Kusuoka-Stroock gradient bounds for the solution of the filtering equation

Dan Crisan* Christian Litterer† Terry Lyons‡

Abstract

We obtain sharp gradient bounds for perturbed diffusion semigroups. In contrast with existing results, the perturbation is here random and the bounds obtained are pathwise. Our approach builds on the classical work of Kusuoka and Stroock [12, 14, 15, 16], and extends their program developed for the heat semi-group to solutions of stochastic partial differential equations. The work is motivated by and applied to nonlinear filtering. The analysis allows us to derive pathwise gradient bounds for the un-normalised conditional distribution of a partially observed signal. It uses a pathwise representation of the perturbed semigroup following Ocone [21]. The estimates we derive have sharp small time asymptotics.

MSC 2010: 60H30 (60G35; 60H35; 93E11).

Keywords: Stochastic partial differential equation; Filtering; Zakai equation; Randomly perturbed semigroup, gradient bounds, small time asymptotics.

1 Introduction

In the eighties, Kusuoka and Stroock [12, 14, 15, 16] analysed the smoothness properties of the (perturbed) semigroup associated to a diffusion process. More precisely, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which we have defined a d_1-dimensional standard Brownian motion B and $X^x = \{X_t^x,\ t \geq 0\}$, $x \in \mathbb{R}^N$ be the stochastic flow

$$X_t^x = x + \int_0^t V_0(X_s^x)ds + \sum_{i=1}^{d_1} \int_0^t V_i(X_s^x) \circ dB_s^i, \quad t \geq 0, \tag{1}$$

where the vector fields $\{V_i, \ \ i = 0, ..., d_1\}$ are in $C_b^\infty(\mathbb{R}^N, \mathbb{R}^N)$, by which we mean that they are smooth and bounded with bounded derivatives of all orders, and the stochastic integrals in (1) are of Stratonovich type. The corresponding perturbed diffusion semigroup is then given by

*Department of Mathematics, Imperial College London, 180 Queen's Gate, London, SW7 2AZ, United Kingdom, d.crisan@imperial.ac.uk, Tel no: +442075948489.

†Centre de Mathématiques Appliquées, École Polytechnique, Route de Saclay, 91128 Palaiseau Cedex, France

‡University of Oxford and Oxford-Man Institute Eagle House, Walton Well Road, Oxford. OX2 6ED

$$(P_t^c\varphi)(x) = \mathbb{E}\left[\varphi(X_t^x)\exp\left(\int_0^t c\left(X_s^x\right)ds\right)\right], \quad t \geq 0, \quad x \in \mathbb{R}^N,$$

where $c \in C_b^\infty(\mathbb{R}^N)$ and $\varphi : \mathbb{R}^N \to \mathbb{R}$ is an arbitrary bounded measurable function. In the following, if $c = 0$ we will also write P_t in place of P_t^c. The vector fields $\{V_i, \quad i = 0, ..., d_1\}$ are assumed to satisfy Kusuoka's so-called UFG condition (see Definition 1 below). This condition essentially states that the $C_b^\infty(\mathbb{R}^N)$-module $\mathcal{W}$ generated by the vector fields $\{V_i, \quad i = 1, ..., d_1\}$ within the Lie algebra generated by $\{V_i, \quad i = 0, ..., d_1\}$ is finite dimensional. In particular, the condition does not require that the vector space $\{W(x)|W \in \mathcal{W}\}$ is isomorphic to $\mathbb{R}^N$ for all $x \in \mathbb{R}^N$. Hence, in this sense, the UFG condition is weaker than the uniform Hörmander condition.

Kusuoka and Stroock prove under the UFG condition that $P_t^c\varphi$ is differentiable in the direction of any vector field W belonging to $\mathcal{W}$. Moreover, they deduce sharp gradient bounds of the following form: Given vector fields $W_i \in \mathcal{W}$, $i = 1, ..., m+n$ there exist constants $C > 0$, $l > 0$ such that

$$\|W_1 \dots W_m P_t^c(W_{m+1} \dots W_{m+n}\varphi)\|_p \leq Ct^{-l}\|\varphi\|_p, \tag{2}$$

holds for any $\varphi \in C_0^\infty\left(\mathbb{R}^N\right)$, $t \in (0, 1]$ and $p \in [1, \infty]$ (see [16], Corollary 2.19 and [12], Theorem 2 under the UH and UFG condition respectively). The constant l depends explicitly on the vector fields W_i, $i = 1, ..., m+n$ and the small time asymptotics (2) are sharp. In this paper we deduce a similar result for the randomly perturbed semigroup. More precisely, let $Y = \{\left(Y_t^i\right)_{i=1}^{d_2}, t \geq 0\}$ be a d_2-dimensional standard Brownian motion independent of X, and define

$$\rho_t^{Y(\omega)}(\varphi)(x) = \mathbb{E}\left[\varphi(X_t^x)Z_t^x \,|\, \mathcal{Y}_t\right](\omega), \; t \geq 0, \; x \in \mathbb{R}^N, \tag{3}$$

where $Z^x = \{Z_t^x, t \geq 0\}$, $x \in \mathbb{R}^N$ is the stochastic process

$$Z_t^x = \exp\left(\sum_{i=1}^{d_2}\int_0^t h_i\left(X_s^x\right)dY_s^i - \frac{1}{2}\sum_{i=1}^{d_2}\int_0^t h_i\left(X_s^x\right)^2 ds\right), \; t \geq 0, \tag{4}$$

$h_i \in C_b^\infty(\mathbb{R}^N)$, $i = 1, ..., d_2$ and φ is an arbitrary bounded measurable function on $\mathbb{R}^N$. In the following, we prove that the mapping $x \longrightarrow \rho_t^{Y(\omega)}(\varphi)(x)$ has the property that there exists a $\mathbb{P}$-almost surely finite random variable $\omega \to C(\omega)$ such that

$$\|W_1 \dots W_m \rho_t^{Y(\omega)}\left(W_{m+1} \dots W_{m+n}\varphi\right)\|_p \leq C(\omega)\, t^{-l}\|\varphi\|_p, \tag{5}$$

for any $\varphi \in C_0^\infty\left(\mathbb{R}^N\right)$, $t \in (0, 1], p \in [1, \infty]$ and l being that same constant as in (2)

We are interested in this particular perturbation as it provides the Feynman-Kac representation for solutions of linear parabolic stochastic partial differential equations (SPDEs)[1].

[1] We expect the methodology presented here can be extended to handle a wider class of random perturbations. We chose this particular perturbation because the corresponding randomly perturbed semigroup provides the Feynman-Kac representation for the solution of the filtering problem. See the Kallianpur-Striebel formula (11) below.

To make this more precise, let $\rho^x = \{\rho^x_t,\ t \geq 0\}$, $x \in \mathbb{R}^N$ be the measure valued process defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ by the formula

$$\left(\rho^x_t(\omega)\right)(\varphi) = \rho_t^{Y(\omega)}(\varphi)(x), \tag{6}$$

where φ is an arbitrary Borel measurable function. Then ρ^x is the solution of the following linear parabolic SPDE (written here in its weak form):

$$\rho^x_t(\varphi) = \rho^x_0(\varphi) + \int_0^t \rho^x_s(A\varphi)ds + \sum_{k=1}^{d_2} \int_0^t \rho^x_s(h_k\varphi)dY^k_s \tag{7}$$

where $\rho^x_0 = \delta_x$ is the Dirac delta distribution centered at $x \in \mathbb{R}^N$, $A = V_0 + \frac{1}{2}\sum_{i=1}^{d_1} V_i^2$ is the infinitesimal generator of X, and φ is a suitably chosen test function. Equation (7) is called the Duncan-Mortensen-Zakai equation (cf. [8, 22, 23]). It plays a central rôle in nonlinear filtering: The normalized solution of (7) gives the conditional distribution of a partially observed stochastic process. We give details of this intrinsic connection in the second section.

As already stated, in this paper we study the mapping $x \longrightarrow \rho_t^{Y(\omega)}(\varphi)(x)$ for a fixed (Brownian) path $Y(\omega)$ and a suitably chosen test function φ. In [3], the authors look at the application $Y(\omega) \longrightarrow \rho_t^{Y(\omega)}(\varphi)(x)$ for a fixed $x \in \mathbb{R}^N$ and any suitably chosen test function φ and show that it is a (locally) Lipschitz continuous function as defined on the space of continuous paths[2].

Note that regularity properties for the non-linear filtering problem have previously been obtained by Kusuoka-Stroock [17] using the techniques of the partial Malliavin calculus, see also earlier work by Bismut-Michel [2] and subsequently Nualart, Zakai [20]. Our present approach frequently makes use of the fact that we are dealing with the uncorrelated filtering problem. Previous work using rough paths in the context of filtering [4], also [7] considers the setting where the noises in the signal and observation are correlated.

The paper is structured as follows: In Section 2 we introduce the filtering problem and explain the connection with the randomly perturbed semigroup (RPS). In section 3 we state the main results of the paper, that is, we introduce the corresponding sharp gradient bounds of the type (5) for the RPS. In addition, we also give direct corollaries on the smoothness properties of the solution of the filtering problem.

In Section 4, we derive an expansion of the RPS in terms of a classical perturbation series. The expansion is in terms of a series of (iterated) integrals with respect to the Brownian motion Y and is derived by exploiting the intrinsic connection between the RPS and the mild solution to the Zakai equation. We then proceed to prove the main theorem. The proof of the main theorem is contingent on two non-trivial regularity estimates for the terms appearing in the perturbation expansion of $\rho_t^{Y(\omega)}$ (Propositions 9 and 10), which we prove in the remainder of the paper.

[2]Here we consider the space of continuous paths defined on $[0,\infty)$ with values $\mathbb{R}^{d_2}$ endowed with the topology of convergence in the supremum norm on compacts. The choice of the norm is important. See [4] for further details.

In a first step towards proving these two propositions, in Section 5 we re-write the terms of the perturbation expansion iteratively using integration by parts to derive a pathwise representation of the RPS. We then prove a priori regularity estimates for the terms in the perturbation series in Subsection 5.2. For this, we state Hölder type regularity estimates for each term in the pathwise representation of the perturbation expansion. These estimates in turn are later proved in the appendix by carefully leveraging the gradient estimates for heat semi-groups due to Kusuoka and Stroock. Although these a priori estimates are asymptotically sharp for the lower order terms in the expansion, they are not summable.

Finally, in Section 6 we rely on both the a prior estimates derived in Section 5.2 and arguments underlying the Extension Theorem - a fundamental result from rough path theory (see, e.g. [18, 19]) - to deduce factorially decaying Hölder type bounds for the terms in the perturbation expansion. To this end, we observe that the terms of the original series (as derived in Section 4), when regarded as bounded linear operators between suitable spaces that encode the derivatives, are multiplicative functionals. Such multiplicative functionals are more general than ordinary rough paths but arise similarly for example also in the context of the work of Deya, Gubinelli, Tindel et al (see e.g. [6]) where they analyse rough heat equations. The paper is completed with an appendix containing several useful lemmas and the proof of the regularity estimates stated in Subsection 5.2.

Acknowledgements. We would like to thank the anonymous referee for carefully reading the paper and providing us with insightful comments that allowed us to improve the presentation of our results significantly. The work of D. Crisan and C. Litterer was partially supported by the EPSRC Grant No: EP/H0005500/1. The work of T. Lyons was partially supported by the EPSRC Grant No: EP/H000100/1.

2 The non-linear filtering problem

The nonlinear filtering problem is stated on the probability space $(\Omega, \mathcal{F}, \tilde{\mathbb{P}})$, where the new probability measure $\tilde{\mathbb{P}}$ is related to the probability measure $\mathbb{P}$ under which the triple[3] (X, Y, B) has been introduced in the previous section. More precisely, the probability measure $\tilde{\mathbb{P}}$ is absolutely continuous with respect to $\mathbb{P}$ and its Radon-Nikodym derivative is given by

$$\left.\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = Z_t, \quad t \geq 0,$$

where $Z = \{Z_t, t \geq 0\}$ is the exponential martingale defined in (4), that is,

$$Z_t = \exp\left(\sum_{i=1}^{d_2} \int_0^t h_i(X_s)\, dY_s^i - \frac{1}{2}\sum_{i=1}^{d_2}\int_0^t h_i(X_s)^2\, ds\right),\ t \geq 0.$$

[3]Throughout this section, we will omit the dependence on the initial condition $x \in \mathbb{R}^N$ for the processes X^x. The same appllies to all other processes (Z, W, ρ etc).

Under $\tilde{\mathbb{P}}$ the law of the process X remains the same as under $\mathbb{P}$. That is, under both $\mathbb{P}$ and $\tilde{\mathbb{P}}$, X satisfies the stochastic differential equation

$$dX_t = V_0(X_t)dt + \sum_{i=1}^{d_1} V_i(X_t) \circ dB_t^i, \quad X_0 = x \in \mathbb{R}^N, \quad t \geq 0. \tag{8}$$

Let $C_b^\infty(\mathbb{R}^N)$ denote the space of smooth bounded functions on $\mathbb{R}^N$ with bounded derivatives of all orders and $C_0^\infty(\mathbb{R}^N)$ the space of compactly supported smooth functions on $\mathbb{R}^N$.As in the previous section, we assume that the vector fields $\{V_i, \ i = 0, ..., d_1\}$ are smooth and bounded with bounded derivatives, i.e. $V_i = (V_i^j)_{j=1}^N$, where $V_i^j \in C_b^\infty(\mathbb{R}^N)$ for all $j = 1, ..., N$ and that the stochastic integrals in (8) are of Stratonovich type. We denote by π_0 the initial distribution of X and, from (8) we have that we that $\pi_0 = \delta_x$, $x \in \mathbb{R}^N$

The process Y is no longer a Brownian motion under $\tilde{\mathbb{P}}$, but becomes a semi-martingale. More precisely, Y satisfies the following evolution equation

$$Y_t = \int_0^t h(X_s)ds + W_t, \tag{9}$$

where W is a standard $\mathcal{F}_t$-adapted d_2-dimensional Brownian motion (under $\tilde{\mathbb{P}}$) independent of X. Let $\{\mathcal{Y}_t, \ t \geq 0\}$ be the usual filtration associated with the process Y, that is $\mathcal{Y}_t = \sigma(Y_s, \ s \in [0, t])$.

Within the filtering framework, the process X is called the *signal* process and the process Y is called the *observation* process. The filtering problem consists in determining π_t, the conditional distribution of the signal X at time t given the information accumulated from observing Y in the interval $[0, t]$, that is, for any φ Borel bounded function, computing

$$\pi_t(\varphi) = \mathbb{E}[\varphi(X_t) \mid \mathcal{Y}_t]. \tag{10}$$

The connection between π_t, the conditional distribution of X_t, and the randomly perturbed semigroup is given by the Kallianpur-Striebel formula. More precisely, we have

$$\pi_t(\varphi) = \frac{\rho_t^{Y(\omega)}(\varphi)}{\rho_t^{Y(\omega)}(\mathbf{1})}, \quad \mathbb{P} - \text{a.s.}, \tag{11}$$

where $\mathbf{1}$ is the constant function $\mathbf{1}(x) = 1$ for any $x \in \mathbb{R}^N$. Equivalently, the Kallianpur-Striebel formula can be stated as

$$\pi_t = \frac{1}{c_t}\rho_t \quad \mathbb{P} - \text{a.s.},$$

where ρ_t is the measure valued process which solves the Duncan-Mortensen-Zakai equation (7) and $c_t = \rho_t(\mathbf{1})$. The Kallianpur-Striebel formula explains the usage of the term unnormalised for ρ_t as the denominator $\rho_t(\mathbf{1})$ can be viewed as the normalizing factor for ρ_t. For further details of the filtering framework see, for example, [1] and the references therein.

3 The main theorems

Let $\mathbb{A}$ the set of multi-indices

$$\mathbb{A} = \{\emptyset,\ (\alpha_1, \ldots, \alpha_k),\ k \geq 1,\ a_j \in \{0, \ldots, d_1\}, j = 1, \ldots, k\}\,.$$

Following Kusuoka [12] we define a multiplication/concatenation operation on $\mathbb{A}$ by setting

$$\alpha * \beta = (\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_l)$$

for multi-indices $\alpha = (\alpha_1, \ldots, \alpha_k), \beta = (\beta_1, \ldots, \beta_l) \in \mathbb{A}$. Furthermore we define the degree of the multi-index $\alpha = (\alpha_1, \ldots, \alpha_k)$ by $\|\alpha\| = k + card(j : \alpha_j = 0)$. Let $A_0 = \mathbb{A} \setminus \{0\}$, $A_1 = \mathbb{A} \setminus \{\emptyset, (0)\}$ and $A_1(j) = \{\alpha \in \mathbb{A}_1 : \|\alpha\| \leq j\}$. We inductively define a family of vector fields indexed by A by taking

$$V_{[\emptyset]} = Id, \quad V_{[i]} = V_i, \quad 0 \leq i \leq d_1$$

$$V_{[\alpha * i]} = [V_{[\alpha]}, V_i], \quad 0 \leq i \leq d_1, \alpha \in A.$$

The following condition was introduced by Kusuoka and is weaker than the usual (uniform) Hörmander condition imposed on the vector fields defining the signal diffusion (see Kusuoka [12]).

Definition 1 *The family of vector fields V_i, $i = 0, \ldots, d_1$ is said to satisfy the UFG condition if there exists a positive integer k such that for all $\alpha \in A_1$ there exist $u_{\alpha,\beta} \in C_b^\infty(\mathbb{R}^N)$ satisfying*

$$V_{[\alpha]} = \sum_{\beta \in A_1(k)} u_{\alpha,\beta} V_{[\beta]}. \tag{12}$$

In essence, the UFG conditions states that, eventually, all higher order Lie brackets can be expressed as a linear combination Lie brackets of order k or lower. The uniform Hörmander condition implies the UFG condition, but not vice versa as we can see from the following example due to Kusuoka [12]:

Example 2 *Assume $d = 1$ and $N = 2$. Let V_0, V_1 be given by*

$$V_0(x_1, x_2) = \sin x_1 \frac{\partial}{\partial x_1} \qquad V_1(x_1, x_2) = \sin x_1 \frac{\partial}{\partial x_2}.$$

Then $\{V_0, V_1\}$ do not satisfy the Hörmander condition. However the UFG condition is satisfied with $k = 4$.

From now on, we assume that the family of vector fields V_i, $i = 0, \ldots, d_1$ satisfies the UFG condition. We will assume, in the following that $\bar{k}$ denote the *minimal* integer k for which condition (12) holds. We are ready to formulate the main theorem.

Theorem 3 *Suppose the family of vector fields V_i, $i = 0, \ldots, d_1$ satisfies the UFG condition. Let $m \geq j \geq 0$, $\alpha_1, \ldots, \alpha_j, \ldots, \alpha_m \in A_1(\bar{k})$ and $h \in C_b^\infty(\mathbb{R}^N)$. Then there exists a random variable $C(\omega)$ almost surely finite such that the randomly perturbed semigroup $\rho_t^{Y(\omega)}$ satisfies*

$$\left\| \left(V_{[\alpha_1]} \cdots V_{[\alpha_j]} \rho_t^{Y(\omega)} \left(V_{[\alpha_{j+1}]} \cdots V_{[\alpha_m]} \varphi \right) \right)(x) \right\|_\infty \leq C(\omega) t^{-l} \|\varphi\|_\infty$$

for any $\varphi \in C_b^\infty(\mathbb{R}^N)$, $t \in (0,1]$, where $l = (\|\alpha_1\| + \cdots + \|\alpha_m\|)/2$. If in addition $h \in C_0^\infty(\mathbb{R}^N)$ then there exists a random variable $C(\omega)$ almost surely finite such that

$$\left\| \left(V_{[\alpha_1]} \cdots V_{[\alpha_j]} \rho_t^{Y(\omega)} \left(V_{[\alpha_{j+1}]} \cdots V_{[\alpha_m]} \varphi \right) \right)(x) \right\|_p \leq C(\omega) t^{-l} \|\varphi\|_p \tag{13}$$

for all $\varphi \in C_0^\infty(\mathbb{R}^N)$, $t \in (0,1]$ and $p \in [1, \infty]$, where $l = (\|\alpha_1\| + \cdots + \|\alpha_m\|)/2$.

Remark 4 *The dependence of the constant $C(\omega)$ on the observation path $Y(\omega)$ in Theorem 3 can be made explicit in terms of a rough Hölder norm of Y. More precisely, let $\gamma \in (2,3)$ then for all $M > 0$ there exists C_M such that*

$$\left\| \left(V_{[\alpha_1]} \cdots V_{[\alpha_j]} \rho_t^{Y(\omega)} \left(V_{[\alpha_{j+1}]} \cdots V_{[\alpha_m]} \varphi \right) \right)(x) \right\|_\infty \leq C_M t^{-l} \|\varphi\|_\infty$$

for all $\varphi \in C_b^\infty(\mathbb{R}^N)$, $t \in (0,1]$ and

$$\omega \in \left\{ \omega : \|Y(\omega)\|_{RP(\gamma)} < M \right\},$$

where

$$\|Y(\omega)\|_{RP(\gamma)} := \sup_{0 \leq s \leq t \leq 1} \frac{|Y_t(\omega) - Y_s(\omega)|}{|t-s|^\gamma} + \max_{i,j \in \{1, \ldots d_2\}} \sup_{0 \leq s \leq t \leq 1} \frac{\left| \int_s^t \int_s^{t_2} dY_{t_1}^i(\omega) \, dY_{t_2}^j(\omega) \right|}{|t-s|^{2\gamma}}.$$

An analogous estimate holds for the bound in (13).

Before we begin the proof of our main theorem we explore some immediate consequences of the result. We first observe that we can obtain similar estimates for the conditional distribution.

Corollary 5 *Under the assumptions of Theorem 3, there exists a random variable $\bar{C}(\omega)$ almost surely finite such that the conditional distribution π_t satisfies*

$$\left\| \left(V_{[\alpha_1]} \cdots V_{[\alpha_j]} \pi_t \left(V_{[\alpha_{j+1}]} \cdots V_{[\alpha_m]} \varphi \right) \right)(x) \right\|_\infty \leq \bar{C}(\omega) t^{-l} \|\varphi\|_\infty$$

for any $\varphi \in C_b^\infty(\mathbb{R}^N)$, $t \in (0,1]$, where $l = (\|\alpha_1\| + \cdots + \|\alpha_m\|)/2$.

Proof. We show the results for $j = 1$, and $m = 2$, the general case being done by using the Leibniz rule for the n-th derivative. We have

$$\begin{aligned} V_{[\alpha_1]} \pi_t \left(V_{[\alpha_2]} \varphi \right)(x) &= V_{[\alpha_1]} \left[\rho_t^{Y(\omega)} \left(V_{[\alpha_2]} \varphi \right) / \rho_t^{Y(\omega)}(1) \right](x) \qquad (14) \\ &= V_{[\alpha_1]} \rho_t^{Y(\omega)} \left(V_{[\alpha_2]} \varphi \right) (\rho_t^{Y(\omega)}(1))^{-1}(x) + \rho_t^{Y(\omega)} \left(V_{[\alpha_2]} \varphi \right) V_{[\alpha_1]}((\rho_t^{Y(\omega)}(1))^{-1})(x). \end{aligned}$$

Since, almost surely, (see the Appendix for a proof)

$$\sup_{x\in\mathbb{R}^N}\left(1/\rho_t^{Y(\omega)}(1)\right)<\infty, \tag{15}$$

we deduce that, for any $x\in\mathbb{R}^N$, we have

$$\left|V_{[\alpha_1]}\pi_t\left(V_{[\alpha_2]}\varphi\right)(x)\right|\leq\bar{C}(\omega)\,t^{-(\|\alpha_1\|+\|\alpha_2\|)/2}\,\|\varphi\|_\infty\,,$$

where $\bar{C}(\omega)=C(\omega)\,((\rho_t^{Y(\omega)}(1))^{-1}+(\rho_t^{Y(\omega)}(1))^{-2})$. ■

Note that if the vector fields $V_i, i=0,...,d_1$ satisfy the UFG condition, we can in general not guarantee the existence of a density of the unnormalised conditional distribution of the signal with respect to the Lebesgue measure given any starting point. However, just as for the unperturbed diffusion semigroup, $\rho_t^{Y(\omega)}$ will have a density $y\to\bar{\rho}_t^{x,Y(\omega)}(y)$ with respect to the Lebesgue measure provided we assume that there exists a positive integer k such that for *any* vector field V with coefficients in $C_b^\infty(\mathbb{R}^N)$, there exist $u_{V,\beta}\in C_b^\infty(\mathbb{R}^N)$ satisfying

$$V=\sum_{\beta\in A_1(k)}u_{\alpha,\beta}V_{[\beta]}. \tag{16}$$

The above assumption is equivalent to the existence of a positive integer k such that for $i=1,...,N$, there exist $u_{i,\beta}\in C_b^\infty(\mathbb{R}^N)$ satisfying

$$\partial_i=\sum_{\beta\in A_1(k)}u_{i,\beta}V_{[\beta]}. \tag{17}$$

In particular this means that

$$\mathrm{Span}\{V_{[\alpha]}(x):\alpha\in\mathcal{A}_1(k)\}=\mathbb{R}^N$$

holds for all $x\in\mathbb{R}^N$. Following from [5], under condition (16), the law of the signal X_t^x has a smooth density $y\to p_t^x(y)$ with respect to the Lebesgue measure for all $t>0$.[4] Under this assumption, we can deduce also deduce gradient estimates for the density $y\to\bar{\rho}_t^{x,Y(\omega)}(y)$:

Corollary 6 *Assume that the vector fields $V_i, i=0,...,d_1$ satisfy condition (16) and that $\pi_0=\delta_x$ is the Dirac measure at x and $h\in C_b^\infty(\mathbb{R}^N)$ Then, for all $t>0$, the unnormalised conditional distribution of the signal $\rho_t^{Y(\omega)}$ has a smooth density $y\to\bar{\rho}_t^{x,Y(\omega)}(y)$ with respect to the Lebesgue measure. Moreover for any multi-index $\iota=(i_1,...,i_n)\in\{1,\ldots,N\}^n$ there exists a random variable $\bar{C}_\iota(\omega)$ almost surely finite such that*

$$\left\|\partial_{i_1}...\partial_{i_n}\bar{\rho}_t^{x,Y(\omega)}\right\|_1\leq\bar{C}_\iota(\omega)\,t^{-\frac{kn}{2}},\quad t\in(0,1]. \tag{18}$$

If in addition $h\in C_0^\infty\left(\mathbb{R}^N\right)$ then for any multi-index $\iota=(i_1,...,i_n)\in\{1,\ldots,N\}^n$ and any $p\in[1,\infty]$, there exists a random variable $\bar{C}_\iota(\omega)$ almost surely finite such that

$$\left\|\partial_{i_1}...\partial_{i_n}\bar{\rho}_t^{x,Y(\omega)}\right\|_p\leq\bar{C}_\iota(\omega)\,t^{-\frac{kn}{2}},\quad t\in(0,1]. \tag{19}$$

[4] See [5] for the connection between condition (16) and the uniform Hörmander condition and the corresponding extensions for the smoothness results.

Proof. As already stated, following from [5], under condition (16), the law of the signal X_t^x has a density $y \to p_t^x(y)$ with respect to the Lebesgue measure for all $t > 0$. Moreover from the definition (6) of the measure $\rho_t^{Y(\omega)}$ in terms of the randomly perturbed semigroup (3) it follows that $\rho_t^{Y(\omega)}$ is absolutely continuous with respect to the law of the signal X_t^x and its density is given by the function $y \to \Psi_t^x(y)$ defined as

$$\Psi_t^x(y) = \tilde{\mathbb{E}}[Z_t^x | X_t = y, \mathcal{Y}_t]$$

and called the likelihood function in the context of stochastic filtering. Therefore, the unnormalised conditional distribution of the signal $\rho_t^{Y(\omega)}$ has, indeed, a density $y \to \bar{\rho}_t^{x,Y(\omega)}(y)$ with respect to the Lebesgue measure and $\bar{\rho}_t^{x,Y(\omega)}(y) = \Psi_t^x(y) p_t^x(y)$ for all $y \in \mathbb{R}^N$. In particular,

$$\rho_t^{Y(\omega)}(\varphi) = \int_{\mathbb{R}^N} \varphi(y) \Psi_t^x(y) p_t^x(y)\, dy,$$

for any bounded measurable test function φ. From Theorem 3 we then deduce that for any multi-index $\iota = (i_1, ..., i_n) \in \mathbb{A}$, there exists a random variable $C_\iota(\omega)$ almost surely finite such that

$$\left| \left(\rho_t^{Y(\omega)} (\partial_{i_1} ... \partial_{i_n} \varphi) \right) \right| \leq C_\iota(\omega)\, t^{-\frac{km}{2}} \|\varphi\|_\infty \tag{20}$$

for any $\varphi \in C_b^\infty(\mathbb{R}^N)$. The smoothness of $\bar{\rho}_t^{Y(\omega)}$ follows immediately as in [5] by classical results. The inequality (18) follows from (20) and the bound (19) follows in a similar manner. ■

4 Proof of the main theorem

As a first step in the proof of our main theorem we expand the unnormalised conditional distribution of the signal using its representation as the mild solution of the Zakai equation. This is a standard result, see for example [21] and [11]. For completeness we include it in Lemma 7 below. We define the set of operators $R_{\bar{t},\bar{\imath}}$, where $\bar{t} = (t_1, t_2, \ldots, t_k)$ is a non-empty multi-index with entries $t_0, t_1, \ldots, t_k \in [0, \infty)$ that have increasing values $t_0 < t_1 < ... < t_k$ and $\bar{\imath} = (i_1, ..., i_{k-1})$ is a multi-index with entries $i_1, ..., i_{k-1} \in \{1, 2, ..., d_2\}$ as follows

$$R_{(t_0,t_1),\varnothing}(\varphi) = P_{t_1 - t_0}(\varphi)$$

and, inductively, for $k > 1$,

$$\begin{aligned} R_{(t_0,t_1,t_2,\ldots,t_k),(i_1,\ldots,i_{k-1})}(\varphi) &= R_{(t_0,t_1,\ldots,t_{k-1}),(i_1,\ldots,i_{k-2})}\left(h_{i_{k-1}} P_{t_k - t_{k-1}}(\varphi)\right) \\ &= P_{t_1 - t_0}\left(h_{i_1} P_{t_2 - t_1} \ldots \left(h_{i_{k-1}} P_{t_k - t_{k-1}}(\varphi)\right)\right) \\ &= P_{t_1 - t_0}\left(h_{i_1} R_{(t_1,t_2,\ldots,t_k),(i_2,\ldots,i_{k-1})}(\varphi)\right). \end{aligned}$$

Note that the length of the multi-index $\bar{t}$ is always two units more than $\bar{\imath}$. Also let $S(m)$ to denote the set of all multi-indices

$$S(m) = \{(i_1, ..., i_m) \mid 1 \leq i_j \leq d_2, \quad 1 \leq j \leq m\}.$$

and let $S = \bigcup_{m=1}^\infty S(m)$.

Lemma 7 *Let ρ_t be the unnormalised conditional density defined in* (3) *and $\varphi \in C_b^\infty(\mathbb{R}^N)$. Then we have in L^2 and almost surely that*

$$\rho_t^{Y(\omega)}(\varphi) = P_t(\varphi) + \sum_{m=1}^{\infty} \sum_{\bar{\imath} \in S(m)} R_{0,t}^{m,\bar{\imath}}(\varphi) \tag{21}$$

where, for $\bar{\imath} = (i_1, ..., i_m)$,

$$R_{0,t}^{m,\bar{\imath}}(\varphi) = \underbrace{\int_0^t \int_0^{t_m} \cdots \int_0^{t_2}}_{m \quad \text{times}} R_{(0,t_1,\ldots,t_m,t),\bar{\imath}}(\varphi) dY_{t_1}^{i_1} \ldots dY_{t_m}^{i_m}.$$

Proof. The measure $\rho_t^{Y(\omega)}$ admits the following (mild) representation:

$$\rho_t^{Y(\omega)}(\varphi) = P_t(\varphi) + \sum_{i=1}^{d_2} \int_0^t \rho_s^{Y(\omega)}(h_i P_{t-s}(\varphi)) dY_s^i.$$

Arguing by induction it is easy to see that

$$\rho_t^{Y(\omega)}(\varphi)(x) = P_t(\varphi)(x) + \sum_{m=1}^{k} \sum_{\bar{\imath} \in S(m)} R_{0,t}^{m,\bar{\imath}}(\varphi) + \sum_{\bar{\imath} \in S(k+1)} \mathrm{Rem}_{0,t}^{k+1,\bar{\imath}}(\varphi),$$

where

$$\mathrm{Rem}_{0,t}^{k+1,\bar{\imath}}(\varphi) = \underbrace{\int_0^t \int_0^{t_{k+1}} \cdots \int_0^{t_2}}_{k+1 \quad \text{times}} \rho_{t_1}^{Y(\omega)}(h_{i_1} P_{t_2-t_1} h_{i_2} \cdots h_{i_{k+1}} P_{t-t_{k+1}}(\varphi))(x)\, dY_{t_1}^{i_1} \cdots dY_{t_{k+1}}^{i_{k+1}}.$$

Using iteratively Jensen's inequality and the Itô isometry we see that

$$\begin{aligned}
&\mathbb{E}\left[\mathrm{Rem}_{0,1}^{k+1,\bar{\imath}}(\varphi)^2\right] \\
&\le \int_0^1 \int_0^{t_{k+1}} \cdots \int_0^{t_2} \mathbb{E}\left[\rho_{t_1}^{Y(\omega)}(h_{i_1} P_{t_2-t_1} h_{i_2} \cdots h_{i_{k+1}} P_{t-t_{k+1}}(\varphi))^2\right] dt_1 \cdots dt_{k+1} \\
&\le e^{t||h||_\infty} \frac{||h||_\infty^{2(k+1)}}{(k+1)!} \|\varphi\|_\infty^2,
\end{aligned}$$

since, by Jensen's inequality

$$\mathbb{E}\left[\rho_{t_1}^{Y(\omega)}(h_{i_1} P_{t_2-t_1} h_{i_2} \cdots h_{i_{k+1}} P_{t-t_{k+1}}(\varphi))^2\right] \le ||h||_\infty^{2k+2}\, \mathbb{E}\left[(Z_t^x)^2\right] \le e^{t||h||_\infty} ||h||_\infty^{2(k+1)}.$$

hence $\mathrm{Rem}_{0,t}^{k+1,\bar{\imath}}(\varphi)$ converges to 0 as k tends to ∞. As the convergence is factorially fast a.s. convergence holds and the claim follows. ■

Before we can prove the main theorem we require three non-trivial estimates for the regularity of the terms appearing in the expansion (21) of $\rho_t^{Y(\omega)}(\varphi)$. The first is the aforementioned gradient estimate due Kusuoka and Stroock for the heat semi-group. The following theorem is due to Kusuoka-Stroock [16] (Corollary 2.19) under the uniform Hörmander condition and Kusuoka [12] (Theorem 2) under the UFG assumption.

Theorem 8 *Suppose the family of vector fields V_i, $i = 0, \ldots, d_1$ satisfies the UFG condition. Let $m \geq j \geq 0$, $\alpha_1, \ldots, \alpha_j, \ldots, \alpha_m \in A_1\left(\bar{k}\right)$ then there exists a constant C such that*

$$\left\| V_{[\alpha_1]} \cdots V_{[\alpha_j]} P_t \left(V_{[\alpha_{j+1}]} \cdots V_{[\alpha_m]} \varphi \right) \right\|_p \leq C t^{-(\|\alpha_1\| + \cdots + \|\alpha_m\|)/2} \|\varphi\|_p$$

for any $\varphi \in C_0^\infty\left(\mathbb{R}^N\right)$, $t \in (0, 1]$ and $p \in [1, \infty]$.

The second ingredient for the proof of the main theorem are the following regularity estimates for the terms $R_{0,t}^{m,\bar{\imath}}$.

Proposition 9 *Under the assumptions of Theorem 3, let $l \geq j \geq 0$, $\alpha_1, \ldots, \alpha_j, \ldots, \alpha_l \in A_1\left(\bar{k}\right)$ and $\gamma \in (1/3, 1/2)$. Then, for any $m \in \mathbb{N}$ there exist a random variable $C = C(\omega, m, l, \gamma) > 0$ almost surely finite such that*

$$\left\| V_{[\alpha_1]} \cdots V_{[\alpha_j]} R_{0,t}^{m,\bar{\imath}} (V_{[\alpha_{j+1}]} \cdots V_{[\alpha_l]} \varphi) \right\|_\infty \leq C t^{-(\|\alpha_1\| + \cdots + \|\alpha_l\|)/2 + m\gamma} \|\varphi\|_\infty$$

for all $\bar{\imath} \in S(m)$, $\varphi \in C_b^\infty(\mathbb{R}^N)$ and $t \in (0, 1]$.

The preceding proposition suggests that the short term asymptotics of the regularity of ρ_t are determined by the leading term of the expansion - the heat semi-group $P_t f$ itself. The estimate is unfortunately not summable in m and will therefore only be used to control the regularity of $R_{0,t}^{m,\bar{\imath}}$ for small m. Before we proceed we state a second set of a priori estimates that capture the regularity of the $R_{0,t}^{m,\bar{\imath}}$ in terms of operator norms on some carefully chosen spaces. These estimate do not lead to sharp short small time asymptotics and will therefore only be used to estimate the regularity of $R_{0,t}^{m,\bar{\imath}}$ for sufficiently large values of m.

To derive the second set of factorially decaying estimates we regard the $R_{0,t}^{m,\bar{\imath}}$ as linear operators acting on smooth functions endowed with suitable norms. Since both the heat kernels and the multiplication operators defined by the sensor functions h_i map $C_b^\infty(\mathbb{R}^N)$ functions to $C_b^\infty(\mathbb{R}^N)$ functions we see that $R_{0,t}^{m,\bar{\imath}}$ maps $C_b^\infty(\mathbb{R}^N)$ to $C_b^\infty(\mathbb{R}^N)$. We first define a distribution space appropriate for our problem. For $\varphi \in C_b^\infty(\mathbb{R}^N)$ let

$$\|\varphi\|_{H^{-1}} := \inf \left\{ \sum_{\alpha \in A_0\left(\bar{k}\right)} \|\varphi_\alpha\|_\infty : \varphi = \sum_{\alpha \in A_0\left(\bar{k}\right)} V_{[\alpha]} \varphi_\alpha, \ \varphi_a \in C_b^\infty(\mathbb{R}^N) \right\}.$$

Then $\|\cdot\|_{H^{-1}}$ defines a norm on $C_b^\infty(\mathbb{R}^N)$ that is bounded above by $\|\varphi\|_\infty$, but potentially smaller. Similarly we may define a Sobolev type norm on $C_b^\infty(\mathbb{R}^N)$ by letting

$$\|\varphi\|_{H^1} := \sum_{\alpha \in A_0\left(\bar{k}\right)} \left\| V_{[\alpha]} \varphi \right\|_\infty .$$

Recall in this context that the index set $A_0\left(\bar{k}\right)$ contains the empty set and we have set $V_{[\emptyset]} = Id$.

Proposition 10 *Under the assumptions of Theorem 3 there exist constants $\theta > 0$, $\gamma' \in (1/3, 1/2)$, $m_0 = m_0(\gamma') \in \mathbb{N}$ and a random variable $c(\gamma', \omega)$, almost surely finite, such that*

$$\left\| R_{0,t}^{m,\bar{\imath}} \right\|_{H^{-1}\to H^1} \leq \frac{(c(\gamma',\omega)\, t)^{m\gamma'}}{\theta\,(m\gamma')!}$$

for all $m \geq m_0$, $\bar{\imath} \in S(m)$ and $t \in (0,1]$.

The proofs of Proposition 9 and 10 are non-trivial and will be given in Section 5.2 and 6 respectively. Combining the previous estimates we are ready to prove our main theorem.

Proof of Theorem 3. We prove the result for $j = 1$, and $m = 2$, the general case being done by using the corresponding estimates for the higher order derivatives of the integral kernels and the higher order Sobolev and distribution spaces corresponding to H^1 and H^{-1} to accommodate higher order derivatives. For the first part of the theorem, we are going to show that there exists a positive random variable C almost surely finite such that

$$\left\| V_{[\alpha]} \rho_t^{Y(\omega)}(V_{[\beta]}\varphi) \right\|_\infty \leq C(\omega)\, t^{-(\|\alpha\|+\|\beta\|)/2} \left|\left|\varphi\right|\right|_\infty$$

for any $t \in (0,1]$ and $\varphi \in C_b^\infty(\mathbb{R}^N)$. Fix $\gamma \in (1/3, 1/2)$ and let γ', θ and m_0 as in Proposition 10. We have by Lemma 7

$$\begin{aligned} \left\| V_{[\alpha]} \rho_t^{Y(\omega)}(V_{[\beta]}\varphi) \right\|_\infty \leq{}& \left\| V_{[\alpha]} P_t(V_{[\beta]}\varphi) \right\|_\infty \\ &+ \sum_{k=1}^{m_0} \sum_{\bar{\imath}\in S(k)} \left\| V_{[\alpha]} R_{0,t}^{k,\bar{\imath}}\left(V_{[\beta]}\varphi\right) \right\|_\infty + \sum_{k=m_0+1}^{\infty} \sum_{\bar{\imath}\in S(k)} \left\| V_{[\alpha]} R_{0,t}^{k,\bar{\imath}}\left(V_{[\beta]}\varphi\right) \right\|_\infty . \end{aligned} \tag{22}$$

Now

$$\begin{aligned} \left\| V_{[\alpha]} R_{0,t}^{k,\bar{\imath}}\left(V_{[\beta]}\varphi\right) \right\|_\infty &\leq \left\| R_{0,t}^{k,\bar{\imath}}\left(V_{[\beta]}\varphi\right) \right\|_{H^1} \\ &\leq \left\| R_{0,t}^{k,\bar{\imath}} \right\|_{H^{-1}\to H^1} \left\| V_{[\beta]}\varphi \right\|_{H^{-1}} \\ &\leq \left\| R_{0,t}^{k,\bar{\imath}} \right\|_{H^{-1}\to H^1} \left\| \varphi \right\|_\infty . \end{aligned}$$

Therefore using Theorem 8 for the first, Proposition 9 for the second and Proposition 10 for the third term in the sum on the right hand side of (22) we see that

$$\begin{aligned} \left\| V_{[\alpha]} \rho_t^{Y(\omega)}(V_{[\beta]}\varphi) \right\|_\infty &\leq t^{-(\|\alpha\|+\|\beta\|)/2} \left|\left|\varphi\right|\right|_\infty + \sum_{k=1}^{m_0} c_k t^{-(\|\alpha\|+\|\beta\|)/2+k\gamma} \left|\left|\varphi\right|\right|_\infty \\ &\quad + \sum_{k=m_0+1}^{\infty} t^{k\gamma'} \frac{c(\gamma',\omega,d_2)^k}{\theta\,(k\gamma')!} \left|\left|\varphi\right|\right|_\infty \\ &\leq c(\omega)\, t^{-(\|\alpha\|+\|\beta\|)/2} \left|\left|\varphi\right|\right|_\infty \end{aligned}$$

where

$$c(\omega) = 1 + \sum_{k=1}^{m_0} c_k + \sum_{k=m_0+1}^{\infty} \frac{c(\gamma', \omega, d_2)^k}{\theta(k\gamma')!}.$$

Clearly, the constants and the parameter m_0 in equation (22) will depend on the number of derivatives required.

For the proof of the second part of the theorem, the general L^p estimate, we follow Kusuoka [12]. First observe that

$$\|\varphi\|_1 = \sup_{\{g \in C_0^\infty(\mathbb{R}^N),\ \ \|g\|_\infty \le 1\}} \left| \int_{\mathbb{R}^N} \varphi(x) g(x) dx \right|. \tag{23}$$

Let P_t^* be the (formal) adjoint operator of P_t, that is, let P_t^* be defined as

$$P_t^* \varphi(x) := E\left(\exp\left(\int_0^t \widetilde{c}\left(\widetilde{X}_s^x\right) ds \right) \varphi\left(\widetilde{X}_t^x\right) \right), \quad x \in \mathbb{R}^N,$$

where

$$\widetilde{c} = \operatorname{div}(V_0) + \frac{1}{2} \sum_{j=1}^d V_j (\operatorname{div}(V_j)) + \frac{1}{2} \sum_{j=1}^d (\operatorname{div}(V_j))^2$$

and $\widetilde{X}_t$ be the diffusion associated to the vector fields $\left(\widetilde{V}_0, V_1, \ldots, V_d\right)$ and

$$\widetilde{V}_0 = -V_0 + \frac{1}{2} \sum_{j=1}^d V_j (\operatorname{div}(V_j)).$$

Then P_t^* satisfies

$$\int P_t \varphi(x) g(x) dx = \int \varphi(x) P_t^* g(x) dx, \tag{24}$$

for any $\varphi, g \in C_0^\infty(\mathbb{R}^N)$ (see Kusuoka, Stroock [16] for a more general result).

By Lemma 7 we may write

$$\rho_t^{Y(\omega)} = P_t + \sum_{m=1}^{\infty} \sum_{\overline{\imath} \in S(m)} \int_{\Delta_{0,t}^m} P_{t_1} H_{i_1} P_{t_2 - t_1} H_{i_2} \cdots H_{i_m} P_{t - t_m} dY_{t_1}^{i_1} \cdots dY_{t_m}^{i_m},$$

where H_i are the (self-adjoint) multiplication operators corresponding the (compactly supported) h_i. Iteratively applying identity (24) to the expansion of $\rho_t^{Y(\omega)}$ to identify its formal adjoint ρ_t^* as

$$\rho_t^* = P_t^* + \sum_{m=1}^{\infty} \sum_{\overline{\imath} \in S(m)} \int_{\Delta_{0,t}^m} P_{t-t_m}^* H_{i_m} P_{t_m - t_{m-1}}^* H_{i_{m-1}} \cdots H_{i_1} P_{t_1}^* dY_{t_1}^{i_1} \cdots dY_{t_m}^{i_m}. \tag{25}$$

Using (23) and (25) we see that

$$\begin{aligned}\left\|V_{[\alpha]}\rho_t^x\left(V_{[\beta]}\varphi\right)(Y)\right\|_1 &= \sup_{\{g\in C_0^\infty(\mathbb{R}^N),\ \ \|g\|_\infty\le 1\}}\left|\int g(x)\,V_{[\alpha]}\rho_t^{Y(\omega)}\left(V_{[\beta]}\varphi(x)\right)dx\right| \\ &= \sup_{\{g\in C_0^\infty(\mathbb{R}^N),\ \ \|g\|_\infty\le 1\}}\left|\int V_{[\beta]}^*\rho_t^*\left(V_{[\alpha]}^*g(x)\right)\varphi(x)\,dx\right| \\ &\le \sup_{\{g\in C_0^\infty(\mathbb{R}^N),\ \ \|g\|_\infty\le 1\}}\left\|V_{[\beta]}^*\rho_t^*\left(V_{[\alpha]}^*g(x)\right)\right\|_\infty\|\varphi\|_1,\end{aligned}$$

where the formal adjoint of a vector field $V_{[\alpha]}$ is given by

$$V_{[\alpha]}^* = -V_{[\alpha]} - \sum_{i=1}^N \frac{\partial}{\partial x^i}V_{[\alpha]}^i.$$

The arguments in the proof of Proposition 9 can be easily extended to allow us to deduce the relevant estimates for the terms in the expansion (25). Extending the proof of Proposition 10 requires some small modifications that are discussed in Remark 20. Going through the steps in the proof of the first part of the proof with ρ_t^* in place of ρ_t we deduce that

$$\left\|V_{[\beta]}^*\rho_t^*\left(V_{[\alpha]}^*g\right)\right\|_\infty \le c(\omega)\,t^{-(\|\alpha\|+\|\beta\|)/2}\,||g||_\infty,$$

and the case of general $p\in[1,\infty]$ is a now consequence of classical Riesz-Thorin interpolation. ■

The proof of the main result is now complete. The remainder of the paper is dedicated to the proof of Propositions 9 and 10.

5 Pathwise representation of the perturbation expansion and some preliminary estimates

5.1 A pathwise perturbation expansion

For the first step towards a proof of Proposition 9 we derive the pathwise representation for the multiple stochastic integrals $R_{0,t}^{m,\bar{\imath}}(\varphi)$ as a sum of Riemann integrals with integrands depending on the Brownian motion Y. We will require the following notation. For $k\in\mathbb{N}$ let $\Delta_{s,t}^k$ denote the simplex defined by the relation $s<t_1<\cdots<t_k<t$ and let $d\bar{t}_k := dt_1\cdots dt_k$.For $\bar{\imath}=(i_1,...,i_k)\in S(k)$ we set $dY_t^{\bar{\imath}} = dY_{t_1}^{i_1}...dY_{t_k}^{i_k}$ and define iterated integrals $q_{s,t}^{\bar{\imath}}(Y)$ by setting

$$q_{s,t}^{\bar{\imath}}(Y) := \int_{\Delta_{s,t}^k} dY_t^{\bar{\imath}} = \underbrace{\int_s^t\int_s^{t_k}\cdots\int_s^{t_2}}_{k\ \text{ times}} dY_{t_1}^{i_1}...dY_{t_k}^{i_k}.$$

Let $q_{s,\bar{t}}^{\bar{k}_1,...,\bar{k}_r}(Y)$, $\bar{k}_1,...,\bar{k}_r\in S$, $\bar{t}=(t_1,...t_r)$ be the products of iterated integrals

$$q_{s,\bar{t}}^{\bar{k}_1,...,\bar{k}_r}(Y) = \prod_{i=1}^r q_{s,t_i}^{\bar{k}_i}(Y).$$

Next define the sets $\Theta(k)$

$$\Theta(k) = \mathrm{sp}\left\{ q_{s,\bar{t}}^{\bar{k}_1,...,\bar{k}_r}(Y), \ \ \bar{k}_1,...,\bar{k}_r \in S, \bar{t} = (t_1,...t_r), \sum_{i=1}^{r} \left|\bar{k}_i\right| = k \right\}$$

and let $\Theta := \bigcup_{k\in\mathbb{N}} \Theta(k)$. For $q \in \Theta$ we define its formal degree by setting $\deg(q) := r$, where r is the unique number such that $q \in \Theta(r)$. For $\bar{\imath} = (i_1,...,i_k) \in S(k)$ define $\Phi_{\bar{\imath}}$, $\Psi_{\bar{\imath}}$, be the following operators

$$\begin{aligned}
\Phi_{\bar{\imath}}\varphi &= h_{i_1}...h_{i_k}\varphi \\
\Psi_{\bar{\imath}}\varphi &= [\Phi_{\bar{\imath}}, A](\varphi) = A(h_{i_1}...h_{i_k})\varphi + \sum_{i=1}^{d} V_i(h_{i_1}...h_{i_k})V_i\varphi.
\end{aligned}$$

and $\Gamma_{\bar{\imath}}$ be the set of operators $\Gamma_{\bar{\imath}} = \{\Phi_{\bar{\imath}}, \Psi_{\bar{\imath}}, \Psi_{\bar{\imath}}\Phi_{\bar{\imath}}\}$. In the following proposition we obtain a pathwise representation of the terms in our expansion of the un-normalised conditional density. The proof will exploit integration by parts formulas of the form

$$\int_0^t q_{0,s}^{(i_1,...,i_k)}(Y)\left(\int_0^s Z_r dr\right) dY_s^{i_{k+1}} = q_{0,t}^{(i_1,...,i_{k+1})}(Y)\int_0^t Z_s ds - \int_0^t q_{0,s}^{(i_1,...,i_{k+1})}(Y) Z_s ds,$$

where Z is a suitably chosen process.

Proposition 11 *Let $\bar{\imath} = (i_1,..,i_m) \in S(m)$. Then we have, almost surely, that*

$$\begin{aligned}
R_{s,t}^{m,\bar{\imath}}(\varphi) &= P_{t-s}(h_{i_1}...h_{i_m}\varphi)(x)q_{s,t}^{\bar{\imath}}(Y) \\
&\quad + \sum_{k=1}^{m-1} \sum_{\bar{\jmath}_1,...,\bar{\jmath}_k;\,\bar{\imath}=\bar{\jmath}_1*...*\bar{\jmath}_k} a_{(s,t)}^{m,\bar{\jmath}_1,...,\bar{\jmath}_k}(Y)\int_{\Delta_{s,t}^k} b_{(s,t_1,...,t_k)}^{m,\bar{\jmath}_1,...,\bar{\jmath}_k}(Y)\,\bar{R}_{(s,t_1,...,t_k,t)}^{m,\bar{\jmath}_1,...,\bar{\jmath}_k}(\varphi)(x)d\bar{t}_k \\
&\quad + \sum_{k=1}^{m} \sum_{\bar{\jmath}_1,...,\bar{\jmath}_k;\,\bar{\imath}=\bar{\jmath}_1*...*\bar{\jmath}_k} \int_{\Delta_{s,t}^k} c_{(s,t_1,...,t_k)}^{m,\bar{\jmath}_1,...,\bar{\jmath}_k}(Y)\,\hat{R}_{(s,t_1,...,t_k,t)}^{m,\bar{\jmath}_1,...,\bar{\jmath}_k}(\varphi)(x)d\bar{t}_k, \qquad (26)
\end{aligned}$$

and $a_{(s,t)}^{m,\bar{\jmath}_1,...,\bar{\jmath}_k}(Y), b_{(s,t_1,...,t_k)}^{m,\bar{\jmath}_1,...,\bar{\jmath}_k}(Y), c_{(s,t_1,...,t_k)}^{m,\bar{\jmath}_1,...,\bar{\jmath}_k}(Y) \in \Theta$ are linear combinations of (products of) iterated integrals of Y and $\bar{R}_{(s,t_1,...,t_k,t)}^{m,\bar{\jmath}_1,...,\bar{\jmath}_k}(\varphi)$, respectively $\hat{R}_{(s,t_1,...,t_k,t)}^{m,\bar{\jmath}_1,...,\bar{\jmath}_k}(\varphi)$ are of the form

$$P_{t_1-s}\left(\bar{\Phi}_1 P_{t_2-t_1}\dots\left(\bar{\Phi}_k P_{t-t_k}(\varphi)\right)\right),$$

where $\bar{\Phi}_p \in \Gamma_{\bar{\jmath}_p}$, $p = 1,..,k$. Moreover we have

$$\deg\left(a_{(s,t)}^{m,\bar{\jmath}_1,...,\bar{\jmath}_k}(Y)\right) + \deg\left(b_{(s,t_1,...,t_k)}^{m,\bar{\jmath}_1,...,\bar{\jmath}_k}(Y)\right) = \deg\left(c_{(s,t_1,...,t_k)}^{m,\bar{\jmath}_1,...,\bar{\jmath}_k}(Y)\right) = m. \qquad (27)$$

Before we begin the proof note that $\bar{R}_{(s,t_1,...,t_k,t)}^{m,\bar{\jmath}_1,...,\bar{\jmath}_k}(\varphi)$ and $\hat{R}_{(s,t_1,...,t_k,t)}^{m,\bar{\jmath}_1,...,\bar{\jmath}_k}(\varphi)$ are in general different, but have the same structure as they both can be written as iterated compositions of the heat semi-group and operators drawn from the sets $\Gamma_{\bar{\jmath}_p}$.

Proof. The proof follows by induction. For $m=1$ observe that

$$R^{1}_{s,t}(\varphi) = \int_s^t P_{t_1-s}(h_{i_1}P_{t-t_1}(\varphi))(x)dY^{i_1}_{t_1}$$
$$= P_{t-s}(h_{i_1}\varphi)(x)\int_s^t dY^{i_1}_r - \int_s^t \left(\int_s^{t_1} dY^{i_1}_r\right)\frac{d}{dt_1}P_{t_1-s}(h_{i_1}P_{t-t_1}(\varphi))(x)dt_1,$$

where

$$\frac{d}{dt_1}P_{t_1-s}(h_{i_1}P_{t-t_1}(\varphi))(x) = P_{t_1-s}(A(h_{i_1}P_{t-t_1}(\varphi)))(x) - P_{t_1-s}(h_{i_1}AP_{t-t_1}(\varphi))(x)$$
$$= P_{t_1-s}(\Psi_{(i_1)}P_{t-t_1}(\varphi)))(x).$$

so (26) holds true with

$$c^{1,(i_1)}_{(s,t_1)}(Y) = \int_s^{t_1} dY^{i_1}_r$$

and, obviously, (27) is satisfied. For the induction step, observe that for $\bar{\imath} * i_{m+1}$

$$R^{m+1,\bar{\imath}*i_{m+1}}_{s,t}(\varphi) = \int_s^t R^{m,\bar{\imath}}_{s,t_{m+1}}(h_{i_{m+1}}P_{t-t_{m+1}}(\varphi))(x)\,dY^{i_{m+1}}_{t_{m+1}}.$$

Hence, assuming that $R^{m,\bar{\imath}}_{s,t_{m+1}}$ has an expansion of as in (26), it follows that

$$R^{m+1,\bar{\imath}*i_{m+1}}_{s,t}(\varphi) = R^{1,m+1,\bar{\imath}*i_{m+1}}_{s,t}(\varphi) + R^{2,m+1,\bar{\imath}*i_{m+1}}_{s,t}(\varphi) + R^{3,m+1,\bar{\imath}*i_{m+1}}_{s,t}(\varphi), \qquad (28)$$

where

$$R^{1,m+1,\bar{\imath}*i_{m+1}}_{s,t}(\varphi) = \int_s^t P_{t_{m+1}-s}(h_{i_1}...h_{i_{m+1}}P_{t-t_{m+1}}(\varphi))(x)\int_{\Delta^m_{s,t_{m+1}}} dY^{\bar{\imath}}_t dY^{i_{m+1}}_{t_{m+1}}$$

$$R^{2,m+1,\bar{\imath}*i_{m+1}}_{s,t}(\varphi) = \sum_{k=1}^{m-1}\sum_{\bar{\jmath}_1,\dots,\bar{\jmath}_k;\,\bar{\imath}=\bar{\jmath}_1*\dots*\bar{\jmath}_k}\int_s^t a^{m,\bar{\jmath}_1,\dots,\bar{\jmath}_k}_{(s,t_{m+1})}(Y)$$
$$\int_{\Delta^k_{s,t_{m+1}}} b^{m,\bar{\jmath}_1,\dots,\bar{\jmath}_k}_{(s,t_1,\dots,t_k)}(Y)\,\bar{R}^{k,\bar{\jmath}_1,\dots,\bar{\jmath}_k}_{(s,t_1,\dots,t_k,t_{m+1})}(h_{i_{m+1}}P_{t-t_{m+1}}(\varphi))(x)d\bar{t}_k dY^{i_{m+1}}_{t_{m+1}}$$

$$R^{3,m+1,\bar{\imath}*i_{m+1}}_{s,t}(\varphi) =$$
$$\sum_{k=1}^{m}\sum_{\bar{\jmath}_1,\dots,\bar{\jmath}_k;\,\bar{\imath}=\bar{\jmath}_1*\dots*\bar{\jmath}_k}\int_s^t\int_{\Delta^k_{s,t_{m+1}}} c^{m,\bar{\jmath}_1,\dots,\bar{\jmath}_k}_{(s,t_1,\dots,t_k)}(Y)\,\hat{R}^{k,\bar{\jmath}_1,\dots,\bar{\jmath}_k}_{(s,t_1,\dots,t_k,t_{m+1})}(h_{i_{m+1}}P_{t-t_{m+1}}(\varphi))(x)d\bar{t}_k dY^{i_{m+1}}_{t_{m+1}}.$$

We expand each of the three terms in (28). For the first term we have

$$
\begin{aligned}
R^{1,m+1,\bar{\imath}*i_{m+1}}_{s,t}(\varphi) &= P_{t-s}(h_{i_1}...h_{i_{m+1}}\varphi)(x)\int_{\Delta^{m+1}_{s,t}} dY^{\bar{\imath}*i_{m+1}}_t \\
&\quad -\int_s^t\left(\int_s^{t_{m+1}}\int_{\Delta^m_{s,r}} dY^{\bar{\imath}}_t dY^{i_{m+1}}_r\right)\frac{d}{dt_{m+1}}P_{t_{m+1}-s}(h_{i_1}...h_{i_{m+1}}P_{t-t_{m+1}}(\varphi))(x)dt_{m+1}
\end{aligned}
$$

$$
\begin{aligned}
&= P_{t-s}(h_{i_1}...h_{i_{m+1}}\varphi)(x)\int_{\Delta^{m+1}_{s,t}} dY^{\bar{\imath}*i_{m+1}}_t \\
&\quad -\int_s^t\left(\int_s^{t_{m+1}}\int_{\Delta^m_{s,r}} dY^{\bar{\imath}}_t dY^{i_{m+1}}_r\right)P_{t_{m+1}-s}(\Psi_{\bar{\imath}*i_{m+1}}P_{t-t_{m+1}}(\varphi))(x)dt_{m+1}. \quad (29)
\end{aligned}
$$

so the first term in the expansion of $R^{1,m+1,\bar{\imath}*i_{m+1}}_{s,t}(\varphi)$ gives us the first term in the expansion of (28) and the second term in the expansion of $R^{1,m+1,\bar{\imath}*i_{m+1}}_{s,t}(\varphi)$ can be incorporated in the last term in the expansion of (28). Obviously,

$$
\deg\left(\int_{\Delta^{m+1}_{s,t}} dY^{\bar{\imath}*i_{m+1}}_t\right) = m+1
$$

so (27) is satisfied. For the second term we have

$$
\begin{aligned}
R^{2,m+1,\bar{\imath}*i_{m+1}}_{s,t}(\varphi) &= \sum_{k=1}^{m-1}\sum_{\bar{\jmath}_1,\ldots,\bar{\jmath}_k;\,\bar{\imath}=\bar{\jmath}_1*\ldots*\bar{\jmath}_k}\int_s^t a^{m,\bar{\jmath}_1,\ldots,\bar{\jmath}_k}_{(s,t_{m+1})}(Y) \\
&\int_s^{t_{m+1}}\int_{\Delta^{k-1}_{s,t_k}} b^{m,\bar{\jmath}_1,\ldots,\bar{\jmath}_k}_{(s,t_1,\ldots,t_k)}(Y)\,\bar{R}^{m,\bar{\jmath}_1,\ldots,\bar{\jmath}_k}_{(s,t_1,\ldots,t_k,t_{m+1})}(h_{i_{m+1}}P_{t-t_{m+1}}(\varphi))(x)d\bar{t}_{k-1}dY^{i_{m+1}}_{t_{m+1}} \\
&= \sum_{k=1}^{m-1}\sum_{\bar{\jmath}_1,\ldots,\bar{\jmath}_k;\,\bar{\imath}=\bar{\jmath}_1*\ldots*\bar{\jmath}_k}\int_s^t a^{m,\bar{\jmath}_1,\ldots,\bar{\jmath}_k}_{(s,t_{m+1})}(Y)\,dY^{i_{m+1}}_{t_{m+1}}\int_s^t S^{2,m+1,\bar{\jmath}_1,\ldots,\bar{\jmath}_k,i_{m+1}}_{s,t_{m+1}}(\varphi)dt_{m+1} \\
&-\sum_{k=1}^{m-1}\sum_{\bar{\jmath}_1,\ldots,\bar{\jmath}_k;\,\bar{\imath}=\bar{\jmath}_1*\ldots*\bar{\jmath}_k}\int_s^t\left(\int_s^{t_{m+1}} a^{m,\bar{\jmath}_1,\ldots,\bar{\jmath}_k}_{(s,r)}(Y)\,dY^{i_{m+1}}_r\right)S^{2,m+1,\bar{\jmath}_1,\ldots,\bar{\jmath}_k,i_{m+1}}_{s,t_{m+1}}(\varphi)dt_{m+1} \quad (30)
\end{aligned}
$$

where

$$
\begin{aligned}
&S^{2,m+1,\bar{\jmath}_1,\ldots,\bar{\jmath}_k,i_{m+1}}_{s,t_{m+1}}(\varphi) \\
&\qquad = \frac{d}{dt_{m+1}}\int_{\Delta^k_{s,t_{m+1}}} b^{m,\bar{\jmath}_1,\ldots,\bar{\jmath}_k}_{(s,t_1,\ldots,t_k)}(Y)\,\bar{R}^{m,\bar{\jmath}_1,\ldots,\bar{\jmath}_k}_{(s,t_1,\ldots,t_k,t_{m+1})}(h_{i_{m+1}}P_{t-t_{m+1}}(\varphi))(x)d\bar{t}_k
\end{aligned}
$$

$$= \int_{\Delta^{k-1}_{s,t_{m+1}}} b^{m,\bar{j}_1,\ldots,\bar{j}_k}_{(s,t_1,\ldots,t_k)}(Y)\,\bar{R}^{m,\bar{j}_1,\ldots,\bar{j}_k}_{(s,t_1,\ldots,t_{m+1},t_{m+1})}(h_{i_{m+1}}P_{t-t_{m+1}}(\varphi))(x)d\bar{t}_{k-1}$$

$$+\int_{\Delta^{k}_{s,t_{m+1}}} b^{m,\bar{j}_1,\ldots,\bar{j}_k}_{(s,t_1,\ldots,t_k)}(Y)\,\bar{R}^{m,\bar{j}_1,\ldots,\bar{j}_k}_{(s,t_1,\ldots,t_k,t_{m+1})}(A\left(h_{i_{m+1}}P_{t-t_{m+1}}(\varphi)\right)-h_{i_{m+1}}A\left(P_{t-t_{m+1}}(\varphi)\right))(x)d\bar{t}_k$$

$$= \int_{\Delta^{k-1}_{s,t_{m+1}}} b^{m,\bar{j}_1,\ldots,\bar{j}_k}_{(s,t_1,\ldots,t_k)}(Y)\,\bar{R}^{m,\bar{j}_1,\ldots,\bar{j}_k}_{(s,t_1,\ldots,t_{m+1},t_{m+1})}(\Phi_{(i_{m+1})}P_{t-t_{m+1}}(\varphi))(x)d\bar{t}_{k-1}$$
$$+\int_{\Delta^{k}_{s,t_{m+1}}} b^{m,\bar{j}_1,\ldots,\bar{j}_k}_{(s,t_1,\ldots,t_k)}(Y)\,\bar{R}^{m,\bar{j}_1,\ldots,\bar{j}_k}_{(s,t_1,\ldots,t_k,t_{m+1})}(\Psi_{(i_{m+1})}P_{t-t_{m+1}}(\varphi))(x)d\bar{t}_k \quad (31)$$

The first term in the expansion of $R^{2,m+1,\bar{\imath}*i_{m+1}}_{s,t}(\varphi)$ contributes to the second term in the expansion of (28). The identity (27) is also satisfied as each of the terms $a^{m,\bar{j}_1,\ldots,\bar{j}_k}_{(s,t_{m+1})}(Y)$ is replaced by

$$\int_s^t a^{k,m,\bar{\imath}}_{(s,t_{m+1})}(Y)\,dY^{i_{m+1}}_r$$

so the degree for each term increases by 1. Similarly, the second term in the expansion of $R^{2,m+1,\bar{\imath}*i_{m+1}}_{s,t}(\varphi)$ contributes to the third term in the expansion of (28), whilst the identity (27) is also satisfied as each of the terms $a^{m,\bar{j}_1,\ldots,\bar{j}_k}_{(s,r)}(Y)$ is replaced by

$$\int_s^{t_{m+1}} a^{m,\bar{j}_1,\ldots,\bar{j}_k}_{(s,r)}(Y)\,dY^{i_{m+1}}_r$$

so, again, the degree for each term increases by 1. Similarly,

$$R^{3,m+1,\bar{\imath}*i_{m+1}}_{s,t}(\varphi)$$
$$=\sum_{k=1}^{m}\sum_{\bar{j}_1,\ldots,\bar{j}_k;\,\bar{\imath}=\bar{j}_1*\ldots*\bar{j}_k}\int_s^t\int_{\Delta^k_{s,t_{m+1}}} c^{m,\bar{j}_1,\ldots,\bar{j}_k}_{(s,t_1,\ldots,t_k)}(Y)\,\hat{R}^{m,\bar{j}_1,\ldots,\bar{j}_k}_{(s,t_1,\ldots,t_k,t_{m+1})}(h_{i_{m+1}}P_{t-t_{m+1}}(\varphi))(x)d\bar{t}_k dY^{i_{m+1}}_{t_{m+1}} \quad (32)$$

$$=\sum_{k=1}^{m}\sum_{\bar{j}_1,\ldots,\bar{j}_k;\,\bar{\imath}=\bar{j}_1*\ldots*\bar{j}_k}\int_s^t dY^{i_{m+1}}_{t_{m+1}}\int_s^t S^{3,m,\bar{j}_1,\ldots,\bar{j}_k,i_{m+1}}_{s,t_{m+1}}(\varphi)\,dt_{m+1}$$
$$-\int_s^t\int_s^{t_{m+1}} dY^{i_{m+1}}_r S^{3,m,\bar{j}_1,\ldots,\bar{j}_k,i_{m+1}}_{s,t_{m+1}}(\varphi)\,dt_{m+1}$$

where

$$S_{s,t_{m+1}}^{3,m,\bar{j}_1,\dots,\bar{j}_k,i_{m+1}}(\varphi) = \frac{d}{dt_{m+1}} \int_{\Delta_{s,t_{m+1}}^k} c_{(s,t_1,\dots,t_k)}^{m,\bar{j}_1,\dots,\bar{j}_k}(Y)\, \hat{R}_{(s,t_1,\dots,t_k,t_{m+1})}^{m,\bar{j}_1,\dots,\bar{j}_k}(h_{i_{m+1}}P_{t-t_{m+1}}(\varphi))(x)d\bar{t}_k$$

$$= \int_{\Delta_{s,t_{m+1}}^{k-1}} c_{(s,t_1,\dots,t_k)}^{m,\bar{j}_1,\dots,\bar{j}_k}(Y)\, \hat{R}_{(s,t_1,\dots,t_{m+1},t_{m+1})}^{m,\bar{j}_1,\dots,\bar{j}_k}(\Phi_{(i_{m+1})}P_{t-t_{m+1}}(\varphi))(x)d\bar{t}_{k-1}$$

$$+ \int_{\Delta_{s,t_{m+1}}^k} c_{(s,t_1,\dots,t_k)}^{m,\bar{j}_1,\dots,\bar{j}_k}(Y)\, \hat{R}_{(s,t_1,\dots,t_k,t_{m+1})}^{m,\bar{j}_1,\dots,\bar{j}_k}(\Psi_{(i_{m+1})}P_{t-t_{m+1}}(\varphi))(x)d\bar{t}_k \quad (33)$$

The first term in the expansion of $R_{s,t}^{3,m+1,\bar{\imath}*i_{m+1}}(\varphi)$ contributes to the second term in the expansion of (28). Identity (27) is again satisfied as we add $\int_s^t dY_{t_{m+1}}^{i_{m+1}}$ to each of the terms so the total degree increases by 1. Similarly, the second term in the expansion of $R_{s,t}^{3,m+1,\bar{\imath}*i_{m+1}}(\varphi)$ contributes to the third term in the expansion of (28), whilst the identity (27) is again satisfied as we add $\int_s^t dY_{t_{m+1}}^{i_{m+1}}$ to each term.

The result now follows from (29), (30), (31), (32) and (33).

■

We will require a pathwise control of the iterated (Itô) integrals $q_{s,t}^{\bar{\imath}}(Y)$ of the Brownian motion Y. It is well known that the Itô lift of Brownian motion is a Hölder controlled rough path (see e.g. [19] or [9]), which immediately implies the following lemma.

Lemma 12 *For any $1/3 < \gamma < 1/2$ there exists a positive random variable $c = c(\omega,\gamma)$ and some constant $\theta > 0$ such that, almost surely,*

$$\left|q_{s,t}^{\bar{\imath}}(Y)\right| \leq \frac{(c(\omega,\gamma)|s-t|)^{k\gamma}}{\theta(k\gamma)!}$$

for all $0 \leq s \leq t \leq 1$, $\bar{\imath} \in S(k)$.

It is important to note that the operators Φ that arise when we recursively apply the integration by parts in the Proposition 11 only involve the vector fields V_i, $i = 1,\dots,d_1$ (but not the vector field V_0) and these vector fields do not change if we consider the Itô or Stratonovich versions of the SDE defining the signal.

We have already seen that the $R_{s,t}^{m,\bar{\imath}}$ may be regarded as bounded linear operators. The following two lemmas show us how to deduce regularity estimates on $R_{s,t}^{m,\bar{\imath}}$ from regularity estimates on the integral kernels $\bar{R}$ and $\hat{R}$, provided these operators make sense as bounded linear operators over suitable function spaces. More specifically, let W and $\tilde{W}$ denote two function spaces and let $\left(L\left(W,\tilde{W}\right),\|\cdot\|\right)$ be the space of bounded linear operators from W to $\tilde{W}$. In the following the function spaces W and $\tilde{W}$ will be taken to be either H^1 or H^{-1} and we will (later) see that the $R_{s,t}^{m,\bar{\imath}}$ are indeed bounded linear operators on these spaces and thus satisfy the hypothesis of Lemma 13.

Lemma 13 *With the notation of Lemma 11. Let* $\left(L\left(W,\tilde{W}\right),\|\cdot\|\right)$ *be the space of bounded linear operators discussed above,* $\bar{\imath}\in S\left(m\right)$ *and suppose that* P_t, $R_{s,t}^{m,\bar{\imath}}$, $\bar{R}_{(s,t_1,\ldots,t_k,t)}^{m,\bar{\jmath}_1,\ldots,\bar{\jmath}_k\cdot}$, $\hat{R}_{(s,t_1,\ldots,t_k,t)}^{m,\bar{\jmath}_1,\ldots,\bar{\jmath}_k}\in L\left(W,\tilde{W}\right)$ *for all* $0\le s<t\le 1$. *Then for any* $1/3<\gamma<1/2$ *there exist random variables* $c(\gamma,\omega,m)$ *such that, almost surely*

$$\left\|R_{s,t}^{m,\bar{\imath}}\right\|\le\left(c\left(\gamma,\omega,m\right)|t-s|\right)^{m\gamma}\left(\left\|P_{t-s}\left(h_{i_1}\cdots h_{i_m}\cdot\right)\right\|\right. \tag{34}$$
$$\left.+\sum_{k=1}^{m}\sum_{\bar{\jmath}_1,\ldots,\bar{\jmath}_k;\,\bar{\imath}=\bar{\jmath}_1*\ldots*\bar{\jmath}_k}\int_{\Delta_{s,t}^k}\left\|\bar{R}_{(s,t_1,\ldots,t_k,t)}^{m,\bar{\jmath}_1,\ldots,\bar{\jmath}_k}\right\|+\left\|\hat{R}_{(s,t_1,\ldots,t_k,t)}^{m,\bar{\jmath}_1,\ldots,\bar{\jmath}_k}\right\|d\bar{t}_k\right).$$

Proof. It follows immediately from combining the Hölder estimates for the iterated integrals $q_{s,t}^{\bar{\imath}}\left(Y\right)$ obtained in Lemma 12 and Proposition 11 that

$$\left\|R_{s,t}^{m,\bar{\imath}}\right\|\le\left(c\left(\gamma,\omega,m\right)|t-s|\right)^{m\gamma}\left(\left\|P_{t-s}\left(h_{i_1}\cdots h_{i_m}\cdot\right)\right\|\right.$$
$$\left.+\sum_{k=1}^{m}\sum_{\bar{\jmath}_1,\ldots,\bar{\jmath}_k;\,\bar{\imath}=\bar{\jmath}_1*\ldots*\bar{\jmath}_k}\left\|\int_{\Delta_{s,t}^k}\bar{R}_{(s,t_1,\ldots,t_k,t)}^{m,\bar{\jmath}_1,\ldots,\bar{\jmath}_k}d\bar{t}_k\right\|+\left\|\int_{\Delta_{s,t}^k}\hat{R}_{(s,t_1,\ldots,t_k,t)}^{m,\bar{\jmath}_1,\ldots,\bar{\jmath}_k}d\bar{t}_k\right\|\right).$$

■

In the following lemma we assume that the integral kernels $\bar{R}$ and $\hat{R}$ have bounds with integrable singularities. The control of the constants in the lemma is actually stronger than we will later require.

Lemma 14 *Under the assumptions of Lemma 13. Let* $\bar{\imath}\in S\left(m\right)$, $m\ge 1$. *Suppose there exists a constant* c *such that for all* $\bar{\jmath}_1,\ldots,\bar{\jmath}_k\in S$ *satisfying* $\bar{\imath}=\bar{\jmath}_1*\ldots*\bar{\jmath}_k.$, $t_0=0<t_1<\cdots<t_k<t$ *we have both*

$$\left\|\bar{R}_{(0,t_1,\ldots,t_k,t)}^{m,\bar{\jmath}_1,\ldots,\bar{\jmath}_k}\right\|\le ct^{k_0}\frac{1}{\sqrt{t_1-t_0}}\cdots\frac{1}{\sqrt{t_k-t_{k-1}}} \tag{35}$$

and

$$\left\|\hat{R}_{(0,t_1,\ldots,t_k,t)}^{m,\bar{\jmath}_1,\ldots,\bar{\jmath}_k}\right\|\le ct^{k_0}\frac{1}{\sqrt{t_1-t_0}}\cdots\frac{1}{\sqrt{t_k-t_{k-1}}}$$

for some $k_0\in\mathbb{R}$. *Then*

$$\int_{\Delta_{s,t}^k}\left\|\bar{R}_{(s,t_1,\ldots,t_k,t)}^{m,\bar{\jmath}_1,\ldots,\bar{\jmath}_k}\right\|+\left\|\hat{R}_{(s,t_1,\ldots,t_k,t)}^{m,\bar{\jmath}_1,\ldots,\bar{\jmath}_k}\right\|d\bar{t}_k\le a_k\left|t-s\right|^{k/2+k_0} \tag{36}$$

where

$$a_k=\frac{4\left(2\sqrt{\pi}\right)^k c}{k\Gamma\left(\frac{k}{2}\right)}.$$

Proof. First observe that

$$\underbrace{\int_s^t\int_s^{t_k}\cdots\int_s^{t_2}}_{k\ \text{ times}}\bar{R}_{(s,t_1,\ldots,t_k,t)}^{m,\bar{\jmath}_1,\ldots,\bar{\jmath}_k}(\varphi)dt_1\ldots dt_k=\underbrace{\int_0^{t-s}\int_0^{t_k}\cdots\int_0^{t_2}}_{k\ \text{ times}}\bar{R}_{(0,t_1,\ldots,t_k,t-s)}^{m,\bar{\jmath}_1,\ldots,\bar{\jmath}_k}(\varphi)dt_1\ldots dt_k.$$

Hence, it is sufficient to prove the result for $s=0$. Writing $t-s=u$ let Λ_u be the simplex

$$\Lambda_u = \left\{ (a_1, ..., a_k) \in \mathbb{R}_+^k | \sum_{i=1}^k a_i \leq u \right\}.$$

We have

$$\left\| P_{a_1}\left(\bar{\Phi}_1 P_{a_2} \dots \bar{\Phi}_{k-1} P_{a_k}\left(\bar{\Phi}_k P_{u-\left(\sum_{j=1}^k a_j\right)}\right)\right)\right\| \leq c_k u^{k_0} \frac{1}{\sqrt{a_1}} \frac{1}{\sqrt{a_2}} \cdots \frac{1}{\sqrt{a_k}}$$

and introduce the change of variable $a_i = u z_i^2, \quad i = 1, ..., k$ with the determinant of its Jacobian being $2^k u^k z_1 z_2 ... z_k$. Then

$$\int_{\Lambda_t^{k+1}} \left\| P_{a_1}\left(\bar{\Phi}_1 P_{a_2} \dots \bar{\Phi}_{k-1} P_{a_k}\left(\bar{\Phi}_k P_{u-\left(\sum_{j=1}^k a_j\right)}\right)\right)\right\| da \leq c_k 2^k u^{\frac{k}{2}+k_0} l\left(\Lambda_1^{k+1}\right),$$

where

$$\Lambda_1^{k+1} \subset \left\{ (z_1, ..., z_k) \in \mathbb{R}_+^k | \sum_{i=1}^k z_i^2 \leq 1 \right\}.$$

In other words Λ_1^{k+1} is a subset of the unit hypersphere hence its volume less the volume of the sphere so

$$l\left(\Lambda_1^{k+1}\right) \leq \frac{2\pi^{\frac{k}{2}}}{k\Gamma\left(\frac{k}{2}\right)}.$$

A similar argument using $\hat{R}^{m,\bar{j}_1,\dots,\bar{j}_k}_{(s,t_1,\dots,t_k,t)}$ in place of $\bar{R}^{m,\bar{j}_1,\dots,\bar{j}_k}_{(s,t_1,\dots,t_k,t)}$ completes the proof. ■

5.2 Kusuoka-Stroock regularity estimates for the integral kernels and the proof of Proposition 9

In this subsection we state regularity estimates for the integral kernels $\hat{R}$ and $\bar{R}$ that arise in the pathwise representation of the expansion of the unnormalised conditional density. The results are subsequently proved in Appendix 7.2. The use of these bounds is twofold. Firstly, they will allow us to control directly the lower order terms in the expansion derived in Section 5 (as in Proposition 9) and, secondly, they provide us via Lemma 13 with bounds on the operator norms of the operators $R^{m,\bar{\imath}}$ acting on the spaces H^1 and H^{-1} respectively.

Proposition 15 *Let $\alpha, \beta \in A_1\left(\bar{k}\right)$ and $\bar{R}^{m,\bar{j}_1,\dots,\bar{j}_k}_{(s,t_1,\dots,t_k,t)}(\varphi)$ and $\hat{R}^{m,\bar{j}_1,\dots,\bar{j}_k}_{(s,t_1,\dots,t_k,t)}(\varphi)$ as in Proposition 11. Then, for any $m \in \mathbb{N}$ there exist a constant c_m such that for all $\bar{\imath} \in S(m)$, $\bar{j}_1, \dots, \bar{j}_k \in S$ satisfying $\bar{\imath} = \bar{j}_1 * ... * \bar{j}_k$. and $t_0 = 0 < t_1 < \dots < t_k < t \leq 1$ we have both*

$$\left\| V_{[\alpha]} \bar{R}^{m,\bar{j}_1,\dots,\bar{j}_k}_{(0,t_1,\dots,t_k,t)} V_{[\beta]} \varphi \right\|_\infty \leq c_m t^{-(\|\alpha\|+\|\beta\|)/2} \frac{1}{\sqrt{t_1 - t_0}} \cdots \frac{1}{\sqrt{t_k - t_{k-1}}} \|\varphi\|_\infty \tag{37}$$

and

$$\left\| V_{[\alpha]} \hat{R}^{m,\bar{j}_1,\dots,\bar{j}_k}_{(0,t_1,\dots,t_k,t)} V_{[\beta]} \varphi \right\|_\infty \leq c_m t^{-(\|\alpha\|+\|\beta\|)/2} \frac{1}{\sqrt{t_1 - t_0}} \cdots \frac{1}{\sqrt{t_k - t_{k-1}}} \|\varphi\|_\infty .$$

for all $\varphi \in C_b^\infty(\mathbb{R}^N)$.

The proof of the proposition is non-trivial, but closely follows the ideas and techniques of Kusuoka [12]. We therefore defer the proof of Proposition 15 and all the lemmas required on the way to Appendix 7.2. Finally, we record that the proof of Proposition 9 is an immediate consequence.

Proof of Proposition 9. Proposition 15 provides regularity estimates for the kernels $\bar{R}$ and $\hat{R}$. Theorem 8 may be used to estimate the regularity of $P_{t-s}\left(h_{i_1}\cdots h_{i_m}\cdot\right)$. The corresponding bounds for $R_{0,t}^{m,\bar{\imath}}$ in Proposition 9 follow by arguing exactly as in the proof of Lemma 14. ■

6 Proof of Proposition 10: Factorial decay of the integral summands via rough path techniques

6.1 Some preliminary estimates

Before we can proceed with the proof of Proposition 10 we explore some of the consequences of the estimates derived in Proposition 15, which were already used in the proof of Proposition 9. It turns out the same estimate can be used to control various operator norms of the terms in the perturbation expansion. Recall that the constant $\bar{k}$ was defined to be the minimal number of Lie brackets required to satisfy the UFG condition.

Lemma 16 *With the notation of Lemma 7 for any $0<\gamma<1/2$, $m>0$ there exist random variables $c(\gamma,m,\omega)$ such that, almost surely*

$$\left\|R_{s,t}^{m,\bar{\imath}}\right\|_{H^{-1}\to H^{-1}} \leq c\left(\gamma,m,\omega\right)\left|t-s\right|^{m\gamma}. \tag{38}$$

$$\left\|R_{s,t}^{m,\bar{\imath}}\right\|_{H^{1}\to H^{1}} \leq c\left(\gamma,m,\omega\right)\left|t-s\right|^{m\gamma}. \tag{39}$$

and finally

$$\left\|R_{s,t}^{m,\bar{\imath}}\right\|_{H^{-1}\to H^{1}} \leq c\left(\gamma,m,\omega\right)\left|t-s\right|^{m\gamma-\bar{k}}. \tag{40}$$

for all $\bar{\imath}\in S\left(m\right)$, $0<s<t<1$.

The proof of Lemma 16 is once again deferred to the appendix (Section 7.3), as it relies on the same Malliavin calculus techniques employed in the proof of Proposition 15.

So far, we have established a priori Hölder type estimates for $R_{s,t}^{m,\bar{\imath}}\left(\varphi\right)$, but the estimates in their current form are not yet summable. The following proof of Proposition 10 relies on a fundamental rough path technique to improve on these bounds and demonstrate that the operator norms of $R_{s,t}^{m,\bar{\imath}}$ decay in fact factorially in m.

6.2 Proof of Proposition 10

To make the presentation more transparent we introduce some additional notations for the following arguments. Recall that $\Delta_{s,t}^{k}$ denotes the simplex defined by the relation $s<t_1<\cdots<t_k<t$ and the H_i are the operators corresponding to multiplication by the sensor

function h_i. For any $0 \leq s < t \leq 1$ define $R^{\emptyset}_{s,t} := 1$ and recall the linear operators $R^{n,\bar{\imath}}_{s,t}$ may be written as

$$R^{n,\bar{\imath}}_{s,t} = \int_{\Delta^k_{s,t}} P_{t_1 - s} H_{i_1} P_{t_2 - t_1} H_{i_2} \cdots H_{i_n} P_{t - t_n} dY^{i_1}_{t_1} \cdots dY^{i_n}_{t_n}$$

for all $\bar{\imath} = (i_1, \ldots, i_n) \in S$, $n \geq 1$.

Let $W := \mathbb{R}^{d_2}$ and $\varepsilon_1, \ldots, \varepsilon_{d_2}$ be a basis for W. For $\bar{\imath} = (i_1, \ldots i_j) \in S(j)$ let $\varepsilon_{\bar{\imath}} = \varepsilon_{i_1} \otimes \cdots \otimes \varepsilon_{i_j}$ and note that the $\varepsilon_{\bar{\imath}}$ are a basis for the space $W^{\otimes j}$. Finally, let V be a Banach algebra (i.e. a Banach space with a multiplication and a submultiplicative norm). We define $\mathcal{P}_{d_2,k}(V)$ the space of non-commutative polynomials in d_2 variables of degree at most k over V by letting

$$\mathcal{P}_{d_2,k}(V) := \left\{ \sum_{j=0}^{k} \sum_{\bar{\imath} \in S(j)} c_{\bar{\imath}} \varepsilon_{\bar{\imath}} : c_{\bar{\imath}} \in V \right\}.$$

Define a multiplication for $a = \sum_{j=0}^{k} a_j$, $a_j = \sum_{\bar{\imath} \in S(j)} a_{\bar{\imath}} \varepsilon_{\bar{\imath}}$ and $b = \sum_{j=0}^{k} b_j$, $b_j = \sum_{\bar{\imath} \in S(j)} b_{\bar{\imath}} \varepsilon_{\bar{\imath}}$ by setting

$$ab := \sum_{v=0}^{k} \sum_{j=0}^{v} a_j b_{v-j} := \sum_{v=0}^{k} \sum_{j=0}^{v} \sum_{\bar{\imath} \in S(j)} \sum_{\bar{l} \in S(v-j)} a_{\bar{\imath}} b_{\bar{l}} \varepsilon_{\bar{\imath} * \bar{l}}. \tag{41}$$

Further note that

$$\sum_{j=0}^{v} \sum_{\bar{\imath} \in S(j)} \sum_{\bar{l} \in S(v-j)} a_{\bar{\imath}} b_{\bar{l}} \varepsilon_{\bar{\imath} * \bar{l}} = \sum_{\bar{\imath} \in S(v)} \sum_{\bar{m} * \bar{l} = \bar{\imath}} a_{\bar{m}} b_{\bar{l}} \varepsilon_{\bar{m} * \bar{l}} \tag{42}$$

and define for $k \geq i \geq 1$ the projection π_i by setting $\pi_i(a) = a_i$. We impose a norm on $\mathcal{P}_{d_2,k}(V)$ by setting

$$\left\| \sum_{j=0}^{k} \sum_{\bar{\imath} \in S(j)} c_{\bar{\imath}} \varepsilon_{\bar{\imath}} \right\| = \sup \left\{ \|c_{\bar{\imath}}\| : j \in \{0, \ldots, k\}, \bar{\imath} \in S(j) \right\}$$

Let $Q^0_{s,t} = 1$ and $Q^j_{s,t}$ for $j \in \mathbb{N}$ be given by

$$Q^j_{s,t} = \sum_{\bar{\imath} \in S(j)} R^{j,\bar{\imath}}_{s,t} \varepsilon_{\bar{\imath}}.$$

Finally, we may set

$$Q^{[n]}_{s,t} = \sum_{i=0}^{n} Q^i_{s,t}.$$

Observe that for any $s < u < t$ and $k \in \mathbb{N}$ and $\bar{\imath} = (i_1, \ldots i_k) \in S(k)$, we have partitioning the simplex $\Delta^k_{s,t}$

$$\begin{aligned}
R^{k,\bar{\imath}}_{s,t} &= \int_{\Delta^k_{s,u}} P_{t_1-s}H_{i_1}P_{t_2-t_1}H_{i_2}\cdots H_{i_k}P_{t-t_k}dY^{i_1}_{t_1}\cdots dY^{i_k}_{t_k} \\
&+ \int_{\Delta^k_{u,t}} P_{t_1-s}H_{i_1}P_{t_2-t_1}H_{i_2}\cdots H_{i_k}P_{t-t_k}dY^{i_1}_{t_1}\cdots dY^{i_k}_{t_k} \\
&+ \sum_{j=1}^{k-1}\int_{\Delta^k_{s,u}} P_{t_1-s}H_{i_1}\cdots P_{t_j-t_{j-1}}H_{i_j}P_{u-t_j}dY^{i_1}_{t_1}\cdots dY^{i_j}_{t_j} \\
&\int_{\Delta^{k-j}_{u,t}} P_{t_{j+1}-u}H_{i_{j+1}}P_{t_{j+2}-t_{j+1}}\cdots H_{i_k}P_{t-t_k}dY^{i_{j+1}}_{t_{j+1}}\cdots dY^{i_k}_{t_k} \\
&= \sum_{j=0}^{k} R^{j,(i_1,\ldots i_j)}_{s,u} R^{k-j,(i_{j+1},\ldots i_k)}_{u,t} \\
&= \sum_{\bar{m}*\bar{l}=\bar{\imath}} R^{|\bar{m}|,\bar{m}}_{s,u} R^{|\bar{l}|,\bar{l}}_{u,t} \qquad (43)
\end{aligned}$$

and therefore using (42)

$$\begin{aligned}
Q^{[k]}_{s,t} &= \sum_{v=0}^{k}\sum_{\bar{\imath}\in S(v)} R^{v,\bar{\imath}}_{s,t}\varepsilon_{\bar{\imath}} \\
&= \sum_{v=0}^{k}\sum_{\bar{\imath}\in S(v)}\sum_{\bar{m}*\bar{l}=\bar{\imath}} R^{|\bar{m}|,\bar{m}}_{s,u} R^{|\bar{l}|,\bar{l}}_{u,t}\varepsilon_{\bar{m}}\varepsilon_{\bar{l}} \\
&= \sum_{v=0}^{k}\sum_{j=0}^{v}\sum_{\bar{\imath}\in S(j)}\sum_{\bar{l}\in S(v-j)} R^{j,\bar{\imath}}_{s,u} R^{v-j,\bar{l}}_{u,t}\varepsilon_{\bar{\imath}*\bar{l}} \\
&= \sum_{v=0}^{k}\sum_{j=0}^{v} Q^{j}_{s,u} Q^{v-j}_{u,t}
\end{aligned}$$

or equivalently

$$Q^{[k]}_{s,t} = Q^{[k]}_{s,u} Q^{[k]}_{u,t}. \qquad (44)$$

Analogous to the corresponding rough path concept we will refer to (44) as the multiplicative property. We recall that by Lemma 7

$$\rho_t = P_t + \sum_{n=1}^{\infty}\sum_{\bar{\imath}\in S(n)} R^{n,\bar{\imath}}_{0,t}. \qquad (45)$$

The following proposition demonstrates that it suffices to obtain Holder type controls on finitely many of the $Q^n_{s,t}$ to control the infinite series in (45) . The proof utilizes techniques of the classical extension theorem for rough paths due to Lyons (see e.g. [19] p.45f) and exploits the multiplicative structure of the operator valued integrands.

Lemma 17 *Let $q \geq 1$ and let $\lfloor q \rfloor$ denote the integer part of q and V be a Banach algebra with norm $\|\cdot\|$. Suppose $Q^{[\lfloor q \rfloor]} = \sum_{j=0}^{\lfloor q \rfloor} Q^j \in \mathcal{P}_{d_2, \lfloor q \rfloor}(V)$ satisfies the multiplicative property (44). Suppose there exists a constant $C > 0$ such that for all $(s,t) \in \Delta_{[0,1]}$, $j = 1, \cdots \lfloor q \rfloor$,*

$$\left\| Q_{s,t}^j \right\| \leq \frac{(C|t-s|)^{j/q}}{\theta (j/q)!}, \tag{46}$$

where $\theta = \left(q^2 + \sum_{r=3}^{\infty} \left(\frac{2}{r-2} \right)^{\frac{\lfloor q \rfloor + 1}{q}} \right)$. Then for all $m > \lfloor q \rfloor$ there exists a multiplicative extension $1 + Q_{s,t}^1 + \cdots + Q_{s,t}^{\lfloor q \rfloor} + \widetilde{Q}_{s,t}^{\lfloor q \rfloor + 1} + \cdots + \widetilde{Q}_{s,t}^m$ on $\mathcal{P}_{d_2,m}(V)$ such that (46) holds for all $j \in \{1, \ldots, m\}$, $(s,t) \in \Delta_{[0,1]}^2$. Moreover if $\overline{Q}_{s,t}^j$ is another multiplicative extension such that $\left\| \overline{Q}_{s,t}^j \right\| \leq C(j)(|t-s|)^{j/q}$ for all $(s,t) \in \Delta_{[0,1]}^2$, then $\overline{Q}_{s,t}^j = \widetilde{Q}_{s,t}^j$ for all $j \in \{1, \ldots, m\}$.

Before we begin the proof of the lemma we recall the neo-classical inequality due to [18] (Lemma 2.2.2). The slightly stronger form of the inequality we state below is due to Hara-Hino [10].

Theorem 18 (Neo-classical inequality, Lyons 98, Hara-Hino 2010) *For any $q \in [1, \infty)$, $n \in \mathbb{N}$ and $s, t \geq 0$*

$$\frac{1}{q} \sum_{i=0}^{n} \frac{s^{\frac{i}{q}} t^{\frac{n-i}{q}}}{\left(\frac{i}{q}\right)! \left(\frac{n-i}{q}\right)!} \leq \frac{(s+t)^{n/q}}{(n/q)!}.$$

Proof of Lemma 17 . We will inductively construct $Q_{s,t}^{[n]}$ for $n > \lfloor q \rfloor$, the base case of the induction following from the assumption on the $Q_{s,t}^j$, $j = 1, \ldots, \lfloor q \rfloor$. The proof closely follows the proof of the classical extension theorem for rough paths (see [19] p.45f). To extend from $n - 1 \geq \lfloor q \rfloor$ to n first let on $\mathcal{P}_{d_2,n}(V)$

$$\widehat{Q}_{s,t} := \sum_{j=1}^{n-1} Q_{s,t}^j.$$

Given any finite partition $\mathcal{D}$ of the interval $[s,t]$ define $Q_{s,t}^{[n],\mathcal{D}}$ by setting

$$Q_{s,t}^{[n],\mathcal{D}} := \prod_{\mathcal{D}} \widehat{Q}_{t_i, t_{i+1}}.$$

By the pigeonhole principle there exists t_j such that

$$(t_{j+1} - t_{j-1}) \leq \frac{2}{|\mathcal{D}| - 1}(t-s)$$

and we may coarsen the partition by dropping t_j and write $\mathcal{D}' := \mathcal{D} \backslash \{t_j\}$. Then

$$Q_{s,t}^{[n],\mathcal{D}} - Q_{s,t}^{[n],\mathcal{D}'} = \widehat{Q}_{s,t_1} \cdots \left(\widehat{Q}_{t_{j-1},t_j} \widehat{Q}_{t_j,t_{j+1}} - \widehat{Q}_{t_{j-1},t_{j+1}} \right) \cdots \widehat{Q}_{t_{|\mathcal{D}|-1},t}$$

and noting that $\widehat{Q}_{t_{j-1},t_j}\widehat{Q}_{t_j,t_{j+1}} - \widehat{Q}_{t_{j-1},t_{j+1}}$ is a homogeneous polynomial of degree n we see that

$$Q_{s,t}^{[n],\mathcal{D}} - Q_{s,t}^{[n],\mathcal{D}'} = \sum_{i=1}^{n-1} Q_{t_{j-1},t_j}^{i} Q_{t_j,t_{j+1}}^{n-i}.$$

Therefore using the submultiplicative property for the norm, the inductive hypothesis and finally the neo-classical inequality we see that

$$\begin{aligned} \left\| \pi_n \left(Q_{s,t}^{[n],\mathcal{D}} - Q_{s,t}^{[n],\mathcal{D}'} \right) \right\| &= \left\| \sum_{i=1}^{n-1} Q_{t_{j-1},t_j}^{i} Q_{t_j,t_{j+1}}^{n-i} \right\| \leq \sum_{i=1}^{n-1} \left\| Q_{t_{j-1},t_j}^{i} \right\| \left\| Q_{t_j,t_{j+1}}^{n-i} \right\| \qquad (47) \\ &\leq \sum_{i=1}^{n-1} \left(\frac{(C\,|t_j - t_{j-1}|)^{i/q}}{\theta\,(i/q)!} \right) \left(\frac{(C\,|t_{j+1} - t_j|)^{(n-i)/q}}{\theta\,((n-i)\,/q)!} \right) \\ &\leq \frac{q^2}{\theta} \left(\frac{2}{|\mathcal{D}| - 1} \right)^{\frac{n}{q}} \frac{(C\,|t-s|)^{\frac{n}{q}}}{\theta\,(n/q)!}. \end{aligned}$$

Successively dropping points from the partition until $\mathcal{D} = \{s,t\}$ we see that

$$\left\| \pi_n \left(Q_{s,t}^{[n],\mathcal{D}} - \widehat{Q}_{s,t} \right) \right\| \leq \frac{q^2}{\theta} \left(1 + 2^{n/q} \left(\zeta \left(\frac{\lfloor q \rfloor + 1}{q} \right) - 1 \right) \right) \frac{(C\,|t-s|)^{\frac{n}{q}}}{\theta\,(n/q)!}.$$

Thus whenever $\theta \geq q^2 \left(1 + 2^{n/q} \left(\zeta \left(\frac{\lfloor q \rfloor + 1}{q} \right) - 1 \right) \right)$ the maximal inequality implies that

$$\left\| \pi_n \left(Q_{s,t}^{[n],\mathcal{D}} \right)^n \right\| \leq \frac{|t-s|^{\frac{n}{q}}}{\theta\,(n/q)!}$$

holds for any partition of $[s,t]$. It remains to verify the existence of the limit $\lim_{|\mathcal{D}| \to 0} Q_{s,t}^{n,\mathcal{D}}$. We proceed as in [19] and exhibit the Cauchy property for the sequence. Suppose $\mathcal{D} = (t_j)$ and $\widetilde{\mathcal{D}}$ are two partitions of mesh size less than δ. Let $\widehat{\mathcal{D}}$ denote the common refinement of the two partitions and let $\widehat{\mathcal{D}}_j = [t_j, t_{j+1}] \cap \widetilde{\mathcal{D}}$. Then

$$Q_{s,t}^{n,\widehat{\mathcal{D}}} - Q_{s,t}^{n,\mathcal{D}} = \sum Q_{t_0,t_1}^{n,\widehat{\mathcal{D}}_0} \ldots Q_{t_{j-1},t_j}^{n,\widehat{\mathcal{D}}_{j-1}} \left(Q_{t_j,t_{j+1}}^{n,\widehat{\mathcal{D}}_j} - \widehat{Q}_{t_j,t_{j+1}} \right) \ldots Q_{t_{|\mathcal{D}|-1},t}^{n,\widehat{\mathcal{D}}_j}.$$

As seen before this is a sum of homogeneous polynomials of degree n and by the maximal inequality

$$\left\| \pi_n \left(Q_{s,t}^{n,\widehat{\mathcal{D}}} - Q_{s,t}^{n,\mathcal{D}} \right) \right\| \leq \sum_{\mathcal{D}} \frac{|t_{j+1} - t_j|^{\frac{n}{q}}}{\theta\,(n/q)!} \leq \frac{|t-s|}{\theta\,(n/q)!} \delta^{\frac{n}{q} - 1}$$

as $\frac{n}{q} - 1 > 0$ we have a uniform estimate in δ independent of the choice of partition. Going through the same argument for the partition $\widetilde{\mathcal{D}}$ and using the triangle inequality the Cauchy property is established and the existence of the limit follows. The uniqueness of the limit follows as in [19]. The difference of two multiplicative functionals that agree up to level $\lfloor q \rfloor$ is additive (see Lyons [18] Lemma 2.2.3) As the difference of the extensions is also a continuous path and by assumption

$$\left\| \overline{Q}_{s,t}^{\lfloor q \rfloor+1} - \widetilde{Q}_{s,t}^{\lfloor q \rfloor+1} \right\| \leq C\left(\lfloor q \rfloor+1\right)|t-s|^{\frac{\lfloor q \rfloor+1}{q}}$$

it follows that $\overline{Q}_{s,t}^{\lfloor q \rfloor+1} - \widetilde{Q}_{s,t}^{\lfloor q \rfloor+1}$ is identically zero. A simple induction now completes the proof. ■

Lemma 19 *For any $1/3 < \gamma < 1/2$ there exist a constant $\theta > 0$ and random variables $c(\gamma, \omega)$, almost surely finite, such that*

$$\left\| R_{s,t}^{n,\bar{\imath}} \right\|_{H^1 \to H^1} \leq \frac{\left(c\left(\gamma, \omega\right)|t-s|\right)^{n\gamma}}{\theta\left(n\gamma\right)!}. \tag{48}$$

and

$$\left\| R_{s,t}^{n,\bar{\imath}} \right\|_{H^{-1} \to H^{-1}} \leq \frac{\left(c\left(\gamma, \omega\right)|t-s|\right)^{n\gamma}}{\theta\left(n\gamma\right)!} \tag{49}$$

for all $\bar{\imath} \in S\left(n\right)$, $n \in \mathbb{N}$, $0 < s < t \leq 1$.

Proof. We now take for V the space of bounded linear operators on (the completion of) H^1 and H^{-1} respectively. From the a priori estimates we know that $Q_{s,t}^{[n]} \in \mathcal{P}_{d_2,n}\left(V\right)$ for all $n \geq 1$. First note that by Lemma 16 $\left(Q_{s,t}^1, Q_{s,t}^2\right)$ satisfies the assumptions of Lemma 17 with $3 > q = 1/\gamma$ and therefore has a multiplicative extension $\widetilde{Q}_{s,t}^j$ controlled in the sense of (46). Once again by Lemma 16 the uniqueness part of Lemma 17 applies and we deduce that $Q_{s,t}^j = \widetilde{Q}_{s,t}^j$ for $j \in \mathbb{N}$. ■

Armed with these two factorially decaying a priori estimates we are finally ready proof a regularity estimate for Q^n that decays factorially in n. When considering $R_{s,t}^{n,\bar{\imath}}, \bar{\imath} \in S\left(n\right)$ as an operator from H^{-1} to H^1 we cannot directly apply Lemma 17 as the a priori bounds in Lemma 16 have singularities for small n. Instead we exploit that there is more than one way to estimate the operator norm of the composition of such operators. Together with the estimates already obtained in Lemma 19 this will be sufficient to prove factorially decaying bounds for n sufficiently large. We recall Proposition 10 and restate it in the notation of the current section.

Proposition 10: *Let $1/3 < \gamma < 1/2$ be fixed. There exists $\theta > 0$, $\gamma' \in (1/3, \gamma)$,* $m_0\left(\gamma'\right) \in \mathbb{N}$ and *random variables $c(\gamma', \omega)$, almost surely finite, such that*

$$\left\| R_{s,t}^{n,\bar{\imath}} \right\|_{H^{-1} \to H^1} \leq \frac{\left(c\left(\gamma', \omega\right)|t-s|\right)^{n\gamma'}}{\theta\left(n\gamma'\right)!} \tag{50}$$

for all $n \geq m_0$, $\bar{\imath} \in S\left(n\right)$ and $t \in (0, 1]$.

Before we begin the proof note that by choosing $\gamma' < \gamma$ we have for n sufficiently large by Lemma 16

$$\left\| R_{s,t}^{n,\bar{\imath}} \right\|_{H^{-1} \to H^1} \leq c\left(\gamma, n, \omega\right)|t-s|^{n\gamma - \bar{k}} \leq c\left(\gamma, n, \omega\right)|t-s|^{n\gamma'}$$

for all $0 < s < t < 1$.

Proof of Proposition 10. Choose m_0 and $0 < \gamma' \leq \gamma$ such that $\gamma n - \bar{k} \geq \gamma' n$ for all $n \geq m_0$. Using Corollary 19 and Lemma 16 (with $\gamma = \gamma'$) we can find $c(\gamma', \omega)$ such that simultaneously (50) holds for all $n \in [m_0, 2m_0]$ and the two inequalities (48) and (49) hold for all $n \in \mathbb{N}$. Note that this also serves as the base case for our induction argument. For this lemma we set V to be the space of bounded linear operators from H^{-1} to H^1.

We argue now exactly as in the proof Lemma 17 to extend the functional from level $n \geq 2m_0$ to $n+1$, with the only difference being that we have no direct control over $\left\| R_{s,t}^{k,\bar{\imath}} \right\|_{H^{-1}\to H^1}$ for $\bar{\imath} \in S(k)$, $k < m_0$. We therefore replace inequality (47) with the following more refined estimate that exploits that the operator norm of a composition of two operators can be estimated in several ways, which allows us to draw on our a priori estimates in Lemma 16. We have

$$\begin{aligned}
\left\| \pi_{n+1}\left(Q_{s,t}^{D,n+1} - Q_{s,t}^{D',n+1} \right) \right\| &\leq \sum_{i=1}^{n} \left\| Q_{t_{j-1},t_j}^{i} Q_{t_j,t_{j+1}}^{n+1-i} \right\| \\
&= \sum_{i=1}^{n} \left\| \sum_{\bar{m}\in S(i)} \sum_{\bar{l}\in S(n+1-i)} R_{t_{j-1},t_j}^{i,\bar{m}} R_{t_j,t_{j+1}}^{n+1-i,\bar{l}} \varepsilon_{\bar{m}*\bar{l}} \right\| \\
&= \sum_{i=1}^{n} \sup_{\bar{m}\in S(i),\bar{l}\in S(n+1-i)} \left\| R_{t_{j-1},t_j}^{i,\bar{m}} R_{t_j,t_{j+1}}^{n+1-i,\bar{l}} \right\|_{H^{-1}\to H^1} \\
&\leq \sum_{i=1}^{m_0-1} \sup_{\bar{m}\in S(i)} \left\| R_{t_{j-1},t_j}^{i,\bar{m}} \right\|_{H^1\to H^1} \sup_{\bar{l}\in S(n+1-i)} \left\| R_{t_j,t_{j+1}}^{n+1-i,\bar{l}} \right\|_{H^{-1}\to H^1} \\
&+ \sum_{i=m_0}^{n} \sup_{\bar{m}\in S(i)} \left\| R_{t_{j-1},t_j}^{i,\bar{m}} \right\|_{H^{-1}\to H^1} \sup_{\bar{l}\in S(n+1-i)} \left\| R_{t_j,t_{j+1}}^{n+1-i,\bar{l}} \right\|_{H^{-1}\to H^{-1}} \\
&\leq \sum_{j=1}^{n} \frac{(C|t-s|)^{\gamma' j}}{\theta(j\gamma')!} \frac{(C|t-s|)^{\gamma' m+1-j}}{\theta((m+1-j)\gamma')!}.
\end{aligned} \tag{51}$$

The bounds for $\left\| R_{t_{j-1},t_j}^{i,\bar{m}} \right\|_{H^1\to H^1}$ and $\left\| R_{t_j,t_{j+1}}^{n+1-i,\bar{l}} \right\|_{H^{-1}\to H^{-1}}$ use inequalities (48) and (49) respectively. The bounds for $\sup_{\bar{l}\in S(n+1-i)} \left\| R_{t_j,t_{j+1}}^{n+1-i,\bar{l}} \right\|_{H^{-1}\to H^1}$ and $\sup_{\bar{m}\in S(i)} \left\| R_{t_{j-1},t_j}^{i,\bar{m}} \right\|_{H^{-1}\to H}$ follow (for the appropriate values of i) from the inductive hypothesis. With this modification in place arguing exactly as in the proof of Lemma 17 yields the result. Note that the extension is only carried out for $n \geq 2m_0$. For $m_0 \leq n < 2m_0$ the estimates use the a priori bounds. ■

Remark 20 *To extend the proof of Proposition 10 to cover the terms in the expansion of* ρ_t^* *we make the following modifications. In place of* $R_{s,t}^{n,\bar{\imath}}$ *we have*

$$X_{s,t}^{n,\bar{\imath}} = \int_{\Delta_{s,t}^{k}} P_{t-t_n} H_{i_n} P_{t_n - t_{n-1}} H_{i_{n-1}} \cdots H_{i_1} P_{t_1 - s} dY_{t_1}^{i_1} \cdots dY_{t_n}^{i_n},$$

i.e. the order of non-commutative product in the integrand is reversed. We therefore define

$\bar{\mathcal{P}}_{d_2,k}(V)$ *as* $\mathcal{P}_{d_2,k}(V)$ *but with the multiplication in (41) replaced by*

$$ab := \sum_{v=0}^{k}\sum_{j=0}^{v}\sum_{\bar{\imath}\in S(j)}\sum_{\bar{l}\in S(v-j)} b_{\bar{l}}a_{\bar{\imath}}\varepsilon_{\bar{\imath}*\bar{l}}. \tag{52}$$

With this modification (43) becomes

$$\begin{aligned}
X_{s,t}^{k,\bar{\imath}} &= \int_{\Delta_{s,u}^{k}} P_{t-t_k}H_{i_k}P_{t_k-t_{k-1}}H_{i_{k-1}}\cdots H_{i_1}P_{t_1-s}dY_{t_1}^{i_1}\cdots dY_{t_k}^{i_k} \\
&+\int_{\Delta_{u,t}^{k}} P_{t-t_k}H_{i_k}P_{t_k-t_{k-1}}H_{i_{k-1}}\cdots H_{i_1}P_{t_1-s}dY_{t_1}^{i_1}\cdots dY_{t_k}^{i_k} \\
&+\sum_{j=1}^{k-1}\int_{\Delta_{u,t}^{k-j}} P_{t-t_k}H_{i_k}P_{t_k-t_{k-1}}\cdots P_{t_{j+2}-t_{j+1}}H_{i_{j+1}}P_{t_{j+1}-u}dY_{t_{j+1}}^{i_{j+1}}\cdots dY_{t_k}^{i_k} \\
&+\int_{\Delta_{s,u}^{k}} P_{u-t_j}H_{i_j}P_{t_j-t_{j-1}}\cdots H_{i_1}P_{t_1-s}dY_{t_1}^{i_1}\cdots dY_{t_j}^{i_j} \\
&=\sum_{j=0}^{k} X_{u,t}^{k-j,(i_{j+1},\ldots i_k)}X_{s,u}^{j,(i_1,\ldots i_j)} \\
&=\sum_{\bar{m}*\bar{l}=\bar{\imath}} X_{u,t}^{|\bar{l}|,\bar{l}}X_{s,u}^{|\bar{m}|,\bar{m}}.
\end{aligned}$$

Combining this identity with the modified multiplication (52) we see that (44) holds on $\bar{\mathcal{P}}_{d_2,k}(V)$*, i.e. our functional* $\bar{Q}_{s,t}^{[n]} = \sum_{j=0}^{n}\sum_{\bar{\imath}\in S(j)} X_{s,t}^{j,\bar{\imath}}\varepsilon_{\bar{\imath}}$ *has the multiplicative property. Going through the same steps as before with these modifications in place the proof of Proposition 10 may now be completed.*

Remark 21 *The arguments in this section may easily be generalised to higher derivatives. For example, one may define spaces* H^2 *and* H^{-2} *analogous to* H^1 *and* H^{-1} *by setting*

$$\|\varphi\|_{H^{-2}} := \inf\left\{\sum_{\alpha\in A_0(\bar{k}),\beta\in A_0(\bar{k})}\|\varphi_{\alpha,\beta}\|_\infty : \varphi = \sum_{\alpha\in A_0(\bar{k}),\beta\in A_0(\bar{k})} V_{[\alpha]}V_{[\beta]}\varphi_{\alpha,\beta},\ \varphi_{a,\beta}\in C_b^\infty(\mathbb{R}^N)\right\}$$

and

$$\|\varphi\|_{H^2} := \sum_{\alpha\in A_0(\bar{k}),\beta\in A_0(\bar{k})}\left\|V_{[\alpha]}V_{[\beta]}\varphi\right\|_\infty.$$

and study the operator norm between H^{-2} *and* H^2*. Note that in this case in the proof of Proposition 10 we choose* m_0 *such that* $\gamma n - 2\bar{k} \geq \gamma' n$ *for all* $n \geq m_0$*, i.e. the cut off between the a priori estimates and the factorially decaying estimates we use in the proof of the main theorem depends explicitly on the number of derivatives we consider.*

7 Appendix:

7.1 Proof of the inequality (15)

As $Z_t^x \geq 0$, we have, by Jensen's inequality, that $\tilde{\mathbb{E}}\left[Z_t^x|\mathcal{Y}_\cdot^x\right]^{-1} \leq \tilde{\mathbb{E}}\left[(Z_t^x)^{-1}|\mathcal{Y}_\cdot^x\right]$. Then observe that, by integration by parts

$$
\begin{aligned}
-\sum_{i=1}^{d_2}\int_0^t h^i(X_s^x)\,dY_s^{x,i} &= \sum_{i=1}^{d_2}\left(-h^i(X_t^x)\,Y_t^{x,i} + \int_0^t Y_s^{x,i}Ah^i(X_s^x)ds + \sum_{j=1}^{d}\int_0^t Y_s^{x,i}V^jh^i(X_s^x)\,dB_s^j\right)\\
&\leq \sum_{i=1}^{d_2}\sup_{s\in[0,t]}\left|Y_s^{x,i}\right|\left(\left|\left|h^i\right|\right|_\infty + t\left|\left|Ah^i\right|\right|_\infty\right)\\
&\quad + \frac{d_2t}{2}\sum_{i=1}^{d_2}\sup_{s\in[0,t]}\left|Y_s^{x,i}\right|^2\left(\sum_{j=1}^{d}\left|\left|V^jh^i\right|\right|^2\right) + \eta_t^x,
\end{aligned}
$$

where

$$
\eta_t^x = \sum_{j=1}^{d}\int_0^t\left(\sum_{i=1}^{d_2}Y_s^{x,i}V^jh^i(X_s^x)\right)dB_s^j - \sum_{j=1}^{d}\frac{1}{2}\int_0^t\left(\sum_{i=1}^{d_2}Y_s^{x,i}V^jh^i(X_s^x)\right)^2 ds.
$$

Since, $\tilde{\mathbb{E}}\left[\exp\eta_t^x|\,\mathcal{Y}_t^x\right] = 1$,we get that

$$
\left(1/\rho_t^{Y(\omega)}(1)\right) < \exp C\left(\sum_{i=1}^{d_2}\sup_{s\in[0,t]}\left|Y_s^i(\omega)\right| + \sup_{s\in[0,t]}\left|Y_s^i(\omega)\right|^2\right),
$$

where C is a constant independent of x,

$$
C = \max_{i=1,\ldots,d}(\left|\left|h^i\right|\right|_\infty + t\left|\left|Ah^i\right|\right|_\infty + \frac{d_2t}{2}\sum_{j=1}^{d}\left|\left|V^jh^i\right|\right|^2).
$$

Inequality (15) follows as $\sup_{s\in[0,t]}\left|Y_s^i(\omega)\right|$ is finite for almost every ω.

7.2 Proof of Proposition 15

The remainder of the paper is dedicated to the proof of Proposition 15, which requires us to prove a number of elementary lemmas in preparation.

Recall that $\Delta_{s,t}^k$ denotes the simplex defined by the relation $s = t_0 < t_1 < \cdots < t_k < t$. In Proposition 15 we would like to obtain estimates of the form

$$
\left|\left|V_{[\alpha]}\bar{R}_{(0,t_1,\ldots,t_k,t)}^{m,\bar{j}_1,\ldots,\bar{j}_k}V_{[\beta]}\varphi\right|\right|_\infty \leq c_m t^{-(\|\alpha\|+\|\beta\|)/2}\frac{1}{\sqrt{t_1-t_0}}\cdots\frac{1}{\sqrt{t_k-t_{k-1}}}\|\varphi\|_\infty
$$

that are (essentially) uniform across the simplex $\Delta_{0,t}^k$. The basic idea is that for any $(t_1,\ldots,t_k)\in\Delta_{0,t}^k$ there exists always at least one time interval $[t_j,t_{j-1}]$ that is of length at least $t/\,(k+1)$.

We then use the Kusuoka-Strock regularity estimates (Theorem 8) to deduce smoothness of the heat semigroup over this particular interval. The proof of Theorem 8 employs the methods of Malliavin calculus. As we will in the following draw on elements of their method we recall some basic concepts of the Malliavin calculus.

Let $(\Theta, \mathcal{H}, \mu)$ be the abstract Wiener space and let $\mathcal{L}$ denote the Ornstein Uhlenbeck operator defined as in Kusuoka [13]. Denote by $G(\mathcal{L})$ the set of arbitrarily often Malliavin differentiable real valued random variables on Θ and denote by D_p^s, the usual Kusuoka-Stroock Sobolev spaces based on the Ornstein-Uhlenbeck operator (see e.g. Kusuoka [13] or [12] for details). The following definition is taken from Kusuoka [12], p.267.

Definition 22 *Let $r \in \mathbb{R}$ and K_r denote the set of functions $f : (0,1] \times \mathbb{R}^N \to G(\mathcal{L})$ satisfying the following conditions*

1. *$f(t,x)$ is smooth in x and $\frac{\partial^\nu f}{\partial^\nu x}$ is continuous in $(t,x) \in (0,1] \times \mathbb{R}^N$ with probability one for any multi-index ν*

2. $$\sup_{t \in (0,1], x \in R^N} t^{-r/2} \left\| \frac{\partial^\nu f}{\partial^\nu x}(t,x) \right\|_{D_p^s} < \infty$$

   for any $s \in \mathbb{R}$, $p \in (1, \infty)$.

For $\Phi \in \mathcal{K}_r$, $\varphi \in C_b^\infty$ define $P_t^\Phi \varphi = E(\Phi(t,x)\varphi(X_t(x)))$. An important ingredient in the proof of Theorem 8 which we will use repeatedly is the following Lemma (see Kusuoka [12] Corollary 9).

Lemma 23 (Kusuoka) *Let $r \in \mathbb{R}$, $\Phi \in \mathcal{K}_r$ and $\alpha \in A_1(\bar{k})$. Then there are $\Phi_{\alpha,1}$, $\Phi_{\alpha,2} \in \mathcal{K}_{r-\|\alpha\|}$ such that*

$$P_t^\Phi V_{[\alpha]} = P_t^{\Phi_{\alpha,1}} \text{ and } V_{[\alpha]} P_t^\Phi = P_t^{\Phi_{\alpha,2}}. \tag{53}$$

Moreover there exists C such that

$$\left\| P_t^\Phi \varphi \right\|_\infty \leq t^{r/2} \left\| \varphi \right\|_\infty$$

for any $\varphi \in C_b^\infty(\mathbb{R}^N)$ and $t \in (0,1]$.

Before we proceed we gather some simple properties of the spaces $\mathcal{K}_r$. The following Lemma may be found in Kusuoka [12] (Lemma 7).

Lemma 24 *Let $r_1, r_2 \in \mathbb{R}$. Then*

1. *If $f_1 \in \mathcal{K}_{r_1}$, and $f_2 \in \mathcal{K}_{r_2}$ then $f_1 f_2 \in \mathcal{K}_{r_1+r_2}$*
2. *If $\varphi \in C_b^\infty(\mathbb{R}^N)$ then $\varphi(X_t(x)) \in \mathcal{K}_0$*

3. *For any* $\alpha, \beta \in A_1\left(\bar{k}\right)$ *there exist* a_α^β, $b_\alpha^\beta \in \mathcal{K}_{(\|\beta\|-\|\alpha\|\vee 0)}$ *such that*

$$\left(\left(X_t^{-1}\right)_* V_{[\alpha]}\right)(x) = \sum_{\beta\in A_1\left(\bar{k}\right)} a_\alpha^\beta (t,x) V_{[\beta]}(x)$$

and

$$V_{[\alpha]}(x) = \sum_{\beta\in A_1\left(\bar{k}\right)} b_\alpha^\beta (t,x) \left(\left(X_t^{-1}\right)_* V_{[\beta]}\right)(x).$$

Proof. The claims (2) and (3) are shown in [12] (Lemma 7). For (1) note that the space $\bigcap_{1<p<\infty} D_p^s(\mathbb{R})$ is an algebra (Kusuoka [13] Lemma 2.13) and $\|fg\|_{D_p^s} \leq \|f\|_{D_r^s}\|g\|_{D_q^s}$ for $\frac{1}{p} = \frac{1}{r} + \frac{1}{q}$, and any $f, g \in \bigcap_{1<p<\infty} D_p^s$ and

$$\begin{aligned} &\sup_{t\in(0,1],x\in R^N} t^{-(r_1+r_2)/2} \left\|\frac{\partial (f_1 f_2)}{\partial x}(t,x)\right\|_{D_p^s} \\ &\leq \sum_{1\leq i,j\leq 2, i\neq j} \sup_{t\in(0,1],x\in R^N} t^{-r_i/2}\left\|\frac{\partial f_i}{\partial x}(t,x)\right\|_{D_r^s} \sup_{t\in(0,1],x\in R^N} t^{-r_j/2}\|f_j(t,x)\|_{D_q^s} < \infty. \end{aligned}$$

The generalisation to higher derivatives is clear and the claim follows. ■

In particular the Lemma implies that for any multi-index γ, $p \in [1,\infty)$

$$\sup_{x\in R^N} E\left[\sup_{t\in(0,1]}\left|\frac{\partial^{|\gamma|}}{\partial x^\gamma} a_\alpha^\beta(t,x)\right|^p\right] < \infty$$

and

$$\sup_{x\in R^N} E\left[\sup_{t\in(0,1]}\left|\frac{\partial^{|\gamma|}}{\partial x^\gamma} b_\alpha^\beta(t,x)\right|^p\right] < \infty.$$

Let $J_t^{ij}(x) = \frac{\partial}{\partial x_i} X^j(t,x)$ and note that for any C_b^∞ vector field W we have

$$\left((X_t)_* W\right)^i (X_t(x)) = \sum_{j=1}^N J_t^{ij}(x) W^j.$$

Suppose $\Phi \in \mathcal{K}_r$. Then

$$V_{[\alpha]} P_t^\Phi \varphi(x) = E\left[V_{[\alpha]}\Phi\varphi(X_t(x)) + \sum_{i,j=1}^N \Phi V_{[\alpha]}^j(x)\left(\frac{\partial}{\partial x^j}\varphi\right)(X_t(x)) J_t^{ij}(x)\right].$$

It is straightforward to see that $V_{[\alpha]}\Phi \in \mathcal{K}_r$ and for the second term in the sum we have

$$\begin{aligned}
&E\left[\sum_{i,j=1}^{N} \Phi V_{[\alpha]}^{j}\left(x\right)\left(\frac{\partial}{\partial x^{i}}\varphi\right)\left(X_{t}\left(x\right)\right)J_{t}^{ij}\left(x\right)\right] \\
&= E\left[\sum_{i,j=1}^{N} \Phi \sum_{\beta\in A_{1}\left(\bar{k}\right)} b_{\alpha}^{\beta}\left(t,x\right)\left(\left(X_{t}^{-1}\right)_{*}V_{[\beta]}\right)^{j}\left(x\right)\left(\frac{\partial}{\partial x^{i}}\varphi\right)\left(X_{t}\left(x\right)\right)J_{t}^{ij}\left(x\right)\right] \\
&= E\left[\Phi \sum_{\beta\in A_{1}\left(\bar{k}\right)} b_{\alpha}^{\beta}\left(t,x\right)\sum_{i=1}^{N}\left(\left(X_{t}\right)_{*}\left(X_{t}^{-1}\right)_{*}V_{[\beta]}\right)^{i}\left(X_{t}\left(x\right)\right)\left(\frac{\partial}{\partial x^{i}}\varphi\right)\left(X_{t}\left(x\right)\right)\right] \\
&= E\left[\Phi \sum_{\beta\in A_{1}\left(\bar{k}\right)} b_{\alpha}^{\beta}\left(t,x\right)\sum_{i=1}^{N}V_{[\beta]}^{i}\left(X_{t}\left(x\right)\right)\left(\frac{\partial}{\partial x^{i}}\varphi\right)\left(X_{t}\left(x\right)\right)\right] \\
&= \sum_{\beta\in A_{1}\left(\bar{k}\right)} P^{\Phi b_{\alpha}^{\beta}}\left(V_{[\beta]}\varphi\right)\left(x\right).
\end{aligned}$$

Note that by Lemma 24 $\Phi b_{\alpha}^{\beta}\left(t,x\right) \in \mathcal{K}_{(\|\beta\|-\|\alpha\|\vee 0)+r}$. We have just proved the following Lemma (see e.g. Kusuoka [12] Corollary 9).

Lemma 25 *Let* $\Phi \in \mathcal{K}_r$ *and* $\alpha \in A_1\left(\bar{k}\right)$ *then* $V_{[\alpha]}\Phi \in \mathcal{K}_r$ *and there exist* $\Phi b_{\alpha}^{\beta} \in \mathcal{K}_{(\|\beta\|-\|\alpha\|\vee 0)+r}$ *such that we have*

$$V_{[\alpha]}P_t^{\Phi}\varphi\left(x\right) = P^{V_{[\alpha]}\Phi}\varphi\left(x\right) + \sum_{\beta\in A_{1}\left(\bar{k}\right)} P^{\Phi b_{\alpha}^{\beta}}\left(V_{[\beta]}\varphi\right)\left(x\right),$$

for all $\varphi \in C_b^{\infty}(\mathbb{R}^N)$.

The following Lemma is an immediate consequence of Lemma 23.

Lemma 26 *Let* $\Phi \in \mathcal{K}_r$ *and* $\alpha \in A_1\left(\bar{k}\right)$ *then there exists* $C > 0$ *such that*

$$\left\|V_{[\alpha]}P_t^{\Phi}\varphi\right\|_{\infty} \leq C \sum_{\beta\in A_{0}\left(\bar{k}\right)} \min\left(t^{r/2}, t^{(\|\beta\|-\|\alpha\|)/2+r/2}\right)\left\|V_{[\beta]}\varphi\right\|_{\infty}$$

for all $t \in (0,1]$, $\varphi \in C_b^{\infty}(\mathbb{R}^N)$. *In particular, if* H *is of the form* $H = u\ V_i + v$ *for some* $u, v \in C_b^{\infty}$, $i \in \{1, \ldots, d_1\}$ *and* $\Phi \in \mathcal{K}_0$ *we have*

$$\left\|V_{[\alpha_1]}P_t^{\Phi}H\varphi\left(x\right)\right\|_{\infty} \leq C \sum_{\beta\in A_{0}\left(\bar{k}\right)} \min\left(t^{-1/2}, t^{(\|\beta\|-\|\alpha\|)/2-1/2}\right)\left\|V_{[\beta]}\varphi\right\|_{\infty}.$$

Proof. By Lemma 25 there exist $\Phi_\beta \in \mathcal{K}_{[(\|\beta\|-\|\alpha\|)\vee 0]+r}$. such that

$$\left\|V_{[\alpha]} P_t^\Phi \varphi(x)\right\|_\infty \leq \sum_{\beta\in A_0(\bar{k})} \left\|P_t^{\Phi_\beta} V_{[\beta]}\varphi\right\|_\infty$$
$$\leq C \sum_{\beta\in A_0(\bar{k})} \min\left(t^{r/2}, t^{(\|\beta\|-\|\alpha\|)/2+r/2}\right) \left\|V_{[\beta]}\varphi\right\|_\infty .$$

The last inequality is a consequence of Lemma 23 (2). To deduce the second claim from the first of the proposition we note that by [12] Corollary 9 (2) if $\Phi \in \Phi_a \in \mathcal{K}_r$ there exists $\Phi_a \in \mathcal{K}_{r-\|\alpha\|}$ such that $P_t^\Phi V_i = P_t^{\Phi_a}$. ■

Intuitively the preceding lemma provides us with a uniform (for small times) bound when we move derivatives through the heat kernel from the outside to the inside.

We now consider the reverse situation in which we move the vector fields from the inside to the outside. We have the following Lemma.

Lemma 27 *Let $\Phi \in \mathcal{K}_r$ and $\alpha \in A_1(\bar{k})$ then there exists $\Phi_\beta \in \mathcal{K}_r$ and $\Phi a_\alpha^\beta \in \mathcal{K}_{(\|\beta\|-\|\alpha\|\vee 0)+r}$ such that*

$$\left(P_t^\Phi V_{[\alpha]}\varphi\right)(x) = \sum_{\beta\in A_1(\bar{k})} \left\{V_{[\beta]} P_t^{a_\alpha^\beta \Phi}\varphi(x) - P_t^{\Phi_\beta}\varphi\right\},$$

for all $\varphi \in C_b^\infty(\mathbb{R}^N)$.

Proof. We have using Lemma 24 part 3. that

$$\begin{aligned}
\left(P_t^\Phi V_{[\alpha]}\varphi\right)(x) &= E\left[\Phi \sum_{i=1}^N V_{[\alpha]}^i(X_t(x)) \left(\frac{\partial}{\partial x_i}\varphi\right)(X_t(x))\right] \\
&= E\left[\Phi \sum_{i=1}^N \left((X_t)_* \left(X_t^{-1}\right)_* V_{[\alpha]}\right)^i (X_t(x)) \left(\frac{\partial}{\partial x_i}\varphi\right)(X_t(x))\right] \\
&= E\left[\Phi \sum_{i,j=1}^N \left(\left(X_t^{-1}\right)_* V_{[\alpha]}(x)\right)^j J_t^{ij}(x) \left(\frac{\partial}{\partial x_i}\varphi\right)(X_t(x))\right] \\
&= E\left[\Phi \sum_{\beta\in A_1(\bar{k})} a_\alpha^\beta(t,x) \sum_{j=1}^N V_{[\beta]}^j(x) \sum_{i=1}^N J_t^{ij}(x) \left(\frac{\partial}{\partial x_i}\varphi\right)(X_t(x))\right]. \\
&= \sum_{\beta\in A_1(\bar{k})} E\left[\Phi a_\alpha^\beta(t,x) \sum_{j=1}^N V_{[\beta]}^j(x) \frac{\partial}{\partial x_j}\varphi(X_t(x))\right],
\end{aligned}$$

where $\Phi a_\alpha^\beta \in \mathcal{K}_{(\|\beta\|-\|\alpha\|\vee 0)+r}$. On the other hand we have

$$\begin{aligned}
&V_{[\beta]} P_t^{a_\alpha^\beta \Phi}\varphi(x) \\
&= E\left[\Phi a_\alpha^\beta(t,x) \sum_{j=1}^N V_{[\beta]}^j(x) \frac{\partial}{\partial x_j}\varphi(X_t(x))\right] + E\left[V_{[\beta]}\left(\Phi a_\alpha^\beta\right)(t,x)\varphi(X_t(x))\right]
\end{aligned}$$

and deduce that

$$\left(P_t^{\Phi}V_{[\alpha]}\varphi\right)(x) = \sum_{\beta\in A_1(\bar{k})} \left\{ V_{[\beta]}P_t^{a_\alpha^\beta\Phi}\varphi(x) - E\left[V_{[\beta]}\left(\Phi a_\alpha^\beta\right)(t,x)\,\varphi(X(t,x))\right]\right\},$$

where $V_{[\beta]}\left(a_\alpha^\beta\Phi\right)(t,x)\in\mathcal{K}_r$ and $a_\alpha^\beta\in\mathcal{K}_{(\|\beta\|-\|\alpha\|\vee 0)}$. ■

The representation obtained in the previous lemma generalises to multiple heat kernels as we observe in the following proposition.

Proposition 28 *Let $k\in\mathbb{N}$, $\Phi_k\in\mathcal{K}_r$, $\Phi_j\in\mathcal{K}_0$ for $1\le j<k$, $\alpha\in A_1\left(\bar{k}\right)$, and $H_j = u_jV_{i_j}+v_j$, where $1\le i_j\le d$, $u_j,v_j\in C_b^\infty(\mathbb{R}^N)$, $j=1,\ldots,k-1$. Then there exist $\Phi_{\beta^1}\in\mathcal{K}_{r_1}$, $\ldots,\Phi_{\beta^k}\in\mathcal{K}_{r_k}$ such that $r_k\ge r$, $r_1,\ldots r_{k-1}\ge -1/2$ and*

$$r_1+r_2+\cdots+r_k \ge \left(\left\|\beta^1\right\|-\|\alpha\|\right)\vee 0-(k-1)/2+r \tag{54}$$

and

$$P_{t_1}^{\Phi_1}H_1P_{t_2}^{\Phi_2}\cdots H_{k-1}P_{t_k}^{\Phi_k}V_{[\alpha]}\varphi(x) = \sum_{\beta^1\in A_0(\bar{k})}\cdots\sum_{\beta^k\in A_0(\bar{k})} V_{[\beta^1]}P_{t_1}^{\Phi_{\beta^1}}P_{t_2}^{\Phi_{\beta^2}}\cdots P_{t_k}^{\Phi_{\beta^k}}\varphi(x)$$

holds for all $\varphi\in C_b^\infty(\mathbb{R}^N)$.

Before we begin the proof of this proposition we examine the meaning of the assumptions on the r_j. The assumptions $r_1,\ldots,r_{k-1}\ge -1/2$ imply that singularities in the bounds

$$\left\|P_t^{\Phi_{r_j}}\varphi\right\|_\infty \le t^{r_j/2}\|\varphi\|_\infty$$

in Lemma 23 are integrable. The inequality (54) can be interpreted as follows: The left hand side is the total regularity of the resulting expression in the proposition. For every application of an operator H we loose 1/2 regularity reflected in the term $-(k-1)/2$. The degree of a singularity introduced by differentiating by $V_{[\alpha]}$ depends on $\|\alpha\|$. Thus if $\|\beta\|>\|\alpha\|$ and we replace a $V_{[\alpha]}$ by $V_{[\beta]}$ we expect a compensating term, which is captured in $\left(\left\|\beta^1\right\|-\|\alpha\|\right)\vee 0$.

Proof. As before it is by linearity sufficient to consider the case $H_j=u_jV_{i_j}$, for some $u_j\in C_b^\infty(\mathbb{R}^N)$ the case of the multiplication operator v_j following by a similar but easier calculation. We argue by induction, the base case being covered by Lemma 27. For the inductive step we note that if $\Phi_k\in\mathcal{K}_0$ then by Lemma 23 there exists $\bar{\Phi}_k\in\mathcal{K}_{-1/2}$ such that $P_t^{\Phi_k}uV_i = P_t^{\bar{\Phi}_k}$. Combining this fact with Lemma 27 we see

$$\begin{aligned}
P_{t_1}^{\Phi_1}H_1P_{t_2}^{\Phi_2}\cdots H_{k-1}P_{t_k}^{\Phi_k}HP_t^{\Phi}V_{[\alpha]}\varphi(x) &= P_{t_1}^{\Phi_1}H_1P_{t_2}^{\Phi_2}\cdots H_{k-1}P_{t_k}^{\Phi_k}uV_iP_t^{\Phi}V_{[\alpha]}\varphi(x)\\
&= P_{t_1}^{\Phi_1}H_1P_{t_2}^{\Phi_2}\cdots H_{k-1}P_{t_k}^{\bar{\Phi}_k}P_t^{\Phi}V_{[\alpha]}\varphi(x)\\
&= \sum_{\beta\in A_0(\bar{k})} P_{t_1}^{\Phi_1}H_1P_{t_2}^{\Phi_2}\cdots H_{k-1}P_{t_k}^{\bar{\Phi}_k}V_{[\beta]}P_t^{\Phi_\beta}\varphi(x),
\end{aligned}$$

where $\Phi_\beta \in \mathcal{K}_{[(\|\beta\|-\|\alpha\|)\vee 0]+r}$ and $\overline{\Phi}_k \in \mathcal{K}_{-1/2}$ Using the inductive hypothesis we get

$$\sum_{\beta\in A_0(\bar{k})} P_{t_1}^{\Phi_1} H_1 P_{t_2}^{\Phi_2}\cdots H_{k-1}P_{t_k}^{\overline{\Phi}_k} V_{[\beta]} P_t^{\Phi_\beta}\varphi(x)$$
$$= \sum_{\beta^1\in A_0(\bar{k})}\cdots\sum_{\beta^k\in A_0(\bar{k})}\sum_{\beta\in A_0(\bar{k})} V_{[\beta^1]} P_{t_1}^{\Phi_{\beta^1}} P_{t_2}^{\Phi_{\beta^2}}\cdots P_{t_k}^{\Phi_{\beta^k}} P_t^{\Phi_\beta}\varphi(x).$$

From the inductive hypothesis we know that $\Phi_{\beta^1} \in \mathcal{K}_{r_1}, \ldots, \Phi_{\beta^k} \in \mathcal{K}_{r_k}$ such that $r_1, \ldots, r_k \geq -1/2$ (using that $\overline{\Phi}_k \in \mathcal{K}_{-1/2}$) and

$$r_1 + r_2 + \cdots + r_k \geq \left(\left\|\beta^1\right\| - \|\beta\|\right)\vee 0 - k/2.$$

Hence, as required

$$[(\|\beta\|-\|\alpha\|)\vee 0] + r + r_1 + r_2 + \cdots + r_k$$
$$\geq [(\|\beta\|-\|\alpha\|)\vee 0] + r + \left(\left\|\beta^1\right\| - \|\beta\|\right)\vee 0 - k/2$$
$$\geq \left(\left\|\beta^1\right\| - \|\alpha\|\right)\vee 0 - k/2 + r.$$

■

We are ready to prove Proposition 15.

Proof of Proposition 15. Note that arguing as in the proof of Lemma 14 it is sufficient to show

$$\left\|V_{[\alpha]}\bar{R}^{m,\bar{j}_1,\ldots,\bar{j}_k}_{(0,t_1,\ldots,t_k,t)}V_{[\beta]}\varphi\right\|_\infty \leq c_m t^{-(\|\alpha\|+\|\beta\|)/2}\frac{1}{\sqrt{t_1-t_0}}\cdots\frac{1}{\sqrt{t_k-t_{k-1}}}\|\varphi\|_\infty$$

for some constant c_m (the bounds on $\hat{R}^{m,\bar{j}_1,\ldots,\bar{j}_k}_{(0,t_1,\ldots,t_k,t)}$ follow by using the same arguments). The functions $V_{[\alpha]}\bar{R}^{m,\bar{j}_1,\ldots,\bar{j}_k}_{(0,t_1,\ldots,t_k,t)}V_{[\beta]}\varphi$ are linear combination of terms of the form

$$V_{[\alpha]}P^\Phi_{t_1}H_1P^\Phi_{t_2-t_1}\cdots P^\Phi_{t_k-t_{k-1}}H_kP^\Phi_{t-t_k}V_{[\beta]}\varphi$$

for some $\Phi\in\mathcal{K}_0$ and $H_j = u_j\ V_{i_j} + v_j$ with $u_j, v_j \in C_b^\infty$. Recall the convention $t = t_{k+1}$. Suppose $[t_{j-1}, t_j]$ is the maximal subinterval, i.e. satisfies

$$t_j - t_{j-1} = \max_{i=1,\ldots k+1}(t_i - t_{i+1}) \geq \frac{t}{k} \tag{55}$$

For notational reasons we have to treat the case $j = k+1$ separately, however it will be clear from the proof that the same arguments apply in this case.

Suppose now that $j \in \{1, \ldots, k\}$, then by Proposition 28 we observe that

$$P^\Phi_{t_{j+1}-t_j}H_{j+1}P^\Phi_{t_{j+2}-t_{j+1}}\cdots H_kP^\Phi_{t-t_k}V_{[\beta]}\varphi(x) = \sum_{\beta^{j+1}\in A_0(\bar{k})}\cdots\sum_{\beta^k\in A_0(\bar{k})} G_{\beta^{j+1},\ldots,\beta^k}(x),$$

where

$$G_{\beta^{j+1},\ldots,\beta^k} := V_{[\beta^{j+1}]}P^{\Phi_{\beta^{j+1}}}_{t_{j+1}-t_j}P^{\Phi_{\beta^{j+2}}}_{t_{j+2}-t_{j+1}}\cdots P^{\Phi_{\beta^{k+1}}}_{t-t_k}\varphi,$$

for some functionals $\Phi_{\beta^{j+1}} \in \mathcal{K}_{r_1}, \ldots, \Phi_{\beta^{k+1}} \in \mathcal{K}_{r_k}$ with $r_{k+1} \geq 0$, $r_{j+1}, \ldots r_k \geq -1/2$ and $r_{j+1} + r_2 + \cdots + r_{k+1} \geq \left(\left\| \beta^{j+1} \right\| - \|\beta\| \right) \vee 0 - (k-j)/2$. It follows from the maximality of $[t_j, t_{j-1}]$ that

$$
\begin{aligned}
&(t_{j+1} - t_j)^{r_{j+1}} \cdots (t_k - t_{k-1})^{r_k} \\
&\leq (t_{j+1} - t_j)^{-1/2} \cdots (t_k - t_{k-1})^{-1/2} (t_j - t_{j-1})^{\left[\left(\left\|\beta^{j+1}\right\| - \|\beta\|\right) \vee 0\right]/2}.
\end{aligned}
$$

On the other hand, to pass the derivative $V_{[\alpha]}$ to $P^{\Phi}_{t_j - t_{j-1}}$ we will iteratively use Lemma 26. Once again by maximality of $[t_j, t_{j-1}]$ it follows that

$$
\begin{aligned}
&(t_1 - t_0)^{-1/2 \vee \left(\left\|\beta^1\right\| - \|\alpha\|\right)/2} \cdots (t_{j-1} - t_{j-2})^{-1/2 \vee \left(\left\|\beta^{j-1}\right\| - \left\|\beta^{j-2}\right\|\right)/2} \\
&\leq (t_1 - t_0)^{-1/2} \cdots (t_{j-1} - t_{j-2})^{-1/2} (t_j - t_{j-1})^{\left[\left(\left\|\beta^{j-1}\right\| - \|\alpha\|\right) \vee 0\right]/2}.
\end{aligned}
$$

Using Lemma 26 iteratively we see from our preceding observations that

$$
\begin{aligned}
&\left\| V_{[\alpha]} P^{\Phi}_{t_1} H_1 P^{\Phi}_{t_2 - t_1} \cdots P^{\Phi}_{t_k - t_{k-1}} H_k P^{\Phi}_{t - t_k} V_{[\beta]} \varphi \right\|_{\infty} \\
&= \left\| \sum_{\beta^{j+1} \in A_0(\bar{k})} \cdots \sum_{\beta^k \in A_0(\bar{k})} V_{[\alpha]} P^{\Phi}_{t_1} H_1 \cdots H_{j-1} P^{\Phi}_{t_j - t_{j-1}} H_j G_{\beta^{j+1}, \ldots, \beta^k} \right\|_{\infty} \\
&\leq \widetilde{C}^j \frac{1}{\sqrt{t_1 - t_0}} \cdots \frac{1}{\sqrt{t_{j-1} - t_{j-2}}} \\
&\quad \sum_{\beta^{j-1}, \ldots, \beta^k \in A_0(\bar{k})} (t_j - t_{j-1})^{\left[\left(\left\|\beta^{j-1}\right\| - \|\alpha\|\right) \vee 0\right]/2} \left\| V_{[\beta^{j-1}]} P^{\Phi}_{t_j - t_{j-1}} H_j G_{\beta^{j+1}, \ldots, \beta^k} \right\|_{\infty} \\
&\leq \widetilde{C}^k \frac{1}{\sqrt{t_1 - t_0}} \cdots \frac{1}{\sqrt{t_k - t_{k-1}}} \|\varphi\|_{\infty} \\
&\quad \sum_{\beta^{j-1}, \beta^{j+1} \in A_0(\bar{k})} (t_j - t_{j-1})^{\left[\left(\left\|\beta^{j-1}\right\| - \|\alpha\|\right) \vee 0\right]/2 + \left[\left(\left\|\beta^{j+1}\right\| - \|\beta\|\right) \vee 0\right]/2 - \left(\left\|\beta^{j-1}\right\| + \left\|\beta^{j+1}\right\|\right)/2} \\
&\leq C^k \frac{1}{\sqrt{t_1 - t_0}} \cdots \frac{1}{\sqrt{t_k - t_{k-1}}} t^{-(\|\alpha\| + \|\beta\|)/2} \|\varphi\|_{\infty},
\end{aligned}
$$

where the penultimate inequality results by using Lemma 23. ■

Remark 29 *It is clear that Proposition 15 may be generalised to allow for multiple derivatives $V_{[\alpha_1]} \cdots V_{[\alpha_j]}$ and $V_{[\beta_1]} \cdots V_{[\beta_l]}$. Note that all lemmas and proofs presented in this section may be generalised using straightforward induction arguments.*

7.3 Proof of Lemma 16.

We are finally ready to prove Lemma 16 that provides a set of a priori estimates for the operator norms of the terms in the perturbation expansion that are not yet summable though.

The proof relies on the lemmas derived in the previous subsection. For the convenience of the reader we begin by restating the lemma.

Lemma 16: *With the notation of Lemma 7 for any* $0 < \gamma < 1/2$, $m > 0$ *there exist random variables* $c(\gamma, m, \omega)$ *such that, almost surely*

$$\left\|R_{s,t}^{m,\bar{\imath}}\right\|_{H^{-1}\to H^{-1}} \leq c\left(\gamma, m, \omega\right)|t-s|^{m\gamma}.$$

$$\left\|R_{s,t}^{m,\bar{\imath}}\right\|_{H^{1}\to H^{1}} \leq c\left(\gamma, m, \omega\right)|t-s|^{m\gamma}.$$

and finally

$$\left\|R_{s,t}^{m,\bar{\imath}}\right\|_{H^{-1}\to H^{1}} \leq c\left(\gamma, m, \omega\right)|t-s|^{m\gamma-\bar{k}}.$$

for all $\bar{\imath} \in S\left(m\right),\ 0 < s < t < 1.$

Proof. For all $\bar{\jmath}_1, \ldots, \bar{\jmath}_k \in S$ such that $\bar{\imath} = \bar{\jmath}_1 * \ldots * \bar{\jmath}_k$ we note that for any $0 < t \leq 1$ we have by iteratively applying Lemma 26

$$\begin{aligned}\left\|\bar{R}_{(t_0,t_1,\ldots,t_k,t)}^{m,\bar{\jmath}_1,\ldots,\bar{\jmath}_k}\varphi\right\|_{H^1} &= \sum_{\alpha\in A_0\left(\bar{k}\right)}\left\|V_{[\alpha]}\left(\bar{R}_{(t_0,t_1,\ldots,t_k,t)}^{m,\bar{\jmath}_1,\ldots,\bar{\jmath}_k}\varphi\right)\right\|_\infty \\ &\leq C^k\frac{1}{\sqrt{t_1-t_0}}\cdots\frac{1}{\sqrt{t_k-t_{k-1}}}\sum_{\beta\in A_0\left(\bar{k}\right)}\left\|V_{[\beta]}\varphi\right\|_\infty.\end{aligned}$$

The bound on $\left\|R_{s,t}^{m,\bar{\imath}}\left(\varphi\right)\right\|_{H^1\to H^1}$ now follows by applying Lemmas 13 and 14 and noting that by Lemma 26

$$\left\|P_t\left(h_{i_1}\cdots h_{i_m}\cdot\right)\right\|_{H^1} \leq c_m\sum_{\beta\in A_0\left(\bar{k}\right)}\left\|V_{[\beta]}\varphi\right\|_\infty.$$

Finally to show inequalities (39) and (40) we let $\varphi \in H^{-1}$. Then there exist for every $\varepsilon > 0$ functions φ^β such that

$$\varphi = \sum_{\beta\in A_0\left(\bar{k}\right)}V_{[\beta]}\varphi^\beta$$

and $\sum_{\beta\in A_0\left(\bar{k}\right)}\left\|\varphi^\beta\right\|_\infty \leq \|\varphi\|_{H^{-1}} + \varepsilon$. First we have

$$\left\|\bar{R}_{(t_0,t_1,\ldots,t_k,t)}^{m,\bar{\jmath}_1,\ldots,\bar{\jmath}_k}\varphi\right\|_{H^{-1}} = \left\|\sum_{\beta\in A_0\left(\bar{k}\right)}\bar{R}_{(t_0,t_1,\ldots,t_k,t)}^{m,\bar{\jmath}_1,\ldots,\bar{\jmath}_k}V_{[\beta]}\varphi^\beta\right\|_{H^{-1}}$$

and by Proposition 28 for each $\beta \in A_0\left(\bar{k}\right)$ there exist functionals $\Phi_{\beta^1} \in \mathcal{K}_{r_1}, \ldots, \Phi_{\beta^k} \in \mathcal{K}_{r_k}$ such that $r_k \geq 0$, $r_1, \ldots r_{k-1} \geq -1/2$ and

$$\bar{R}_{(t_0,t_1,\ldots,t_k,t)}^{m,\bar{\jmath}_1,\ldots,\bar{\jmath}_k}V_{[\beta]}\varphi^\beta = \sum_{\beta\in A_0\left(\bar{k}\right)}\sum_{\beta^1\in A_0\left(\bar{k}\right)}\cdots\sum_{\beta^k\in A_0\left(\bar{k}\right)}V_{[\beta^1]}P_{t_1-t_0}^{\Phi_{\beta^1}}P_{t_2-t_1}^{\Phi_{\beta^2}}\cdots P_{t-t_k}^{\Phi_{\beta^k}}\varphi.$$

We deduce from Lemma 23 that

$$\begin{aligned}\left\|\bar{R}^{m,\bar{j}_1,\dots,\bar{j}_k}_{(t_0,t_1,\dots,t_k,t)}V_{[\beta]}\varphi^{\beta}\right\|_{H^{-1}} &\leq \sum_{\beta\in A_0(\bar{k})}\sum_{\beta^1\in A_0(\bar{k})}\cdots\sum_{\beta^k\in A_0(\bar{k})}\left\|P^{\Phi_{\beta^1}}_{t_1-t_0}P^{\Phi_{\beta^2}}_{t_2-t_1}\cdots P^{\Phi_{\beta^k}}_{t-t_k}\varphi^{\beta}\right\|_{\infty}\\ &\leq C_k\frac{1}{\sqrt{t_1-t_0}}\cdots\frac{1}{\sqrt{t_k-t_{k-1}}}\left\|\varphi^{\beta}\right\|_{\infty}\end{aligned}$$

and consequently

$$\begin{aligned}\left\|\bar{R}^{m,\bar{j}_1,\dots,\bar{j}_k}_{(t_0,t_1,\dots,t_k,t)}\varphi\right\|_{H^{-1}} &\leq c_k\frac{1}{\sqrt{t_1-t_0}}\cdots\frac{1}{\sqrt{t_k-t_{k-1}}}\sum_{\beta\in A_0(\bar{k})}\left\|\varphi^{\beta}\right\|_{\infty}\\ &\leq c_k\frac{1}{\sqrt{t_1-t_0}}\cdots\frac{1}{\sqrt{t_k-t_{k-1}}}\left\|\varphi\right\|_{H^{-1}}+\varepsilon.\end{aligned}$$

A similar, but easier argument leads to

$$\left\|P_t\left(h_{i_1}\cdots h_{i_m}\cdot\right)\right\|_{H^{-1}}\leq c_m\left\|\varphi\right\|_{H^{-1}}.$$

To demonstrate the last inequality observe that arguing exactly as in the proof of Proposition 9 we have,

$$\begin{aligned}\left\|\bar{R}^{m,\bar{j}_1,\dots,\bar{j}_k}_{(t_0,t_1,\dots,t_k,t)}\varphi\right\|_{H^{1}} &= \sum_{\alpha\in A_0(\bar{k})}\left\|V_{[\alpha]}\left(\bar{R}^{m,\bar{j}_1,\dots,\bar{j}_k}_{(t_0,t_1,\dots,t_k,t)}\varphi\right)\right\|_{\infty}\\ &\leq \sum_{\alpha\in A_0(\bar{k})}\sum_{\beta\in A_0(\bar{k})}\left\|V_{[\alpha]}\left(\bar{R}^{m,\bar{j}_1,\dots,\bar{j}_k}_{(t_0,t_1,\dots,t_k,t)}V_{[\beta]}\varphi^{\beta}\right)\right\|_{\infty}\end{aligned}$$

$$\begin{aligned}&\leq c_k\frac{1}{\sqrt{t_1-t_0}}\cdots\frac{1}{\sqrt{t_k-t_{k-1}}}\sum_{\alpha\in A_0(\bar{k})}\sum_{\beta\in A_0(\bar{k})}t^{-(\|\alpha\|+\|\beta\|)/2}\left\|\varphi^{\beta}\right\|_{\infty}\\ &\leq c_k\frac{1}{\sqrt{t_1-t_0}}\cdots\frac{1}{\sqrt{t_k-t_{k-1}}}t^{-\bar{k}}\sum_{\beta\in A_0(\bar{k})}\left\|\varphi^{\beta}\right\|_{\infty}\\ &\leq c_k\frac{1}{\sqrt{t_1-t_0}}\cdots\frac{1}{\sqrt{t_k-t_{k-1}}}t^{-\bar{k}}\left\|\varphi\right\|_{H^{-1}}+\varepsilon,\end{aligned}$$

where c_k are constants changing from line to line. Again, it is easy to see that

$$\left\|P_t\left(h_{i_1}\cdots h_{i_m}\cdot\right)\right\|_{H^{1}}\leq c_m t^{-\bar{k}}\left\|\varphi\right\|_{H^{-1}}.$$

The claim in both cases now follows once again from Lemmas 13 and 14. As before we note that the same estimates apply to $\hat{R}^{m,\bar{j}_1,\dots,\bar{j}_k}_{(t_0,t_1,\dots,t_k,t)}$ in place of $\bar{R}^{m,\bar{j}_1,\dots,\bar{j}_k}_{(t_0,t_1,\dots,t_k,t)}$. ∎

References

[1] A. Bain, D. Crisan, *Fundamentals of Stochastic Filtering*, Stochastic Modelling and Applied Probability, Vol 60, Springer Verlag, 2008.

[2] J.-M. Bismut and D. Michel, *Diffusions conditionnelles, I. Hypoellipticite partielle*, Journal of Functional Analysis 44 (1981), 174–211.

[3] J. M. C. Clark, D. Crisan, *On a robust version of the integral representation formula of nonlinear filtering*, Probab. Theory Related Fields 133, no. 1, 43-56, 2005.

[4] D. Crisan, J. Diehl, P.K. Friz, H. Oberhauser, *Robust Filtering: Correlated noise and multidimensional observation*, The Annals of Applied Probability, Vol. 23, No. 5, 2139-2160, 2013.

[5] D. Crisan, K. Manolarakis, C. Nee, *Cubature Methods and Applications*, in Paris-Princeton Lectures on Mathematical Finance 2013, Eds Vicky Henderson, Ronnie Sircar (Lecture Notes in Mathematics), 2013.

[6] A. Deya, M. Gubinelli, S. Tindel, *Non-linear rough heat equations,* Probab. Theory Related Fields,Vol.153, Issue 1-2, pp 97-147, 2012

[7] J. Diehl, H. Oberhauser, S. Riedel: *A Levy-area between Brownian motion and rough paths with applications to robust nonlinear filtering and rough partial differential equations,* Stochastic Processes and their ApplicationsVol. 125, Issue 1, Jan. 2015, 161–181

[8] T. E. Duncan. *Likelihood functions for stochastic signals in white noise* Information and Control, 16 pages 303-310, 1970.

[9] P. Friz, N. Victoir, *Multidimensional stochastic processes as rough paths.*Theory and applications. Cambridge Studies in Advanced Mathematics,120. Cambridge University Press, Cambridge, 2010

[10] K. Hara, M. Hino, *Fractional order Taylor's series and the neo-classical inequality,* Bull. London Math. Soc. (2010) 42 (3): 467-477.

[11] H. Kunita, *Stochastic partial differential equations connected with non-linear filtering,* Springer Lecture Notes in Mathematics Volume 972, 1982, pp 100-169

[12] S. Kusuoka, *Malliavin calculus revisited*, J. Math. Sci. Univ. Tokyo,10 (2003), 261-277

[13] S. Kusuoka, *Analysis on Wiener spaces*, I. Nonlinear maps, J. Funct. Anal.98 (1991), 122–168.

[14] S. Kusuoka, D. W. Stroock, *Applications of the Malliavin calculus - I*, Stochastic Analysis Katata/Kyoto, pages 271-306, 1982.

[15] S. Kusuoka, D. W. Stroock, *Applications of the Malliavin calculus - II* Journal of the Faculty of Science, Univ. of Tokyo 1, pages 1-76, 1985.

[16] S. Kusuoka, D. W. Stroock, *Applications of Malliavin Calculus III*, J.Fac. Sci. Univ. Tokyo Sect. IA Math. 34 (1987), 391-442.

[17] S. Kusuoka, D. W. Stroock, *The partial malliavin calculus and its application to non-linear filtering,* Stochastics,Volume 12, Issue 2, 1984

[18] T. Lyons, *Differential equations driven by rough signals*, Rev. Mat. Iberoamericana 14 (1998), no. 2, 215-310.

[19] T. Lyons, M. Caruana, T. Levy, *Differential Equations Driven by Rough Paths*, Ecole d'Eté de Probabilités de Saint-Flour XXXIV-2004, Lecture Notes in Mathematics Vol. 1908, Springer, 2007

[20] D. Nualart, M. Zakai, *The partial Malliavin calculus,* Springer Lecture Notes in Mathematics Volume 1372, 1989, pp 362-381

[21] D. Ocone, *Multiple integral expansions for nonlinear filtering*, Stochastics 10 (1983), no. 1, 1-30.

[22] R. E. Mortensen. *Stochastic optimal control with noisy observations*, Internat. J. Control, 1(4) pages 455-464, 1966.

[23] M. Zakai, *On the optimal filtering of diffusion processes,* Z. Wahrs.und Verw. Gebiete, 11 pp 230-243, 1969.