm{Id})$, where as usual, w^i denote the free vector fields in $T^l(\mathbb{R}^{d+1})$. Consider a smooth localising function ξ such that $\xi(S) = 1$ if $|S| \leq \frac{\rho}{2}$ and $\xi(S) = 0$ when $|S| \geq \rho$, where ρ is taken small enough to ensure

$$\det(JF_{t,S}^{ST}(JF_{t,S}^{ST})^T) > \frac{H_L}{2}.$$

Note that JF^{ST} and its derivatives are bounded above by some constant K for any $t \leq 1$ (because $g \geq (2n+2)^2 2^{4L} + n + 3$ and $l = (2n+2)^2 2^{4L}$).

Note that because S_T is the solution at time T of the system $(0, \mathrm{Pr}_{\mathcal{S}}(w^i), \mathrm{Id})$, and this system (for any $T \leq \rho$), restricted to a ball of radius ρ, is uniformly tense and weak Hörmander for some fixed constants depending only on d, n, H_L, L, G, we can apply Theorem 4.4.1 with the localising function $\xi(S)$ to obtain (for any $T \leq \rho$):

$$\mathbb{E}\left(\xi(S_T)\delta(S_T = u)\right) \leq C\frac{e^{-\frac{|u|^2}{MT}}}{T^k},$$

for some constants C, M, k.

Now, for $T \leq \rho$ (for instance, $T = \rho$), for any $t \leq T$, consider the random variable $R = \delta_{\sqrt{T/t}}(S_t)$. Because of Brownian scaling, R_t is distributed in the same way as S_T. As a result,

$$\mathbb{E}\left(\xi(R_{T,t})\delta(R_{T,t} = u)\right) \leq C\frac{e^{-\frac{|u|^2}{MT}}}{T^k}.$$

Thus we can apply Theorem 8.1.1 to obtain:

$$\mathbb{E}\left(\xi(R_{T,t})\delta(S_t = u)\right) \leq C\frac{e^{-\frac{|u|^2_{\mathcal{S}}}{Mt}}}{t^{\frac{\nu}{2}}} \tag{8.2.4}$$

where ν is the homogeneous dimension of $\mathcal{S}$.

Further localisation, and further use of scaling, would yield the global equivalent of the above estimate without a localising function, as is performed in [**37**] (cf. Theorem 3.12). This is not strictly necessary to obtain the final estimate however, and to stay in line with the proof of our other theorems, we don't go in this direction, and stop at the estimate (8.2.4).

Now, we can use the disintegration argument from our theory on models, more specifically, Theorem 5.1.1, to obtain the following:

$$\begin{aligned}\mathbb{E}\left(\xi(R_{T,t})\delta(\bar{Y}_t = y)\right) &\leq C\frac{e^{-\frac{d^l_t(x,y)^2}{Mt}}}{|B_{d^l_t}(x,\sqrt{t})|}\\ &\leq C\frac{e^{-\frac{d_t(x,y)^2}{Mt}}}{|B_{d_t}(x,\sqrt{t})|},\end{aligned}$$

for some (different) C, M changing between the two expressions after an application of Theorem 7.0.1.

Now, we can compare the ($\xi(R_{T,t})$-perturbed) density of the KSTA $\bar{Y}_t$ to that of the actual solution Y_t using Theorem 6.4.1 to get:

$$\mathbb{E}\left(\xi(R_{T,t})\delta(Y_t = y)\right) \leq C\frac{e^{-\frac{d_t(x,y)^2}{Mt}}}{|B_{d_t}(x,\sqrt{t})|} + Mt,$$

for some (different) C, M, D and for any $t, d_t(x,y)^2 \leq D$.

Now, because the Malliavin derivative of $\xi(R_{T,t})$ is bounded regardless of whether $|S_t| \leq \rho$, $1 - \xi(R_{T,t}) = 0$ whenever $|S_t| > \frac{\rho}{2}$ and the system $((x), (\sigma), F)$ is uniformly tense and weak Hörmander, we can use Theorem 4.4.1 with localising function $(1 - \xi(R_{T,t}))$ to obtain:

$$\begin{aligned}\mathbb{E}\left((1-\xi(R_{T,t}))\delta(Y_t = y)\right) &\leq C\frac{e^{-\frac{|*x-y|^2}{Mt}}}{t^k}\sqrt{\mathbb{P}(|S_t| \geq \frac{\rho}{2})}\\ &\leq C\frac{e^{-\frac{|*x-y|^2}{Mt}}}{t^k}e^{-\frac{Q}{t}}\\ &\leq Ce^{-\frac{Q}{t}} \leq Ct,\end{aligned}$$

for some C, M, Q, k changing from line to line and for any $t \leq D$ for some constant D.

Fix a constant $M_1 \leq D$ (for instance, $M_1 = D$). For $t \leq e^{-\frac{d_t(x,y)^2}{2\nu Mt}}$, where ν is the constant in Theorem 8.1.2, using the above estimates, we get the following (where C, M, Q are constants that are allowed to depend on M_1 and to change from line to line):

$$\begin{aligned}\mathbb{E}\left(\delta(Y_t = y)\right) &= \mathbb{E}\left(\xi(R_{T,t})\delta(Y_t = y)\right) + \mathbb{E}\left((1-\xi(R_{T,t}))\delta(Y_t = y)\right)\\ &\leq C\frac{e^{-\frac{d_t(x,y)^2}{Mt}}}{|B_{d_t}(x,\sqrt{t})|} + Mt + Ct\\ &\leq C\frac{e^{-\frac{d_t(x,y)^2}{Mt}}}{|B_{d_t}(x,\sqrt{t})|} + (M+C)e^{-\frac{d_t(x,y)^2}{2\nu Mt}}\\ &\leq C\frac{e^{-\frac{d_t(x,y)^2}{Mt}}}{|B_{d_t}(x,\sqrt{t})|}.\end{aligned} \tag{8.2.5}$$

Now, for $t \geq e^{-\frac{d_t(x,y)^2}{2\nu Mt}}$, using Theorem 8.1.2, we get (for some constant ν and some constants C, M that can change from line to line, M coming from the application of Theorem 8.1.2), for any x, y, t such that $t, d_t(x,y)^2 \leq D$ (for some constant D)

$$\begin{aligned} \mathbb{E}\left(\delta(Y_t = y)\right) &\leq C\frac{e^{-\frac{d_t(x,y)^2}{Mt}}}{t^\nu} \\ &\leq Ce^{-\frac{d_t(x,y)^2}{2Mt}}\frac{e^{-\frac{d_t(x,y)^2}{2Mt}}}{t^\nu} \\ &\leq C\frac{e^{-\frac{d_t(x,y)^2}{2Mt}}}{|B_{d_t}(x,\sqrt{t})|}. \end{aligned} \tag{8.2.6}$$

Note that by 7.0.6 and 7.0.14, we have that $d_t(x,y) \leq K(1 + |*x - y|)$ for some constant K. Noting also that for $d_t(x,y)^2 \geq D$ there is a constant D' such that $|*x - y| \geq D'$ (by Proposition 7.0.7), it follows that $d_t(x,y) \leq K(|*x-y|)$ for some (other K) whenever $d_t(x,y)^2 \geq D$. Finally, if $d_t(x,y)^2 \geq D$, we conclude (by Theorems 4.4.2 and the above remarks) that for some ν,

$$\mathbb{E}(\delta(Y_t = y)) \leq C\frac{e^{-\frac{d_t(x,y)^2}{Mt}}}{|B_{d_t}(x,\sqrt{t})|} \tag{8.2.7}$$

for some constants C, M.

Putting together the estimates (8.2.5), (8.2.6) and (8.2.7) yields the desired estimate. ☐

8.2.1. Time-dependent SDE. As a particular case of the above theorem, we can obtain a full generalisation to time-dependent coefficients of the main upper bound in [**39**], and a partial generalisation of the upper bound result in [**37**].

Let X be the solution to the following SDE:

$$\begin{aligned} X_0 &= x_0 \in \mathbb{R}^n \\ dX_t &= \sum_{i=1}^{d} \sigma^i(t, X_t) \circ dW_t^i + \sigma^0(t, X_t)dt. \end{aligned}$$

Define iterated brackets (in $\mathbb{R}^n$) of the σ^i's via the following iterative formula:

$$\begin{aligned} \sigma^{[i]} &= \sigma^i \qquad \text{and} \forall \alpha \in \mathrm{Multi}(\{0, 1, \ldots, d\}) \\ \sigma^{[i,\alpha]} &= \frac{\partial \sigma^{[\alpha]}}{\partial \sigma^i} - \frac{\partial \sigma^i}{\partial \sigma^{[\alpha]}} \qquad \text{if} \quad i \neq 0 \quad \text{and} \\ \sigma^{[0,\alpha]} &= \frac{\partial \sigma^{[\alpha]}}{\partial \sigma^0} + \frac{\partial \sigma^{[\alpha]}}{\partial t} - \frac{\partial \sigma^0}{\partial \sigma^{[\alpha]}}. \end{aligned}$$

We assume that the vector fields σ^i are in C^∞ for each i, and satisfy the weak Hörmander condition uniformly in the following sense:

There exist constants L, H_L such that for all $x \in \mathbb{R}^n$, for all $t \in \mathbb{R}^+$ and for all unit $v \in \mathbb{R}^n$,

$$\sum_{\substack{|\alpha| \leq L \\ \alpha \neq (0)}} \langle \sigma^{[\alpha]}(t,x), v \rangle^2 \geq H_L.$$

The distances d_t and $\tilde{d}_t$ from the rest of this thesis have the following definition in this particular context: $d_t(x,y)$ is the minimum of

$$\int_0^1 \sqrt{|\dot{\gamma^0}_s| + \sum_{i=1}^{d} |\dot{\gamma^i}_s|^2 ds}$$

over all the possible smooth curves $\gamma \in \mathbb{R}^{d+1} \otimes [0,1]$ such that

$$\gamma_1^0 = t,$$

and y is the solution to the following SDE:

$$\begin{aligned} X_0 &= x \\ dX_s &= \sum_{i=1}^{d} \sigma^i(\gamma_s^0, X_s) d\gamma_s^i + \sigma^0(\gamma_s^0, X_s) dt. \end{aligned}$$

$\tilde{d}_t$ is the minimum of

$$\int_0^1 \sqrt{t + \sum_{i=1}^{d} |\dot{\gamma^i}_s|^2 ds}$$

over all the possible smooth curves $\gamma \in \mathbb{R}^d \otimes [0,1]$ such that y is the solution to the following SDE:

$$\begin{aligned} X_0 &= x \\ dX_s &= \sum_{i=1}^{d} \sigma^i(s, X_s) d\gamma_s^i + \sigma^0(s, X_s) dt. \end{aligned}$$

Now, applying Theorem 8.2.1, we obtain:

THEOREM 8.2.2 (General integrable upper bound for time-dependent SDE). *Let X be the solution to the following SDE:*

$$\begin{aligned} X_0 &= x_0 \in \mathbb{R}^n \\ dX_t &= \sum_{i=1}^{d} \sigma^i(t, X_t) \circ dW_t^i + \sigma^0(t, X_t) dt, \end{aligned}$$

where the σ^i's are supposed to satisfy the C^∞ and weak Hörmander conditions from above. There exist constants $D, C, M > 0$ such that for all $t \leq D$, X_t admits a density $p_t(x,y)$ satisfying the following inequality:

$$p_t(x,y) \leq C \frac{e^{-\frac{d_t(x,y)^2}{Mt}}}{B_{d_t}(x, \sqrt{t})},$$

where d_t is the distance defined above.

In particular, the above implies the following two theorems:

THEOREM 8.2.3 (Time dependent version of Léandre's result from [**39**]). *Let X be the solution to the following SDE:*

$$\begin{aligned} X_0 &= x_0 \in \mathbb{R}^n \\ dX_t &= \sum_{i=1}^{d} \sigma^i(t, X_t) \circ dW_t^i + \sigma^0(t, X_t) dt, \end{aligned}$$

where the σ^i's are supposed to satisfy the C^∞ and weak Hörmander conditions from above. There exist constants $D, C, M > 0$ such that for all $t \leq D$, X_t admits a density $p_t(x,y)$, and for all $xy \in \mathbb{R}^n$ such that $\tilde{d}_t(x,y) \leq \sqrt{t}$, we have

$$p_t(x,y) \leq \frac{C}{|B_{\tilde{d}_t}(x, \sqrt{t})|},$$

where $\tilde{d}_t$ is the distance defined above.

For vector fields satisfying the analogue of the Kusuoka-Stroock condition:

THEOREM 8.2.4. *Let X be the solution to the following SDE:*

$$X_0 = x_0 \in \mathbb{R}^n$$
$$dX_t = \sum_{i=1}^{d} \sigma^i(t, X_t) \circ dW_t^i + \sigma^0(t, X_t)dt.$$

We assume that

(1) *The σ^i's are C^∞*
(2) *For any $t \geq 0$, the time-frozen vector fields $\sigma^1(t, \cdot), \sigma^2(t, \cdot), \ldots, \sigma^d(t, \cdot)$ satisfy the (strong) Hörmander condition uniformly.*
(3) *There exists a constant Λ such that for any $x \in \mathbb{R}^n$, there exist some real numbers $\lambda_1, \ldots, \lambda_d$ such that*

$$\sum_{i=1}^{d} \lambda_i(x)\sigma^i(x) = \sigma^0(x) \qquad \textit{and}$$
$$\sum_{i=1}^{d} |\lambda_i(x)|^2 \leq \Lambda.$$

Define the distance d_t as the minimum of

$$\int_0^1 \sqrt{|\dot{\gamma^0}_s| + \sum_{i=1}^{d} |\dot{\gamma^i}_s|^2 ds}$$

over all the possible smooth curves $\gamma \in \mathbb{R}^{d+1} \otimes [0,1]$ such that

$$\gamma_1^0 = t,$$

and y is the solution to the following SDE:

$$X_0 = x$$
$$dX_s = \sum_{i=1}^{d} \sigma^i(\gamma_s^0, X_s) d\gamma_s^i.$$

Then there exist constants D, C, M such that for any $t \leq D$, X_t admits a density that satisfies, for all $x, y \in \mathbb{R}^n$:

$$p_t(x,y) \leq C \frac{e^{-\frac{d_t(x,y)^2}{Mt}}}{|B_{d_t}(x, \sqrt{t})|}.$$

8.3. Polynomial integrable upper bound for progressive Hörmander systems

We now have everything in place to prove the main result of the thesis:

THEOREM 8.3.1. *Let $\mathcal{A} = (x, \sigma, F)$, be a uniformly $(r, 2L+1, g, \mathcal{G})$-mixed tense*[2]*, uniformly (L, H_L)-progressive Hörmander system such that the σ are (L, H_L)-detailed-Progressively uniformly finitely generated. Suppose that $g \geq 2L + 3 + nL$. There exist polynomial constants C, M, D such that for any $t \leq D$,*

$$p_t(x,y) = \mathbb{E}_x(\delta(Y_t = y)) \leq C \frac{e^{-\frac{d_{t,\log,\infty}(x,y)^2}{Mt}}}{|B_{d_{t,\log,\infty}}(x, \sqrt{t})|}.$$

[2]In particular, F is assumed to be linear

If $\mathcal{A}$ is only $(r, L, g, \mathcal{G})$-mixed tense and (H, L)-detailed weak progressive Hörmander locally[3] *inside a Euclidean ball $B(*x, R) \subset \mathbb{R}^n$, there are proper constants C, M such that the above estimate holds for $y \in B(*x, \frac{R}{2})$.*

REMARK 8.3.1. The proof has the following advantages:

- It works directly with the log-homogeneous distance, and takes advantage of the progressive Hörmander structure of the problem.
- There is no need to use the density of a Kusuoka-Stroock-Taylor approximation to approximate the density of the target random variable, or to use the corresponding Fourier argument from [**37**] or the corresponding Malliavin calculus argument from [**6, 50, 8**].
- Assuming a proof of the strongly polynomial equivalent of Theorem 4.4.1, this proof would yield strongly polynomial constants. (This also applies to the proof of Theorem 8.6.1)

PROOF. The main differences with the proof of Theorem 8.2.1 are that

- We have to be more careful about scaling, and use the full power of Theorem 5.2.6, rather than the obvious scaling-invariance of Brownian motion
- All constants are polynomial because the previous propositions we are using have polynomial constants
- The main 'auxiliary random variable' of interest is now the compensated signature, but the localising random variable is still composed of a function involving the log signature, as well as other components[4].
- Thanks to the full construction of the compensated signature, we don't need the Taylor approximation $\tilde{Y}_t$ and we can work directly on the actual solution Y_t. This means we don't need to us Theorem 6.4.1 or any Fourier argument, and is also the reason why we can have a boundedness assumption as weak as controlling the mixed tension of orders $(2L+1, g)$ rather than $((2n+2)^2 2^{4L} + L + 1, g)$.

We fix $l = L$ (we truncate signatures at order L).

Step I [A computational lemma] The first thing to see is that, for $|\log(S)| \leq \rho$ (for some fixed $\rho \leq \frac{1}{2}$), norm of the projection $\mathrm{Pr}_{T\setminus\mathcal{L}}(S)$ of the signature $S = \exp(\log(S))$ on the orthogonal complement of the log-signature space $\mathcal{L}^l(\mathbb{R}^d, \mathbb{R})$ is bounded by a proper constant.

Indeed, by the series expansion for the exponential, and writing $U = \log(S)$, we have

$$\mathrm{Pr}_{T\setminus\mathcal{L}}(\exp(U)) = \mathrm{Pr}_{T\setminus\mathcal{L}}\left(\sum_{k=2}^{L} \frac{U^{\otimes k}}{k!}\right).$$

Furthermore, distributing the products shows that the Euclidean norm $|U^{\otimes k}|$ in $T^l(\mathbb{R}^d, \mathbb{R})$ of $U^{\otimes k}$ is bounded by a proper constant times $|U|^k$.

Step II [Defining an auxiliary system]

Consider the system $\mathcal{R}$ defined as follows:

$$\begin{aligned}
(S, X, Z) &\in (T^L(\mathbb{R}^d, \mathbb{R})) \oplus \mathbb{R}^m \oplus ((T^{L+1}(\mathbb{R}^d, \mathbb{R}) \otimes T^L(\mathbb{R}^d, \mathbb{R})) \otimes \mathbb{R}^m) \\
(S_0, X_0, Z_0) &= (0, x, 0) \\
dS_s^{(\alpha,i)} &= S_s^\alpha \circ dW_s^i \\
dX_s &= \sum_{i=0}^{d} \sigma^i(X_s) \circ dW_s^i
\end{aligned}$$

[3] It is admittedly quite a strict requirement that the boundedness assumptions must hold for any $\bar{x}$ such that $F(\bar{x}) \in B(*x, R)$, but it is satisfied for instance in the case of a composite system with core SDE satisfying conditions locally but with the conditions on λ being uniform over the background space.

[4] Assuming a larger value of g, we could work with the Taylor approximation first, in which case the localising random variable is just a function of the log signature.

$$\begin{aligned}
dZ_s^{\beta,(i)} &= \sigma^\beta(X_s) \circ dW_s^i \\
dZ_s^{\beta,(\alpha,i)} &= Z_s^{\beta,\alpha} \circ dW_s^i \\
R &= F^{SR_\infty}((S_s, X_s, Z_s))
\end{aligned}$$

where R is the compensated signature defined in (7.1.1). This means (from the construction of the compensated signature), that F^{SR_∞} takes the following form

$$\begin{aligned}
\mathrm{Pr}_{F_i}(F^{SR_\infty}(s,S,x,z)) = \sum_{|\alpha|=i} c_\alpha^i S^\alpha + \sum_{|\alpha|=i+1} c_\alpha^{i+1} S^\alpha + \dots \\
+ \sum_{|\alpha|=L} c_\alpha^L S^\alpha + \sum_{|\alpha|=L+1} c_\alpha^{L+1} Z^{\alpha,\alpha}
\end{aligned}$$

where for $j = i, i+1, \dots, L+1$, the c_α^j's are column vectors in

$$F_i = \mathrm{span}_{\substack{|\alpha|=i \\ \alpha\neq(0)}} e^{[\alpha]}$$

coming from the construction of the compensated signature and are such that F^{SR_∞} is a properly bounded linear operator and $\sum_{|\alpha|=i} c_\alpha^i s^\alpha$ has a properly bounded pseudo-inverse.

By straightforward generalisations of step I and Lemma 8.1.2, there is a proper constant ρ such that the system $\mathcal{R}$, restricted to $|(s,S)| \leq \rho$ is uniformly (L,H)-weak Hörmander, for some proper constant H. Furthermore, a straightforward generalisation of Lemma 8.1.1 ensures that the background space component of $\mathcal{R}$ (still restricted to restricted to $|(s,S)| \leq \rho$) is uniformly (L,g,G)-tense for some proper constant G:

Indeed, write

$$\begin{aligned}
\Omega &= (T^L(\mathbb{R}^d,\mathbb{R})) \oplus \mathbb{R}^m \oplus ((T^{L+1}(\mathbb{R}^d,\mathbb{R}) \otimes T^L(\mathbb{R}^d,\mathbb{R})) \otimes \mathbb{R}^m) \\
&= \mathrm{span}_\alpha\, e^\alpha \oplus \mathrm{span}_i\, e_i \oplus \mathrm{span}_{\alpha,\beta,i}\, e_i^{\alpha,\beta}, \\
\mathcal{M} &= \mathrm{Multi}(\{0,1,\dots,d\}), \\
\forall v &\in \Omega,\ \bar\alpha = (\alpha,\beta,\gamma) \in \mathcal{M} \otimes (\mathcal{M}\otimes\mathcal{M}),\ \bar v_{\bar\alpha} = e^\alpha + \sum_i \bar v_i e_i + \sum_i \bar v_i^{\beta,\gamma} e_i^{\beta,\gamma},
\end{aligned}$$

write (with some notational shortcuts)

$$\begin{aligned}
w^i = \sum_{\alpha,i} z^\alpha \frac{\partial}{\partial z^{(\alpha,i)}} + \sum_i \sigma^i(\mathrm{Pr}_{\mathrm{span}_i e_i}(z))_\bullet \frac{\partial}{\partial z_\bullet} \\
+ \sum_{\alpha,\beta,i} z^{\beta,\alpha} \frac{\partial}{\partial z^{\beta,(\alpha,i)}} + \sum \sigma^\beta(\mathrm{Pr}_{\mathrm{span}_i e_i}(z))_\bullet \frac{\partial}{\partial z_\bullet^{\beta,(i)}}
\end{aligned}$$

for the vector fields driving the above equation, E for the standard basis of $\mathbb{R}^m$, and write components of elements of Ω as $z^\alpha = z_\varnothing^{\alpha,\varnothing}$, $z_i = z_i^{\varnothing,\varnothing}$, $z_i^{\alpha,\beta}$ for $\alpha,\beta \in \mathrm{Multi}(\{0,1,\dots,d\})$, $i \in \{0,1,\dots,d\}$. Using vector field notation, and notational shortcuts such as $\frac{\partial}{\partial z^{\alpha,\beta}}$ for $\sum_i \frac{\partial}{\partial z_i^{\alpha,\beta}}$, writing $\bar\gamma$ for the truncation of the multi-index γ so that $\gamma = (\bar\gamma, \gamma_{\#(\gamma)})$, for any $v \in \Omega$ with $|v| = 1$ any $x \in B(0,R) \subset \Omega$, and any $v^1, v^2 \dots, v^N \in \Omega$ with $|v^1|, |v^2|, \dots, |v^N| = 1$, then we have

$$\begin{aligned}
(8.3.1) \sum_{0\leq N\leq g} \sum_{\substack{\gamma\in\mathcal{M} \\ |\gamma|\leq L}} \langle \frac{\partial^N w_x^\gamma}{\partial v^1 \partial v^2 \dots \partial v^N}, v\rangle^2 \\
\leq \sum_{0\leq N\leq g} \sum_{\substack{\forall k, |\alpha^k|\leq l \\ \alpha^k\in\mathcal{M}\otimes(\mathcal{M}\otimes\mathcal{M})}} \sum_{\substack{\gamma\in\mathcal{M} \\ |\gamma|\leq L}} \langle \frac{\partial^N w_x^\gamma}{\partial v_{\alpha^1}^1 \partial v_{\alpha_2}^2 \dots \partial v_{\alpha^N}^N}, v\rangle^2 \\
\leq \sum_{\substack{\gamma\in\mathcal{M} \\ |\gamma|\leq L}} \langle \sum_\delta \frac{x^\delta \partial}{\partial x^{(\delta,\gamma)}}, v\rangle^2 + \sum_{\substack{\gamma\in\mathcal{M} \\ |\gamma|\leq L}} \sum_\delta \langle \frac{\partial}{\partial x^{(\delta,\gamma)}}, v\rangle^2
\end{aligned}$$

$$+ \sum_{\substack{\gamma\in\mathcal{M}\\|\gamma|\le L}} \langle \frac{\partial^N \sigma^\gamma}{\partial \mathrm{Pr}_{\mathrm{span}_i e_i}(v^1)\partial \mathrm{Pr}_{\mathrm{span}_i e_i}(v^2)\dots\partial \mathrm{Pr}_{\mathrm{span}_i e_i}(v^N)}\frac{\partial}{\partial x_{\bullet}}, v\rangle^2$$

$$\sum_{0\le N\le g}\sum_{\substack{\gamma\in\mathcal{M}\\|\gamma|\le L}}\sum_{|\beta|=L+1}\langle\sum_{\alpha}\frac{x^{\beta,\alpha}\partial}{\partial x^{\beta,(\alpha,\gamma)}}, v\rangle^2 + \sum_{\substack{\gamma\in\mathcal{M}\\|\gamma|\le L}}\sum_{|\beta|=L+1}\sum_{\alpha}\langle\frac{\partial}{\partial x^{\beta,(\alpha,\gamma)}}, v\rangle^2$$

$$+ \sum_{0\le N\le g}\sum_{\substack{\gamma\in\mathcal{M}\\|\gamma|\le L}}\sum_{|\beta|=L+1}\langle\frac{\partial^N\sigma^{(\beta,\bar\gamma)}}{\partial \mathrm{Pr}_{\mathrm{span}_i e_i}(v^1)\partial \mathrm{Pr}_{\mathrm{span}_i e_i}(v^2)\dots\partial \mathrm{Pr}_{\mathrm{span}_i e_i}(v^N)}\frac{\partial}{\partial x^{\beta,\gamma_{\#(\gamma)}}}, v\rangle^2$$

$$\le (L|v|^2|x|^2 + L|v|^2) + \mathcal{G} + (L|v|^2|x|^2 + L|v|^2) + \mathcal{G}\sum_{\beta}\sum_{\gamma}|v_{\bullet}^{\beta,(\gamma_{\#(\gamma)})}|^2$$

$$\le 2\mathcal{G} + 2L(|x|^2+1) \le 2\mathcal{G} + 2L(R^2+1).$$

Since the progressive Hörmander condition ensures that the operator norm of F^{SR_∞} (which is linear) is bounded above by a proper constant, this ensures that the auxiliary system is (L, g, G)-tense for some proper constant G.

Next, note that for any $|\alpha| \le L$, we have

$$w_0^{[\alpha]} = \frac{\partial}{\partial z_{\varnothing}^{[\alpha],\varnothing}},$$

from which it follows, using Proposition 8.1.2, that the weak Hörmander constant of the auxiliary system, evaluated at 0, is a proper constant. Finally, the above calculation (8.3.1), together with Proposition 7.0.9, ensures also that the derivative of the weak Hörmander constant in any unit direction of

$$(T^L(\mathbb{R}^d,\mathbb{R}))\oplus\mathbb{R}^m\oplus((T^{L+1}(\mathbb{R}^d,\mathbb{R})\otimes T^L(\mathbb{R}^d,\mathbb{R}))\otimes\mathbb{R}^m),$$

is bounded by a proper constant.

It follows that there is a proper constant ρ such that the system $\mathcal{R}$, restricted to $|(S, X, Z)| \le \rho$, is uniformly (L, g, G)-tense and (L, H) weak Hörmander for proper constants G, H.

Step III [Scaling and local estimate]

Recall that the compensated signature space $\mathcal{F}$ is equipped with a graded structure, and a family of dilations. For $T \le \rho$ (ρ being defined in step II), and for $t \le T$, we define the random variable

$$\Xi^{t,T} = \delta_{\sqrt{T/t}}(F^{SR_\infty}((t, S_t, X_t)))$$

We write $\tilde W_s$ for $W_{sT/t}$. By Brownian scaling,

$$\tilde W_s \simeq W_s\sqrt{T/s}.$$

So we can now rewrite $\Xi^{t,T} = \delta_{\sqrt{T/t}}(F^{SR_\infty}((t, S_t, X_t))$ as the solution *at time* T to the following system:

$$\begin{aligned}
(s,\tilde S, X, Z) &\in (T^l(\mathbb{R}^d,\mathbb{R}))\oplus\mathbb{R}^m\oplus((T^l(\mathbb{R}^d,\mathbb{R})\otimes T^l(\mathbb{R}^d,\mathbb{R}))\otimes\mathbb{R}^m)\\
(\tilde S_0,\tilde X_0) &= (0,x)\\
d(s,\tilde S_s) &= \sum_{i=0}^d w_s^i((s,\tilde S_s))\circ d\tilde W_s^i\\
d\tilde X_s &= \sum_{i=0}^d \frac{\sigma^i(\tilde X_s)}{\sqrt{T/t}}\circ d\tilde W_s^i\\
d\tilde Z^{\beta,(i)} &= \sigma^\beta(\tilde X_s)\circ d\tilde W_s^i\\
d\tilde Z_s^{\beta,(\alpha,i)} &= \tilde Z_s^{\beta,\alpha}\circ d\tilde W_s^i
\end{aligned}$$

$$
\begin{aligned}
\tilde{\Xi}_s^{t,T} &= \phi(s, \tilde{S}_s, \tilde{X}_s, \tilde{Z}_s) \\
&= \sum_{|\alpha|=i} c_\alpha^i \tilde{S}_s^\alpha + \sum_{|\alpha|=i+1} \frac{1}{\sqrt{T/t}} c_\alpha^{i+1} \tilde{S}_s^\alpha + \dots \\
&\quad + \sum_{|\alpha|=L} \frac{1}{\sqrt{T/t}^{L-i}} c_\alpha^L \tilde{S}_s^\alpha + \sum_{|\alpha|=L+1} \frac{1}{\sqrt{T/t}^{L+1-i}} c_\alpha^{L+1} \tilde{Z}_s^{\alpha,\alpha}.
\end{aligned}
$$

For ξ being a smooth localising function on

$$(T^L(\mathbb{R}^d, \mathbb{R})) \oplus \mathbb{R}^m \oplus ((T^{L+1}(\mathbb{R}^d, \mathbb{R}) \otimes T^L(\mathbb{R}^d, \mathbb{R})) \otimes \mathbb{R}^m)$$

with $\xi(\epsilon) = 0$ if $|\epsilon| \geq \rho$ and $\xi(\epsilon) = 1$ if $|\epsilon| \geq \frac{\rho}{2}$, we define our localising random variable as $\zeta_t = \xi(S_t, Z_t)$. Because the above system is still (L, H)-weak Hörmander, and (L, g, G)-tense for $|(s, S, Z)| \leq \rho$, (with the same G, ρ, H as in step II) we can now apply Lemma 8.1.1 to obtain

$$\mathbb{E}(\delta(R^\infty = z)\zeta_t) \leq C \frac{e^{-\frac{|z|_{\mathcal{F}}^2}{Mt}}}{t^{\nu/2}},$$

valid for all $|z|_{\mathcal{F}}^2, t \leq D$, for some proper constants D, C, M. Then because $F^{RT}(R^\infty) = Y_t$, we can apply Theorem 5.1.1 to obtain

$$\mathbb{E}(\delta(Y_t = y)\zeta_t) \leq C \frac{e^{-\frac{d_{t,\log,L}(x,y)^2}{Mt}}}{|B_{d_{t,\log,L}}(x, \sqrt{t})|},$$

which holds for all $t \leq D$ and for all x, y such that $d_{t,\log,L}(x, y)^2 \leq D$, where D, C, M are all polynomial constants. By Theorem 7.1.1 and the (proper) doubling condition for d_l, changing the constants, but keeping them polynomial, we have that for all $t \leq D$ and for all x, y such that $d_{t,\log,\infty}(x, y)^2 \leq D$,

$$\mathbb{E}(\delta(Y_t = y)\zeta_t) \leq C \frac{e^{-\frac{d_{t,\log,\infty}(x,y)^2}{Mt}}}{|B_{d_{t,\log,l}}(x, \sqrt{t})|}.$$

Now we can proceed as in the proof of Theorem 8.2.1:

Step IV [Patching]

This step is almost exactly the same as in the proof of Theorem 8.2.1 except that the constants are polynomial and the theorems used are the progressive equivalents of the ones used in that proof.

Because the Malliavin derivative of ζ is (properly) bounded regardless of whether $|(s, S, Z)| \leq \rho$, $1 - \zeta = 0$ for $|(s, S, Z)| > \frac{\rho}{2}$, and the original system $\mathcal{A}$ is properly weak Hörmander and tense, the we can apply Theorem 4.4.1 with localising function $1 - \zeta_t$ to obtain:

$$\mathbb{E}(\delta(\Xi^{t,T} = y)(1 - \zeta_t)) \leq C \frac{e^{-\frac{|*x-y|^2}{Mt}}}{t^\mu}$$

for some proper constants μ, M and a polynomial C.

Hence we can write, for $t \leq \exp\left(\frac{d_{t,\log,\infty}^2(x,y)}{2\mu Mt}\right)$ where M, C are polynomial constants that change from line to line:

$$
\begin{aligned}
\mathbb{E}\left(\delta(Y_t = y)\right) &= \mathbb{E}\left(\zeta_t \delta(Y_t = y)\right) + \mathbb{E}\left((1 - \zeta_t)\delta(Y_t = y)\right) \\
&\leq C \frac{e^{-\frac{d_{t,\log,\infty}(x,y)^2}{Mt}}}{|B_{d_{t,\log,\infty}}(x, \sqrt{t})|} + Mt + Ct \\
&\leq C \frac{e^{-\frac{d_{t,\log,\infty}(x,y)^2}{Mt}}}{|B_{d_{t,\log,\infty}}(x, \sqrt{t})|} + (M + C)e^{\frac{d_{t,\log,\infty}^2(x,y)}{2\mu Mt}}
\end{aligned}
$$

$$\leq C\frac{e^{-\frac{d_{t,\log,\infty}(x,y)^2}{Mt}}}{|B_{d_{t,\log,\infty}}(x,\sqrt{t})|}$$

Next, for $t \geq \exp\left(\frac{d^2_{t,\log,\infty}(x,y)}{2\mu Mt}\right)$, we have by Theorem 8.1.3, for some proper constants ν, M, C changing from line to line:

$$\begin{aligned}
p_t(x,y) &\leq C\frac{\exp\left(-\frac{d_{t,\log,\infty}(x,y)^2}{Mt}\right)}{t^\nu} \\
&\leq \exp\left(-\frac{d_{t,\log,\infty}(x,y)^2}{2Mt}\right)\frac{e^{-M_1t/2}}{t^\nu} \\
&\leq C\frac{\exp\left(-\frac{d_{t,\log,\infty}(x,y)^2}{Mt}\right)}{|B_{d_{t,\log,\infty}}(x,\sqrt{t})|}.
\end{aligned}$$

Finally, Proposition 7.0.14 and the local equivalence between d_t^l and $d_{t,\log,\infty}$, there exists a proper constant K such that $d_{t,\log,\infty}(x,y) \leq K(|{*x} - y| + 1)$. Noting also that for $d_t(x,y)^2 \geq D$ there is a proper constant D' such that $|{*x} - y| \geq D'$ (by Proposition 7.1.9), it follows that $d_t(x,y) \leq K(|{*x} - y|)$ for some (other, proper) K whenever $d_t(x,y)^2 \geq D$.

Putting together those three estimates, we obtain the required result.

□

8.4. Polynomial control bounds for detailed-progressive Hörmander systems

One can also remove the UFPG condition on the background vector fields:

THEOREM 8.4.1. *Let $\mathcal{A} = (x, \sigma, F)$, be a uniformly $(r, 2L+1, g, \mathcal{G})$-mixed tense*[5]*, uniformly (L, H_L)-progressive Hörmander system. Suppose that $g \geq 2L + 3 + nL$. There exist polynomial constants C, M, D such that for any $t \leq D$,*

$$p_t(x,y) \leq C\frac{e^{-\frac{d_t(x,y)^2}{Mt}}}{|B_{d_t}(x,\sqrt{t})|}.$$

PROOF. The proof is almost exactly the same as that of Theorem 8.3.1: we only use the first part of Theorem 7.1.1, and obtain the most local part of the estimate immediately. The mid-local part of the estimate is known from 8.1.2 (contrary to the proof of 8.3.1, we use this estimate as it is without changing the distance inside the exponential). And as usual, the '$d_t >> 1$' part of the estimate follows from Euclidean bounds 4.4.2. □

8.5. Lower bounds for situations with zero drift

The result in this section is a generalisation of the lower bound in [**37**] to systems instead of SDE. Unlike the result in [**37**], our result is only local.

We prove the result directly from the density of the log-signature, rather than attempting to use bounds on the density of X_t in its intrinsic space.

THEOREM 8.5.1. *Let $\mathcal{A} = (x, \sigma, F)$ be a uniformly (g, G)-tense*[6]*, uniformly (L, H_L) Hörmander system with $\sigma_z^0 = 0$ for any $z \in \mathbb{R}^m$. Suppose $g \geq (2n+2)^2 2^{4L} + n + 3$, and let $d\colon \mathbb{R}^m \otimes \mathbb{R}^n \to \mathbb{R}^+$ be our 'distance' d_t, which in this case obeys*

$$d(x,y) = \inf_{\tilde{x}\in F^{-1}(\{y\})} d_c(x,\tilde{x})$$

where d_c is the Carnot-Carathéodory distance in $\mathbb{R}^m$ (in particular, $d_c(x,\tilde{x}) = 0$ if $x, \tilde{x}$ are not on the same leaf of the foliation induced by σ).

[5] In particular, F is assumed to be linear

[6] Because we don't aim at controlling the constants here, saying that σ is C^∞ would suffice.

Y_t admits a density $p_t(x,y)$ in $\mathbb{R}^n$ that satisfies, for some constants C, M, D depending only on g, G, L, H_L, n, m, d, for any $t \leq D$ and for any $x, y \in \mathbb{R}^m$ such that $d(x,y)^2 \leq Dt$:

$$p_t(x,y) \geq \frac{C}{|B_d(x,\sqrt{t})|}.$$

REMARK 8.5.1.

- Note that as usual, for $x \in \mathbb{R}^m$, $B_d(x,\sqrt{t})$ is a subset of $\mathbb{R}^n$.
- We have written d instead of d_t, because the assumption on the drift implies that d_t does not depend on t.
- The driftlessness assumption also means that the above does not apply to time-dependent SDE

PROOF. The proof is very similar to the proof of the analogous result in [**37**], (cf. p 426 of said reference).

Fix $l = (2n+2)^2 2^{4L}$. Let $\xi_t \in \mathcal{L}^l(\mathbb{R}^d)$ be the log-signature of W at time t. By the support theorem, and concatenation/scaling, there exist constants C, M such that the density satisfies for $s \in \mathcal{L}^l(\mathbb{R}^d)$

$$p_t(s) \geq C \frac{e^{-\frac{|s|^2_{\mathcal{L}^l(\mathbb{R}^d)}}{Mt}}}{t^{\frac{\nu}{2}}}.$$

(See (up to a constant change in the metric used) Theorem 3.12 in [**37**]) where ν is the homogeneous dimension of $\mathcal{L}^l(\mathbb{R}^d)$.

It follows that we can use Theorem 5.1.2 to obtain the following result for the density of the Taylor approximation $\tilde{Y}_t$:

$$\mathbb{E}(\delta(\bar{Y}_t = y)) \geq C \frac{e^{-\frac{d^l(x,y)^2}{Mt}}}{|B_{d^l}(x,\sqrt{t})|}.$$

Then for $d(x,y)^2, t \leq D$ for some constant D, we have (using Theorem 7.0.1):

$$\mathbb{E}(\delta(\bar{Y}_t = y)) \geq C \frac{e^{-\frac{d(x,y)^2}{Mt}}}{|B_d(x,\sqrt{t})|}.$$

Then using Theorem 6.4.1, we obtain for $t \leq D$ and $d(x,y)^2 \leq M_1 t$ for some fixed $M_1 \leq D$, for C, M being constants changing from line to line:

$$\begin{aligned}
\mathbb{E}(\delta(Y_t = y)) &\geq \mathbb{E}(\delta(\bar{Y}_t = y)) + \big(\mathbb{E}(\delta(Y_t = y)) - \mathbb{E}(\delta(\bar{Y}_t = y))\big) \\
&\geq C \frac{e^{-\frac{d(x,y)^2}{Mt}}}{|B_d(x,\sqrt{t})|} - Mt \\
&\geq C \frac{e^{-\frac{d(x,y)^2}{Mt}}}{|B_d(x,\sqrt{t})|} \\
&\geq \frac{C}{|B_d(x,\sqrt{t})|},
\end{aligned}$$

as expected. □

8.6. Polynomial diagonal estimate in the general case

The main purpose of this section is to show *potential* for proving a proper version of Theorem 8.2.1 with a suitably altered distance or assumptions, most likely with the $d_{t,\infty}$ distance[7] with polynomial constants outside the exponential. The estimate we actually prove is in itself, of course, unsatisfactory.

[7] Clearly $d_{t,\infty}$ and d_t are **not** properly equivalent, except under much stricter assumptions For instance, proper sparsity of the projection of the σ tensor on $\operatorname{span}_{\#(\alpha)>1} e^{\alpha}$.

PROPOSITION 8.6.1 (Polynomial diagonal estimate). *Let $\mathcal{A} = (x, \sigma, F)$ be a uniformly (L, g, G)-tense, uniformly (L, H_L)-weak Hörmander system. Fix $N \in \mathbb{N}$, and suppose that $g \geq (2n+2)^2 2^{4L} + n + 3$. There exist polynomial constants D, C such that for any $t \leq D$ and any $x \in \mathbb{R}^m$, and $i \in \{0, 1, \ldots, d\}$,*

$$p_t(x, *x) \leq \frac{C}{|B_{d_{t,\log,L}}(x, \sqrt{t})|}.$$

PROOF. As usual, without loss of generality, F is linear. Fix $l = (2n+2)^2 2^{4L}$. Because $d_{t,\log,L}$ and $d_{t,\log,l}$ are locally properly equivalent (cf. Proposition 5.1.12 or 5.2.6), we can work with the latter.

We begin with the following auxiliary (Stratonovich) system $\mathcal{R}_1$, where the target random variable is $\tilde{Y}$:

$$\begin{aligned}
X_{t\to 0} &\in \mathbb{R}^{m\times m}, \qquad X_{0\to 0} = \mathrm{Id},\\
X_t &\in \mathbb{R}^m, \qquad X_0 = x,\\
\tilde{X}_t &\in \mathbb{R}^m, \qquad \tilde{X}_0 = 0,\\
dX_t &= \sum_{i=0}^{d} \sigma^i(X_t) \circ dW_t^i,\\
d\tilde{X}_t &= \sum_{i=0}^{d} X_{t\to 0}\sigma^i(X_t) \circ dW_t^i,\\
\tilde{Y}_t &= F(\tilde{X}_t).
\end{aligned}$$

Let $\mathcal{U}$ be the orthogonal complement in $\mathcal{S}$ of the kernel of the F^{SR_r} map. We will approximate $\tilde{Y}_t$ by the random variable $\bar{Y}_t$, solution of the following auxiliary system $\mathcal{R}_2$:

$$\begin{aligned}
S_t &\in \mathcal{T}^l(\mathbb{R}^d, \mathbb{R}), \qquad S_0 = 0,\\
d(s, \tilde{S}_s) &= \sum_{i=0}^{d} w_s^i((s, \tilde{S}_s)) \circ d\tilde{W}^i,\\
\bar{Y}_t &= F^{S,T}(\mathrm{Pr}_{\mathcal{U}}(S_t)) \in \mathbb{R}^n.
\end{aligned}$$

We use the localising function $\phi(\mathcal{G})$ where ϕ is such that $\phi(z) = 0$ when $|z| \geq 2\epsilon$ and $\phi(z) = 1$ when $|z| \leq \epsilon$, and $\mathcal{G} = \sup_{0\leq s\leq t} |\mathrm{Id} - X_{0\to s}|, |\mathrm{Id} - X_{s\to 0}|, |S_t|$, for ϵ smaller than the (proper) constants in Theorems 7.1.1 etc.

By a similar scaling argument to the proof of Theorem 8.2.1, and use of the disintegration formula 5.1.1, we have for any y that

$$\mathbb{E}(\delta(\bar{Y}_t = y)\phi(\mathcal{G})) \leq \frac{C}{|B_{d_{t,\log,l}}(x, \sqrt{t})|},$$

for some polynomial constant C.

Next, recall that for any three times differentiable vector field V on $\mathbb{R}^m$, we have (using Stratonovich integrals)

$$X_{t\to 0}V(X_t) = V(x_0) + \int_0^t \sum_{k=0}^{d} X_{s\to 0}[\sigma^k, V](X_s) \circ dW_s^k.$$

This shows that the $\bar{Y}_t$ is the stochastic Taylor approximation of $\tilde{Y}_t$. Indeed, writing as usual $\mathcal{LL}(E)$ for the set of words $a_1a_2\ldots a_k$ over the set E which are the smallest in lexicographic order out of all the permutations of $a_1a_2\ldots a_k$ obtained by switching elements inside brackets in the expression $[a_1,[a_2,[\ldots a_k\ldots]]]$, we have,

$$\begin{aligned}\tilde{Y}_t &= \sum_{i=0}^{d}\int_0^t *X_{s\to 0}\sigma^i(X_s)dW_s^i\\ &= \sum_{\#(\alpha)\leq l} *\sigma^{[\alpha]}W^\alpha + \sum_{\#(\alpha)=l+1}\int^\alpha *X_{s\to 0}\sigma^{[\alpha]}(X_\bullet)dW^\alpha\\ &= \sum_{\substack{\#(\alpha\leq l,\\ \alpha\in\mathcal{LL}(\{0,1,\ldots,d\})}} *\sigma^{[\alpha]}W^{[\alpha]} + \sum_{\substack{\#(\alpha)=l+1\\ \alpha\in\mathcal{LL}(\{0,1,\ldots,d\})}}\int^\alpha *X_{s\to 0}\sigma^{[\alpha]}(X_\bullet)dW^{[\alpha]}.\end{aligned}$$

The localised rest $\sum_{\#(\alpha)=l+1}\int^\alpha \phi(\mathcal{G}) * X_{s\to 0}\sigma^{[\alpha]}(X_\bullet)dW^\alpha$ is amenable to application of Theorem 2.1.1. Note first that we have for any unit $v\in\mathbb{R}^m$,

$$\sum_{\#(\alpha)=l+1}\int_0^t \langle\phi(\mathcal{G})X_{s\to 0}\sigma^{[\alpha]}(X_s),v\rangle^2 \leq D,$$

for some polynomial D, then Theorem 2.1.1 ensures that for any $R>0$,

$$\mathbb{P}\left(\Big|\sum_{\#(\alpha)=l+1}\int^\alpha \phi(\mathcal{G}) * X_{s\to 0}\sigma^{[\alpha]}(X_\bullet)dW^\alpha\Big|^2 \geq R\right) \leq C_l\exp\left(-\frac{(\frac{R}{C_l})^{\frac{2}{l}}}{8t}\right),$$

with C_l a polynomial constant.

It follows from Theorem 6.4.1 that the following holds with C a polynomial constant: for any y,

$$\mathbb{E}(\delta(\tilde{Y}_t=y)\phi(\mathcal{G})) \leq \frac{C}{|B_{d_{t,\log,l}}(x,\sqrt{t})|},$$

in particular,

$$\mathbb{E}(\delta(\tilde{Y}_t=*x)\phi(\mathcal{G})) \leq \frac{C}{|B_{d_{t,\log,l}}(x,\sqrt{t})|}.$$

Now, let $\gamma:\mathbb{R}^n\to\mathbb{R}^+$ be the following approximation of the delta function:

$$\gamma_\varepsilon(y) = \frac{e^{-\frac{|*x-y|^2}{2\varepsilon}}}{(2\pi\varepsilon)^{n/2}},$$

we have for any $\varepsilon\leq\epsilon$,

$$\begin{aligned}\mathbb{E}(\gamma_\varepsilon(\tilde{Y}_t)\phi(\mathcal{G})) &= \mathbb{E}\left(\frac{e^{-\frac{|*x-\tilde{Y}_t|^2}{2\varepsilon}}}{(2\pi\varepsilon)^{n/2}}\phi(\mathcal{G})\right)\\ &\leq \mathbb{E}\left(\frac{e^{-\frac{|*x-Y_t|^2}{2(1+\epsilon)\varepsilon}}}{(2\pi\varepsilon)^{n/2}}\phi(\mathcal{G})\right)\\ &\leq (1+\epsilon)^{\frac{n}{2}}\mathbb{E}\left(\frac{e^{-\frac{|*x-Y_t|^2}{2(1+\epsilon)\varepsilon}}}{(2\pi\varepsilon(1+\epsilon))^{n/2}}\phi(\mathcal{G})\right) = (1+\epsilon)^{\frac{n}{2}}\mathbb{E}\left(\gamma_{\varepsilon(1+\epsilon)}(\tilde{Y}_t)\right).\end{aligned}$$

Passing to the limit $\varepsilon\to 0$, we obtain:

$$\mathbb{E}(\delta(\tilde{Y}_t=y)\phi(\mathcal{G})) \leq 2^{n/2}\mathbb{E}(\delta(Y_t=y)\phi(\mathcal{G})).$$

Similarly, we obtain:

$$\mathbb{E}(\delta(\tilde{Y}_t=y)\phi(\mathcal{G})) \geq 2^{-n/2}\mathbb{E}(\delta(Y_t=y)\phi(\mathcal{G})).$$

The result follows. □

Part 3

Examples of applications; Lower bounds and the ‘separated progressive Hörmander condition’

CHAPTER 9

Some examples

9.0.1. L=2. The following theorem is a slight variation of the main result proved in [**8**] and in Chapter 4 of [**50**] to (time-dependent) SDE that satisfy the relevant assumptions globally.

THEOREM 9.0.1. *Let $\mathcal{A} = (x, \sigma, F)$ be a uniformly $(2, 4+n, G)$-tense, uniformly $(2, H_2)$-weak Hörmander system such that the background vector fields σ are uniformly finitely generated (with constant H_2). There exist proper constants C, M and polynomial constants $\bar{C}$ and $\bar{M}$ such that for any $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$,*

$$p_t(x,y) \leq C \frac{\exp\left(-\frac{d_{t,\log,\infty}(x,y)^2}{Mt}\right)}{|B_{d_{t,\log,\infty}}(x,\sqrt{t})|} \quad \textit{and} \quad p_t(x,y) \leq \bar{C} \frac{\exp\left(-\frac{d(x,y)^2}{\bar{M}t}\right)}{|B_d(x,\sqrt{t})|},$$

where d is the Carnot-Carathéodory distance.

PROOF. Clearly, any (L, H_L)-weak Hörmander system[1] is also (L, H_L) detailed-Progressive weak Hörmander, therefore, we can use Theorem 8.3.1 to obtain the first estimate. The only extra step to get the second estimate is to use Proposition 7.2.4. Note also that for a time-dependent SDE we can express the Hörmander of order two condition in terms of brackets of the time-frozen coefficients, since derivatives of σ^i's with respect to time are already of order three. □

9.0.2. The Pigato SDE. Consider the following SDE, considered by Pigato in [**50**] (Chapter 3) (here $m = n = 2$, $d = 1$):

$$\begin{aligned} X_t &\in \mathbb{R}^2, \quad X_0 = 0, \\ dX_t &= \sigma^1(X_s) \circ dW_t + \sigma^0(X_s)dt, \end{aligned}$$

with the assumptions being weak Hörmander of order 3, uniform boundedness of coefficients and the following *geometric condition on the variance*: there exists a uniformly bounded $\kappa : \mathbb{R}^2 \to \mathbb{R}$ such that for any unit $v \in \mathbb{R}^2$,

$$\frac{\partial \sigma(x)}{\partial v} = \kappa \sigma(x).$$

This is quite a strong assumption: It makes the system mono-scaled, i.e. setting $E_i = \operatorname{span}_{|\alpha| \leq i} \sigma^\alpha$ ensures that for all i, $\inf_{\substack{|v|=1 \\ v \in \mathbb{E}_i}} \sum_{|\alpha| \leq i} \langle \sigma^{[\alpha]}, v \rangle^2$ is bounded below by a constant H. In mono-scaled situations, it is enough to use an auxiliary object of the same dimension as the original diffusion, and to use the chain rule instead of the disintegration formula. This is equivalent to not using any auxiliary objects and using a homogeneous norm on the target space $\mathbb{R}^n$ after picking an appropriate basis and turning it into a graded space.

This system is progressive Hörmander, in particular, Theorem 8.3.1 applies. The 'distance' $d_{t,\log,\infty}$ has similar balls of radius $\sqrt{t}$ to those of the metric introduced in [**50**].

REMARK 9.0.1. As mentioned in the introduction, the estimate in [**50**] is sharper for $t \lesssim d^2 \lesssim t\ln(t)$, because the 'distance' we use is not as well tailored to the problem. Nevertheless, our estimate is still space-time integrable.

[1] This includes any (L, H_L)-weak Hörmander time-dependent SDE.

9.0.3. Higher-dimensional Pigato type SDE's. Consider the following SDE:

$$X_t \in \mathbb{R}^n, \quad X_0 = 0,$$
$$dX_t = \sigma^1(X_s) \circ dW_t + \sigma^0(X_s)dt,$$

assuming the uniform weak Hörmander condition of order three, that the vector fields are C^∞, and the following generalisation of the geometric condition on the variance:

There exist smooth functions $\lambda_1, \lambda_2, \dots, \lambda_d \colon \mathbb{R}^n \to \mathbb{R}$ and a constant Λ such that for all $x \in \mathbb{R}^n$,

$$\sigma^{(0)}(x) = \sum_{i=1}^{d} \lambda_i(x)\sigma^i(x) \quad \text{and} \quad \sum_{i=1}^{d} |\lambda^i|^2 \le \Lambda,$$

then the resulting vector fields are uniformly (detailed-) progressively finitely generated and the system is progressive Hörmander. In particular, if we have fixed suitable bounds on Λ, the weak Hörmander constant and on the tension, and $F : \mathbb{R}^n \to \mathbb{R}^k$ is a linear map with fixed bonds on the operator norms of F and its Moore-Penrose pseudo-inverse, the density of $F(X_t)$ will admit control bounds in terms of the log-homogeneous distance with constants that do not depend on n or d.

9.0.4. Delarue-Menozzi type chain of Differential equations. Consider, as in [**50**] and [**20**], the following system of SDE:

$$X^1, X^2, \dots, X^l \in \mathbb{R}^n,$$
$$dX_t^1 = \sum_{i=1}^{d} \sigma^i(X^1) \circ dW_t^i,$$
$$dX^i = B_i(X^{i-1}, X^i, \dots, X^l)dt$$

where we assume that the vector fields σ are uniformly (L, H_L) progressive Hörmander in the space $\mathbb{R}^n$ (this is a weaker assumption than [**20, 50**] where in both cases, ellipticity was assumed), σ and B_i (for all i) are C^∞ (this is a stronger assumption than is used in [**20**]) and for any $i \ge 2$, the minimum eigenvalue of the matrix $D_{x_{i-1}} B_i$ is greater than λ (for some fixed constant λ.). This system is $(L+2(l-1), H)$ progressive Hörmander for some constant H depending only on λ and the C^∞ constants in our assumptions. In particular, Theorem 8.3.1 applies.

Note that in the elliptic case (which was treated in [**20**] and [**50**]), Remark 9.0.1 applies as well.

9.0.5. Two-sided version of Delarue-Menozzi chains. The following generalisation of the previous section is also progressive Hörmander:

$$X^1, X^2, \dots, X^l \in \mathbb{R}^n,$$
$$dX_t^1 = \sum_{i=1}^{d} \sigma^i(X^1) \circ dW_t^i,$$
$$dX^i = B_i(X^{i-1}, X^i, X^{i+1})dt \quad (\text{for } i = 2, 3, \dots, l-1), \quad \text{and}$$
$$dX_t^l = \sum_{i=1}^{d} \mu^i(X^1) \circ dW_t^i,$$

with the following assumptions[2]:

- The vector fields σ, μ, B_i are C^∞.
- The vector fields σ and μ are uniformly elliptic with ellipticity constant H.

[2]The initial condition is irrelevant as long as we suppose that the conditions hold uniformly.

- There exists a constant λ such that for $i = 2, \ldots, l-1$, any x and any unit $v \in \mathbb{R}^n$, there exist some μ_k and $\bar{\mu}_k$ ($k = 1, \ldots n$) such that

$$v = \sum_{k=1}^{n} \mu_k \frac{\partial B_i}{\partial \langle x^{i-1}, e^k \rangle} + \sum_{k=1}^{n} \bar{\mu}_k \frac{\partial B_i}{\partial \langle x^{i+1}, e^k \rangle}$$

and

$$\sum (\mu_k)^2 + \sum (\bar{\mu}_k)^2 \leq \lambda^{-2}.$$

For elliptic situations, contrary to the previous example, this system is not mono-scaled.

The progressive Hörmander condition implies that if f is a bounded linear map from $\mathbb{R}^{l \times n}$ to $\mathbb{R}^k$ for some k with a bounded pseudo-inverse, and the driving vector fields satisfy sufficiently strong conditions to make the tension fixed, the density of $f(X^1, X^2, \ldots, X^l)$ will satisfy the bound of Theorem 8.3.1 with constants that depend on l, d but not on n. Furthermore, if the restriction $f(X^1, X^2, \ldots, X^{l_1})$ has a bounded pseudo-inverse, then the constants in the estimate will depend on l_1 but not on l.

9.0.6. Generalised Heisenberg Groups. The following vector fields in $\mathbb{R}^7$ are detailed-progressively finitely generated, and uniformly so in any fixed compact set around zero (in particular, Theorem 8.3.1 applies in any compact set around 0):

$$\begin{aligned} v^1 &= dx_1 + x_2 dx_3 + x_2 x_6 dx_7 \\ v^2 &= dx_2 - x_1 dx_3 - x_1 x_6 dx_7 \\ v^3 &= dx_4 + x_5 dx_6 - x_5 x_3 dx_7 \\ v^4 &= dx_5 - x_4 dx_6 + x_4 x_3 dx_7 \end{aligned}$$

Indeed, writing $(1, 2; 3)$ for the tree with 3 as a root and 1 and 2 as two leaves, and similar notation for other indices, we have:

$$\begin{aligned} v^{(12)} &= -dx_3 - x_6 dx_7 \\ v^{(21)} &= dx_3 + x_6 dx_7 \\ v^{(13)} &= -x_2 x_5 dx_7 \\ v^{(31)} &= x_2 x - 5dx - 7 \\ v^{(14)} &= x_2 x_4 dx_7 \\ v^{(41)} &= -x_2 x_4 dx_7 \\ v^{(23)} &= x_1 x_5 dx_7 \\ v^{(32)} &= -x_1 x_5 dx_7 \\ v^{(34)} &= -dx_6 + x_3 dx_7 \\ v^{(43)} &= dx_6 - x_3 dx_7, \end{aligned}$$

and

$$\begin{aligned} v^{(1,2;3)} &= 0 \\ v^{((12)3)} &= x_5 dx_7 \\ v^{(123)} &= x_5 dx_7 \\ v^{(213)} &= -x - 5dx_7 \\ v^{(1,3;2)} &= -x_5 dx_7 \\ v^{(132)} &= -x_5 dx_7 \\ v^{(2,3;1)} &= x_5 dx_7 \end{aligned}$$

$$
\begin{aligned}
v^{(231)} &= x_5 dx_7 \\
v^{(213)} &= -x_5 dx_7 \\
v^{(321)} &= x_5 dx_7 \\
v^{[123]} &= x_5 dx_7 - (-x_5 dx_7) - x_5 dx_7 + x_5 dx_7 = 2x_5 dx_7,
\end{aligned}
$$

and then similarly

$$
\begin{aligned}
v^{[124]} &= -2x_4 dx_7 \\
v^{[134]} &= -2x_2 dx_7 \\
v^{[234]} &= 2x_1 dx_7
\end{aligned}
$$

(and for any rooted tree τ with vertices $1, 2, 4$ (resp. $1, 3, 4$, resp. $2, 3, 4$), v^τ is a small multiple of $x_4 dx_7$ (resp. $x_2 dx_7$, resp. $x_1 dx_7$). Also, clearly

$$v^{[[12],[34]]} = -8dx_7.$$

Another similar, but non monoscaled example is the following:

$$
\begin{aligned}
v^1 &= dx_1 + x_2 dx_3 + (1/3)x_2 x_6 dx_7 \\
v^2 &= dx_2 - x_1 dx_3 - (1/3)x_1 x_6 dx_7 \\
v^3 &= dx_4 + x_5 dx_6 - (1/3)x_5 x_3 dx_7 \\
v^4 &= dx_5 - x_4 dx_6 + (1/3)x_4 x_3 dx_7 \\
v^5 &= dx_8 + x_9 dx_{10} + (1/3)x_9 x_{13} dx_7 \\
v^6 &= dx_9 - x_8 dx_{10} - (1/3)x_8 x_{13} dx_7 \\
v^7 &= dx_{11} + x_{12} dx_{13} - (1/3)x_{12} x_{10} dx_7 \\
v^8 &= dx_{12} - x_{11} dx_{13} + (1/3)x_{11} x_{10} dx_7 \\
v^9 &= dx_{14} \\
v^{10} &= x_{14} dx_{15} \\
v^0 &= (1/3)x_{15} dx_7.
\end{aligned}
$$

Higher dimensional generalisations are possible, and if L is fixed, Theorem 8.3.1 will yield estimates polynomial in the background dimension for the densities of low-dimensional projections of the solutions.

CHAPTER 10

Lower bounds and the 'separated progressive Hörmander condition'

10.1. Motivation

The following well-known example shows that under weak Hörmander, the density of the solution can be null on large sets. Similar examples are mentioned in both [**50**] and [**39**]:

$$m = n = 2, \quad d = 1, \quad F = \mathrm{Id}, \tag{10.1.1}$$
$$\sigma^0 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ x^2 \end{pmatrix}, \quad \sigma^1 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

It is clear that the density of the solution starting at zero is null at points (x, y) with $y < 0$.

Since this example actually does satisfy the progressive Hörmander condition, it is worth wondering what is different between this situation compared to situations where lower bounds exist (such as the ones in [**20, 50**]). It appears the problem has to do with the positivity of polynomials of even degree, and more precisely, we see that this issue is related to the existence of areas of $\mathcal{L}^l(\mathbb{R}^d, \mathbb{R})$ which are not in the support of $*W$, and we guess that the first (in terms of order) potentially problematic direction is the direction $e^{[1,[1,0]]}$.

Indeed, a simple calculation shows the appearance of boundary conditions for the signature of the path of $*\gamma$ for an arbitrary path γ of bounded variation:

PROPOSITION 10.1.1. *Let γ be a smooth path in $\mathbb{R}$ parametrised over $[0, 1]$ with $\gamma_0 = 0$. The signature S (at time 1) of $*\gamma$ satisfies:*

$$S^{[1,[1,0]]} \geq \frac{-(S^{[1]})^2}{4}.$$

PROOF. We have:

$$\begin{aligned} S^{(1,1,0)} &= \int_0^1 \int_0^{s_3} \int_0^{s_2} \dot{\gamma}_{s_1} ds_1 \dot{\gamma}_{s_2} ds_2 ds_3 \\ &= \int_0^1 \frac{\gamma_s^2}{2} ds, \end{aligned}$$

$$\begin{aligned} S^{(1,0,1)} &= \int_0^1 \int_0^{s_3} \int_0^{s_2} \dot{\gamma}_{s_1} ds_1 ds_2 \dot{\gamma}_{s_3} ds_3 \\ &= \int_0^1 (\gamma_s \gamma_1 - \gamma_s^2) ds, \end{aligned}$$

and

$$\begin{aligned} S^{(0,1,1)} &= \int_0^1 \int_0^{s_3} \int_0^{s_2} ds_1 \dot{\gamma}_{s_2} ds_2 \dot{\gamma}_{s_3} ds_3 \\ &= \int_0^1 s \dot{\gamma}_s (\gamma_1 - \gamma_s) ds \\ &= \gamma_1^2 - \gamma_1 \int_0^1 \gamma_s ds - \frac{\gamma_1^2}{2} + \int_0^1 \frac{\gamma_s^2}{2} ds \\ &= \frac{\gamma_1^2}{2} + \int_0^1 \frac{\gamma_s^2}{2} ds - \gamma_1 \int_0^1 \gamma_s ds. \end{aligned}$$

Therefore we obtain:

$$
\begin{aligned}
S^{[[1,[1,0]]} &= \int_0^1 3\gamma_s^2 - 3\int_0^1 \gamma_s\gamma_1 + \frac{\gamma_1^2}{2} \\
&= \int_0^1 3(\gamma_s - \frac{1}{2}\gamma_1)^2 - \frac{\gamma_1^2}{4} \geq -\frac{\gamma_1^2}{4},
\end{aligned}
$$

as expected. □

In the next sections, we show that one can move freely in all directions $e^{[\alpha]}$ with either $o_1(\alpha) \leq 1$ or $o_0(\alpha) = 0$. The spirit of the 'separated Hörmander condition' is to ensure that the effect of 'problematic' directions such as $e^{[1,[1,0]]}$ or $e^{[1,[1,[1,[1,0]]]]}$ is *relatively* negligible because the diffusion can already move freely at a faster speed along these directions thanks to terms of lower scaling taken from 'nice directions' such as $e^{[1,3]}$ or $e^{[0,[0,1]]}$.

Lemma 10.2.1 below, coupled with Proposition 5.2.6, already shows the local equivalence between the Léandre distance and the log-homogeneous distance (and by extension, the distance d_t) for systems satisfying the separated progressive Hörmander condition. It is also one of the main steps to prove our probabilistic result.

10.2. On the support of the log signature of $*W$

The first aim of this section is to provide a linear subspace of $\mathcal{L}^l(\mathbb{R}^{d+1})$ such that the projection of $\operatorname{logsig}(*W)$ on it has full support. This will be the first step to proving lower bounds under the separated progressive Hörmander condition. We do not worry about polynomial or proper dependence here.

We will use the following notation:

$$
\begin{aligned}
A_l &= \operatorname{span}_{\substack{o_0(\alpha)=0 \\ \#(\alpha)\leq l}} e^{[\alpha]} \\
B_l &= \operatorname{span}_{\substack{o_0(\alpha)\geq 1 \\ o_1(\alpha)=1, \#(\alpha)\leq l}} e^{[\alpha]} \\
C_l &= \operatorname{span}_{\substack{o_0(\alpha)\geq 1 \\ o_1(\sigma)\geq 2, \#(\alpha)\leq l}} e^{[\alpha]} \\
E_l &= A_l \oplus B_l \oplus C_l \\
\bar{B} &= B \oplus \operatorname{span}_{i=1,\dots,d} e^i \\
\bar{A} &= A / \operatorname{span}_{i=1,\dots,d} e^i.
\end{aligned}
$$

We sometimes omit the subscript l. We have that

$$
\begin{aligned}
\mathcal{L}^l(\mathbb{R}^{d+1}) &= \operatorname{span}(e^{[0]}) \oplus A \oplus B \oplus C \\
&= \operatorname{span}(e^{[0]}) \oplus E.
\end{aligned}
$$

LEMMA 10.2.1. *Let $b \in B$ (resp. $b \in \bar{B}$), there exists a piecewise linear curve $\gamma \in \mathcal{P}_1^{d+1}$ and an element $F(b) \in E$ such that*

$$
\begin{aligned}
F(b) &= \operatorname{logsig}(*\gamma) \\
\operatorname{Pr}_{A_\infty}(F(b)) &= 0 \quad (\textit{resp.}\ \operatorname{Pr}_{\bar{A}_\infty}(F(b)) = 0) \\
\operatorname{Pr}_{B_l}(F(b)) &= b \quad (\textit{resp.}\ \operatorname{Pr}_{\bar{B}_l}(F(b)) = b) \quad \textit{and} \\
\langle e^{[0]}, F(b)\rangle &= 1.
\end{aligned}
$$

Furthermore, there exists a constant C such that for any $b \in B$,

$$
\begin{aligned}
|*\gamma|_{L^2} &\leq C|b| = C|b|_{\mathcal{L}^l(\mathbb{R}^{d+1}),1} \\
|F(b)|_{\mathcal{L}^l(\mathbb{R}^{d+1}),1} &\leq C|b| = C|b|_{\mathcal{L}^l(\mathbb{R}^{d+1}),1}
\end{aligned}
$$

where $|\cdot|_{\mathcal{L}^l(\mathbb{R}^{d+1}),t}$ denotes the time scaled homogeneous norm on $\mathcal{L}^l(\mathbb{R}^{d+1})$.

PROOF. By compactness and scaling, it is enough to find γ and $F(b)$ satisfying the first set of conditions, as long as F is continuous, which will clearly be the case from our construction. We will show the statement for $\bar{B}$ and $\bar{A}$, as the one for B and A follows immediately. We will also write an element $b \in \bar{B}$ as $x = (x_1, x_2, \ldots, x_l)$ with $x_i \in \mathbb{R}^d$ so that x is the expression of b in the basis given by the $w^{i,j} = e^{[0,[0,[0,\ldots,j]]]}$ where there are $i-1$ zeros's and one j in the bracketed expression.

We have

$$\bar{B} = \mathrm{span}_{1 \leq i \leq l, j=1,2,\ldots d}\, w^{i,j}$$

In the first constructions below, for $u \in \mathbb{R}^d$, the signature $(\lambda e^{[0]} + u)$ is always achieved with the straight path $*\gamma$ parametrised over $[0, \lambda]$ such that $*\gamma_\lambda = u$.

We now prove the result by induction:

Initial case. For $l = 1$, one can simply take the linear path corresponding to the following expression:

$$(e^{[0]} + x) = (e^{[0]} + x_1).$$

Induction case. Suppose that $x = (x_1, \ldots, x_{l+1}) \in \bar{B}_{l+1}$ is given. For any $z \in \bar{B}_{l+1}$, write $u = (0, z_2, z_3, \ldots, z_{l+1})$ and consider the following element of $\mathcal{L}^{l+1}(\mathbb{R}^d)$:

$$\delta = \delta_1 \otimes \delta_2 = (\frac{e^{[0]}}{2} + z_1) \otimes (\frac{e^{[0]}}{2} + u).$$

All the terms of the form $[\delta_1, [\delta_2, \ldots [\delta_1]]]$ in the Baker-Campbell-Hausdorff (BCH) formula expansion for $\delta_1 \otimes \delta_2$ can further be expanded so that we obtain an expression involving multiple brackets of $\frac{e^{[0]}}{2}$, z_1, and u only. Now, consider $\mathrm{Pr}_{A \oplus B}(\delta)$. Here all the terms in the above expansion involving both z_1, and u are in C. Therefore, we can write

$$\mathrm{Pr}_{\bar{A}_\infty}(\delta) = 0 \quad \text{and} \quad \forall i \leq l+1, \tag{10.2.1}$$

$$\mathrm{Pr}_{\mathrm{span}_{1 \leq j \leq d} w^{i,j}}(\delta) = z_i + \frac{1}{4} z_{i-1} + \sum_{k=2}^{i-1} \mu_{k,i} z_{i-k} = \sum_{k=0}^{i-1} \mu_{k,i} z_{i-k}$$

with $\mu_{0,i} = 1$ and $\mu_{1,i} = \frac{1}{4}$.

Next, for our fixed $x = (x_1, \ldots, x_{l+1}) \in \bar{B}_{l+1}$ we can turn equation (10.2.1) into the following system of $l+1$ ($\mathbb{R}^d$-dimensional) equations with $l+1$ ($\mathbb{R}^d$-dimensional) unknowns:

$$\forall 1 \leq i \leq l+1, \quad x_i = \sum_{k=0}^{i-1} \mu_{k,i} z_{i-k},$$

which has the recursively defined solution

$$z_1 = x_1 \quad \text{and} \quad \forall 2 \leq i \leq l+1,$$
$$z_i = x_i - \sum_{k=1}^{i-1} \mu_{k,i} z_{i-k}.$$

We fix z (and therefore u) given by the solution thus obtained. Now for any $\bar{y} \in \bar{B}_l$, by the induction hypothesis, there is a piecewise linear path γ_y such that $\mathrm{Pr}_{\bar{B}_l \oplus \bar{A}_\infty}(\mathrm{logsig}(*\gamma_y)) = \bar{y}$. Write y for $\mathrm{Pr}_{\bar{A}_\infty \oplus B_{l+1}}(\mathrm{logsig}(*\gamma_{\bar{y}}))$. Because

$$\mathrm{Pr}_{\bar{A}_\infty}(\mathrm{logsig}(*\gamma_y)) = 0,$$

we have that

$$\mathrm{Pr}_{\bar{B}_{l+1} \oplus \bar{A}_\infty}(-\gamma_y) = -y.$$

The following element of $\mathcal{L}^{l+1}(\mathbb{R}^d)$ can be obtained as the concatenation of rescaled versions of γ_y and $-\gamma_y$:

$$\bar{\delta} = \bar{\delta}_1 \otimes \bar{\delta}_2 = (\frac{e^{[0]}}{4} - y) \otimes (\frac{e^{[0]}}{4} + y).$$

Here we have written $\bar{\delta}_1 = (\frac{e^{[0]}}{4} - y)$ and $\bar{\delta}_2 = (\frac{e^{[0]}}{4} + y)$ etc for notational simplicity. We can apply the BCH formula to expand the above expression.

All the terms of the form $[\bar{\delta}_1, [\bar{\delta}_2, \dots [\bar{\delta}_1]]]$in the BCH formula expansion for $\bar{\delta}_1 \otimes \bar{\delta}_2$ can further be expanded so that we obtain an expression involving multiple brackets of $\frac{e^{[0]}}{4}$ and y only.

All the terms of the form $[y, [-y, \dots [y, -y]]]$ (involving only y and $-y$) cancel each other from the symmetry of the BCH formula[1]; then any term left involving more than one y is an element of C (its projections on A and B are zero). Indeed, any non zero such term must involve at least one of $\frac{e^{[0]}}{2}$, on top of at least two of y. This means any such term is in the span of the $e^{[\alpha]}$ where $o_1(\alpha) \geq 2$ and $o_0(\alpha) \geq 1$.

So the only terms left are the terms involving exactly one y. All those terms involve at least one $\frac{e^{[0]}}{2}$, so we have that $\mathrm{Pr}_{\bar{A}_\infty}(\delta) = 0$.

As a conclusion and using the explicit form of the BCH formula for order up to 2, we obtain (with the convention that $y_i = 0$ for $i \leq 0$):

$$\mathrm{Pr}_{\bar{A}_\infty}(\bar{\delta}) = 0 \quad \text{and} \quad \forall i \leq l+1,$$

$$\mathrm{Pr}_{\mathrm{span}_{1\leq j\leq d}\, w^{i,j}}(\bar{\delta}) = \frac{1}{4} y_{i-1} + \sum_{k=2}^{i-1} \lambda_{k,i} y_{i-k} = \sum_{k=0}^{i-1} \lambda_{k,i} y_{i-k}$$

for some fixed $\lambda_k \in \mathbb{R}$ coming from the algebraic BCH expansion only (a priori $\lambda_k = 0$ is possible for $k \geq 2$) and in particular, not depending on y. Here $\lambda_{0,i} = 0$ and $\lambda_{1,i} = \frac{1}{4}$ For our any given $u_2, \dots, u_{l+1}$(obtained above), we can now turn Eq. 10.2 into the following linear system of equations with l $\mathbb{R}^d$-dimensional unknowns $y_1, y_2, \dots, y_l$:

$$\forall\, 2 \leq i \leq l+1, \quad u_i = \sum_{k=1}^{i-1} \lambda_{k,i} y_{i-k}.$$

This is easily solved as a triangular system with recursive solution:

$$y_1 = 4u_2 \qquad \text{and} \quad \forall\, 2 \leq i \leq l,$$

$$y_i = 4u_{i+1} - \sum_{k=2}^{i} 4\lambda_{k,i+1} y_{i+1-k}.$$

It follows that the quantities y, z, u thus defined satisfy the following

$$\mathrm{Pr}_{A\oplus B}\left(\left(\frac{e^{[0]}}{2} + z_1\right) \otimes \left(\frac{e^{[0]}}{4} - y\right) \otimes \left(\frac{e^{[0]}}{4} + y\right)\right) = b.$$

Now recall that $(\frac{e^{[0]}}{2} + z_1)$ can be canonically represented as a linear path $*\gamma_1$ parametrised over $[0, \frac{1}{2}]$, and by the induction hypothesis plus simple rescaling, $(\frac{e^{[0]}}{4} + y)$ and $(\frac{e^{[0]}}{4} - y)$ can be represented as piecewise linear paths $*\gamma_2$ and $*\gamma_3$ parametrised over $[\frac{1}{2}, \frac{3}{4}]$ and $[\frac{3}{4}, 1]$ respectively.

Then the concatenation $*\gamma_1 \otimes *\gamma_2 \otimes *\gamma_3$ satisfies the required conditions. $\square$

Using this lemma, we can finally prove the following proposition, which shows that the projection onto $A \oplus B$ of the support of the log signature is the full space $A \oplus B$.

PROPOSITION 10.2.2. *Let $u = a + b \in A \oplus B$ with $a \in A$ and $b \in B$, there exists a piecewise linear curve $\gamma \in \mathcal{P}_1^{d+1}$ and an element $F(b) \in E$ such that*

$$F(b) = \mathrm{logsig}(*\gamma),$$

[1]Clearly, for any element z of the Lie algebra, $z \otimes (-z) = 0$

$$\begin{aligned}&\mathrm{Pr}_A(F(b)) = a,\\ &\mathrm{Pr}_B(F(b)) = b \quad \textit{and}\\ &\langle e^{[0]}, F(b)\rangle = 1.\end{aligned}$$

Furthermore, there exists a constant C such that for any u,

$$\begin{aligned}|*\gamma|_{L^2} &\le C|u|_{\mathcal{L}^l(\mathbb{R}^{d+1}),1}\\ |F(b)|_{\mathcal{L}^l(\mathbb{R}^{d+1}),1} &\le C|u|_{\mathcal{L}^l(\mathbb{R}^{d+1}),1}\end{aligned}$$

where $|\cdot|_{\mathcal{L}^l(\mathbb{R}^{d+1}),t}$ denotes the time-scaled homogeneous norm on $\mathcal{L}^l(\mathbb{R}^{d+1})$.

PROOF. As in the proof of Lemma 10.2.1, by compactness and scaling, it is enough to find γ and $F(b)$ satisfying the first set of conditions, as long as F is continuous, which will clearly be the case from our construction.

Now, using Lemma 7.0.11, and rescaling, we can pick a path $\gamma_a = \gamma_0$, parametrised over $[0, \frac{1}{l+2}]$ with $\mathrm{Pr}_A(\mathrm{logsig}(*\gamma)) = \mathrm{logsig}(\gamma) = a$. Let $\eta_0 = \mathrm{Pr}_B(\mathrm{logsig}(*\gamma_a))$. For $i = 1, \dots, l$, we define η_i by

$$(\frac{(i+1)e^{[0]}}{l+2} + \eta_i) = (\frac{ie^{[0]}}{l+2} + \eta_{i-1}) \otimes (\frac{e^{[0]}}{l+2} - \mathrm{Pr}_B(\eta_{i-1})).$$

Note that we have $\mathrm{Pr}_{\mathrm{span}_{i\le k, j\in\{1,\dots,d\}} w^{i,j}}(\eta_k) = 0$, and in particular, $\mathrm{Pr}_B(\eta_l) = 0$. We also have $\mathrm{Pr}_A(\eta_i) = a$ for all i.

Now by a simple extension of Lemma 10.2.1, for any $u \in B$, we can find a path γ_u parametrised over $[\frac{i}{l+2}, \frac{i+1}{l+2}]$ with

$$\mathrm{Pr}_{A\oplus B}(\mathrm{logsig}(*\gamma_u)) = \frac{e^{[0]}}{l+2} + u.$$

We apply this to the $(\frac{e^{[0]}}{l+2} - \mathrm{Pr}_B(\eta_{i-1}))$ and find γ_i with $\mathrm{Pr}_{A\oplus B}(\gamma_i) = (\frac{e^{[0]}}{l+2} - \mathrm{Pr}_B(\eta_{i-1}))$. We now have that

$$\mathrm{Pr}_{A\oplus B}\left(\gamma_0 \otimes \gamma_1 \otimes \dots \otimes \gamma_l\right) = a.$$

Finally, we pick a path γ_{l+1}, parametrised over $[\frac{l+1}{l+2}, 1]$ such that

$$\mathrm{Pr}_{A\oplus B}(\mathrm{logsig}(\gamma_{l+1})) = \delta = (y_1, y_2, \dots, y_l) \in B,$$

with $\delta = (y_1, y_2, \dots, y_l) \in B$ solving the system

$$b_i = y_i + \sum_{k=1}^{i-1} \lambda_{k,i} y_{i-1} \quad \text{for } i \ne 2 \quad \text{and}$$

$$b_2 + \frac{\mathrm{Pr}_{\mathrm{span}_{k=1,\dots,d} e^k}(a)}{l+2} = y_2 + \frac{y_1(l+1)}{2(l+2)}$$

where the $\lambda_{k,i}$ are obtained through BCH expansion of the formula

$$\mathrm{Pr}_{A\oplus B}\left(\Big(\frac{(l+1)e^{[0]}}{l+2} + a\Big) \otimes \Big(\frac{e^{[0]}}{l+2} + y\Big)\right) = a + b.$$

(Note that this can be done because cross-brackets involving both elements of A and elements of B cancel, being elements of C.)

We now have

$$\mathrm{Pr}_{A\oplus B}\left(\gamma_0 \otimes \gamma_1 \otimes \dots \otimes \gamma_l \otimes \gamma_{l+1}\right) = a + b,$$

as required. □

While the above is enough for purely deterministic purposes (such as proving equivalence of pseudo-distances), more work needs to be done to produce elements of $\mathcal{S}$ with strictly positive density.

PROPOSITION 10.2.3. *For any* $0 \leq t \leq 1$, *there exists a function* $\phi_t : A \oplus B \to A \oplus B \oplus C$ *such that*

(1) *for all* $u \in A \oplus B$, $\mathrm{Pr}_{A\oplus B}(\phi_t(u)) = u$,
(2) ϕ_t *is differentiable away from the origin,*
(3) ϕ_t *is scale invariant, meaning that for any* $0 \leq t \leq 1$, $\phi_t(u) = \phi_1(\delta_{\frac{1}{t}}(u))$, *where as usual* δ_s *denotes the homogeneous dilations on* $\mathcal{L}^l(\mathbb{R}^{d+1})$ *or any of its subspaces.*
(4) *For any* $M > 0$, *there exists a constant* C_M *such that for any* $u \in A\oplus B$ *such that* $|u|^2_{\mathcal{S}} \leq Mt$, *we have*
$$p_t(\phi_t(u)) \geq \frac{C_M}{t^{\frac{\nu}{2}}},$$
where $p_t : \mathcal{S} \to \mathbb{R}^+$ *is the density function of the random variable* $\mathrm{Pr}_{\mathcal{S}}(\mathrm{logsig}(*W)_t)$, $\mathcal{S} = A \oplus B \oplus C$, *and* ν *and* $|\cdot|_{\mathcal{S}}$ *are the homogeneous dimension and norm on* $\mathcal{S}$ *respectively.*

PROOF. By scaling, we can suppose without loss of generality that $t = 1$, as long as we define ϕ_t from ϕ_1 using the formula $\phi_t(u) = \phi_1(\delta_{\frac{1}{t}}(u))$.

Note that the random variable $\mathrm{Pr}_{\mathcal{S}}(\mathrm{logsig}(*W)_t)$ has a smooth density in $\mathcal{S}$ by (a localised version of) Hörmander's theorem for SDE with time dependent coefficients, or similar results from the first part of this thesis.

Step 1 [Definition of a bi-scaling invariant $\bar{\phi}$]. First, we prove that there exists a $\bar{\phi} : A\oplus B \to A \oplus B \oplus C$ such that for all $u \in A \oplus B$,
$$\mathrm{Pr}_{A\oplus B}(\bar{\phi}_t(u)) = u$$
and
$$p_t(\bar{\phi}_t(u)) > 0.$$

To see this, observe first that for any $u \in A \oplus B$, by Proposition 10.2.2, $F(u)$ belongs to the support of the random variable $\mathrm{Pr}_{\mathcal{S}}(\mathrm{logsig}(*W)_1)$. Because by definition, the support is the closure of its interior, this implies that the projection of the interior of the support of $\mathrm{Pr}_{\mathcal{S}}(\mathrm{logsig}(*W)_t)$ onto $A \oplus B$ is a dense open set.

Now, observe that for $u \in A + B$, writing $p^{A\oplus B}$ for the density of the projected random variable $\mathrm{Pr}_{A\oplus B}(\mathrm{logsig}(*W)_1)$ we have that
$$p_1^{A\oplus B}(u) = \int_{\substack{Pr_{A\oplus B}(v)=u \\ v\in A\oplus B\oplus C}} p_1(v). \tag{10.2.2}$$
It follows from this and the continuity of p_1 that for almost every $u \in A \oplus B$, $p_1^{A\oplus B}(u) > 0$.

We now make use of the algebraic structure of the subspaces A, B, C again: recall that for any $a_1, a_2 \in A$, $b_1, b_2, \in B$, $c_1, c_2 \in C$, $\lambda, \mu \in \mathbb{R}^+$ we have
$$(\lambda e^{[0]} + a_1 + b_1 + c_1) \otimes (\mu e^{[0]} + a_2 + b_2 + c_2) = (\lambda e^{[0]} + a_1 + b_1) \otimes (\mu e^{[0]} + a_2 + b_2) + c_3$$
for some $c_3 \in C$. For any $v, u \in A \oplus B$, we also have a continuous way of choosing an element $w \in A + B$ such that $(\frac{1}{2}e^{[0]} + v) \otimes (\frac{1}{2}e^{[0]} + w) = e^{[0]} + u + c$ for some $c \in C$: for instance, take the projection on $A+B$ of the element $(F_{\frac{1}{2}}(v))^{-1}(F_1(u))$ where the inverse is taken in the Carnot group of step l with $d + 1$ underlying dimensions and F is the function from Proposition 10.2.2 rescaled with the time scale indicated as index). We write this element $w = I(v, u)$. It follows that we can use a similar argument to $p^{A\oplus B}$ as the one used on the density of the log signature of W (not $*W$) in [**37**]: we have
$$p_1^{A+B}(u) \geq \int_{v\in A\oplus B} p_{\frac{1}{2}}^{A+B}(v)p_{\frac{1}{2}}^{A+B}(I(v, u))dv.$$
Since $p_{\frac{1}{2}}^{A+B}(\cdot)$ is non zero almost everywhere, it follows from the above that $p_1^{A+B}(\cdot)$ is non zero everywhere (and by scaling, the same is true of $p_t^{A+B}(\cdot)$ for any t). Using that information to reinterpret formula (10.2.2) above, we see that for every $u \in A \oplus B$, there must exist a $v \in$

$A \oplus B \oplus C$ such that $p_1^{A\oplus B\oplus C}(v) > 0$. We denote that element $v = \bar{\phi}_1(u)$, and extend this definition for other times by scaling.

Note that this construction does not yet guarantee that $\bar{\phi}_t$ is smooth away from the origin. This is the problem we remedy in Step 2.

We now make the following observation: because for fixed t, a path of log signature $te^{[0]} + v$ for some $v \in \mathcal{S}$ can be rescaled, for any s, into a path of log signature $te^{[0]} + \bar{\delta}_s(v)$, where $\bar{\delta}_s$ denotes the homogeneous dilation in $\mathcal{S}$ where the degree of $e^{[\alpha]}$ is assigned to be $o_1(\alpha)$ instead of $|\alpha|$, we can conclude that v is in the *interior of* the support of $\mathrm{Pr}_{\mathcal{S}}(\mathrm{logsig}(*W)_t)$ if and only if $\bar{\delta}_s(v)$ is[2]. It follows that we can define $\bar{\phi}_t$ in a way that is scaling invariant both with respect to dilations δ_s and to dilations $\bar{\delta}_s$ acting only on the non-time components.

Step 2 [Definition of ϕ_t, smooth away from the origin]

We now want to turn $\bar{\phi}_t$ into a smooth function ϕ_t. To do this we must readapt a certain proportion of the construction from the previous proofs.

The first and main step is to reprove Lemma 10.2.1 for $\bar{B}$ (not B), in such a way that the log-signature (minus the time component) of the path constructed is in the interior of the support.

To do that, note that for any i, j with $i + 1 \leq l$, $j = 1, 2, \ldots d$ and $i \geq 1$, writing $w^{i,j}$ for $e^{[0,[0,[0,\ldots,[0,j]]]]}$ where there are i zero's, we have for any $s > 0$ and any $\lambda \in \mathbb{R}$, by the previous part of the proof, an element $\bar{\phi}_s(\lambda w^{i,j}) \in \mathrm{Supp}(\mathrm{Pr}_{\mathcal{S}}(\mathrm{logsig}(*W)_s)^\circ$. Now for any element $u \in \bar{B}_l$, develop via BCH the following equation (the order of concatenation is assumed to be fixed!):

$$x(\lambda) = \otimes_{\substack{1\leq i\leq l-1 \\ 1\leq j\leq d}} \bar{\phi}_{1/(l-1)}(\lambda_{ij} w^{i,j}) \backsimeq u,$$

where $\backsimeq$ means equal up to an element of C.

By definition of $\bar{\phi}$, we have that $\mathrm{Pr}_{\bar{A}}(x(\lambda)) = 0$ for any choice of λ. The above system is a linear triangular system in λ, and therefore has a solution which is a smooth function of u. For the λ that solves the system, we define $\phi(u) = \mathrm{logsig}(x(\lambda))$. Then $\phi(u)$ is a continuous function of u because λ is a continuous function of u and $\bar{\phi}_s(\lambda_{ij} w^{i,j})$ is continuous on the set $\{(i,j) : 1 \leq i \leq l-1, 1 \leq j \leq d\} \otimes \mathbb{R} \otimes \mathbb{R}$. This proves the equivalent of Lemma 10.2.1 further ensuring that ϕ is smooth away from the origin, and that $\phi(u) \in \mathrm{Supp}(\mathrm{logsig}(*W)_t)^\circ$ for any $u \in \bar{B}_l$. A simple extension of Proposition 7.0.11 that there exist some numbers $L(l) \in \mathbb{N}$, $\kappa_1, \ldots, \kappa_{L(l)} \in \{1, 2, \ldots, d\}$,$t_0 = 0 \leq t_1 \leq \ldots, t_{L(l)} = s$ such that for any element $a \in A$, and any s, it is possible to express $se^{[0]} + a$ as

$$\begin{aligned} &se^{[0]} + a \\ &= \mathrm{Pr}_A\left((t_1 + \lambda_1 e^{\kappa_1}) \otimes ((t_2 - t_1)e^{[0]} + \lambda_2 e^{\kappa_2}) \otimes \ldots \otimes ((t_{L(l)} - t_{L(l)-1}) + \lambda_{L(l)} e^{\kappa_{L(l)}})\right), \end{aligned}$$

with the choice of λ being smooth away from the origin and respecting all the usual scaling properties. Indeed, just pick a suitable path for the initial case, and the induction case is then unchanged. In the above concatenation, we now replace each element of the form $(t_i - t_{i-1})e^{[0]} + \lambda_i e^{\kappa_i}$ by $((t_i - t_{i-1})e^{[0]} + \phi_{t_i - t_{i-1}}(\lambda_i e^{\kappa_i}))$. This yields a smoothly defined element $\psi(a)$ with

$$\psi(a) \in \mathrm{Supp}(\mathrm{Pr}_{\mathcal{S}}(\mathrm{logsig}(*W)_s))^\circ \quad \text{and} \quad \mathrm{Pr}_A(\psi(a)) = a.$$

Now, for any $u \in A \oplus B$, we can proceed through the proof of Proposition 10.2.2, using the above constructions instead of the original piecewise linear ones, to obtain our element $\phi(u)$ which satisfies the same properties as F does in Proposition 10.2.2, plus smoothness and the property that $\phi(u) = \mathrm{Supp}(\mathrm{Pr}_{\mathcal{S}}(\mathrm{logsig}(*W)_t))^\circ$. Indeed, after having expressed $te^{[0]} + \phi(u) \in \mathcal{S}$ as a concatenation $\otimes_{i=1}^{L}(l)(s_i e^{[0]} + u_i)$ for some fixed $L(l)$ and some $u_1, \ldots, u_{L(l)}$ depending on the construction, we apply again the classic integration argument that

$$p_{\sum s_i}(u)$$

[2] For a short proof of the support theorem, see [**46**]

$$= \int_{(s_1 e^0 + u_1)\otimes(s_2 e^0 + u_2)\otimes \ldots \otimes (s_n + u_N) = te^0 + u} p_{s_1}(u_1) p_{s_2}(u_2) \ldots p_N(u_N) du_1 du_2 \ldots d_{u_N}$$

and

$$\forall i, \quad p_{s_i}(u_i) > 0$$

imply

$$p_{\sum s_i}(u) > 0.$$

As usual, this argument uses the continuity of the density, which is known from (the time-dependent version of) Hörmander's theorem, or alternatively, the previous parts of this thesis.

Because for any u, $\phi(u) > 0$, restricting u to the compact $|u|^2_{\mathcal{S}} \leq M$ (for $t = 1$) ensures we have a constant C_M such that

$$p(\phi_1(u)) \geq C_M,$$

The required upper bound now follows by Brownian Scaling. This concludes the proof. □

10.3. Application: Lower bounds for systems satisfying the separated weak progressive Hörmander condition

DEFINITION. We say that the system $\mathcal{A} = (x, \sigma, F)$ (for F linear) satisfies the separated weak progressive Hörmander condition with constants (L, H_L) if the following conditions are satisfied:

- $\mathcal{A}$ is (L, H_L)-detailed weak progressive Hörmander
- The vector fields σ are (L, H_L)-detailed weakly uniformly progressively finitely generated
- For any multi-index α such that $o_1(\alpha) \geq 2$, $o_0(\alpha) \geq 1$ and $|\alpha| \leq L$, there exist some λ such that

$$\sum_{|\beta| < |\alpha|} \lambda_\beta \sigma^{[\beta]} = \sigma^{[\alpha]} \quad \text{and}$$
$$\sum_{\beta} |\lambda_\beta|^2 \leq H_L^{-1}.$$

First, we must slightly adapt our construction of the compensated signature to this setting. Define, similarly to the development of showed for the compensated signature,

$$\bar{\Psi} : T^L(\mathbb{R}^d, \mathbb{R}) / \operatorname{span}(e^{(0)}) \to \mathcal{S} / \operatorname{Ker}(F^{SR_r})$$

such that

$$\sum_{h \in H} \bar{\Psi}^h(s) \sigma^h = \sum_{|\alpha| \leq L} \sigma^\alpha s^\alpha$$
$$|\bar{\Psi}(s)|^2 \leq \bar{H}_L^{-1} |s|^2 \quad \text{and} \quad \forall i \leq L$$
$$s \in \operatorname{span}_{\substack{|\alpha| \leq i \\ \alpha \neq (0)}} e^\alpha \implies \bar{\Psi}(s) \in \operatorname{span}_{\substack{|\alpha| \leq i; \alpha \neq (0) \\ o_1(\alpha) \leq 1 \vee o_0(\alpha) = 0}} e^{[\alpha]}$$

for some constant $\bar{H}_L$. Then define the following function:

$$\bar{\Psi}_1 : T^L(\mathbb{R}^d, \mathbb{R}) / \operatorname{span}(e^{(0)}) \to \mathcal{S} / \operatorname{Ker}(F^{SR_r}),$$
$$s \mapsto \operatorname{Pr}_{\mathcal{S}/\operatorname{Ker}(F^{SR_r})}(s) + \bar{\Psi}(s - \operatorname{Pr}_{\mathcal{S}}(s)).$$

Writing S for the l^{th} order signature, we can now define the l^{th} order *strictly compensated signature* $\bar{R}_l$ by

$$\tilde{S} := \operatorname{Pr}_{T^L(\mathbb{R}^d, \mathbb{R})/\operatorname{span}(e^{(0)})}(S)$$
$$\bar{R}_l := \bar{\Psi}_1(\tilde{S}) + \bar{\Psi}_2(F^{ST}(S - \tilde{S}) - \bar{x}_t).$$

We write $\bar{\mathcal{F}}_l$ (or $\bar{\mathcal{F}}$, omitting the index l) for the strictly compensated signature space

$$\operatorname{span}_{\substack{\alpha \neq (0); |\alpha| \leq L \\ o_0(\alpha) = 0 \vee o_1(\alpha) \leq 1}} e^{[\alpha]} / \operatorname{Ker}(F^{SR_r}),$$

and $F^{\bar{R}T}$ for the restriction of F^{ST} to $\bar{\mathcal{F}}$.

We now have the following:

THEOREM 10.3.1. *Let $\mathcal{A}$ be a uniformly (L,g,G)-tense, (L,H_L)-separated weak progressive Hörmander system. Suppose $g \geq (2n+2)^2 2^{4L} + n + 3$. There exists a proper constant D such that for any $M > 0$, there exists a constant C_M such that and for any x, y and t such that $d_{t,\log,\infty}(x,y) + \sqrt{t} \leq D$ and $d_{t,\log,\infty}(x,y) \leq Mt$, we have*

$$p_t(x,y) \geq \frac{C_M}{|B_{d_{t,\log,\infty}}(\sqrt{t})|}.$$

If the system is only (L,g,G)-tense, (L,H_L)-separated weak progressive Hörmander in a compact set $\mathcal{K}$ around $*x$, the same result holds for any $*x, y$ in a fixed compact set $\mathcal{K}'$ of non zero distance to $\mathcal{K}$.

PROOF. Fix $l \geq (2n+2)^2 2^{4L}$. Start by picking D small enough that the following conditions are satisfied:

- $F^{\log(S),R,l}$, restricted to $u \in \mathcal{S}$ with $|u|_{\mathcal{S}} \leq 2D$ and any $t \leq D$, has a C^∞ constant bounded by a proper constant.
- For $|u|_{\mathcal{S}} \leq 2D$ and any $t \leq D$, each diagonal bloc of $JF^{\log(S),R,l}(\phi(u) + \cdot)$, relative to components of the same scaling has a pseudo-inverse bounded by a proper constant.
- D is small enough to apply Theorem 7.1.1.

Recall that it is a fact about the construction of $F^{\log(S),R,l}$ for progressive Hörmander systems that this can be done for D being a proper constant. Note that $F^{\log(S),R,l}$ is the composition of the exponential function with a linear (x dependent) map. Because we don't aim to show anything about the dependence of our constants in the final estimate apart from uniformity in space-time, we do not need the fact that D is proper, but we do need the fact that the constant does not depend on the initial point x.

We note the following consequence of Theorem 4.4.1:
the density $p_t(s)$ for $s \in \mathcal{S}$ of $\mathrm{Pr}_{\mathcal{S}}(\mathrm{logsig}(*W))$ is differentiable and for time $t = 1$ and $|s|^2_{\mathcal{S}} \leq M$ (for any fixed M), the derivatives of given order are bounded above by some constant $\bar{C}_M$. By Brownian scaling, this immediately gives that for arbitrary $t \leq 1$, $|s|^2_{\mathcal{S}} \leq tM$ and $\alpha \in \mathrm{Multi}(\{0,1,\ldots,d\}) \setminus \{e^0\}$, we have for some fixed $\bar{C}_M$:

$$|\frac{\partial p_t(s)}{\partial e^\alpha}| \leq \frac{\bar{C}_M}{t^{\frac{\nu}{2}+|\alpha|}}.$$

Therefore if $|s+s'|^2_{\mathcal{S}} \leq Mt$; we have that[3]

$$|p_t(s+s') - p_t(s)| \leq \frac{K_M |s'|^2_{\mathcal{S}}/t}{t^{\frac{\nu}{2}}}$$

for some K_M. Using Proposition 10.2.3, this shows that if $u \in A \oplus B$, $|u|^2_{\mathcal{S}} \leq Mt$, $|\phi(u)+s'|^2_{\mathcal{S}} \leq Mt$ and $|s'|^2_{\mathcal{S}} \leq C_M/2K_M$ (where C_M is the constant from Proposition 10.2.3), we have the following:

$$p_t(\phi(u)+s') > \frac{C_M}{t^{\frac{\nu}{2}}} - \frac{K_M|s'|^2_{\mathcal{S}}/t}{t^{\frac{\nu}{2}}} \geq \frac{C_M}{2t^{\frac{\nu}{2}}}.$$

Replacing C_M and K_M by C_{4M} and K_{4M} in the above statements and using Lemma 5.1.2 (triangle inequality up to constants in graded spaces with homogeneous norms) we obtain that, still for any fixed M, there exist constants R_M and K_M such that for any $u \in A \oplus B$ with $|u|^2_{\mathcal{S}} \leq tM$, and any $s \in \mathcal{S}$ with $|s|^2_{\mathcal{S}} \leq tR_M$, we have

$$p_t(\phi(u)+s) > \frac{K_M}{t^{\frac{\nu}{2}}}.$$

[3] Note expressions such as $s+s'$ really mean the sum of s and s' as elements of the vector space $\mathcal{S}$, not the group multiplication $(e^0+s) \otimes (e^0+s')$

Consider any v inside the strictly compensated signature space $\bar{\mathcal{F}}$. The key point is that the function $F^{\log(S),R,l} \circ \phi : A \oplus B \to \mathcal{F}$ satisfies the conditions of Proposition 5.2.6, as a consequence of the separated progressive Hörmander condition. Indeed, the M_j have good Jacobians by construction of $F^{S,\bar{R}_l}$, the ψ^j come from the exponential function applied to the lower components of u, and the ϕ^i_k's come from the construction of $\phi(u)$, which require all of the u_i's to construct, but have strictly lower scaling because of the strict requirement $|\beta| < |\alpha|$ in the definition of the separated progressive Hörmander condition. Therefore there exist proper constants D_1, C_1 such that if $|v|^2_{\mathcal{F}}, t \leq D_1$, there exists a $u \in A \oplus B$ such that $|u|^2_{\mathcal{S}} \leq C_1 |v|^2_{\mathcal{F}}$ and $F^{\log(S),R,l}(\phi(u)) = v$.

Now, replace D by $\bar{D} = \min(D_1, D, D/C_1)$. We want to use the disintegration formula 2.2.1 on the neighborhood of $\phi(u)$ in the fiber $(F^{\log(S),R,l})^{-1}(\{v\})$. Because the isoscaled diagonal blocs of the Jacobian of $F^{\log(S),R,l}$ continue to have good pseudo-invertibility properties (second condition in 10.3) the area of the relevant neighborhoods in each fiber is greater than $Q_M t^{\frac{\nu-\bar{\nu}}{2}}$ for some constant Q_M (still proper, at this point). If it is furthermore the case that $|v|^2_{\mathcal{F}} \leq t D_M$ (where D_M comes from the application of Proposition 5.2.6 to make sure that $|u|^2_{\mathcal{S}}, |\phi(u)|^2_{\mathcal{S}} \leq Mt$), we can use Eq. 10.3 and the disintegration formula 2.2.1 to finally obtain, for some new C_M, D, D_M, and for any $v \in \mathcal{F}$ with $|v|^2_{\mathcal{S}} \leq D_M t$ and $|v|^2_{\mathcal{S}} \leq D$ and any $t \leq D$,

$$p_t(v) \geq \frac{C_M}{t^{\frac{\bar{\nu}}{2}}}$$

(Here $p_t(v)$ denotes the density at v of the random variable $\bar{R}_l$)

This puts us in a position to apply the results of Part 2 on Models, more specifically, Theorem 5.1.2, to obtain that for p^l_t being the density of the lth order KST approximation $\bar{Y}_t$of Y_t, we have for any y and t such that $d_{t,\log,\infty}(x,y) + \sqrt{t} \leq D$ and $d_{t,\log,\infty}(x,y) \leq Mt$,

$$\begin{aligned}
p^l_t(x,y) &> C_M \frac{e^{-d_{\log,l,t}(x,y)^2}}{|B_{d_{t,\log,l}}(x,\sqrt{t})|} \\
&\geq \frac{C_M}{|B_{d_{t,\log,l}}(x,\sqrt{t})|} \\
&\geq \frac{C_M}{|B_{d_{t,\log,\infty}}(x,\sqrt{t})|},
\end{aligned}$$

where C_M changes from line to line, and we have used doubling conditions and Theorem 7.1.1 at the last line.

Because the above line of reasoning obviously works exactly the same way if we replace p^l_t by a localised version of it, localised by ξ such that $\xi = 0$ whenever $|\operatorname{logsig}(*W)| \geq 1/2$, we can now use Theorem 6.4.1, with with the same localising function ξ, to obtain for some proper constant C and some constants C_M (depending on M but not x) that change from line to line,

$$\begin{aligned}
p_t(x,y) \geq p^{\xi}_t(x,y) &\geq p^{l,\xi}_t(x,y) - Ct \\
&> \frac{C_M}{|B_{d_{t,\log,\infty}}(x,\sqrt{t})|} - Ct \\
&\geq \frac{C_M}{|B_{d_{t,\log,\infty}}(x,\sqrt{t})|},
\end{aligned}$$

as required. The last statement is the consequence of another similar localizing argument. □

We can now concatenate the above result to show that the density of a uniformly separated weak Hörmander system is strictly positive everywhere:

THEOREM 10.3.2. *Let $\mathcal{A}$ be a uniformly (L, g, G)-tense, (L, H_L)-separated weak progressive Hörmander SDE ($F = \mathrm{Id}$). Suppose $g \geq (2n+2)^2 2^{4L} + n + 3$. we have, for any y, any x and* *any t,*

$$p_t(x, y) > 0.$$

PROOF. Let θ be a path in $\mathcal{L}^L(\mathbb{R}^{d+1})$ such that the solution to the SDE

$$x_0 = x$$
$$dx_t = \sum_{a \in H} (\sigma)^a(x_t) \frac{\partial \theta_t^a}{\partial t} dt = \sum_{a \in H} *\sigma^a(x_t) \frac{\partial \theta_t^a}{\partial t} dt$$

has $x_1 = y$. We can find times $0 < s_1 < s_2 < \dots s_N$ such that $x_{s_1}, x_{s_2}, \dots, x_{s_N}$ satisfy for any $i = 1, \dots, N-1$, $d_{s_i - s_{i-1}}(x_{i-1}, x_i) + \sqrt{s_i - s_{i-1}} \leq D$ where D is the constant from Theorem 10.3.1. Pick:

$$M = 2 \sup_i \frac{d_{s_i - s_{i-1}}(x_{i-1}, x_i)^2}{s_i - s_{i-1}}.$$

Then, we can apply Theorem 10.3.1 to obtain the strict positivity of the density of x_{s_i} conditional on the position of $x_{s_{i-1}}$. The classic integration argument $p_t(x, y) = \int_{x_1, \dots x_N} \Pi_{i=1}^N p_t(x_{i-1}, x_i)$, coupled with the continuity of the density (which as usual we know even from Hörmander's theorem), allows us to conclude that $p_t(x, y) > 0$, as expected. □

REMARK 10.3.1. Of course, the optimal possible M to choose above is a multiple of $d_t(x, y)$, but unlike the zero drift situation, this information isn't very easy to make the most of because the tail behaviour of sums of integrals of the form $\sum_{|\alpha|=i} W^\alpha$ for $o_1(\alpha)$ and $o_0(\alpha)$ non fixed is more difficult to study. In [**20**], indices of same order but different o_1 and o_0 don't mix, which helps in getting a global estimate.

REMARK 10.3.2. In [**30**] (Section 4), the authors manage to prove the strict positivity of the density for a different specific class of examples of diffusion which does not necessarily satisfy the separated progressive Hörmander condition.

CHAPTER 11

Examples of systems satisfying the separated weak progressive Hörmander condition

The Pigato SDE 9.0.2 satisfies the separated weak progressive Hörmander condition. The Delarue-Menozzi chain of differential Eqs. 9.0.4 from Chapter 9, and the two-sided version 9.0.5, *with an elliptic assumption on the σ's* also satisfies the separated weak progressive Hörmander condition. Below are more elaborate examples.

11.1. Multiple interacting (elliptic) Delarue-Menozzi chains

A more general class of examples would be the following further extension of 9.0.4:

Let $G = G_b \cup G_y \cup G_n$ a graph composed of a set of blue vertices G_b, a set of yellow vertices G_y, and a set of black edges G_n, such that the following conditions are satisfied:

- each yellow vertex is of degree two
- For each yellow vertex v_1, there are at least two paths of the form $(v_1, v_2, \dots v_N)$ such that for any $i \leq N$, (v_{i-1}, v_i) is an edge, $v_1, \dots v_{N-1}$ are yellow, and v_N is blue.
- For each black vertex v_1, there is only one path of the form $(v_1, v_2, \dots v_N)$ such that for any $i \leq N$, (v_{i-1}, v_i) is an edge, $v_1, \dots v_{N-1}$ are yellow, and v_N is not black. Furthermore, for the only such path, it is the case that v_N is blue.

Now, let $\mathcal{S}$ be the set of maps from G to $\mathbb{R}^n$, our state space.

For each yellow vertex $v \in G_y$, define $D(v)$ to be the set composed of the two immediate neighbours of v. Set also $\bar{D}(v) = D(v)$

For each black vertex v, define $D(v)$ to be the union of the set of all vertices that can be reached from v, but are not on the path $(v = v_1, v_2, \dots v_N)$ and the set $\{v, v_2\}$. Set also $\bar{D} = \{v_2\}$.

Consider the following SDE in $\mathcal{S}$:

$$
\begin{aligned}
S &\in \mathcal{S} \\
S_0 &= 0 \\
\forall v \in G_n \cup G_y, \quad dS_t^v &= \sum_{w \in D(v)} B^{v,W}(S) dt \\
\forall v \in G_b, \quad dS_t^v &= \sigma_v(S) \circ dW_t^v,
\end{aligned}
$$

where for any s, v, w, $B^{v,w}, \sigma^v$ are matrices in $\mathbb{R}^n \otimes \mathbb{R}^n$ satisfying the following conditions:

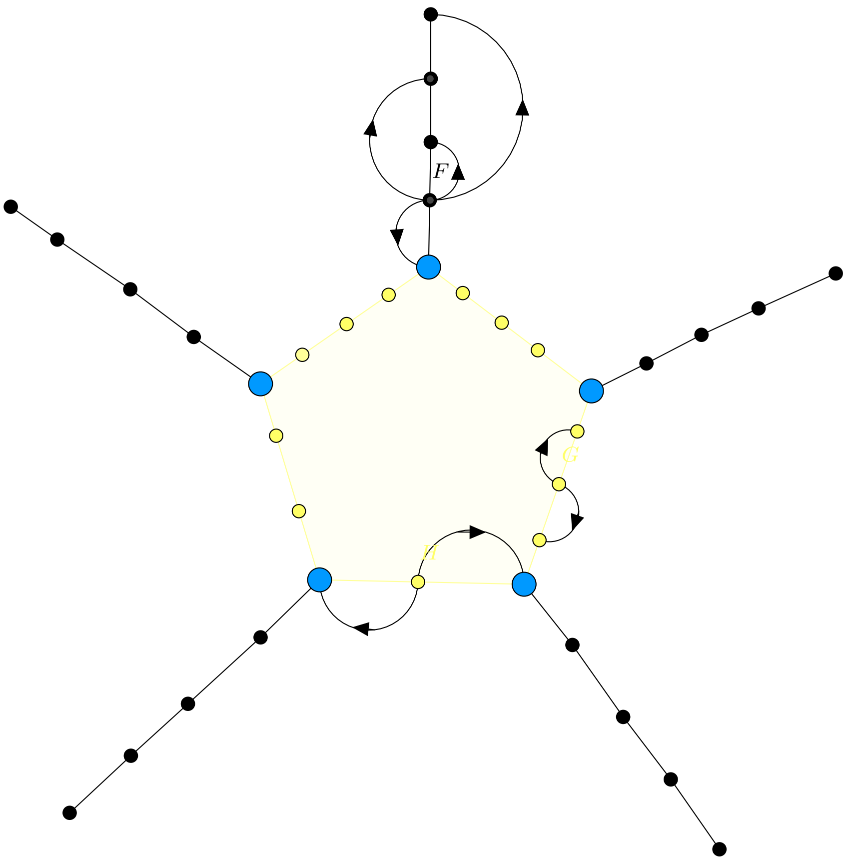
- All operator norms of $B^{v,w}, \sigma^v$ are uniformly bounded by a constant C.
- $\forall v \in G_n \cup G_y$, and for any $\omega \in \mathbb{R}^n$, there exists a $\lambda \in \mathbb{R}^{n\#(\bar{D}(v))}$ such that $|\lambda|^2 \leq H^{-1}$ (for some given constant H) and (where w runs over all elements of $\bar{D}(v)$:

$$(\partial_{s^{w_1}} B^{v,w_1} \partial_{s^{w_2}} B^{v,w_2} \dots)\lambda = \omega$$

- $\det(\sigma^v(\sigma^v)^T)$ is bounded below uniformly over s and v.

Such a system is uniformly weak separated progressive Hörmander, in particular, Theorems 10.3.1 and 10.3.2 apply.

The following graph shows an example. An arrow from a to b signifies that B_a depends on S^b. All arrows from points F, G, H have been drawn.



11.2. More specific examples

1. The following SDE is not mono-scaled, and is one of the most simple 'completely non-trivial' examples of an separated weak progressive Hörmander system.

Consider the following SDE in $\mathbb{R}^4$: $X_0 = 0; \quad dX_t = \sigma(X_t) \circ dW_t$ with

$$\sigma^1 \begin{pmatrix} x \\ y \\ z \\ u \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -y \\ 0 \end{pmatrix}, \qquad \sigma^2 \begin{pmatrix} x \\ y \\ z \\ u \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ x \\ 0 \end{pmatrix}, \qquad \sigma^0 \begin{pmatrix} x \\ y \\ z \\ u \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ x^2 \\ xu+1 \end{pmatrix}.$$

The non zero brackets and derivatives of order less than 5 are:

$$\sigma^{[12]} = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 0 \end{pmatrix}, \qquad \sigma^{(12)} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \sigma^{(21)} = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \end{pmatrix},$$

$$\sigma^{[10]} = \begin{pmatrix} 0 \\ 0 \\ 2x \\ u \end{pmatrix}, \qquad \sigma^{(01)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad \sigma^{(10)} = \begin{pmatrix} 0 \\ 0 \\ 2x \\ u \end{pmatrix},$$

$$\sigma^{[1,[1,0]]} = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 0 \end{pmatrix}, \qquad \sigma^{[0,[1,0]]} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1+xu \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 0 \\ xu \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

We see that the separated progressive weak Hörmander condition is satisfied. In particular, Theorems 10.3.1 and 10.3.2 apply.

The intuitive explanation of the separated weak Hörmander philosophy using this particular example is: by our construction in 10.2.2 etc, 'the diffusion moves freely', at speed $t^{5/2}$ in the direction $\sigma^{[0,[1,0]]} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$. On the other hand, fourth order (=at speed t^2) movement in the direction of $\sigma^{[1,[1,0]]} = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 0 \end{pmatrix}$ is constrained and problematic because of issues mentioned such as the phenomenon observed in 10.1.1 and in Léandre's example (10.1.1). However, unconstrained movement in the same direction is occurring at speed t from $\sigma^{[1,2]}$. Theorem 5.2.6 helps show that the 'free' order two movement wins over the constrained order four movement.

Part 4

Löcherbach systems

CHAPTER 12

Introduction and first properties

The notation in the chapter on 'composite systems', the notation in the rest of Part 4, and the notation in the rest of this thesis, are three independent sets of notations. For instance, here we will use l to denote the number of particles, as in [**42**] *and* [**10**] *not the order of any Taylor approximation, as in the rest of the thesis and in* [**37**].

12.1. The independent case

In [**10**], Bally and Löcherbach consider a finite system of branching diffusions with immigration, where the movement of each particle is independent from the others.

More precisely, the model is the following (as explained in [**10**], with different notation): Write $S = \cup_{l\geq 0}\mathbb{R}^{n\times l}$ for the set of all configurations of particles in $\mathbb{R}^n$. Write elements of S as $x = (x^1, x^2, \dots, x^l)$.

We consider the following a process $\eta_t \in S$ whose sample paths are piecewise continuous $\mathbb{R}^{n\times l}$-valued functions, for varying l, with the following assumptions:

- During the life time of an l- particle configuration, each particle ξ^k (for $k = 1, 2, \dots, l$) evolves, independently of other particles, according to the following SDE:

$$d\xi_t^k = \sum_{i=1}^{d} \sigma^i(\xi^k) \circ dW_t + \sigma^0(\xi^k)dt = \sum_{i=0}^{d} \sigma^i(\xi^k) \circ dW_t. \tag{12.1.1}$$

- Particles branch at a position dependent Poisson rate $\kappa(\cdot)$ according to the reproduction law $(p_k(\cdot))_{k\in\mathbb{N}_0}$. Each new particle then evolves according to the SDE (12.1.1).
- New particles appear at rate $c > 0$. The immigration measure r satisfies $r(dx) = r(x)dx$ (other situations are also considered in [**10**]).
- At time 0, the system starts at the void configuration with no particles.

The assumptions are:

- σ, κ are C^∞ and bounded, and there exist constants a_1, a_2 such that $a_1 \leq \kappa(z) \leq a_2$ for any $z \in \mathbb{R}^m$;
- σ satisfies the weak Hörmander condition uniformly with constant H_L;
- there exists a fixed probability measure $\bar{p}$ on $\mathbb{N}_0$ such that $\sum_0^\infty \bar{p}_k < 1$, and for any $x \in \mathbb{R}^n$, there exists a probability measure $\hat{p}$ on $\mathbb{N}_0$ such that $p \star \hat{p} = \bar{p}$, where $\star$ denotes convolution of probability measures.

We very briefly summarize the strategy used for the proof in [**10**], since it is useful to understand the proof in the interacting case. See [**10**] for further details.

Following [**10**], we write $\bar{m}$ for the invariant measure (for a Borel set $A \subset \mathbb{R}^n$, $\bar{m}(A)$ is the expected number of particles in A for a random configuration following the invariant probability of the process in S), Δ for the void configuration (with no particles), and R for the first return time to Δ. For a configuration η and a Borel set $A \subset \mathbb{R}^n$, we also write $\eta(A)$ for the number of particles in A.

Write, for $x, y \in \mathbb{R}^n$ and $A \subset \mathbb{R}^n$:

$$U(x, A) = \mathbb{E}_x \left(\int_0^\infty e^{-\int_0^s \kappa(\xi_r)dr} 1_A(\xi_s)ds \right),$$

$$K(x, dy) = U(x, dy)(\kappa\rho)(y) = \mathbb{E}_x \left(\int_0^\infty e^{-\int_0^s \kappa(\xi_r)dr} (\kappa\rho)(y)\delta_{\xi_s}(dy)ds \right).$$

Roughly speaking, $U(x, A)$ is the expected amount of time that a particle started at x will spend in A before death, and $K(x, dy)$ is the expected number of first generation offspring that will be born at y supposing we start with a configuration of one particle at x.

It is proved in [**10**] and [**29**] that

$$\bar{m}(A) = \frac{1}{\mathbb{E}_\Delta(R)}\mathbb{E}_\Delta(\int_0^R \eta_s(A)ds) = c\sum_{N=0}^{\infty} rK^N U(A). \tag{12.1.2}$$

Classical results from [**36**], or the rest of this thesis, show that $U(x, dy)$ has a smooth density. If one is to show that $\bar{m}$ has a density, the main difficulties will be to show that U^N also has a density, and that the summation involved in (12.1.2) converges in a sufficiently strong sense.

Using the inverse stochastic flow $\phi_{t,0}(\cdot)$, we can rewrite K in a form closer to the expression of a density function: For a test function ψ,

$$rK\psi = \mathbb{E}\left(\int_0^\infty r(\phi_{t,0}(y))e^{-\int_0^s \kappa(\phi_{t,s}(x))ds}\psi(y)(\kappa\rho)(y)J(\phi_{t,0}(y))dydt\right),$$

where $J(\cdot)$ denotes the Jacobian.

Next define, for any $\epsilon > 0$:

$$\theta(r)(x) = \theta_\epsilon(r)(x) = (\kappa\rho)(x)\mathbb{E}\left(\int_0^\epsilon r(\phi_{t,0}(x))J(\phi_{t,0}(x))e^{-\int_0^s \kappa(\phi_{t,s(x)})dr}dt\right).$$

Note that θ is the invariant density of a branching process with $c = 1$, no possibility of branching (only death), and where death is forced to occur at time ϵ if it hasn't occurred before. Because, by classical results, $\mathbb{E}(J(\phi_{t,0}(x)))$ is locally space-time integrable, we can pick ϵ so that θ is a contraction. Then we can define:

$$\Theta(r)(x) = \Theta_\epsilon(r)(x) = \sum_{N=1}^{\infty}\frac{\theta^N(y)(x)}{(\kappa\rho)(x)}.$$

Write also $\gamma_t(x, \cdot)$ for the density of the κ-killed diffusion started at x, and

$$T(x, y) = T_\epsilon = \Theta_\epsilon(\gamma_\epsilon(x, \cdot)\kappa\rho)(y) + \gamma_\epsilon(x, y)$$

Then the key formula to understand the proof in [**10**] is the following, where $p(r, x)$ denotes the density of the invariant measure $\bar{m}$ (in particular, $\bar{m}$ has a density):

$$p(r, x) = c\Theta_\epsilon(r)(x) + \int p(r, y)T_\epsilon(y, x)dy.$$

The first term corresponds to particles arriving in the small set dx coming from a line of descendants which all died (branched) after a life of less than ϵ. The existence of the density, and the above formula, are established via dominated convergence applied to formula (12.1.2) after changes of variables along the lines described above.

12.2. Description of the general case.

In [**42**], Löcherbach investigates the situation where each particle's movement, branching rate and distribution are influenced by the configuration of other particles. The existence of an invariant density is established under assumptions that include uniform ellipticity of the driving vector fields σ. This is in contrast with the independent situation, where only the weak Hörmander condition was required. In this section, we both:

- Reproduce the result under weaker assumptions (getting rid of assumption 'H6' from [**42**]), and in a more general situation where the rates are allowed to depend in a certain way on the path of the process rather than only the position of other particles at time t, and
- Prove an extension of the result to a weak Hörmander situation, at the cost of imposing the so called 'No Degeneracy from Interaction' (NDI) condition, *and assuming a proof of a strongly polynomial version of Theorem 4.4.1.*

Both situations will be treated simultaneously due to the similarities between the two proofs. (In fact one could even see the elliptic case as a particular case of the NDI case)

The model is the following, and includes 'latent particles', meant to model path-dependence of the branching rate and diffusion coefficients:

– The configuration space is $S = \cup_{l\in\mathbb{N}_0}\mathbb{R}^{n\times l}\otimes\mathbb{R}^{n\times\mathcal{F}(l)}$. The l first particles are the 'real' particles, and the $\mathcal{F}(l)$ others are the 'latent particles'. $\mathcal{F}:\mathbb{N}_0\to\mathbb{N}_0$ is a fixed function satisfying $\mathcal{F}(x)\le C_P x^P$ for some fixed P, C_P and such that
$$\mathcal{F}(0) = 0.$$
We write a complete configuration in the form
$$(\eta_t,\bar\eta_t) = (\eta^1,\eta^2,\ldots,\eta^l,\bar\eta^1,\ldots,\bar\eta^{\mathcal{F}(l)})\in S.$$
– During the lifetime of each configuration, particles $\xi\in\mathbb{R}^{n\times l}$ and $\zeta\in\mathbb{R}^{n\times\mathcal{F}(l)}$ move according to the following SDE
$$\begin{aligned} d\xi^i &= \sum_{k=0}^{d}\sigma^k(i,\xi,\zeta)\circ dW^k_{i,t} \quad \forall i\in\{1,2,\ldots,l\}\\ d\zeta^i &= \sum_{k=0}^{d}\Sigma^k(i,\xi,\zeta)\circ dW^k_{i,t} \quad \forall i\in\{1,2,\ldots,\mathcal{F}(l)\}, \end{aligned}$$
where each value of i corresponds to an independent Brownian motion
$$*W_{i,\cdot} = (t, W^1_{i,t}, W^2_{i,t},\ldots,W^d_{i,t}).$$
We will often omit the subscript i when it is clear from context.

– Real particles branch at branching rate κ according to a uniformly bounded reproduction law p, and latent particles adapt their configuration and branch accordingly. More specifically,
  - A particle x^i with $i\in\{1,2,\ldots,l\}$ which belongs to a configuration
$$(x,y) = (x^1,x^2,\ldots,x^l,y^1,y^2,\ldots,y^{\mathcal{F}(l)})\in\mathbb{R}^{n(l+\mathcal{F}(l))}$$
will die during interval $[t,t+h]$ with probability $\kappa(i,x,y)h+o(h)$.
  - At its death, a particle x^i is replaced by k particles, where $k\in\{0,2,3,\ldots,K\}$ (for some fixed $K\in\mathbb{N}$) is chosen according to a probability distribution
$$p(i,x,y) = (p_0(i,x,y),p_2(i,x,y),\ldots,p_K(i,x,y)).$$
  - At each branching event T_N, where particle η^i branches into k new particles, the latent particles $\bar\eta$ change position according to the following equation:
$$\bar\eta_{T+} = \bar{\mathcal{F}}_{1,l,k,i}(\eta_{T-},\bar\eta_{T-}),$$
where $\bar{\mathcal{F}}_{1,l,k,i}:\mathbb{R}^{n\times l}\otimes\mathbb{R}^{n\times\mathcal{F}(l)}\to\mathbb{R}^{n\times\mathcal{F}(l+k-1)}$ is a fixed function (no assumptions on $\bar{\mathcal{F}}_1$ are required).
  - Particles enter the configuration at constant rate $c>0$. Each immigrant chooses its position according to the an immigration law $\pi(dx)=\pi(x)dx$. We suppose $\pi\in C_b(\mathbb{R}^n)$.
  - At each immigration event T_N, where particle η^{l+1} appears at position x^{l+1}, the latent particles $\bar\eta$ change position according to the following equation:
$$\bar\eta_{T+} = \bar\eta_T = \bar{\mathcal{F}}_{2,l}(\eta_T,\bar\eta_{T-}),$$
where $\bar{\mathcal{F}}_{2,l}:\mathbb{R}^{n\times(l+1)}\otimes\mathbb{R}^{n\times\mathcal{F}(l)}\to\mathbb{R}^{n\times\mathcal{F}(l+1)}$, is a fixed function (no assumptions on $\bar{\mathcal{F}}_2$ are required).

The subcriticality and boundedness assumptions are the following:
- (SC1) $\kappa(\cdot)$ satisfies $a\le\kappa\le b$ for some fixed $0<a<b$.
- (SC2) The probability distribution $p(i,x,y)$ satisfies $p_k\le\bar p_k$ for all i,x,y,k and for some fixed $(\bar p_0,\bar p_2,\ldots,\bar p_K)$ such that
$$\bar p_0+\bar p_2+\ldots+\bar p_K<1.$$

- (BO) For each q, there exists a constant C_q, independent of l, such that for any $k \in \{0,1,\ldots,d\}$, $i \in \mathbb{N}$, $x \in \mathbb{R}^{n\times l}$, $y \in \mathbb{R}^{n\times\mathcal{F}(l)}$, $u,u_1,u_2,\ldots,u_q \in \{1,2,\ldots,n\}$, $v_1,v_2,\ldots,v_q \in \{1,2,\ldots,l+\mathcal{F}(l)\}$, and any x,y, writing $z=(x,y)\in\mathbb{R}^{n\times(l+\mathcal{F}(l))}$, we have:

$$|\frac{\partial^q(\sigma^k(i,x,y))_u}{\partial z_{u_1}^{v_1}\partial z_{u_2}^{v_2}\ldots\partial z_{u_q}^{v_q}}| \leq C_q$$
$$|\frac{\partial^q(\Sigma^k(i,x,y))_u}{\partial z_{u_1}^{v_1}\partial z_{u_2}^{v_2}\ldots\partial z_{u_q}^{v_q}}| \leq C_q$$
$$|\frac{\partial^q(\kappa(i,x,y))_u}{\partial z_{u_1}^{v_1}\partial z_{u_2}^{v_2}\ldots\partial z_{u_q}^{v_q}}| \leq C_q.$$

The non-degeneracy assumption is either one of the following:

- (Ellipticity, E) There exists a fixed constant E such that for any unit $v\in\mathbb{R}^n$, any l,i,x,y, we have:

$$\sum_{k=1}^{d}\langle\sigma^k(i,x,y),v\rangle^2 > E,$$

or, the combination of both of the following conditions:

- (No Degeneracy from Interactions, NDI) For each l,i, there exists a function

$$\lambda:\mathbb{N}\otimes\mathbb{R}^{n\times l}\otimes\mathbb{R}^{n\times\mathcal{F}(l)}\to\mathbb{R}^{d\times d},$$

  a function

$$\beta:\mathbb{N}\otimes\mathbb{R}^{n\times l}\otimes\mathbb{R}^{n\times\mathcal{F}(l)}\to\mathbb{R}$$

  and d vector fields $\underline{\sigma}^k:\mathbb{R}^n\to\mathbb{R}^n$ such that for any l,x,y,i,k,q, the derivatives, in any axis direction, of order less than q, of $\underline{\sigma}^k$ and all the entries of λ are uniformly bounded by C_q, and the following holds for any i,l,x,y for some fixed constant $E>0$:

$$\begin{aligned}
\sigma(i,x,y)^k &= \sum_{j=1}^{d}\underline{\sigma}(x^i)^j\lambda_{j,k}(i,x,y) \quad (\forall k\in\{1,2,\ldots,d\}),\\
\sigma^0(i,x,y) &= \beta(i,x,y)\underline{\sigma}^0(x^i),\\
E &\leq \lambda_*(\lambda(i,x,y)),\\
E &< \beta < \frac{1}{E}.
\end{aligned}$$

- (Weak Hörmander condition on the underlying, H) There exist constants (L,H_L) such that for any $u\in\mathbb{R}^n$, and any unit $v\in\mathbb{R}^n$,

$$\sum_{\substack{\alpha\in\mathrm{Multi}(\{0,1,\ldots,d\})\\|\alpha|\leq L,\alpha\neq(0)}}\langle\underline{\sigma}^{[\alpha]}(u),v\rangle^2 > H_L.$$

As explained in the next section, the assumptions SC1 and SC2 ensure subcriticality and that the zero configuration Δ is a recurrent point. Following [**42**], we are interested in the intensity measure $\bar{m}$ associated to this branching diffusion, i.e., writing R for the return time to the null configuration Δ and for any set A, writing $\eta_t(A)$ for the number of ('real') particles in the set A at time t:

$$\bar{m}(A)=\frac{\mathbb{E}(\int_0^R\eta_t(A)dt)}{\mathbb{E}(R)}.$$

We will show that $\bar{m}$ admits a continuous density in the elliptic case and that it admits a continuous density in the NDI case assuming a proof of a strongly polynomial version of Theorem 4.4.1. The first step is the recurrence of the null configuration, which was already established in [**42**] and [**29**] (the differences between the settings in terms of non-degeneracy are not relevant to that aspect), then the main technical step involving density estimation will be the estimate in the next chapter.

12.3. Consequence of subcriticality

Here we explain the point made in Lemma 4.1 from [**42**]. Note that this lemma does not require any non-degeneracy assumption, and can be used in our context.

LEMMA 12.3.1. *Under the subcriticality assumptions SC1 and SC2, we have that*

$$\mathbb{P}(T_k < R) \leq Ce^{-ck},$$

for some constants C, c*.*

SKETCH OF PROOF. The idea (see [**42**] and [**29**]) is the following:

To alleviate notation, we assume there are no latent particles, and $K = 2$, the general case is similar. Begin by constructing a coupled process $\bar{\eta}$ such that at every branching event, the probability of splitting into two particles is exactly $\bar{p}_2$. The immigration process stays the same. This can be done whilst making sure that at every time, the number of particles $\ell(\bar{\eta})$ in $\bar{\eta}$ is more than the number of particles $\ell(\eta)$ in η.

Next, we define a process $\tilde{\eta}$ with state space $\mathbb{N} \otimes \{0\}$ as branching diffusion where particles do not move, in a coupled way with $\bar{\eta}$: (taking some notational liberties) if the lifetime of a given particle p_N in $\bar{\eta}$ and all its descendants $p_1, p_2, \ldots, p_{N-1}$ (with positions $\xi_t^1, \xi_t^2, \ldots, \xi_t^n \in \mathbb{R}^n \cup \varnothing$) are

$$[T_0, T_1], [T_1, T_2], \ldots, [T_{N-1}, T_N]$$

respectively, then the lifetime of the corresponding particle in $\tilde{\eta}$ is

$$[\sum_{i=0}^{N-1} \int_{T_i}^{T_i+1} \frac{\kappa(i, \xi_t^i, \xi_t)}{a} dt, \sum_{i=0}^{N} \int_{T_i}^{T_i+1} \frac{\kappa(\xi_t^i, \xi_t)}{a} dt].$$

This new coupled process $\bar{\eta}$ satisfies the following properties:

- The respective numbers of generations $\tilde{N}$ (for $\tilde{\eta}$) and $\bar{N}$ (for $\bar{\eta}$) before return to the zero configuration Δ satisfy $\tilde{N} \leq \bar{N}$, leading to $\mathbb{P}(N \geq k) \leq \mathbb{P}(\tilde{N} \geq k)$.
- The configurations of $\tilde{\eta}$ at successive jump times are a subcritical discrete time Markov chain on a finite state-space. (Indeed, for large enough l, $\frac{al\bar{p}_2+c}{al+c} - \frac{al\bar{p}_0}{al+c} < 0$). Furthermore, the upper bound K on the number of offspring ensures the exponential decay of the return probabilities to the zero configuration.

The result then follows from classical discrete-time Markov chain theory. □

REMARK 12.3.2. The argument above and the assumptions on the branching rate also show that $\mathbb{E}(R) < \sum ce^{-ck}(1/a) < \infty$.

CHAPTER 13

Polynomial upper bounds for a class of systems

This chapter contains the main technical step required to prove an extension of the Löcherbach to an NDI situation with weak Hörmander assuming a strongly polynomial version of Theorem 4.4.1.

Warning: The results in this chapter assume a strongly polynomial version of Theorem 4.4.1.

The notation used in this chapter is separate from the notation used in the rest of this part of the thesis. The notion of 'proper constant' is also less strict, as d *(but not* D*) is considered 'small' in this context.*

We consider the following class of systems:

DEFINITION 13.0.1. A K_λ-composite system is a system of the following form

$$
\begin{aligned}
&(X,Y) \in \mathbb{R}^n \otimes \mathbb{R}^m,\\
&(X_0,Y_0) = (x,y),\\
&dY_t = \sum_{i=0}^{D} \Sigma(X_t,Y_t)^i \circ dW_t^i,\\
&dX_t = \sum_{j=1}^{d}\sum_{i=1}^{d} \sigma(X_t)^j \lambda(X_t,Y_t)^i_j \circ dW_t^i + b(X_t,Y_t)dt,
\end{aligned}
$$

where $\sigma^i \in \mathbb{R}^n$, $\Sigma^i \in \mathbb{R}^m$, $b \in \mathbb{R}$, $\lambda \in \mathbb{R}^{d\times d}$ satisfy the following condition uniformly:

$$
\begin{gathered}
K_\lambda^{-1} \le \lambda_*(\lambda) \le \lambda^*(\lambda) \le K_\lambda,\\
K_\lambda^{-1}\sigma(X_t,Y_t)^0 \le b(X_t,Y_t) \le K_\lambda \sigma(X_t,Y_t)^0.
\end{gathered}
$$

DEFINITION 13.0.2. We call constants relative to a K_λ-composite, (L,H_L)-weak Hörmander, (L,g,G)-tense system *proper* if they depend only on $d,n,G,g,L,H_L,K_\lambda,D$ (but not m) and are polynomial in H_L,G,K_λ,D.

We note the following result:

PROPOSITION 13.0.3. *Let* $\mathcal{A} = ((x,y),(\sigma,\Sigma),\mathrm{Pr})$ *be a* K_λ*-composite,* (L,g,G)*-tense system such that* σ *satisfies the weak Hörmander condition of order* L *with constant* H_L *at* x*. There is a constant* H_L'*, depending only on* K_λ,d,G,H_L*, such that* $\mathcal{A}$ *is* (L,H_L')*-weak Hörmander at* (x,y)*.*

PROOF. In light of Lemma 3.5.1, it is enough to show that for any α with $|\alpha| = k \le L$, there exist some $\mu_{\beta,\alpha}$ such that $\sigma^{[\alpha]} = \sum_{|\beta|\le|\alpha|}(\sigma\lambda)^{[\beta]}\mu_{\beta,\alpha}$ and $\sum_\beta |\mu_{\beta,\alpha}|^2 \le (H_L')^{-1}$. For simplicity, we assume zero drift. The general case is similar.

We can prove the result by induction over k. For $k=1$, simply pick $\mu_{i,j} = (\lambda^{-1})^j_i$. Now, assuming the result holds for a given k, any multi-index of order $k+1$ can now be written in the form (i,α) for some α of order k, and some $1 \le i \le d$. Then we have

$$
\sigma^{[i,\alpha]} = [\sum_j (\lambda^{-1})^i_j \sigma^j, \sum_\beta \mu_{\beta,\alpha}(\sigma\lambda)^{[\beta]}] = \sum_{\beta,j} \mu_{\beta,\alpha}(\lambda^{-1})^i_j \sigma^{[i,\beta]}.
$$

Next, we have clearly:

$$\sum_{\beta,j} |\mu_\beta(\lambda^{-1})^i_j|^2 \leq dK_\lambda \sum_\beta |\mu_{\beta,\alpha}|^2.$$

The result follows. □

As a result of the above proposition, there is no problem, when discussing whether constants are proper, interchanging the weak Hörmander constant relative to the σ and the weak Hörmander constant relative to the whole system $\mathcal{A}$.

We note the following preliminary results on an Taylor approximation that takes interactions into account:

13.1. Taylor approximation for composite systems

We have the following stronger version of Theorem 2.1.1:

LEMMA 13.1.1. *Suppose we are given a multi-index α, a real number a, and stochastic processes $(A^i_t)_{0\leq t\leq T}$ for $i = 0, 1, 2, \ldots, d$ such that $\forall j \in \{1, 2, .., d\}$,*

$$\sup_{t\in[0,T]} |A^j_t| \leq a$$

and

$$\sup_{t\in[0,T]} |A^0_t| \leq a^2.$$

Define the iterated integrals $J^\alpha(A,T)$ by (here $\alpha = (\bar{\alpha}, \alpha_k)$)

$$J^{(0)}(A,T) = \int_0^T A^0_t dt$$
$$J^{(k)}(A,T) = \int_0^T A^k_t dW^k_t$$
$$J^\alpha(A,T) = \int_0^T J^{\bar{\alpha}}(A,t) A^{\alpha_k} dW^{\alpha_k}_t.$$

For $C_{|\alpha|}$ as in Theorem 2.1.1, we have

$$\mathbb{P}\left(\sup_{0\leq t\leq T} |J^\alpha| > R\right) \leq 2C_{|\alpha|} \exp\left(-\frac{(\frac{R}{a})^{\frac{2}{|\alpha|}}}{2T}\right).$$

PROOF. The proof is the same as that of Theorem 2.1.1 except that $B = a^{\frac{-|\bar{\alpha}|}{|\bar{\alpha}|+1}} R^{\frac{|\bar{\alpha}|}{|\bar{\alpha}|+1}}$. □

Of course, a composite system is a system, and we can consider the standard KST approximation on it. This, however, does not allow us to obtain polynomial dependence of the constants. Therefore, we will also consider the Taylor expansion obtained by treating λW as a driving path and σ as the driving vector field, rather than treating $\sigma\lambda$ as the driving vector field and W as the driving path.

Let $\mathcal{A}$ be a composite system, the interactive KST approximation of Y_t of order l is

$$T^l_t = y + \sum_{|\alpha|\leq l} (\sigma)^\alpha (\lambda W)^\alpha$$

PROPOSITION 13.1.2. *Let $\mathcal{A}$ be a K_λ-composite, uniformly (L, g, G)-tense system. For $R \geq 1$ and for any $l \leq g - 1$, there exist constants C_1, C_2, depending only on K_λ, G, d, g, l, D, and depending polynomially on K_λ, G, d, g, D, such that*

$$\mathbb{P}(|Y_t - T^l_t| \geq R) \leq C_1 e^{-\frac{R^{2/(l+1)}}{tC_2}}.$$

PROOF. The theorem follows upon applying Theorem 13.1.1:

$$\begin{aligned}
\mathbb{P}(|Y_t - T_t^l| \geq R) &\leq \sum_{|\alpha|=l+1} \mathbb{P}(|\sigma^\alpha \int^\alpha d(\lambda W)| \geq R/d^{l+1}) \\
&\leq \sum_{|\alpha|=l+1} 2C_{l+1} e^{-\frac{R^{2/(l+1)}}{2t(GK_\lambda)^{2/(l+1)}}} \\
&\leq (d+1)^{l+1} 2C_{l+1} e^{-\frac{R^{2/(l+1)}}{2t(GK_\lambda)^{2/(l+1)}}} \\
&\leq C_1 e^{-\frac{R^{2/(l+1)}}{tC_2}},
\end{aligned}$$

where C_{l+1} denotes the constant from Lemma 13.1.1 for order $l+1$. □

13.2. Proof of upper bounds

The main theorem of this chapter is the following:

THEOREM 13.2.1. *Let $\mathcal{A} = ((x,y),(\sigma,\Sigma),\mathrm{Pr})$ be a uniformly (L,g,G)-tense K_λ-composite system such that σ satisfies the weak Hörmander condition with constants (L,H_L) uniformly in the compact $B(x,R)$ for some $R>0$. Suppose $g \geq (2n+2)^2 2^{4L} + n + 3$. There exist proper constants C_1, C_2, C_3 (depending on R) such that for any $t \leq C_1$ and any $\bar{x} \in \mathbb{R}^n$, X_t admits a density $p_t(x,\bar{x})$ such that*

$$p_t(x,\bar{x}) \leq C_2 \frac{e^{-\frac{d_t(x,\bar{x})^2}{C_3 t}}}{|B_{d_t}(x,\sqrt{t})|},$$

where d_t is the usual time-dependent distance defined for the system (x,σ,Id).

PROOF. By a standard localising argument as in 8.3.1, we can suppose that the σ's are uniformly (L,H_L) weak Hörmander over the whole of $\mathbb{R}^n$.

Since the system $\mathcal{A}$ is uniformly (L,g,G)-tense and uniformly (L,H_L)-weak Hörmander, Theorems 8.1.2 and 4.5.2 apply. Since the 'distance' d_t is trivially properly locally equivalent (with equivalence constant K_λ) to the 'distance' $\bar{d}_t$ defined for the whole system $\mathcal{A}$, we already can deduce that, for some proper constants C_1, C_2, C_3, ν, $\forall t \leq C_1$, $d_t(x,\bar{x}) \leq C_1$,

$$p_t(x,\bar{x}) \leq C_2 \frac{e^{-\frac{d_t(x,\bar{x})^2}{C_3 t}}}{t^\nu},$$

and also that for any $\bar{x} \in \mathbb{R}^n$ and any $t \leq C_1$,

$$p_t(x,\bar{x}) \leq C_2 \frac{e^{\frac{-|x-\bar{x}|^2}{C_3 t}}}{t^\nu}.$$

As in the proofs of Theorems 8.2.1, 8.3.1 etc., it then follows we only have to show the most local part of the estimate, i.e. that $p_t(x,\bar{x}) \leq C_2 \frac{e^{-\frac{d_t(x,\bar{x})^2}{C_3 t}}}{|B_{d_t}(x,\sqrt{t})|}$ for $t \leq e^{-\frac{d_t(x,\bar{x})^2}{2\nu C_3 t}}$.

Fix $l = (2n+2)^2 2^{4L}$ Now, by Theorems 13.1.2 and 6.4.1, it is enough to show the estimate for the stochastic Taylor approximation $x + \sum_{|\alpha|\leq l} \sigma^\alpha (\lambda W)^\alpha$.

We write F^{ST} for the map $T^l(\mathbb{R}^d,\mathbb{R}) \to R^n,\ s \mapsto \sum_{|\alpha|\leq l} s^\alpha \sigma^\alpha$ (i.e. the F^{ST} map relative to the system (x,σ,Id).). We also have the map $F^{\log(S)T} = F^{ST} \circ \exp$. Writing $S^\alpha = (\lambda W)^\alpha$, we take $\mathrm{Pr}_{\mathcal{S}}(\log(S)) \in \mathcal{S}$ as our auxiliary random variable. S is the solution to the following system:

$$\begin{aligned}
&S_0 = 0, \\
&\log(S) = (t, \mathrm{Pr}_{\mathcal{S}}(S)), \\
&S \in T^l(\mathbb{R}^d,\mathbb{R}),
\end{aligned}$$

$$
\begin{aligned}
&dS^{(i,\alpha)} = S^{\alpha}_t \lambda(X_t, Y_t) \circ dW^i_t,\\
&(X, Y) \in \mathbb{R}^n \otimes \mathbb{R}^m,\\
&(X_0, Y_0) = (x, y),\\
&dY_t = \sum_{i=0}^{D} \Sigma(X_t, Y_t)^i \circ dW^i_t,\\
&dX_t = \sum_{j=1}^{d}\sum_{i=1}^{d} \sigma(X_t)^j \lambda(X_t, Y_t)^i_j \circ dW^i_t + b(X_t, Y_t)dt.
\end{aligned}
$$

As usual, $\mathcal{S}$ is endowed with a family of dilations δ_s. We can now write for any $s, t \leq \rho$ for some fixed constant ρ, $\log(\tilde{S}_s) = \delta_{\sqrt{t/s}}(\log(S)_s)$ as the solution at time t to the following de-scaled system, where $\tilde{W}$ is a de-scaled Brownian motion $\tilde{W}_t = \sqrt{t/s}W_{t/s}$, with target random variable $\mathrm{Pr}_{\mathcal{S}}(\tilde{S})$. Then

$$
\begin{aligned}
S_0 &= 0\\
\log(\tilde{S}) &= (u, \mathrm{Pr}_{\mathcal{S}}(\tilde{S}))\\
\tilde{S} &\in T^l(\mathbb{R}^d, \mathbb{R})\\
dS^{(i,\alpha)}_u &= \tilde{S}^{\alpha}_u \lambda(X_{u/\sqrt{t/s}}, Y_{u/\sqrt{t/s}}) \circ d\tilde{W}_u\\
&= \tilde{S}^{\alpha}_u \lambda(\tilde{X}_u, \tilde{Y}_u) \circ d\tilde{W}_u\\
(\tilde{X}, \tilde{Y}) &\in \mathbb{R}^n \otimes \mathbb{R}^m\\
(\tilde{X}_0, \tilde{Y}_0) &= (x, y)\\
d\tilde{Y}_u &= \sum_{i=0}^{D} \frac{1}{\sqrt{t/s}} \Sigma(\tilde{X}_u, \tilde{Y}_u)^i \circ d\tilde{W}^i_u\\
d\tilde{X}_t &= \sum_{j=1}^{d}\sum_{i=1}^{d} \sigma(\tilde{X}_u)^j \lambda(\tilde{X}_u, \tilde{Y}_u)^i_j \circ d\frac{1}{\sqrt{t/s}}\tilde{W}^i_u + b(\tilde{X}_u, \tilde{Y}_u)\frac{1}{\sqrt{t/s}}du.
\end{aligned}
$$

By our assumptions and Lemma 8.1.1, this system is $(L, g, \tilde{G})$-tense for a proper constant $\tilde{G}$. By Proposition 8.1.2, inside a ball of proper radius ρ, the system is also uniformly (L, H'_L) weak Hörmander for some proper constant H'_L. It follows that we are in a position to apply Theorem 8.1.1 to conclude that there exist proper constants C_1, C_2, C_3 such that for any $z \in \mathcal{S}$ and $t \in \mathbb{R}+$ such that $t, |z|^2_{\mathcal{S}} \leq C_1$,

$$
\mathbb{E}(\phi(\sup_{s\leq t} |S|_{\mathcal{S}})\delta(\mathrm{Pr}_{\mathcal{S}}(S) = z)) \leq \frac{C_2 e^{-\frac{|S|^2_{\mathcal{S}}}{C_3 t}}}{t^{\nu/2}},
$$

where ν is the homogeneous dimension of $\mathcal{S}$ and $\phi(\cdot)$ is a localising function such that $\phi(z) = 0$ for $|z|_{\mathcal{S}} \geq \rho$ and $\phi(z) = 1$ for $|z|_{\mathcal{S}} \leq \rho/2$.

Just as in the proof of Theorem 8.2.1, this now allows us to use Theorem 5.1.1 to deduce the following bound on X_t: for all $t, \bar{x}$ with $t, d_t(x, \bar{x})^2 \leq C_1$,

$$
\mathbb{E}(\phi(\sup_{s\leq t} |S|_{\mathcal{S}})\delta(\bar{X}_t = z)) \leq C_2 \frac{e^{-\frac{d_t(x,\bar{x})}{C_3 t}}}{|B_{d_t}(x, \sqrt{t})|} + C_4 t
$$

where $\bar{X}_t$ is the stochastic Taylor approximation described above, and C_1, C_2, C_3, C_4 are proper constants. As mentioned in the beginning of the proof, the result now follows using same arguments as in the proof of Theorem 8.2.1. □

We now have the immediate consequence of the above together with Theorem 7.3.1, which is the result we will use for the proof of the Löcherbach theorem under NDI:

THEOREM 13.2.2. *Let $\mathcal{A} = ((x,y),(\sigma,\Sigma),\mathrm{Pr})$ be a K_λ-composite system and $\mathcal{G}$ be a localising random variable such that σ satisfies the weak Hörmander condition with constants (L, H_L) uniformly in the compact $B(x,R)$ for some $R > 0$, and such that there exists a constant C such that for all $0 \leq k \leq d$, $0 \leq \bar{k} \leq D$ $N \leq (2n+2)^2 2^{4L} + n + 3$, $\alpha \in \mathrm{Multi}(\{1,2,\ldots,n\})$, $\#(\alpha) = N$, $\beta \in \mathrm{Multi}(\{1,2,\ldots,m\})$, $\#(\beta) = N$, and for any $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$,*

$$|\frac{\partial^N \sigma^k}{\partial x^{\alpha_1}\partial x^{\alpha_2}\ldots\partial x^{\alpha_N}}(x)| \leq C$$
$$|\frac{\partial^N \Sigma^k}{\partial y^{\alpha_1}\partial y^{\alpha_2}\ldots\partial y^{\alpha_N}}(y)| \leq C, \qquad \textit{and}$$
$$|D^N(\mathcal{G})| \leq C.$$

Then there exit constants C_1, C_2, C_3, polynomial in K_λ, m, D, C, such that for any $t \leq C_1$,

$$p_t(x,\bar{x}) \leq C_2 \frac{e^{-\frac{d_t(x,\bar{x})^2}{C_3 t}}}{|\overleftarrow{B}_{d_t}(\bar{x},\sqrt{t})|}.$$

PROOF. This result follows immediately from the previous theorem, and Theorem 7.3.1, as the assumptions imply that the tension of order $(2n+2)^2 2^{4L} + n + 3$ is polynomial. ☐

CHAPTER 14

On the absolute continuity of $\bar{m}$

Here, we use the results from the previous chapter to prove that the measure

$$\bar{m}(A) = \frac{\mathbb{E}(\int_0^R \eta_t(A)dt)}{\mathbb{E}(R)}$$

admits a continuous density. The method is similar to the proof in [**42**], with just the density estimate being different.

Let $T_1, T_2, \ldots$ be the jump times of the whole configuration. We write $\mathcal{L}_t^i$ for the probability distribution at time t of the i^{th} particle and $\mathcal{L}_{t,-}^i$ for the probability distribution of the i^{th} particle immediately before t. We also write $\alpha(x,y) = c + \sum_i \kappa(i,x,y)$. As in [**42**], there are two steps in the proof: First, we show that $\mathcal{L}_{T_N}^i$ admits a continuous and bounded density $g_{T_N}^i(\cdot)$ for all N (and i), with an upper bound that is polynomial in N. Then, we use this to show that $\bar{m}$ admits a continuous and bounded density.

PROPOSITION 14.0.1. *Assuming either*

- *the elliptic (E) version of our assumptions, or,*
- *the weak Hörmander version, the NDI condition, and a strongly polynomial version of Theorem 4.4.1,*

$\mathcal{L}_{T_N}^i$ admits a continuous and bounded density $g_{T_N}^i(\cdot)$ for all N (and i). Further more, there exists a function $\mathcal{G} : \mathbb{N} \to \mathbb{R}$, of polynomial growth, such that for any $u \in \mathbb{R}^n$,

$$g_{T_N}^i(u) \leq \mathcal{G}(N).$$

PROOF. We prove the result by induction. The cases $N = 0$ and $N = 1$ are trivial. Indeed, $\mathcal{L}_{T_1}^1 = \pi$, and in particular, $\mathcal{L}_{T_1}^1$ admits a density $g_{T_1}^1(\cdot) = \pi(\cdot)$, which is bounded by assumption.

Now, suppose that the result is established for some value of N, and we have $g_{T_N}^i \leq D$ for some D polynomial in N. We have, for any $u \in \mathbb{R}^n$, any approximation r_u of the delta function δ_u, and any $\epsilon > 0$,

$$\begin{aligned}
(14.0.1)\quad \mathbb{E}(r_u(\xi_{T_{N+1},-}^i)) &= \mathbb{E}\left(\int_0^\infty r_u(\xi_{T_N+t}^i)e^{-\int_0^t \alpha(\xi_{T_N+s},\zeta_{T_N+s})ds}\alpha(\xi_{T_N+t},\zeta_{T_N+t})dt\right) \\
&= \mathbb{E}\left(\int_0^\epsilon r_u(\xi_{T_N+t}^i)e^{-\int_0^t \alpha(\xi_{T_N+s},\zeta_{T_N+s})ds}\alpha(\xi_{T_N+t},\zeta_{T_N+t})dt\right) \\
&+ \mathbb{E}\left(\int_\epsilon^\infty r_u(\xi_{T_N+t}^i)e^{-\int_0^t \alpha(\xi_{T_N+s},\zeta_{T_N+s})ds}\alpha(\xi_{T_N+t},\zeta_{T_N+t})dt\right).
\end{aligned}$$

We use the following shortcuts:

$$\mathcal{G} = e^{-\int_0^t \alpha(\xi_{T_N+s},\zeta_{T_N+s})ds}\alpha(\xi_{T_N+t},\zeta_{T_N+t})$$

$$\bar{\mathcal{G}} = \alpha(\xi_{T_N+t},\zeta_{T_N+t})$$

Now, by Theorem 13.2.2, there exist constants C_1, C_2, C_3, polynomial in N indeed, note that $l + \mathcal{F}(l)$ is a polynomial constant in N by the assumptions on $\mathcal{F}$ and the reproduction law p) such that for any $t \leq C_1$, any $u \in \mathbb{R}^n$ and any $x, y \in \mathbb{R}^{n \times l(\eta_{T_N})} \otimes \mathbb{R}^{n \times \mathcal{F}(l(\eta_{T_N}))}$ the perturbed density

$$p_t^{\mathcal{G}}(x,y,u) = \mathbb{E}_{(x,y)}(\delta_u(\xi_t^i)e^{-\int_0^t \alpha(\xi_{T_N+s},\zeta_{T_N+s})ds}\alpha(\xi_{T_N+t},\zeta_{T_N+t})dt)$$

satisfies both

$$p_t^{\bar{\mathcal{G}}}(x,y,u) \leq C_2 \frac{e^{-\frac{|x^i-u|^2}{C_3 t}}}{t^{\nu/2}}$$

(for some constant ν independent of N, l, etc),
and either:

$$p_t^{\mathcal{G}}(x,y,u) \leq C_2 \frac{e^{-\frac{|x^i-u|^2}{C_3 t}}}{t^{n/2}} \tag{14.0.2}$$

(under the ellipticity assumption E),
or:

$$p_t^{\mathcal{G}}(x,y,u) \leq C_2 \frac{e^{-\frac{d_t(x^i,u)^2}{C_3 t}}}{B_{d_t}(x^i,\sqrt{t})}. \tag{14.0.3}$$

(under the assumptions NDI and H). Here d_t is the time-dependent 'distance' associated with the vector fields $\underline{\sigma}$. In particular, d_t doesn't depend on N, i, x, y, l, etc.

That means we can use Theorem 7.3.1 to turn Eq. (14.0.3) into :

$$p_t^{\mathcal{G}}(x,y,u) \leq C_2 \frac{e^{-\frac{d_t(x^i,u)^2}{C_3 t}}}{\overleftarrow{B}_{d_t}(u,\sqrt{t})} \tag{14.0.4}$$

whilst keeping the constants C_1, C_2, C_3 polynomial in N.

Next, fix $\Xi < 1$ and then pick $\epsilon_N > 0$ such that $\epsilon_N \leq C_1$ and

$$\int_0^{\epsilon_N} \int_{v\in\mathbb{R}^n} C_2 \frac{e^{-\frac{|v-u|^2}{C_3 t}}}{t^{n/2}} dvdt \leq \Xi / \sup_{h=0,1,\ldots,N} (h + \mathcal{F}(h)), \quad \text{and} \tag{14.0.5}$$

$$\int_0^{\epsilon_N} \int_{v\in\mathbb{R}^n} C_2 \frac{e^{-\frac{d_t(v,u)^2}{C_3 t}}}{\overleftarrow{B}_{d_t}(u,\sqrt{t})} dvdt \leq \Xi / \sup_{h=0,1,\ldots,N} (h + \mathcal{F}(h)).$$

Note that ϵ_N is still a polynomial quantity in N.

Going back to Eq. (14.0.1), we obtain:

$$\begin{aligned}
\mathbb{E}(r_u(\xi^i_{T_{N+1},-})) &= \mathbb{E}\left(\int_0^{\epsilon_N} r_u(\xi^i_{T_N+t}) e^{-\int_0^t \alpha(\xi_{T_N+s},\zeta_{T_N+s})ds} \alpha(\xi_{T_N+t},\zeta_{T_N+t})dt\right) \\
&\quad + \mathbb{E}\left(\int_{\epsilon_N}^{\infty} r_u(\xi^i_{T_N+t}) e^{-\int_0^t \alpha(\xi_{T_N+s},\zeta_{T_N+s})ds} \alpha(\xi_{T_N+t},\zeta_{T_N+t})dt\right) \\
&\leq D\Xi / \sup_{h=0,1,\ldots,N} (h+\mathcal{F}(h)) \\
&\quad + \mathbb{E}\left(\int_{\epsilon_N}^{\infty} r_u(\xi^i_{T_N+t}) e^{-\int_0^t \alpha(\xi_{T_N+s},\zeta_{T_N+s})ds} \alpha(\xi_{T_N+t},\zeta_{T_N+t})dt\right) \\
&\qquad \text{by either (14.0.2) or (14.0.4), and by (14.0.5)} \\
&\leq D(N)\Xi + C_2 \frac{1}{\epsilon_N^{\nu/2}} \int_{\epsilon_N}^{\infty} e^{-ct} dt \\
&\leq D(N) + \mathcal{R}(N).
\end{aligned}$$

Letting r tend to the delta function, we obtain that $\mathcal{L}^i_{T_{N+1},-}$ has density $g^i_{T_N,-}$ such that

$$g^i_{T_{N+1},-} \leq D(N+1) = D(N)\Xi + \mathcal{R}(N).$$

Because the density of a particle η^i at time $T_{N+1,+}$ is either taken from the density of η^i at time $T_{N+1,-}$ (if T_{N+1} is a branching event), or taken from π (if T_{N+1} is an immigration event), we can conclude

$$g^i_{T_{N+1},+} \leq D(N+1) = D(N)\Xi + \mathcal{R}(N) \tag{14.0.6}$$

where $\mathcal{R}$ is a polynomial function. Iterating (14.0.6), we get, writing $\mathcal{R}(1)$ for $D(1)$,

$$g^i_{T_{N+1},+} = \sum_{k=1}^{N} \Xi^{N-k}\mathcal{R}(k) \leq \frac{\sup_{k=1,2,\ldots,N}\mathcal{R}(k)}{1-\Xi},$$

which is indeed polynomial in N. □

Using the above result, we can prove the main result of this part of the thesis:

PROPOSITION 14.0.2. *Assuming either*

- *the elliptic (E) version of our assumptions, or,*
- *the weak Hörmander version, the NDI condition, and a strongly polynomial version of Theorem 4.4.1,*

the measure $\bar{m}$ admits a continuous and bounded Lebesgue density g.

PROOF. Write

$$\bar{g}^i_N = \mathbb{E}(\delta_u(\xi^i_{T_{N+1},-})) = \mathbb{E}\left(\int_0^\infty \delta_u(\xi^i_{T_N+t})e^{-\int_0^t \alpha(\xi_{T_N+s},\zeta_{T_N+s})ds}\right),$$

and picking the same ϵ_N as in the proof of 14.0.1, we have, similarly to 14.0.1,

$$\bar{g}^i_N \leq D\Xi/\sup_{h=0,1,\ldots,N}(h+\mathcal{F}(h)) + \mathcal{R}(N).$$

Next, we have

$$\begin{aligned}
\sum_i \bar{g}^i_N &\leq \frac{D\Xi}{\sup_{h=0,1,\ldots,N}(h+\mathcal{F}(h))}\sup_{h=0,1,\ldots,N}(h+\mathcal{F}(h)) + \mathcal{R}(N)\sup_{h=0,1,\ldots,N}(h+\mathcal{F}(h)) \\
&\leq \Xi D(N) + \mathcal{R}(N)\sup_{h=0,1,\ldots,N}(h+\mathcal{F}(h)) \\
&\leq \sum_{k=1}^{N}\Xi^{N-k}\mathcal{R}(k) \leq \frac{\sup_{k=1,2,\ldots,N}\mathcal{R}(k)}{1-\Xi} \qquad \text{(by iterating).}
\end{aligned}$$

Now note that, writing $\bar{m}_N$ for the measure with density $\bar{g}^i_N$, we have

$$\begin{aligned}
\bar{m} &= \frac{1}{\mathbb{E}(R)}\sum_{N=0}^{\infty}\mathbb{P}(T_N < R)\bar{m}_N \\
&< \frac{1}{\mathbb{E}(R)}\sum_{N=0}^{\infty} C_{\text{subcriticality}}e^{-c_{\text{subcriticality}}N}\bar{m}_N \\
&< b\sum_{N=0}^{\infty} C_{\text{subcriticality}}e^{-c_{\text{subcriticality}}N}\bar{m}_N,
\end{aligned}$$

where $C_{\text{subcriticality}}, c_{\text{subcriticality}}$ are the constants from Lemma 12.3.1.

Writing $\bar{m}_{N\uparrow} = \sum_{h=0}^{N}\mathbb{P}(T_N < R)\bar{m}_h$, we have that $\bar{m}_{N\uparrow}$ has density

$$g_{N\uparrow} \leq \sum_{h=0}^{N} bC_{\text{subcriticality}}e^{-c_{\text{subcriticality}}h}\frac{\sup_{k=1,2,\ldots,N}\mathcal{R}(k)}{1-\Xi}.$$

Because $\sup_{k=1,2,\ldots,N}\mathcal{R}(k)$ is polynomial in N, the above series converges as $N\to\infty$, and the theorem follows by dominated convergence. □

Bibliography

1. Gérard Ben Arous, *Développement asymptotique du noyau de la chaleur hypoelliptique sur la diagonale*, Annales de l'institut Fourier **39** (1989), no. 1, 73–99 (french).
2. Vlad Bally, *An elementary introduction to the Malliavin calculus*, (2003).
3. ______, *Lower bounds for the density of locally elliptic Itô processes*, Ann. Probab. **34** (2006), no. 6, 2406–2440. MR 2294988
4. Vlad Bally and Lucia Caramellino, *Positivity and lower bounds for the density of Wiener functionals*, ArXiv e-prints (2010).
5. ______, *Riesz transform and integration by parts formulas for random variables*, Stochastic Process. Appl. **121** (2011), no. 6, 1332–1355. MR 2794979
6. ______, *Tubes estimates for diffusion processes under a local Hörmander condition of order one*, arXiv:1202.4771 (2012).
7. ______, *On the distance between probability density functions*, ArXiv e-prints (2013).
8. Vlad Bally, Lucia Caramellino, and Paolo Pigato, *Diffusions under a local strong Hörmander condition. part I: density estimates*, arXiv:1607.04542 (2016).
9. ______, *Diffusions under a local strong Hörmander condition. part II: tube estimates*, arXiv:1607.04544 (2016).
10. Vlad Bally and Eva Löcherbach, *On the invariant density of branching diffusions*, arXiv:1607.04542 (2001).
11. Fabrice Baudoin, *Rough paths theory*.
12. Horatio Boedihardjo, Xi Geng, Terry Lyons, and Danyu Yang, *The signature of a rough path: uniqueness*, Adv. Math. **293** (2016), 720–737. MR 3474333
13. Marco Bramanti, Luca Brandolini, Ermanno Lanconelli, and Francesco Uguzzoni, *Non-divergence equations structured on Hörmander vector fields: heat kernels and Harnack inequalities*, Mem. Amer. Math. Soc. **204** (2010), no. 961, vi+123. MR 2604962
14. J. C. Butcher, *Trees and numerical methods for ordinary differential equations*, Numerical Algorithms **53** (2010), no. 2, 153–170.
15. Patrick Cattiaux and Laurent Mesnager, *Hypoelliptic non-homogeneous diffusions*, Probab. Theory Related Fields **123** (2002), no. 4, 453–483. MR 1921010
16. Kuo-Tsai Chen, *Iterated path integrals*, Bull. Amer. Math. Soc. **83** (1977), no. 5, 831–879.
17. Mu Fa Chen and Xian Yin Zhou, *Applications of Malliavin calculus to stochastic differential equations with time-dependent coefficients*, Acta Math. Appl. Sinica (English Ser.) **7** (1991), no. 3, 193–216. MR 1132060
18. Ilya Chevyrev and Andrey Kormilitzin, *A primer on the signature method in machine learning*, CoRR **abs/1603.03788** (2016).
19. Dan Crisan, Christian Litterer, and Terry Lyons, *Kusuoka–stroock gradient bounds for the solution of the filtering equation*, Journal of Functional Analysis **268** (2015), no. 7, 1928–1971 (English).
20. François Delarue and Stéphane Menozzi, *Density estimates for a random noise propagating through a chain of differential equations*, J. Funct. Anal. **259** (2010), no. 6, 1577–1630. MR 2659772
21. Peter K. Friz and Nicolas B. Victoir, *Multidimensional stochastic processes as rough paths*, Cambridge U. P., Cambridge, 2010.
22. Xi Geng, *Reconstruction for the signature of a rough path*, arXiv:1508.06890 (2016).
23. Benjamin Graham, *Sparse arrays of signatures for online character recognition*, CoRR **abs/1308.0371** (2013).
24. M. Gubinelli, *Ramification of rough paths*, ArXiv Mathematics e-prints (2006).
25. Lajos Gergely Gyurkó, Terry Lyons, Mark Kontkowski, and Jonathan Field, *Extracting information from the signature of a financial data stream*, arXiv:1307.7244 (2013).
26. M. Hairer, *A theory of regularity structures*, Inventiones Mathematicae **198** (2014), 269–504.
27. B. Hambly and T. Lyons, *Uniqueness for the signature of a path of bounded variation and the reduced path group*, ArXiv Mathematics e-prints (2005).
28. David P. Herzog and Jonathan C. Mattingly, *A practical criterion for positivity of transition densities*, Nonlinearity **28** (2015), no. 8, 2823–2845. MR 3382587
29. R. Höpfner, M. Hoffmann, and E. Löcherbach, *Non-parametric estimation of the death rate in branching diffusions*, Scandinavian Journal of Statistics **29** (2002), no. 4, 665–692.
30. R. Höpfner, E. Löcherbach, and M. Thieullen, *Strongly degenerate time inhomogeneous sdes: Densities and support properties. application to hodgkin–huxley type systems*, Bernoulli **23** (2017), no. 4A, 2587–2616.
31. Y. Inahama, *Rough path theory and stochastic calculus*, ArXiv e-prints (2016).

32. Yuzuru Inahama, *Large deviation principle for certain spatially lifted gaussian rough path*, Tohoku Math. J. (2) **67** (2015), no. 3, 433–463.
33. V. Jurdjevic and I. Kupka, *Polynomial control systems.*, Mathematische Annalen **272** (1985), 361–368.
34. Arturo Kohatsu-Higa, *Lower bounds for densities of uniformly elliptic random variables on wiener space*, Probability Theory and Related Fields **126** (2003), no. 3, 421–457.
35. Shigeo Kusuoka, *Malliavin calculus revisited*, J. Math. Sci. Univ. Tokyo **10** (2003), no. 2, 261–277. MR 1987133
36. Shigeo Kusuoka and Daniel W. Stroock, *Applications of the Malliavin calculus. II*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. **32** (1985), no. 1, 1–76. MR 783181
37. ———, *Applications of the Malliavin calculus. III*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. **34** (1987), no. 2, 391–442. MR 914028
38. ———, *Precise asymptotics of certain Wiener functionals*, J. Funct. Anal. **99** (1991), no. 1, 1–74. MR 1120913
39. R. Léandre, *Uniform upper bounds for hypoelliptic kernels with drift*, Kyoto Journal of Mathematics **34** (1994), no. 2, 263–271.
40. Antoine Lejay, *An Introduction to Rough Paths*, Seminaire de Probabilités XXXVII (J. Azéma; M. Ledoux; M. Emery; M. Yor, ed.), vol. 1832, Lecture Notes in Mathematics, no. 1832, Springer Berlin / Heidelberg, 2003, pp. 1–59.
41. C. Litterer and T. Lyons, *High order recombination and an application to cubature on wiener space*, The Annals of Applied Probability **22** (2012), no. 4, 1301–1327.
42. Eva Löcherbach, *Smoothness of the intensity measure density for interacting branching diffusions with immigrations*, J. Funct. Anal. **215** (2004), no. 1, 130–177. MR 2085113
43. T. Lyons and Z. Qian, *System control and rough paths*, Oxford mathematical monographs, Clarendon Press, 2002.
44. Terry Lyons, *Rough paths, signatures and the modelling of functions on streams*, arXiv:1405.4537 (2014).
45. Terry Lyons and Nicolas Victoir, *Cubature on wiener space*, Proceedings: Mathematical, Physical and Engineering Sciences **460** (2004), no. 2041, 169–198.
46. Annie Millet and Marta Sanz-Solé, *A simple proof of the support theorem for diffusion processes*, Séminaire de Probabilités, XXVIII, Lecture Notes in Math., vol. 1583, Springer, Berlin, 1994, pp. 36–48. MR 1329099
47. Alexander Nagel, Elias M. Stein, and Stephen Wainger, *Balls and metrics defined by vector fields i: Basic properties*, Acta Mathematica **155** (1985), no. 1, 103–147.
48. Andreas Neuenkirch, Ivan Nourdin, Andreas Roessler, and Samy Tindel, *Trees and asymptotic developments for fractional stochastic differential equations*, 46 pages, November 2006.
49. David Nualart, *The Malliavin calculus and related topics*, Probability and its Applications (New York), Springer-Verlag, New York, 1995. MR 1344217
50. Paolo Pigato, *Tube estimates for hypoelliptic diffusions and scaling properties of stochastic volatility models*, PhD Thesis (2015).
51. Jeremy Reizenstein, *Calculation of iterated-integral signatures and log signatures*, (2015).
52. C. Reutenauer, *Free lie algebras*, LMS monographs, Clarendon Press, 1993.
53. Setsuo Taniguchi, *Applications of Malliavin's calculus to time-dependent systems of heat equations*, Osaka J. Math. **22** (1985), no. 2, 307–320. MR 800974