Sharp Derivative Bounds for Solutions of Degenerate Semi-Linear Partial Differential Equations

Dan Crisan, François Delarue

SHARP DERIVATIVE BOUNDS FOR SOLUTIONS OF DEGENERATE SEMI-LINEAR PARTIAL DIFFERENTIAL EQUATIONS

DAN CRISAN AND FRANÇOIS DELARUE

ABSTRACT. The paper is a continuation of the Kusuoka-Stroock programme of establishing smoothness properties of solutions of (possibly) degenerate partial differential equations by using probabilistic methods. We analyze here a class of semi-linear parabolic partial differential equations for which the linear part is a second order differential operator of the form $V_0+\sum_{i=1}^N V_i^2$, where $V_0,\ldots,V_N$ are first order differential operators that satisfy the so-called UFG condition (see [18]), which is weaker than the Hörmander one. Specifically, we prove that the bounds of the higher order-derivatives of the solution along the vector fields coincide with those obtained in the linear case when the boundary condition is Lipschitz continuous, but that the asymptotic behavior of the derivatives may change because of the simultaneity of the nonlinearity and of the degeneracy when the boundary condition is of polynomial growth and measurable only.

KEYWORDS. Degenerate semi-linear parabolic PDE; Second-order differential operator satisfying the Uniformly Finitely Generated condition; Derivative estimates; Backward SDE; Malliavin calculus

AMS CLASSIFICATION (MSC 2010). 60H10, 60H07, 35K58, 35B45

1. INTRODUCTION

In a series of papers [16, 17, 18, 19], Kusuoka and Stroock have analyzed the smoothness properties of solutions of linear parabolic partial differential equations of the form

$$\partial_t u(t,x) = \frac{1}{2}\sum_{i=1}^N V_i^2 u(t,x) + V_0 u(t,x), \quad (t,x)\in(0,\infty)\times\mathbb{R}^d, \tag{1}$$

with initial condition $u(0,x) = h(x), \quad x\in\mathbb{R}^d$. The condition (called the UFG condition) imposed on the vector fields $\{V_i, i=0,\ldots,N\}$ under which they prove their results is weaker than the Hörmander condition. This condition states that the $C_b^\infty(\mathbb{R}^d)$-module $\mathcal{W}$ generated by the vector fields $\{V_i, i=1,\ldots,N\}$ within the Lie algebra generated by $\{V_i, i=1,\ldots,N\}$ is finite dimensional. In particular, the condition does not require that the vector space $\{W(x)|W\in\mathcal{W}\}$ is homeomorphic to $\mathbb{R}^d$ for any $x\in\mathbb{R}^d$. Hence, in this sense, the UFG condition is weaker than the Hörmander condition. It is important to emphasize that, under the UFG condition, the dimension of the space $\{W(x)|W\in\mathcal{W}\}$ is not required to be constant over $\mathbb{R}^d$. Such generality makes any Frobenius type approach to prove smoothness of the solution very difficult. Indeed the authors are not aware of any alternative proof of the smoothness results of the solution of (1) (under the UFG condition) other than that given by Kusuoka and Stroock.

Kusuoka and Stroock use a probabilistic approach to deduce their results. To be more precise, they use the Feynman-Kac representation of the solution of the PDE in terms of the semigroup associated

to a diffusion process. Let $X = \{X_t^x,\ \ (t,x) \in [0,\infty) \times \mathbb{R}^d\}$ be the (time homogeneous) stochastic flow

$$X_t^x = x + \int_0^t V_0(X_s^x)ds + \sum_{i=1}^N \int_0^t V_i(X_s^x) \circ dB_s^i, \quad t \geq 0, \tag{2}$$

where the vector fields $(V_i)_{0 \leq i \leq N}$ are smooth and bounded and the stochastic integrals in (2) are of Stratonovich type. The corresponding diffusion semigroup is then given by

$$\big[P_t g\big](x) = \mathbb{E}\big[g(X_t^x)\big], \quad t \geq 0, \quad x \in \mathbb{R}^d,$$

for any given bounded measurable function $g : \mathbb{R}^d \to \mathbb{R}$. When the boundary condition h in (1) is continuous, the following representation holds true:

$$u(t,x) = P_t h(x), \quad \forall (t,x) \in [0,\infty) \times \mathbb{R}^d.$$

Kusuoka and Stroock prove that, under the UFG condition, $P_t h$ is differentiable in the direction of any vector field W belonging to $\mathcal{W}$. Moreover they deduce sharp gradient bounds of the form:

$$\|W_1 \dots W_k P_t h\|_p \leq C^{p,k} t^{-l} \|h\|_p, \quad p \in [1,\infty], \tag{3}$$

where l is a constant that depends explicitly on the vector fields $W_i \in \mathcal{W}$, $i = 1, \dots, k$. Their results raise a number of fundamental questions related to the PDE (1). For example, the differentiability of $P_t h$ in the V_0 direction is not recovered. This is one of the fundamental differences between the UFG case and the Hörmander case where $P_t h$ is shown to be differentiable in any direction, including V_0. So whilst, in the Hörmander case, it is straightforward to show that $P_t h$ is indeed the (unique) classical solution of (1), the situation is more delicate in the absence of the Hörmander condition. As explained in [21], it turns out that $P_t h$ remains differentiable in the direction $\mathcal{V}_0 = \partial_t - V_0$ when viewed as a function $(t,x) \to P_t h(x)$ over the product space $(0,\infty) \times \mathbb{R}^d$. This together with the continuity at $t = 0$ implies that $P_t h$ is the unique (classical) solution of the equation

$$\mathcal{V}_0 u(t,x) = \frac{1}{2} \sum_{i=1}^N V_i^2 u(t,x), \quad (t,x) \in (0,\infty) \times \mathbb{R}^d. \tag{4}$$

The introduction of a new class of numerical methods for approximating the law of solutions of SDE (and, implicitly, the solution of PDEs as computed by means of the Feynman-Kac formula) has brought a renewed interest in the work of Kusuoka and Stroock. Their fundamental results form the theoretical basis of a recently developed class of high accuracy numerical methods. In the last ten years, Kusuoka, Lyons, Ninomiya and Victoir [15, 20, 22, 23, 24] developed several numerical algorithms based on Chen's iterated integrals expansion (see [7] for a unified approach for the analysis of these methods). These new algorithms generate an approximation of the solution of the SDE in the form of the empirical distribution of a cloud of particles with deterministic trajectories. The particles evolve only in directions belonging to $\mathcal{W}$. This ensures that the particles remain within the support of the limiting diffusion, leading to more stable schemes. The global error of numerical schemes depends intrinsically on the smoothness of $P_t h$ *but only in directions belonging to* $\mathcal{W}$. As a result they work under the (weaker) UFG condition rather than the ellipticity/Hörmander condition. By contrast, the classical Euler based numerical method (combined with a Monte-Carlo procedure) sends the component particles in any direction, hence they require the Hörmander condition.

In recent works [5, 6] the applicability of these schemes has been extended to semilinear PDEs. One of the major hurdles in obtaining convergence results for these schemes has been the absence of smoothness results of the type (3), again under the UFG condition. The authors are not aware of the existence of such bounds proved under the Hörmander condition either. In the following we will consider semilinear PDEs of the form:

$$\partial_t u(t,x) = \frac{1}{2}\sum_{i=1}^{N} V_i^2 u(t,x) + V_0 u(t,x) + f\big(t,x,u(t,x),(Vu(t,x))^\top\big), \quad (t,x)\in(0,\infty)\times\mathbb{R}^d, \tag{5}$$

with initial condition $u(0,x) = h(x), \quad x \in \mathbb{R}^d$. In (5) we used the notation $Vu(t,x)$ to denote the row vector $(V_1u(t,x),\dots,V_Nu(t,x))$. ($(Vu)^\top$ stands for the transpose of Vu.) As we shall see, $u(t,x)$ is differentiable in any direction $W \in \mathcal{W}$ when h is continuous just as in the linear case. If, for example, the vectors V_i, $i = 1,\dots,N$, satisfy the uniform ellipticity condition, then $u(t,x)$ is differentiable in any direction and the analysis covers semilinear PDEs written in the 'standard' format

$$\partial_t u(t,x) = \frac{1}{2}\sum_{i=1}^{N} V_i^2 u(t,x) + V_0 u(t,x) + f\big(t,x,u(t,x),(\nabla_x u(t,x))^\top\big), \quad (t,x)\in(0,\infty)\times\mathbb{R}^d,$$

where $\nabla_x u$ is the usual gradient of u in x, i.e., the row vector of partial derivatives $(\partial_{x_1}u,\dots,\partial_{x_N}u)$.

Following the tradition of Kusuoka and Stroock, we analyze the smoothness of the solution of the semilinear PDE using probabilistic methods. The basis of the analysis is the corresponding Feynman-Kac representation for the solution of (5). This representation was introduced by Pardoux and Peng in [26, 27] and involves the solution of a backward stochastic differential equation (see Section 2.1 below).

1.1. **The UFG condition.** Let $(V_i)_{0\le i\le N}$ be $N+1$ vector fields, V_0 belonging to $\mathcal{C}_b^K(\mathbb{R}^d,\mathbb{R}^d)$ and V_i, $1\le i\le N$, to $\mathcal{C}_b^{K+1}(\mathbb{R}^d,\mathbb{R}^d)$, $K\ge 0$, $\mathcal{C}_b^n(\mathbb{R}^d,\mathbb{R}^d)$ standing for the set of bounded and continuous functions from $\mathbb{R}^d$ to $\mathbb{R}^d$ that are n-times differentiable, with bounded and continuous partial derivatives up to order n. We will make use of the standard notation introduced in [19], (see also [21] and [7]):

$$V_{[i]} = V_i,\ V_{[\alpha\star i]} = \big[V_{[\alpha]},V_i\big], \quad i\in\{0,\dots,N\},$$

where $[\cdot,\cdot]$ stands for the Lie bracket of two vector fields, that is $[V,W] = V\cdot\nabla W - W\cdot\nabla V$ and $\alpha\star i$ stands for the multi-index $(\alpha_1,\dots,\alpha_n,i)$ when α is given by $(\alpha_1,\dots,\alpha_n)$ with $\alpha_j\in\{0,\dots,N\}$, $j=1,\dots,n$. The following "lengths" of a multi-index $\alpha=(\alpha_1,\dots,\alpha_n)$ will be used:

$$|\alpha| = |(\alpha_1,\dots,\alpha_n)| = n, \quad \|\alpha\| = \|(\alpha_1,\dots,\alpha_n)\| = n + \sharp\{i:\alpha_i=0\}.$$

The set of all multi-indices is denoted by $\mathcal{A}$, the set of all multi-indices α different from (0) is denoted by $\mathcal{A}_0$ and the set of non-empty multi-indices α in $\mathcal{A}_0$ for which $\|\alpha\|\le m$ is denoted by $\mathcal{A}_0(m)$.

For n multi-indices $\alpha_1,\dots,\alpha_n$, $n\ge 1$, we often denote the n-tuple $(\alpha_1,\dots,\alpha_n)$ by $\boldsymbol{\alpha}$ and then set $\|\boldsymbol{\alpha}\| = \|\alpha_1\|+\cdots+\|\alpha_n\|$.

Definition 1.1. *Let $m\in\mathbb{N}^*$ be a positive integer and assume that $K\ge m+3$. The vector fields $\{V_i, 0\le i\le N\}$ satisfy* the UFG condition of order m *if, for any $\alpha\in\mathcal{A}_0$ such that $\|\alpha\| = m+1$ or*

$\alpha = \alpha' \star 0$ with $\|\alpha'\| = m$, there exists $\varphi_{\alpha,\beta} \in \mathcal{C}_b^{K+1-|\alpha|}(\mathbb{R}^d)$, with $\beta \in \mathcal{A}_0(m)$, such that

$$V_{[\alpha]}(x) = \sum_{\beta \in \mathcal{A}_0(m)} \varphi_{\alpha,\beta}(x) V_{[\beta]}(x), \quad x \in \mathbb{R}^d.$$

Remark 1.2. *In* [21]*, the constant K is required to be greater than $m+1$. We here need $K \geq m+3$ to ensure the existence of classical solutions to the nonlinear PDE, see Theorem 1.4 below.*

The following example illustrates the difference between the UFG and the Hörmander condition (see [18]):

Example 1.3. *Assume $N = 1$ and $d = 2$. Let V_0 and V_1 be given by*

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

The vector fields $\{V_0, V_1\}$ satisfy the UFG condition of order $m = 4$, but not the Hörmander condition.

The vector fields $\{V_i, 0 \leq i \leq N\}$ satisfy the *uniform* Hörmander condition if there exists $m > 0$ such that

$$\inf_{\{x,\xi \in \mathbb{R}^d | \, |\xi|=1\}} \sum_{\beta \in \mathcal{A}_0(m)} (V_{[\beta]}(x), \xi)^2 > 0.$$

Obviously, if the vector fields $\{V_i, 0 \leq i \leq N\}$ satisfy the uniform Hörmander condition then they satisfy the UFG condition. In particular if the vector fields $\{V_i, 1 \leq i \leq N\}$ satisfy the uniform ellipticity condition then they satisfy the UFG condition.

Definition 1.1 is a (slight) generalization of the corresponding one given in [19]. In [19], both the vector fields $\{V_i, 0 \leq i \leq N\}$ and the coefficients $\varphi_{\alpha,\beta}$ are assumed to be smooth (infinitely differentiable). If the smoothness assumption is imposed then $V_{[\alpha]}$ is well defined for any $\alpha \in \mathcal{A}$ and one can interpret the UFG condition in the following manner. Let $\mathcal{W}$ be the $C_b^\infty(\mathbb{R}^d)$-module generated by the vector fields $\{V_i, i = 1, \ldots, N\}$ within the Lie algebra generated by $\{V_i, i = 1, \ldots, N\}$. Then $\mathcal{W}$ is finitely generated *as a vector space* and $\{V_{[\alpha]}, \alpha \in \mathcal{A}_0(m)\}$ is a finite set of generators for $\mathcal{W}$. In addition, the functions $\varphi_{\alpha,\beta}$ appearing in the decomposition of any vector field $V \in \mathcal{W}$ as a linear combination of the elements of the set $\{V_{[\alpha]}, \alpha \in \mathcal{A}_0(m)\}$ are assumed to be smooth and uniformly bounded over $\mathbb{R}^d$. These are salient properties that are essential to make the proof of Kusuoka and Stroock work and justify the use of the acronym UFG - uniformly finitely generated - for the assumed property.

As shown in [21] the smoothness assumption on the vector fields $\{V_i, 0 \leq i \leq N\}$ and the coefficients $\varphi_{\alpha,\beta}$ is not necessary. The level of differentiability is dictated by the order of the UFG condition assumed. In other words, the vector fields have to be sufficiently many times differentiable for the repeated brackets to make sense up to the required level. Of course, in this case, we can no longer talk about the $C_b^\infty(\mathbb{R}^d)$-module $\mathcal{W}$ or about the Lie algebra generated by $\{V_i, i = 0, \ldots, N\}$ as not all the Lie brackets will make sense (due to the reduced differentiability). Then, we will denote by $\mathcal{W}$ the space generated by the vector fields $V_{[\alpha]}$, with $|\alpha| \leq K+1$, for which there exist $\varphi_{\alpha,\beta} \in \mathcal{C}_b^{K+1-|\alpha|}(\mathbb{R}^d)$, with $\beta \in \mathcal{A}_0(m)$, such that

$$V_{[\alpha]}(x) = \sum_{\beta \in \mathcal{A}_0(m)} \varphi_{\alpha,\beta}(x) V_{[\beta]}(x), \quad x \in \mathbb{R}^d.$$

Definition 1.1 then states that $\{V_{[\alpha]}; \alpha \in \mathcal{A}_0(m+1)\} \cup \{V_{[\alpha]}; \alpha = \alpha' \star 0, \alpha' \in \mathcal{A}_0(m)\} \subset \mathcal{W}$. This extension allows us to identify the *minimal level of differentiability* that we need to impose on the coefficients of the PDE so as to deduce the desired gradient bounds.

1.2. **The Main Results.** Under the UFG condition (see [21] and [19]) the solution of the linear equation (1) is differentiable in any direction $V \in \mathcal{W}$. Moreover, if h is a smooth bounded function, the following gradient bound holds true:

$$\left|V_{[\alpha_1]} \dots V_{[\alpha_n]} u(t,x)\right| \leq C\|h\|_\infty t^{-\|\boldsymbol{\alpha}\|/2}, \tag{6}$$

for $\alpha_1, \dots, \alpha_n \in \mathcal{A}_0(m)$, where C is a constant independent of h and (t,x), and $\|\boldsymbol{\alpha}\| = \|\alpha_1\| + \cdots + \|\alpha_n\|$. If h is Lipschitz continuous function with Lipschitz constant

$$\|h\|_{\mathrm{Lip}} = \sup_{\{x,y \in \mathbb{R}^d, x \neq y\}} \frac{|h(x) - h(y)|}{|x-y|},$$

then there exists a constant C independent of h such that for all (t,x)

$$\left|V_{[\alpha_1]} \dots V_{[\alpha_n]} u(t,x)\right| \leq C\|h\|_{\mathrm{Lip}} t^{(1-\|\boldsymbol{\alpha}\|)/2}. \tag{7}$$

In the current paper we investigate the counterpart of these results for the solution of the semilinear PDE (5). The results are summarized in the following:

Theorem 1.4. *Assume that the vector fields $\{V_i, 0 \leq i \leq N\}$ satisfy* the UFG condition of order m. *Then, if h is of polynomial growth and continuous and if f satisfies additional conditions that are specified below, the semilinear PDE* (5) *is uniquely solvable in a suitable space of* classical solutions *and the solution is differentiable in any direction $V \in \mathcal{W}$. Moreover, if h is a Lipschitz continuous function, then, for any $T > 0$, there exists a constant C such that, for all $(t,x) \in (0,T] \times \mathbb{R}^d$,*

$$\left|V_{[\alpha_1]} \dots V_{[\alpha_n]} u(t,x)\right| \leq C t^{(1-\|\boldsymbol{\alpha}\|)/2}, \quad n \leq K - m - 1, \tag{8}$$

with $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n) \in [\mathcal{A}_0(m)]^n$. (See Footnote[1]) If h is a continuous function of polynomial growth, but not necessarily Lipschitz, then there exists a constant C such that, for all $(t,x) \in (0,T] \times \mathbb{R}^d$,

$$\left|V_{[\alpha_1]} \dots V_{[\alpha_n]} u(t,x)\right| \leq C t^{-\|\boldsymbol{\alpha}\|/2} \tag{9}$$

if $n \leq 2$ or $n = 3$ and $\min\{\|\alpha_i\|, i = 1,2,3\} = 1$. However, if $3 \leq n \leq K - m - 1$, then, for any $\delta > 0$, there exists a constant $C(\delta)$ such that, for all $(t,x) \in (0,T] \times \mathbb{R}^d$,

$$\left|V_{[\alpha_1]} \dots V_{[\alpha_n]} u(t,x)\right| \leq C(\delta) t^{-\|\boldsymbol{\alpha}\|/2} \left[1 + t^{-n/2+1+\min[1/\|\alpha_{(1)}\|, 1/2+1/(2\|\alpha_{(2)}\|)]-\delta}\right], \tag{10}$$

where $\|\alpha_{(1)}\| \leq \|\alpha_{(2)}\|$ stand for the two smallest elements among $\|\alpha_1\|, \dots, \|\alpha_n\|$. If h is of polynomial growth and measurable only, the semilinear PDE (5) *is uniquely solvable as well, but in a suitable space of* generalized solutions. *The solution admits generalized derivatives in any direction $V \in \mathcal{W}$ and satisfies* (9) *and* (10) *almost everywhere. (And Footnote[1] applies as well.)*

[1]The reader now understands why K is chosen to be greater than $m+3$: (8) holds at least for $n = 1, 2$, so that the partial derivatives in space in (5) make sense.

The details of the assumptions imposed on the function f are given in Sections 3 and 4 below. We make explicit the dependence of the constants appearing in equations (8), (9) and (10) on the initial condition h in Theorems 3.1 and 4.1. Theorems 3.1 and 4.1 also contain certain (nonlinear) Feynman-Kac representations for the derivatives $V_{[\alpha_1]} \dots V_{[\alpha_n]} u(t,x)$. Similar bounds and representations are also valid for $V_{[\alpha_1]} \dots V_{[\alpha_n]} V_i u(t,x)$, $i = 1, \dots, N$. These representations are important for the analysis of numerical algorithms for the approximation of the solution of (5).

Let us comment on the bounds contained in (8), (9) and (10). Despite the introduction of the nonlinear term in (5), the solution of the semilinear PDE has the same small time asymptotics as the solution of the linear PDE (1) when the initial condition h is a Lipschitz continuous function. The same applies for the case when h is a measurable function of polynomial growth as long as we differentiate no more than two times. For derivatives of order 3 or more the asymptotics may deteriorate according to the degeneracy: when $n = 3$ and $\|\alpha_{(1)}\| = 1$, the asymptotic rates in (10) are similar to the ones in the linear case; when $n = 3$ and $\|\alpha_{(1)}\| = 2$ or $n = 4$ and $\|\alpha_{(1)}\| = \|\alpha_{(2)}\| = 1$, it is almost the same as in the linear case up to the additional δ; in all the other cases, the asymptotic rates are strictly worse. In particular, the small time asymptotic behavior of the derivatives up to the fourth order are the same as in the linear case when the operator is uniformly elliptic (up to the additional δ for the fourth derivatives). In Section 5, examples, both in the uniformly elliptic and degenerate cases, are given where the announced bound in (10) is attained (up to the additional δ). This shows the sharpness of the bound. As a consequence, it turns out that the simultaneity of the nonlinearity and of the degeneracy will lead to a faster explosion (as $t \to 0$) of the higher derivatives above a certain threshold.

1.3. **Structure of the article.** The article is structured as follows. In Section 2, we collect a number of preliminary results required for the proof of the main theorems. The Feynman-Kac formula for the solution of the equation (5) is presented. It relates the solution of the PDE to the solution of a backward stochastic differential equation. We also give the rigorous definitions of a solution of (5). In Sections 3 and 4, we analyze the smoothness of the solution of (5) in the case when h is a Lipschitz continuous function and, respectively, when h is a measurable function of polynomial growth. In Section 5, we study two examples that show that we cannot expect the same asymptotic behaviour for the case when h is bounded, but not necessarily Lipschitz continuous, as in the linear case. Finally, in Section 6, we relax the Lipschitz condition imposed on the function f appearing in (5) and treat the case when f has quadratic growth and h is bounded. This is an important case with applications in optimisation problems appearing in mathematical finance (see, e.g., [11, 28] and the references therein).

2. Preliminary results

2.1. **The Feynman-Kac representation.** Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space endowed with an $(\mathcal{F}_t)_{t\geq 0}$-adapted Brownian motion $(B_t)_{t\geq 0}$. On $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ we consider the triplet $(X, Y, Z) = \{(X_t, Y_t, Z_t), t \in [0,T]\}$ of $\mathcal{F}_t$-adapted stochastic processes satisfying the following system of equations

$$\left\{ \begin{array}{rl} dX_t & = V_0(X_t)dt + \sum_{j=1}^N V_j(X_t) \circ dB_t^j \\ -dY_t & = f(T-t, X_t, Y_t, Z_t)dt - \langle Z_t, dB_t \rangle \end{array} \right. . \tag{11}$$

The system (11) is called a forward-backward stochastic differential equation (FBSDE). The process X, called the forward component of the FBSDE, is a d-dimensional diffusion satisfying a stochastic differential equation driven by $V_i : \mathbb{R}^d \to \mathbb{R}^d$, $i = 0, 1, \dots, N$. The notation "$\circ$" indicates that the stochastic term in the equation satisfied by X is a Stratonovitch integral. The process Y, called the backward component of the SDE is a one-dimensional stochastic process with final condition $Y_T = h(X_T)$, where $h : \mathbb{R}^d \to \mathbb{R}$ is a measurable function of polynomial growth. The function $f : [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$, referred to as " the driver", is assumed to be of polynomial growth in x, of linear growth in (y, z), being bounded in time t and Lipschitz continuous[2] in y and z, uniformly in time t and space x.

The existence and uniqueness question for the system (11) was first addressed by Pardoux and Peng in [26, 27] and, since then, a large number of papers have been dedicated to the study of FBSDEs. Pardoux and Peng proved that the stochastic flow $\left(X^{t,x}, Y^{t,x}, Z^{t,x}\right)$, $t \in [0, T]$, $x \in \mathbb{R}^d$ associated to the system (11), in other words, the solution of the system

$$\left\{\begin{array}{rl} dX_s^{t,x} & = V_0(X_s^{t,x})ds + \sum_{j=1}^N V_j(X_s^{t,x}) \circ dB_s^j, \\ -dY_s^{t,x} & = f(T-s, X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})ds - \langle Z_s^{t,x}, dB_s\rangle, \quad s \in [t, T] \\ X_t^{t,x} & = x, \quad Y_T^{t,x} = h(X_T^{t,x}), \end{array}\right. \tag{12}$$

provides a non-linear Feynman-Kac representation for the solution of the semilinear PDE (5). More precisely they showed that when the functions f and h are continuous, then the function

$$u(T - t, x) = Y_t^{t,x}, \tag{13}$$

is a continuous solution of (5) *in viscosity sense*. When the coefficients f and h are smooth, it is a solution *in classical sense* and $Z_s^{t,x} = (Vu)^\top(T - s, X_s^{t,x})$. Therefore, the results in this paper represent a strengthening of the results of Pardoux of Peng as we identify conditions under which the stochastic flow $Y_t^{t,x}$ generates a *classical* solution, and respectively, a generalized solution (in Sobolev sense) of (5), the terminal condition h being possibly non-smooth.

We remark that the triplet $\left(X^{t,x}, Y^{t,x}, Z^{t,x}\right)$, $t \in [0, T]$, $x \in \mathbb{R}^d$, which solves the system (12) is adapted to the (augmented) filtration generated by the increments $(B_s - B_t)_{t \le s \le T}$ so that $Y_t^{t,x}$ has a deterministic value (up to a zero-measure event).

2.2. **Properties of the Flow.** When u is continuous on $(0, T] \times \mathbb{R}^d$, the relationship between the deterministic mapping u and the pair (Y, Z) extends as $Y_t = u(T - t, X_t)$, $t \in [0, T)$. Given $X_t = x$, for some $t \in [0, T)$, this relationship reads: $Y_s^{t,x} = u(T - s, X_s^{t,x})$, $s \in [t, T)$. Moreover, (13) reads

$$u(T - t, x) = \mathbb{E}\left[h(X_T^{t,x}) + \int_t^T f\big(T - s, X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x}\big)ds\right]. \tag{14}$$

2.2.1. *Shift Operator.* Eq. (14) is the keystone for the probabilistic analysis of the regularity of u. Since X is a homogeneous diffusion process, we emphasize that $(X_s^{t,x})_{t \le s \le T}$, $t \in [0, T]$, may be understood as a shifted version of $(X_{s-t}^{0,x})_{0 \le s-t \le T-t}$. Specifically, we can choose the canonical Wiener space for $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \ge 0}, \mathbb{P})$ and thus introduce the shift operator $(\theta_t : \omega \mapsto \theta_t(\omega) = \omega(t + \cdot) - \omega(t))_{t \ge 0}$. Then, $(X_s^{t,x})_{t \le s \le T}$ reads as $(X_{s-t}^{0,x} \circ \theta_t)_{0 \le s-t \le T-t}$, or simply as $(X_{s-t}^{x} \circ \theta_t)_{0 \le s-t \le T-t}$, with the convention $X^x = X^{0,x}$.

[2]This assumption will be relaxed in Section 6.

As basic application, we discuss below how to transfer differentiation at starting point into differentiation along the flow. To do so, we first remind the reader of so-called Kusuoka-Stroock functions (see [21] and [19]).

2.2.2. *Kusuoka-Stroock Functions.* In the following, let E be a separable Hilbert space and $\mathbb{D}^{n,\infty}(E)$ be the space of E-valued functionals admitting Malliavin derivatives up to order n, see the monograph by Nualart [25, Chapter 1, Section 2] for details.

Definition 2.1 (Kusuoka-Stroock functions)**.** *Given $r \in \mathbb{R}$ and $n \in \mathbb{N}$, we denote by $\mathcal{K}_r^T(E,n)$ the set of functions: $g : (0,T] \times \mathbb{R}^d \to \mathbb{D}^{n,\infty}(E)$ satisfying the following:*

(1) *$g(t,.)$ is n-times continuously differentiable and $[\partial^\alpha g/\partial x^\alpha](.,.)$ is continuous in $(t,x) \in (0,T] \times \mathbb{R}^d$ a.s., for any tuple α of elements of $\{1,\dots,d\}$ of length $|\alpha| \le n$.*

(2) *For all $k \in \mathbb{N}$, $p \in [1,\infty)$, and $k \le n - |\alpha|$, $\sup_{t\in(0,T],x\in\mathbb{R}^d} t^{-r/2} \left\| \frac{\partial^\alpha g}{\partial x^\alpha}(t,x) \right\|_{\mathbb{D}^{k,p}(E)} < \infty$.*

Define $\mathcal{K}_r^T(n) := \mathcal{K}_r^T(\mathbb{R},n)$.

The functions belonging to the set $\mathcal{K}_r^T(E,n)$ satisfy the following properties which form the basis of our analysis (see [21] for details).

Lemma 2.2 (Properties of Kusuoka-Stroock functions)**.** *Within the framework of Definition 1.1, the followings hold*

(1) *The function $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d \mapsto X_t^x$ belongs to $\mathcal{K}_0^T(K)$, for any $T > 0$.*

(2) *Suppose $g \in \mathcal{K}_r^T(n)$, where $r \ge 0$. Then, for $i = 1,\dots,d$,*

$$\int_0^\cdot g(s,x)dB_s^i \in \mathcal{K}_{r+1}^T(n) \quad \text{and} \quad \int_0^\cdot g(s,x)ds \in \mathcal{K}_{r+2}^T(n).$$

(3) *If $g_i \in \mathcal{K}_{r_i}^T(n_i)$ for $i = 1,\dots,N$, then*

$$\prod_{i=1}^N g_i \in \mathcal{K}_{r_1+\dots+r_N}^T(\min_i n_i) \quad \text{and} \quad \sum_{i=1}^N g_i \in \mathcal{K}_{\min_i r_i}^T(\min_i n_i).$$

2.2.3. *Transport of Differentiation.* As announced, we claim as a consequence of Lemmas 2.2 and 3.9 in [21] (see also page 265 in [19]):

Lemma 2.3. *Define $J_{t,x} = [\partial(X_t^x)_i/\partial x_j]_{1\le i,j\le d}$, $t \ge 0$. Then, there exist two families of random functions $(a_{\alpha,\beta} : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R})_{\alpha,\beta\in\mathcal{A}_0(m)}$ and $(b_{\alpha,\beta} : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R})_{\alpha,\beta\in\mathcal{A}_0(m)}$, $a_{\alpha,\beta}, b_{\alpha,\beta} \in \cap_{T>0}\mathcal{K}_{(\|\beta\|-\|\alpha\|)^+}^T(K-m)$, such that for any $x \in \mathbb{R}^d$ and $\alpha \in \mathcal{A}_0(m)$,*

$$\begin{aligned} V_{[\beta]}\big(X_s^{t,x}\big) &= \theta_t^*[J_{s-t,x}] \sum_{\alpha\in\mathcal{A}_0(m)} \theta_t^*\big[a_{\beta,\alpha}\big](s-t,x)V_{[\alpha]}(x), \\ \theta_t^*[J_{s-t,x}]V_{[\alpha]}(x) &= \sum_{\beta\in\mathcal{A}_0(m)} \theta_t^*\big[b_{\alpha,\beta}\big](s-t,x)V_{[\beta]}\big(X_s^{t,x}\big), \end{aligned} \tag{15}$$

where $\theta_t^*[J_{s-t,x}] = J_{s-t,x} \circ \theta_t$ and $\theta_t^*[a_{\beta,\alpha}](s-t,x) = [a_{\beta,\alpha} \circ \theta_t](s-t,x)$ (and the same for $\beta_{\alpha,\beta}$).

As we will see below, Lemma 2.3 is a key ingredient of the analysis.

2.3. **Classical Solutions for the PDE (5).** We now define the notion of classical solutions in Theorem 1.4. A classical solution u of the PDE (5) will be twice continuously differentiable in the directions of the vector fields V_i, $i = 1, \dots, d$ and once continuously differentiable in the direction $\mathcal{V}_0 = \partial_t - V_0$, when viewed as a function $(t, x) \mapsto u(t, x)$ over the product space $(0, \infty) \times \mathbb{R}^d$.

2.3.1. *Space of Classical Solutions.* For an open ball $\mathbb{B} \subset \mathbb{R}^d$ and for a function φ in $\mathcal{C}_b^\infty(\mathbb{B})$, that is a bounded (real-valued) function φ with bounded derivatives of any order on $\mathbb{B}$, we set

$$\|\varphi\|_{\mathbb{B},\infty}^{V,1} = \|\varphi\|_{\mathbb{B},\infty} + \sum_{\alpha \in \mathcal{A}_0(m)} \|V_{[\alpha]}\varphi\|_{\mathbb{B},\infty}$$

and then define $\mathcal{D}_V^{1,\infty}(\mathbb{B})$ as the closure of $\mathcal{C}_b^\infty(\mathbb{B})$ in $\mathcal{C}_b(\bar{\mathbb{B}})$ w.r.t. $\|\cdot\|_{\mathbb{B},\infty}^{V,1}$. (See Footnote[3] for the closability argument.) More generally, for $1 \le k \le K - m + 2$, we can define by induction

$$\|\varphi\|_{\mathbb{B},\infty}^{V,k} = \|\varphi\|_{\mathbb{B},\infty}^{V,k-1} + \sum_{\alpha_1,\dots,\alpha_k \in \mathcal{A}_0(m)} \|V_{[\alpha_1]} \dots V_{[\alpha_k]}\varphi\|_{\mathbb{B},\infty}, \quad \varphi \in \mathcal{C}_b^\infty(\mathbb{B}).$$

We emphasize that $V_{[\alpha_1]} \dots V_{[\alpha_k]}\varphi$ makes sense for any smooth function because of the bound $k \le K - m + 2$: each $V_{[\alpha_i]}$ is at least $K - m + 1$ times continuously differentiable, so that the last vector field $V_{[\alpha_k]}$ in $V_{[\alpha_1]} \dots V_{[\alpha_k]}$ can be differentiated $K - m + 1$ times.

We then define $\mathcal{D}_V^{k,\infty}(\mathbb{B})$ as the closure of $\mathcal{C}_b^\infty(\mathbb{B})$ in $\mathcal{C}_b(\bar{\mathbb{B}})$ w.r.t. $\|\cdot\|_{\mathbb{B},\infty}^{V,k}$. (The closability argument is the same as above.) In particular, we can define $\mathcal{D}_V^{k,\infty}(\mathbb{R}^d)$ as

$$\mathcal{D}_V^{k,\infty}(\mathbb{R}^d) = \bigcap_{r \ge 1} \mathcal{D}_V^{k,\infty}(\mathbb{B}(0, r)), \quad 1 \le k \le K - m + 2,$$

where $\mathbb{B}(0, r)$ stands for the d-dimensional ball of center 0 and radius r. For $v \in \mathcal{D}_V^{k,\infty}(\mathbb{R}^d)$, $1 \le k \le K - m + 2$, $V_{[\alpha_1]} \dots V_{[\alpha_k]}v$ is understood as the derivative of v in the directions $V_{[\alpha_1]} \dots V_{[\alpha_k]}$, with $\alpha_1, \dots, \alpha_k \in \mathcal{A}_0(m)$.

Similarly, for $\varphi \in \mathcal{C}_b^\infty(\mathbb{B})$ and $0 \le k \le K - m + 1$, we set

$$\|\varphi\|_{\mathbb{B},\infty}^{V,k+1/2} = \|\varphi\|_{\mathbb{B},\infty}^{V,k} + \sum_{i=1}^N \sum_{\alpha_1,\dots,\alpha_k \in \mathcal{A}_0(m)} \|V_{[\alpha_1]} \dots V_{[\alpha_k]}V_i\varphi\|_{\mathbb{B},\infty}.$$

(Above, $\|\cdot\|_{\mathbb{B},\infty}^{V,0} = \|\cdot\|_{\mathbb{B},\infty}$.) We then define $\mathcal{D}_V^{k+1/2,\infty}(\mathbb{B})$ as the closure of $\mathcal{C}_b^\infty(\mathbb{B})$ in $\mathcal{C}_b(\bar{\mathbb{B}})$ w.r.t. $\|\cdot\|_{\mathbb{B},\infty}^{V,k+1/2}$ and we set

$$\mathcal{D}_V^{k+1/2,\infty}(\mathbb{R}^d) = \bigcap_{r \ge 1} \mathcal{D}_V^{k+1/2,\infty}(\mathbb{B}(0, r)), \quad 0 \le k \le K - m + 1.$$

Remark 2.4. *Note that any function in $\mathcal{D}_V^{1,\infty}(\mathbb{R}^d)$ is differentiable along the solutions of the ordinary differential equation $\dot{\gamma}_t = V(\gamma_t)$, $t \ge 0$, for $V \in \mathcal{A}_0(m)$. In particular, any function in $\mathcal{D}_V^{1,\infty}(\mathbb{R}^d)$ is continuously differentiable on $\mathbb{R}^d$ when the uniform Hörmander condition is satisfied.*

[3] We emphasize that the closure is well-defined: if $(\varphi_n, (V_{[\alpha]}\varphi_n)_{\alpha \in \mathcal{A}_0(m)})_{n \ge 1}$ tends to $(0, (G_\alpha)_{\alpha \in \mathcal{A}_0(m)})$ uniformly on $\mathbb{B}$ as n tends to $+\infty$, then for any test function $\psi \in \mathcal{C}^\infty(\mathbb{R}^d)$ with compact support included in $\mathbb{B}$, $\int_{\mathbb{R}^d} G_\alpha^i(x)\psi(x)dx = \lim_{n \to +\infty} -\int_{\mathbb{R}^d} \varphi_n(x)\partial_{x_i}(V_{[\alpha]}^i \times \psi)(x)dx = 0$, $i = 1, \dots, N$, so that G_α is zero.

2.3.2. *Typical Example.* A typical example of function in $\mathcal{D}_V^{n,\infty}(\mathbb{R}^d)$, $1 \le n \le K-m$, is $x \in \mathbb{R}^d \mapsto (P_t\varphi)(x)$, for $t > 0$ and $\varphi \in \mathcal{C}_b(\mathbb{R}^d)$. For this we need to recall the following integration by parts formula (see Corollaries 3.13 and 3.18 in [21])

Theorem 2.5. *Let $(V_i)_{0\le i\le N}$ satisfy the assumptions in Definition 1.1. Then, for any $T > 0$, $n \le K-m$ and $\alpha_1, \ldots, \alpha_n \in \mathcal{A}_0(m)$, there exists $\Phi_{\alpha_1,\ldots,\alpha_n} \in \mathcal{K}_0^T(K-m-n)$ such that*

$$V_{[\alpha_1]} \ldots V_{[\alpha_n]}(P_t h)(x) = t^{-\|\boldsymbol{\alpha}\|/2}\mathbb{E}\left[\Phi_{\alpha_1,\ldots,\alpha_n}(t,x)h(X_t^x)\right], \tag{16}$$

for any $h \in \mathcal{C}_b^\infty(\mathbb{R}^d)$, $t \in (0,T]$, $x \in \mathbb{R}^d$, with $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$. In particular, the following gradient bound holds true:

$$\|V_{[\alpha_1]} \ldots V_{[\alpha_n]} P_t h\|_\infty \le C\|h\|_\infty t^{-\|\alpha\|/2}, \tag{17}$$

where $C = \sup_{0<t\le T}\sup_{x\in\mathbb{R}^d}\mathbb{E}\left[|\Phi_{\alpha_1,\ldots,\alpha_n}(t,x)|\right] < \infty$. In addition, for any $n \le K-m$ and $\alpha_1, \ldots, \alpha_n \in \mathcal{A}_0(m)$ there exist $\Phi^i_{\alpha_1,\ldots,\alpha_n} \in \mathcal{K}_0^T(K-m-n+1)$, $i = 1, \ldots, d$ such that

$$V_{[\alpha_1]} \ldots V_{[\alpha_n]}(P_t h)(x) = t^{-(\|\alpha_1\|+\ldots+\|\alpha_{n-1}\|)/2}\sum_{i=1}^d \mathbb{E}\left[\Phi^i_{\alpha_1,\ldots,\alpha_n}(t,x)\partial_{x_i}h(X_t^x)\right], \tag{18}$$

for any $h \in \mathcal{C}_b^\infty(\mathbb{R}^d)$, $t \in (0,T]$, $x \in \mathbb{R}^d$. Hence, in particular, the following gradient bound holds true:

$$\|V_{[\alpha_1]} \ldots V_{[\alpha_n]} P_t h\|_\infty \le CT^{(m-1)/2}\|\nabla h\|_\infty t^{(1-\|\boldsymbol{\alpha}\|)/2}, \tag{19}$$

where $C = \max_{i=1,..,d}\sup_{0<t\le T}\sup_{x\in\mathbb{R}^d}\mathbb{E}\left[|\Phi^i_{\alpha_1,\ldots,\alpha_n}(t,x)|\right] < \infty$.[4]

To prove that the mapping $x \in \mathbb{R}^d \mapsto (P_t\varphi)(x)$, for $t \in (0,T]$ and $\varphi \in \mathcal{C}_b^0(\mathbb{R}^d)$, is in $\mathcal{D}_{V,\infty}^n(\mathbb{R}^d)$, $1 \le n \le K-m$, it is sufficient to consider a sequence $(\varphi_\ell)_{\ell\ge1}$ of functions in $\mathcal{C}_b^\infty(\mathbb{R}^d)$ converging towards φ uniformly on compact subsets of $\mathbb{R}^d$ as ℓ tends to $+\infty$. Then, from the above theorem, we have that

$$\left[V_{[\alpha_1]} \ldots V_{[\alpha_n]} P_t\varphi_\ell\right](x) = t^{-\|\boldsymbol{\alpha}\|/2}\mathbb{E}\left[\varphi_\ell(X_t^x)\psi(t,x)\right], \tag{20}$$

with $\psi \in \mathcal{K}_0^T(K-m-n)$ is independent of ℓ. Clearly, on every compact subsets of $\mathbb{R}^d$, the right-hand side in (20) converges towards the continuous function $x \in \mathbb{R}^d \mapsto t^{-\|\boldsymbol{\alpha}\|/2}\mathbb{E}[\varphi(X_t^x)\psi(t,x)]$. Therefore, the sequence $(V_{[\alpha_1]} \ldots V_{[\alpha_n]} P_t\varphi_\ell)_{\ell\ge1}$ is Cauchy in any space $\mathcal{C}(\bar{\mathbb{B}}(0,r))$, $r > 0$, so that $P_t\varphi$ belongs to $\mathcal{D}_V^{n,\infty}(\mathbb{R}^d)$ for $1 \le n \le K-m$ and (20) holds for φ as well.

2.3.3. *Definition of Classical Solutions.* To define the notion of a classical solution to (5), we will need to introduce the set of functions that are continuously differentiable in the direction $\mathcal{V}_0 = \partial_t - V_0$. Again, we proceed by a closure argument. For any $r \ge 1$ and any time-space function $\varphi \in \mathcal{C}_b^\infty([1/r,r]\times\mathbb{B}(0,r))$ with bounded derivatives of any order, we set

$$\|\varphi\|^{\mathcal{V}_0,1}_{[1/r,r]\times\mathbb{B}(0,r),\infty} = \|\varphi\|_{[1/r,r]\times\mathbb{B}(0,r),\infty} + \|\mathcal{V}_0\varphi\|_{[1/r,r]\times\mathbb{B}(0,r),\infty}.$$

[4] To be exact one has $\|V_{[\alpha_1]} \ldots V_{[\alpha_n]} P_t h\|_\infty \le C\|\nabla h\|_\infty t^{-(\|\alpha_1\|+\cdots+\|\alpha_{n-1}\|)/2}$ and inequality (17) follows as $t^{-(\|\alpha_1\|+\cdots+\|\alpha_{n-1}\|)} \le T^{m-1}t^{1-\|\boldsymbol{\alpha}\|}$ (recall that $t \le T$).

We then define $\mathcal{D}^{1,\infty}_{\mathcal{V}_0}([1/r,r]\times\mathbb{B}(0,r))$ as the closure of $\mathcal{C}^\infty_b([1/r,r]\times\mathbb{B}(0,r))$ w.r.t. $\|\cdot\|^{\mathcal{V}_0,1}_{[1/r,r]\times\mathbb{B}(0,r),\infty}$ and then define $\mathcal{D}^{1,\infty}_{\mathcal{V}_0}((0,+\infty)\times\mathbb{R}^d)$ as the intersection of the spaces $\mathcal{D}^{1,\infty}_{\mathcal{V}_0}([1/r,r]\times\mathbb{B}(0,r))$ over $r\geq 1$. (As above, the closability property is easily checked.)

We are now in position to define a classical solution to the PDE:

Definition 2.6. *We call a function $v=\{v(t,x),(t,x)\in[0,+\infty)\times\mathbb{R}^d\}$ a* classical *solution of the PDE (5) if the followings are satisfied*

(1) *v belongs to $\mathcal{D}^{1,\infty}_{\mathcal{V}_0}((0,+\infty)\times\mathbb{R}^d)$ and, for any $t>0$, $v(t,\cdot)$ is in $\mathcal{D}^{2,\infty}_V(\mathbb{R}^d)$ such that, for any $\alpha_1,\alpha_2\in\mathcal{A}_0(m)$, the function $(t,x)\in(0,+\infty)\times\mathbb{R}^d\mapsto\big(V_{[\alpha_1]}v(t,x),V_{[\alpha_1]}V_{[\alpha_2]}v(t,x)\big)$ is continuous,*

(2) *for any $(t,x)\in(0,+\infty)\times\mathbb{R}^d$, it holds*

$$\mathcal{V}_0v(t,x)=\frac{1}{2}\sum_{i=1}^N V_i^2v(t,x)+f\big(t,x,v(t,x),(Vv(t,x))^\top\big),$$

(3) *the boundary condition $\lim_{(t,y)\to(0,x)}v(t,y)=h(x)$ holds as well for any $x\in\mathbb{R}^d$.*

Remark 2.7. *We emphasize that we do not assume that a classical solution of the PDE (5) must be differentiable in the time direction or in the direction V_0. However this is the case if vector fields satisfy the uniform Hörmander condition. In this case the above definition coincides with the standard definition of a classical solution.*

As announced, here is the connection between the PDE and the BSDE (the proof is postponed to Section 7):

Proposition 2.8. *Under the standing assumption, if h is a continuous function of polynomial growth and f is bounded in (t,x), uniformly in (y,z), and twice continuously differentiable w.r.t. (x,y,z) with bounded derivatives, the function u given by* (13) *for a given $T>0$ is a classical solution to the PDE* (5) *on $(0,T]\times\mathbb{R}^d$.*

Moreover, any other classical solution v of the semilinear PDE (5) that has polynomial growth matches u. "Polynomial growth" means that there exist $C,r\geq 0$ such that

$$\forall(t,x)\in[0,T]\times\mathbb{R}^d,\quad |v(t,x)|\leq C(1+|x|^r). \tag{21}$$

2.4. **Generalized Solutions to the PDE** (5)**.** We now specify the notion of generalized solutions. A generalized solution u of the PDE (5) will be a function that is p-locally-integrable and that has p-locally-integrable generalized derivatives of second-order in the directions of the vector fields V_i, $i=1,\dots,d$, and a p-locally-integrable generalized derivative of first-order in the direction $\mathcal{V}_0=\partial_t-V_0$, when viewed as a function $(t,x)\mapsto u(t,x)$ over the product space $(0,\infty)\times\mathbb{R}^d$.

2.4.1. *Space of Generalized Solutions.* As we defined $\mathcal{D}^{k,\infty}_V(\mathbb{B})$ as the closure of $\mathcal{C}^\infty_b(\mathbb{B})$ in $\mathcal{C}_b(\bar{\mathbb{B}})$ w.r.t. $\|\cdot\|^{V,k}_{\mathbb{B},\infty}$ for a given ball $\mathbb{B}$, we can define $\mathcal{D}^{k,p}_V(\mathbb{B})$, for a given real $p\geq 1$ and for $1\leq k\leq$

$K-m+2$, as the closure of $\mathcal{C}_b^\infty(\mathbb{B})$ in $L^p(\mathbb{B})$ w.r.t. $\|\cdot\|_{\mathbb{B},p}^{V,k}$, where

$$\|\varphi\|_{\mathbb{B},p}^{V,1} = \|\varphi\|_{\mathbb{B},p} + \sum_{\alpha\in\mathcal{A}_0(m)} \|V_{[\alpha]}\varphi\|_{\mathbb{B},p},$$
$$\|\varphi\|_{\mathbb{B},p}^{V,k} = \|\varphi\|_{\mathbb{B},p}^{V,k-1} + \sum_{\alpha_1,\dots,\alpha_k\in\mathcal{A}_0(m)} \|V_{[\alpha_1]}\dots V_{[\alpha_k]}\varphi\|_{\mathbb{B},p}, \quad \varphi\in L^p(\mathbb{B}),$$

the notation $\|\cdot\|_{\mathbb{B},p}$ here standing for the L^p norm over $\mathbb{B}$. Then, we can define $\mathcal{D}_V^{k,p}(\mathbb{R}^d)$ as the intersection of all the $\mathcal{D}_V^{k,p}(\mathbb{B}(0,r))$, $r>0$. Similarly, we can define $\mathcal{D}_V^{k+1/2,p}(\mathbb{B})$ and $\mathcal{D}_V^{k+1/2,p}(\mathbb{R}^d)$ for $1\le k\le K-m+1$, $\mathcal{D}_V^{1,p}([\frac{1}{r},r]\times\mathbb{B})$ for $r>0$, and $\mathcal{D}_V^{1,p}((0,+\infty)\times\mathbb{R}^d)$.

Remark 2.9. *If the uniform Hörmander condition is satisfied, then $\mathcal{D}_V^{k,p}(\mathbb{R}^d)$ is the set of functions φ that belong to the Sobolev space $W^{k,p}(\mathbb{B}(0,r))$ for any $r>0$.*

2.4.2. *Typical Example.* A typical example of function in $\mathcal{D}_V^{k,n}(\mathbb{R}^d)$, $1\le n\le K-m$, is $x\in\mathbb{R}^d\mapsto(P_t\varphi)(x)$, for $t>0$ and $\varphi\in L^p_{\rm loc}(\mathbb{R}^d)$, φ being at most of polynomial growth at the infinity. The proof is almost the same as in the case when $p=+\infty$. The point is to consider an approximating sequence $(\varphi_\ell)_{\ell\ge1}$, converging towards φ in $L^p_{\rm loc}(\mathbb{R}^d)$ (that is in any $L^p(\mathbb{B}(0,R))$, $R>0$) and then to prove that the right-hand side in (20) is Cauchy in $L^p_{\rm loc}(\mathbb{R}^d)$. To prove it, we claim that for any $R>0$ and $\ell,k\ge0$,

$$\int_{|x|<R}\big|\mathbb{E}\big[\big(\varphi_{\ell+k}-\varphi_\ell\big)(X_t^x)\psi(t,x)\big]\big|^p dx \le C\int_{|x|<R}\mathbb{E}\big[|\varphi_{\ell+k}-\varphi_\ell|^p(X_t^x)\big]dx, \tag{22}$$

with $C=\sup_{x\in\mathbb{R}^d}\mathbb{E}\big[\psi(t,x)^{p'}\big]^{p/p'}<\infty$, $1/p+1/p'=1$. Now, the result follows from

Lemma 2.10. *Let θ_1 and θ_2 be two functions belonging to $L^p_{\rm loc}(\mathbb{R}^d)$, $p\ge1$, and at most of polynomial growth of exponent $r\ge0$ (that is $|\theta_i(x)|\le C(1+|x|^r)$, $i=1,2$, for some constant $C\ge0$), then, for any $A,R>0$,*

$$\int_{|x|<R}\mathbb{E}\big[|\theta_1-\theta_2|^p(X_t^x)\big]dx \le C'\int_{|y|<A}|\theta_1-\theta_2|^p(y)dy + C'A^{-1/2}(1+R^{rp+1/2}),$$

the constant C' being independent of A and R and depending on θ_1 and θ_2 through C and r only.

The proof of Lemma 2.10 is left to the reader: the two terms in the right-hand side are obtained by splitting the left-hand side along the events $\{|X_t^x|\le A\}$ and $\{|X_t^x|>A\}$; the first term in the right-hand side then follows from the boundedness of the inverse of the Jacobian matrix of X_t^x in any $L^q(\mathbb{P})$, $q\ge1$; the second one follows from the polynomial growth property of θ_1 and θ_2 and from Cauchy-Schwarz and Markov inequalities. Choosing $\theta_1=\varphi_{\ell+k}$ and $\theta_2=\varphi_\ell$ therein, we deduce that the right-hand side in (20) is indeed Cauchy in $L^p_{\rm loc}(\mathbb{R}^d)$. (Clearly, we can assume the $(\varphi_\ell)_{\ell\ge1}$ to be of polynomial growth, uniformly in ℓ.)

2.4.3. *Definition of Generalized Solutions.* We are now in position to define the notion of generalized solution to the PDE (5). Following Definition 2.6, we set

Definition 2.11. *We call a function $v=\{v(t,x),(t,x)\in[0,+\infty)\times\mathbb{R}^d\}$ a* generalized *solution of the PDE (5) if the followings are satisfied*

(1) *v is in $\cap_{p\geq 1}\mathcal{D}^{1,p}_{\mathcal{V}_0}((0,+\infty)\times\mathbb{R}^d)$ and, for any $t>0$, $v(t,\cdot)$ is in $\cap_{p\geq 1}\mathcal{D}^{2,p}_{V}(\mathbb{R}^d)$ such that, for any $\alpha_1,\alpha_2\in\mathcal{A}_0(m)$, the function $(t,x)\in(0,+\infty)\times\mathbb{R}^d\mapsto\big(V_{[\alpha_1]}v(t,x),V_{[\alpha_1]}V_{[\alpha_2]}v(t,x)\big)$ is measurable and in any $L^p_{\mathrm{loc}}((0,+\infty)\times\mathbb{R}^d)$, $p\geq 1$,*
(2) *for almost every $(t,x)\in(0,+\infty)\times\mathbb{R}^d$, it holds*

$$\mathcal{V}_0v(t,x)=\frac{1}{2}\sum_{i=1}^N V_i^2v(t,x)+f\big(t,x,v(t,x),(Vv(t,x))^\top\big),$$

(3) *on any compact set, $v(t,\cdot)\to h$ in Lebesgue-measure as $t\searrow 0$, that is, for any ball $\mathbb{B}\subset\mathbb{R}^d$, for any $\varepsilon>0$, $\lim_{t\searrow 0}|\{x\in\mathbb{B}:|v(t,x)-h(x)|\geq\varepsilon\}|=0$, where $|A|$ denotes the Lebesgue volume of A for a Borel subset $A\subset\mathbb{R}^d$.*

In Section 7, we will show the following

Proposition 2.12. *Under the standing assumption, if h is measurable and of polynomial growth and f is bounded in (t,x), uniformly in (y,z), and twice continuously differentiable w.r.t. (x,y,z) with bounded derivatives, the function u given by* (13) *is a generalized solution to the PDE* (5) *on $(0,T]\times\mathbb{R}^d$ for any $T>0$. Moreover, any other generalized solution v of the semilinear PDE (5) that has polynomial growth matches u almost everywhere.*

2.5. **Generalized Gronwall Lemma.** In the following we will make use of the following:

Lemma 2.13. *Consider two bounded measurable functions $g_1,g_2:[0,T]\to\mathbb{R}_+$ such that*

$$g_1(t)\leq C_1+C_2\int_t^T\frac{g_2(s)}{\sqrt{s-t}}ds, \tag{23}$$

for some constants $C_1,C_2\geq 0$. Then there exist $\lambda,\mu>0$, depending on C_2 and T only, such that

$$\begin{aligned}\int_0^T g_1(t)\exp(\lambda t)dt&\leq\mu C_1+\frac{1}{2}\int_0^T g_2(t)\exp(\lambda t)dt,\\ \sup_{0\leq t\leq T}\big[g_1(t)\big]&\leq\mu C_1+2C_2^2\int_0^T g_2(t)dt+\frac{1}{2}\sup_{0\leq t\leq T}\big[g_2(t)\big].\end{aligned} \tag{24}$$

In particular, if $g_1=g_2$, then g_1 is bounded by $\mu' C_1$, for a constant μ' depending on C_2 and T only.

Remark 2.14. *By an obvious change of variable, the result also applies in the forward sense, that is when $g_1(t)\leq C_1+C_2\int_0^t(t-s)^{-1/2}g_2(s)ds$.*

Proof. Integrating (23) w.r.t $\exp(\lambda t)$, we obtain

$$\begin{aligned}\int_0^T g_1(t)\exp(\lambda t)dt&\leq C_1\int_0^T\exp(\lambda t)dt+C_2\int_0^T\left[g_2(s)\int_0^s\frac{\exp(\lambda t)}{(s-t)^{1/2}}dt\right]ds\\ &=C_1\int_0^T\exp(\lambda t)dt+C_2\int_0^T\left[g_2(s)\exp(\lambda s)\int_0^s\frac{\exp(\lambda(t-s))}{(s-t)^{1/2}}dt\right]ds\\ &=C_1\int_0^T\exp(\lambda t)dt+C_2\int_0^T\left[g_2(s)\exp(\lambda s)\int_0^s\frac{\exp(-\lambda t)}{t^{1/2}}dt\right]ds\\ &\leq C_1\int_0^T\exp(\lambda t)dt+C_2\int_0^T\frac{\exp(-\lambda t)}{t^{1/2}}dt\int_0^T g_2(s)\exp(\lambda s)ds.\end{aligned}$$

Choosing λ large enough, this proves the first inequality in (24).

Prove now the second inequality. For any $\varepsilon > 0$, (23) yields

$$g_1(t) \le C_1 + C_2 \int_t^{(t+\varepsilon)\wedge T} \frac{g_2(s)}{(s-t)^{1/2}} ds + C_2 \varepsilon^{-1/2} \int_{(t+\varepsilon)\wedge T}^T g_2(s) ds$$
$$\le C_1 + C_2 \varepsilon^{-1/2} \int_0^T g_2(s) ds + C_2 \varepsilon^{1/2} \sup_{0\le s\le T} \big[g_2(s)\big].$$

Choosing $\varepsilon^{1/2} = 1/(2C_2)$, we complete the proof of (24).

When $g_1 = g_2$, the first inequality in (24) yields $\int_0^T \exp(\lambda t) g_1(t) dt \le 2\mu C_1$ so that $\int_0^T g_1(t) dt \le 2\mu C_1$. By the second inequality in (24),

$$\sup_{0\le t\le T} \big[g_1(t)\big] \le C_1 + 4\mu C_1 C_2^2 + \frac{1}{2} \sup_{0\le s\le T} \big[g_1(s)\big]. \qquad \square$$

3. Lipschitz Boundary Condition

3.1. **Setting and Main Result.** In the whole section, we assume that the boundary condition is Lipschitz continuous. We also assume that $|f(t,x,y,z)| \le \Lambda(1 + |x| + |y| + |z|)$, $x \in \mathbb{R}^d$, $y \in \mathbb{R}$, $z \in \mathbb{R}^N$, and that $f(t,\cdot)$ is $(K-m-1)$-times continuously differentiable, the derivatives up to any order $1 \le n \le K-m-1$ being bounded by some constant $\Lambda_n \ge 0$. (Since $K \ge m+3$, $f(t,\cdot)$ is at least twice differentiable.) To simplify things, we will assume that $\Lambda_n \ge \Lambda$.

In the following, $\boldsymbol{\alpha}$ stands for a tuple of multi-indices $(\alpha_1, \ldots, \alpha_n)$ and $\|\boldsymbol{\alpha}\|$ for $\|\alpha_1\| + \cdots + \|\alpha_n\|$. We write $\sharp(\boldsymbol{\alpha}) = n$ to say that $\boldsymbol{\alpha}$ is an n-tuple of multi-indices and denote $M_n^0(m) = \sharp\{(\beta_1, \ldots, \beta_k) \in [\mathcal{A}_0(m)]^k,\ 1 \le k \le n\}$. In the case when ∇h does not exist at point X_t^x, $|\nabla h(X_t^x)|$ will be understood as $|\nabla h(X_t^x)| = \limsup_{\varepsilon\to 0, \varepsilon\neq 0} \Gamma_d^{-1} |\varepsilon|^{-d} \int_{\{|y|\le\varepsilon\}} |\nabla h(X_t^x + y)| dy$, where Γ_d stands for the volume of the d-dimensional ball of radius 1.

We will analyse the properties of our candidate u for the solution of the PDE as defined in (14). That is

$$u(T-t,x) = \mathbb{E}\left[h(X_T^{t,x}) + \int_t^T f\big(T-s, X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x}\big) ds \right], \quad 0 \le t \le T, \quad x \in \mathbb{R}^d.$$

(Note that by time homogeneity, $u(T-t,x)$ depends on the pair (t,T) through the difference $T-t$ only, as indicated by the notation.) The objective is to prove

Theorem 3.1. *Let $(V_i)_{0\le i\le N}$ be $N+1$ vector fields satisfying Definition 1.1. Then, for any $t > 0$, $u(t,\cdot)$ belongs to $\mathcal{D}_V^{K-m-1/2,\infty}(\mathbb{R}^d)$ and is Lipschitz continuous; $u(t,\cdot)$ is continuously differentiable if h is continuously differentiable, i.e. $\nabla_x u(t,\cdot)$ exists as a continuous function.*

Moreover, for any $T > 0$, $n \le K-m-1$ and $\alpha_1, \ldots, \alpha_n \in \mathcal{A}_0(m)$, there exists a constant $C_n(p)$, depending on Λ_n, n, p, T and the vector fields $V_0, \ldots, V_N$ only, such that, for all $(t,x) \in (0,T] \times \mathbb{R}^d$,

$$\begin{aligned} &\big|V_{[\alpha_1]} \ldots V_{[\alpha_n]} u(t,x)\big| \le C_n(p) t^{(1-\|\boldsymbol{\alpha}\|)/2} \big[1 + \mathbb{E}\big[\big|\nabla h(X_t^x)\big|^{np}\big]^{1/p}\big], \\ &\big|V_{[\alpha_1]} \ldots V_{[\alpha_n]} V_i u(t,x)\big| \le C_n(p) t^{-\|\boldsymbol{\alpha}\|/2} \big[1 + \mathbb{E}\big[\big|\nabla h(X_t^x)\big|^{np}\big]^{1/p}\big], \quad 1 \le i \le N, \end{aligned} \tag{25}$$

Moreover, given $0 \le t < T$, the derivative processes by indexed $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in [\mathcal{A}_0(m)]^n$

$$\big(\big(Y_s^{\boldsymbol{\alpha}} = (V_{[\alpha_1]} \ldots V_{[\alpha_n]} u)(T-s, X_s^{t,x}), Z_s^{\boldsymbol{\alpha}} = ((V_{[\alpha_1]} \ldots V_{[\alpha_n]} V_i u)(T-s, X_s^{t,x}))_{1\le i\le N}\big)_{t\le s<T}\big)_{\boldsymbol{\alpha}}$$

are continuous and satisfy a generalized BSDE of the form

$$
\begin{aligned}
Y_s^{\boldsymbol{\alpha}} &= (S-s)^{[1-\|\boldsymbol{\alpha}\|]/2}\mathbb{E}\big[\nabla_x u(T-S,X_S^{t,x})\theta_s^*[\phi_{\boldsymbol{\alpha}}](S-s,X_s^{t,x})|\mathcal{F}_s\big] \\
&\quad + \mathbb{E}\bigg[\int_s^S \boldsymbol{F}_{\boldsymbol{\alpha}}\big(\omega,s,r,x,Y_r^{t,x},Z_r^{t,x},(Y_r^{\boldsymbol{\beta}})_{\sharp(\boldsymbol{\beta})\le n},(Z_r^{\boldsymbol{\beta}})_{\sharp(\boldsymbol{\beta})\le n}\big)dr\big|\mathcal{F}_s\bigg], \\
(Z_s^{\boldsymbol{\alpha}})_i &= (S-s)^{-\|\boldsymbol{\alpha}\|/2}\mathbb{E}\big[\nabla_x u(T-S,X_S^{t,x})\theta_s^*[\psi_{\boldsymbol{\alpha}}^i](S-s,X_s^{t,x})|\mathcal{F}_s\big] \\
&\quad + \mathbb{E}\bigg[\int_s^S (r-s)^{-1/2}\boldsymbol{G}_{\boldsymbol{\alpha}}^i\big(\omega,s,r,x,Y_r^{t,x},Z_r^{t,x},(Y_r^{\boldsymbol{\beta}})_{\sharp(\boldsymbol{\beta})\le n},(Z_r^{\boldsymbol{\beta}})_{\sharp(\boldsymbol{\beta})\le n}\big)dr\big|\mathcal{F}_s\bigg],
\end{aligned}
\tag{26}
$$

where $1\le i\le N$, $t\le s<S<T$, $\phi_{\boldsymbol{\alpha}}$ *and* $(\psi_{\boldsymbol{\alpha}}^i)_{1\le i\le N}$ *are* $\mathbb{R}^d$*-valued Kusuoka-Stroock functions in* $\mathcal{K}_0^T(K-m-n-1)$*, and* $\boldsymbol{F}_{\boldsymbol{\alpha}}(\omega,s,r,x,y,z,\xi,\zeta)$ *and* $(\boldsymbol{G}_{\boldsymbol{\alpha}}^i(\omega,s,r,x,y,z,\xi,\zeta))_{1\le i\le N}$ *are jointly measurable random functionals from* $\Omega\times[0,T]^2\times\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^N\times\mathbb{R}^{M_n^0(m)}\times\mathbb{R}^{NM_n^0(m)}$ *into* $\mathbb{R}$*, such that, a.s.,*

$$
\begin{aligned}
&\big|\big(\boldsymbol{F}_{\boldsymbol{\alpha}},(\boldsymbol{G}_{\alpha}^i)\big)(\omega,s,r,x,y,z,0,0)\big| \le \Phi(\omega,s,r,x)(1+|y|+|z|), \\
&\big|\big(\boldsymbol{F}_{\boldsymbol{\alpha}},(\boldsymbol{G}_{\boldsymbol{\alpha}}^i)_i\big)(\omega,s,r,x,y',z',\xi',\zeta') - \big(\boldsymbol{F}_{\boldsymbol{\alpha}},(\boldsymbol{G}_{\boldsymbol{\alpha}}^i)_i\big)(\omega,s,r,x,y,z,\xi,\zeta)\big|\mathbf{1}_{\{|y|,|y'|,|z|,|z'|\le R\}} \\
&\quad\le \Phi(\omega,s,r,x)\big(\Theta(\xi,\xi',\zeta,\zeta')+R\big)\big[m_R(y-y',z-z')+|\xi'-\xi|+|\zeta'-\zeta|\big], \quad R>0,
\end{aligned}
$$

where $\Phi(\omega,s,r,x)$ *is a jointly measurable functional, such that, for any* $p\ge 1$, $\mathbb{E}[|\Phi(\omega,s,r,x)|^p]$ *is uniformly bounded in* x *in compact subsets of* $\mathbb{R}^d$ *and in* $0\le s<r\le T$, $\Theta(\xi,\xi',\zeta,\zeta')$ *is a (deterministic) polynomial function and* $m_R(y,z)$ *is a (deterministic) continuous function matching* 0 *at* $(0,0)$*. In* (26)*,* $\nabla_x u(T-S,X_S^{t,x})$ *stands for a bounded* $\mathcal{F}_S$*-measurable random variable when* $\nabla_x u(T-S,\cdot)$ *doesn't exist as a true function.*

Equation (26) provides the stochastic dynamics of the derivative processes when the forward equation is initialized at x at time 0. It must be seen as a *non-linear integration by parts*, that is the equivalent to the integration by parts formula exhibited in the linear case. It must be also compared with the pathwise differentiation result in [27]. The difference between (26) and the result in [27] lies in the lack of well-defined boundary condition in (26): it would be the higher-order derivatives of h if they were well-defined. Here they don't exist as h is assumed to be Lipschitz only. As a consequence, the derivative processes are only defined up to any time $S\in[0,T)$ and the boundary like type condition is expressed as a conditional expectation: the first-order term therein is bounded in s and S so that the leading coefficient $(S-s)^{[1-\|\boldsymbol{\alpha}\|]/2}$ stands for the typical order of the boundary condition in the neighborhood of T.

A straightforward application of Lemma 2.13 shows that $(Y^{\boldsymbol{\alpha}},Z^{\boldsymbol{\alpha}})$ is the unique solution to (26) with continuous paths such that $\mathbb{E}[\sup_{t\le s\le S}|Y_s^{\boldsymbol{\alpha}}|^p+\sup_{t\le s\le S}|Z_s^{\boldsymbol{\alpha}}|^p]<+\infty$ for any $S\in[t,T)$ and for any $p\ge 1$. This is done via a standard fixed point argument similar to that used in the classical proof of the unique solvability of BSDEs driven by Z-independent drivers.

The strategy of the proof of Theorem 3.1 consists in proving the result first for the case when the boundary condition of the equation (5) is smooth and then relax the assumption via a mollification argument. Hence below, we will assume first that h is smooth in x.

3.2. **One-Step Differentiation.** The following one-step differentiation lemma permits the switch from one derivative to another:

Lemma 3.2. *Let F be a continuously differentiable function from $\mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^N$ into $\mathbb{R}$ and φ be in $\mathcal{D}_V^{3/2,\infty}(\mathbb{R}^d)$. Then, setting $\Theta(X_s^x) = (X_s^x, \varphi(X_s^x), ((V_i\varphi)(X_s^x))_{1\le i\le N})$, $0 \le s \le T$, the mapping $x \mapsto F(\Theta(X_s^x))$ is in $\mathcal{D}_V^1(\mathbb{R}^d)$ and, for any $\alpha \in \mathcal{A}_0(m)$,*

$$V_{[\alpha]}\big[F\big(\Theta(X_s^x)\big)\big] = \sum_{\beta\in\mathcal{A}_0(m)} \Big\{ b_{\alpha,\beta}(s,x)\big[V_{[\beta]}(X_s^x)\cdot\nabla_x F\big(\Theta(X_s^x)\big) + \nabla_y F\big(\Theta(X_s^x)\big)\big(V_{[\beta]}\varphi\big)(X_s^x) + \sum_{\ell=1\ldots N} \nabla_{z_\ell} F\big(\Theta(X_s^x)\big)\big(V_{[\beta]}V_\ell\varphi\big)(X_s^x)\big]\Big\}.$$

(Here, $V_{[\alpha]}$ is understood as $V_{[\alpha]}(x)\cdot\nabla$.)

Proof. When φ is a smooth function, we can write

$$V_{[\alpha]}\big[F\big(\Theta(X_s^x)\big)\big] = \sum_{i=1}^d V_{[\alpha]}^i(x) \sum_{j=1}^d \frac{\partial F\circ\Theta}{\partial x_j}(X_s^x)\frac{\partial (X_s^x)^j}{\partial x_i} = \sum_{i=1}^d\sum_{j=1}^d (J_{s,x})_{j,i} V_{[\alpha]}^i(x)\frac{\partial F\circ\Theta}{\partial x_j}(X_s^x).$$

Applying Lemma 2.3 with $t=0$, the result easily follows (when φ is smooth). By a closure argument, the result is still valid when φ is in $\mathcal{D}_V^{3/2,\infty}(\mathbb{R}^d)$. □

In the following, for any $\boldsymbol{a} = (a_1,\ldots,a_n) \in [\mathcal{A}_0(m)]^n$, we denote $\|\boldsymbol{a}\| = \sum_{i=1}^n \|a_i\|$ and we define $\mathcal{I}_k(n)$ as the set of non-decreasing sequences of (possibly zero) integers $i_1,\ldots,i_k$ such that $i_1+\cdots+i_k \le n$. For any $k \in \{0,\ldots,n\}$, we also define $\mathcal{U}_k(\varphi)$ as the set of k-tuples of functions of the form $(v_1,\ldots,v_k)$, with v_i being equal either to φ or $V_\ell\varphi$, $1\le\ell\le N$. (When $k=0$, we set $\mathcal{U}_k(\varphi)=\emptyset$). We deduce the following:

Corollary 3.3. *Let F be a $(K-m-1)$-times differentiable function from $\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^N$ into $\mathbb{R}$, with bounded derivatives of any order $1\le k\le K-m-1$, and φ be in $\mathcal{D}_V^{n+1/2}(\mathbb{R}^d)$, $n \le K-m-1$. Then, for any n-tuple of indices $\boldsymbol{\alpha} = (\alpha_1,\ldots,\alpha_n) \in [\mathcal{A}_0(m)]^n$*

$$V_{[\alpha_n]}\ldots V_{[\alpha_1]}\mathbb{E}\big[F\big(\Theta(X_s^x)\big)\big] = \sum_{k=0}^n \sum_{\boldsymbol{i}\in\mathcal{I}_k(n)} \sum_{\boldsymbol{v}\in\mathcal{U}_k(\varphi)} \sum_{\boldsymbol{\beta}=(\beta_{\ell,j})_{1\le\ell\le i_j,1\le j\le k}\in\prod_{1\le j\le k}[\mathcal{A}_0(m)]^{i_j}} \mathbb{E}\Big[\Big(\prod_{j=1}^k \big(V_{[\beta_{1,j}]}\ldots V_{[\beta_{i_j,j}]}v_j\big)(X_s^x)\Big)\phi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}(s,x)\psi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}\big(\Theta(X_s^x)\big)\Big],$$

where $\phi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}} \in \mathcal{K}^T_{(\|\boldsymbol{\beta}\|-\|\boldsymbol{\alpha}\|)^+}(K-m-n)$ and $\psi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}$ is bounded and $(K-m-n-1)$-times differentiable with bounded derivatives.

Proof. We proceed by induction. The case when $n=1$ follows from Lemma 3.2. Assume then that the result holds true for a given $n\ge 1$. Then, for a given $\alpha_{n+1}\in\mathcal{A}_0(m)$, we are to consider for any $(k,\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta})$ as above

$$V_{[\alpha_{n+1}]}\mathcal{E} \quad\text{with}\quad \mathcal{E} = \Big(\prod_{j=1}^k \big(V_{[\beta_{1,j}]}\ldots V_{[\beta_{i_j,j}]}v_j\big)(X_s^x)\Big)\phi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}(s,x)\psi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}\big(\Theta(X_s^x)\big).$$

Clearly, the term obtained by letting $V_{[\alpha_{n+1}]}$ act on $\phi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}$ gives a new Kusuoka-Stroock function belonging to $\mathcal{K}^T_{(\|\boldsymbol{\beta}\|-\|\boldsymbol{\alpha}\|)^+}(K-m-(n+1))$, which is included in $\mathcal{K}^T_{(\|\boldsymbol{\beta}\|-\|\boldsymbol{\alpha}\|-\|\alpha_{n+1}\|)^+}(K-m-(n+1))$. To differentiate $\psi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}(\Theta(X^x_s))$, we apply Lemma 3.2. There are two cases: (i) the first term in Lemma 3.2 does not add a new term of the form $V_{[\beta]}v$; (ii) the two last terms in Lemma 3.2 add new terms of the form $V_{[\beta]}v$. It is clear that (i) keeps the general form of the formula but the new ψ is $(K-m-n-2)$-times differentiable. We explain now what happens for (ii). Following Lemma 3.2, the function $\psi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}$ is differentiated; for any $\beta_{1,k+1}\in\mathcal{A}_0(m)$, the term $\mathcal{E}$ at rank n is multiplied by $V_{[\beta_{1,k+1}]}v_{k+1}$ for v_{k+1} being either φ or one of the $(V_\ell\varphi)_{1\le\ell\le N}$ and the sum is then performed over all the $\beta_{1,k+1}\in\mathcal{A}_0(m)$. It means that k is increased into $k+1$ and that $\phi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}$ is changed into $\phi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}b_{\alpha_{n+1},\beta_{1,k+1}}$. Now, $b_{\alpha_{n+1},\beta_{1,k+1}}$ is in $\mathcal{K}^T_{(\|\beta_{1,k+1}\|-\|\alpha_{n+1}\|)^+}(K-m)$. In particular, we can say that $\phi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}b_{\alpha_{n+1},\beta_{1,k+1}}$ belongs to $\mathcal{K}^T_{(\|\boldsymbol{\beta}\|-\|\boldsymbol{\alpha}\|)^+ +(\|\beta_{1,k+1}\|-\|\alpha_{n+1}\|)^+}(K-m-n)$. Since the positive part is sub-additive, that is $(x+y)^+\le x^+ + y^+$, we deduce that $\phi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}b_{\alpha_{n+1},\beta_{1,k+1}}$ belongs to $\mathcal{K}^T_{(\|\boldsymbol{\beta}\|+\|\beta_{1,k+1}\|-\|\boldsymbol{\alpha}\|-\|\alpha_{n+1}\|)^+}(K-m-n)$.

It remains to say what happens when differentiating each of the terms $(V_{[\beta_{1,j}]}\dots V_{[\beta_{i_j,j}]}v_j)(X^x_s)$. We use Lemma 2.3 with $(t,\alpha)=(0,\alpha_{n+1})$, i.e. $J_{s,x}V_{[\alpha_{n+1}]}(x)=\sum_{\beta\in\mathcal{A}_0(m)}b_{\alpha_{n+1},\beta}(s,x)V_{[\beta]}(X^x_s)$. The result is that we are increasing the length i_j for some $1\le j\le k$ from i_j to i_j+1, all the other lengthes being preserved, and that the Kusuoka-Stroock function $\phi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}$ is changed into $\phi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}b_{\alpha_{n+1},\beta}$ for any $\beta\in\mathcal{A}_0(m)$, which as we already argued belongs to $\mathcal{K}^T_{(\|\boldsymbol{\beta}\|+\|\beta\|-\|\boldsymbol{\alpha}\|-\|\alpha_{n+1}\|)^+}(K-m-n)$. (Note that some of the weight functions $\phi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}$ and $\psi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}$ in the formula at rank $n+1$ may be zero so that the sums therein run over all the possible indices.) □

3.3. **Proof of Theorem 3.1 in the Smooth Setting.** As announced, we assume first that the boundary condition h in (12) is a $\mathcal{C}^\infty_b$ function. For any $1\le n\le K-m-1$, we denote by Λ_n the common bound for the Lipschitz constant of h and for the derivatives of the coefficients up to the order n. We will make us of the following results whose proofs are postponed for the next subsection.

Lemma 3.4. *In the smooth setting, the mappings u and Vu are $(K-m-1)$-times continuously differentiable on $(0,+\infty)\times\mathbb{R}^d$ with respect to the variable x; moreover, for any $T>0$ and $1\le n\le K-m-1$, $\nabla^n_x u$ and $\nabla^n_x Vu$ are bounded on $[0,T]\times\mathbb{R}^d$.*

Proposition 3.5. *In the smooth setting, for any $p>1$ and $1\le n\le K-m-1$, there exists a constant $C_n(p)$, depending on Λ_n, n, p, T and the vector fields only, such that, for any $(\alpha_1,\dots,\alpha_n)\in$*

$(\mathcal{A}_0(m))^n$ and any $(t,x) \in (0,T] \times \mathbb{R}^d$,

$$\begin{aligned}
\big|V_{[\alpha_1]} \dots V_{[\alpha_n]} u(t,x)\big| &\le C_n(p) \Big[1 + t^{(1-\|\boldsymbol{\alpha}\|)/2} \mathbb{E}\big[\big| \nabla h(X_t^x) \big|^p \big]^{1/p} \Big] \\
&\quad + \int_{t/2}^t \sum_{k=1}^n \sum_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}} (t-s)^{(\|\boldsymbol{\beta}\|-\|\boldsymbol{\alpha}\|)^+/2} \prod_{j=1}^k \mathbb{E}\Big[\big| \big(V_{[\beta_{1,j}]} \dots V_{[\beta_{i_j,j}]} v_j \big)(s, X_{t-s}^x) \big|^{np/i_j} \Big]^{i_j/(np)} ds, \\
\big|V_{[\alpha_1]} \dots V_{[\alpha_n]} V_i u(t,x)\big| &\le C_n(p) \Big[1 + t^{-\|\boldsymbol{\alpha}\|/2} \mathbb{E}\big[\big| \nabla h(X_t^x) \big|^p \big]^{1/p} \Big] \\
&\quad + \int_{t/2}^t \sum_{k=1}^n \sum_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}} \Big[(t-s)^{[(\|\boldsymbol{\beta}\|-\|\boldsymbol{\alpha}\|)^+ - 1]/2} \\
&\qquad\qquad \times \prod_{j=1}^k \mathbb{E}\Big[\big| \big(V_{[\beta_{1,j}]} \dots V_{[\beta_{i_j,j}]} v_j \big)(s, X_{t-s}^x) \big|^{np/i_j} \Big]^{i_j/(np)} \Big] ds \Big].
\end{aligned}$$

Above, both sums run over the indices $\boldsymbol{i} = (i_1, \dots, i_k) \in \mathcal{I}_k(n)$, $\boldsymbol{v} = (v_1, \dots, v_k) \in \mathcal{U}_k(u(s,\cdot))$ and $\boldsymbol{\beta} = ((\beta_{1,j}, \dots, \beta_{i_j,j}) \in [\mathcal{A}_0(m)]^{i_j})_{1 \le j \le k}$.

We prove Theorem 3.1 by induction. For every $1 \le n \le K - m - 1$, we denote by $\mathcal{P}_n$ the following property: for any $p > 1$, there exists a constant $C_n(p)$, depending on Λ_n, n, p, T and the vector fields only, such that, for any $(\alpha_1, \dots, \alpha_n) \in (\mathcal{A}_0(m))^n$ and any $(t,x) \in (0,T] \times \mathbb{R}^d$,

$$(\mathcal{P}_n) \qquad \begin{aligned}
&\big|V_{[\alpha_1]} \dots V_{[\alpha_n]} u(t,x)\big| \le C_n(p) t^{(1-\|\boldsymbol{\alpha}\|)/2} \big[1 + \mathbb{E}\big[\big| \nabla h(X_t^x) \big|^{np} \big]^{1/p} \big], \\
&\big|V_{[\alpha_1]} \dots V_{[\alpha_n]} V_i u(t,x)\big| \le C_n(p) t^{-\|\boldsymbol{\alpha}\|/2} \big[1 + \mathbb{E}\big[\big| \nabla h(X_t^x) \big|^{np} \big]^{1/p} \big], \quad i \in \{1, \dots, N\},
\end{aligned}$$

with $\|\boldsymbol{\alpha}\| = \sum_{i=1}^n \|\alpha_i\|$.

We first prove $\mathcal{P}_1$. For a given $p > 1$, we set for any $\beta_1 \in \mathcal{A}_0(m)$

$$\mathcal{Q}_{\beta_1}^1(s,t,x) = \mathbb{E}\big[\big| \big(V_{[\beta_1]} u \big)(s, X_{t-s}^x) \big|^p \big]^{1/p} + s^{1/2} \sum_{j=1}^N \mathbb{E}\big[\big| \big(V_{[\beta_1]} V_j u \big)(s, X_{t-s}^x) \big|^p \big]^{1/p}.$$

Choose $n = 1$ in Proposition 3.5 and $\alpha_1 \in \mathcal{A}_0(m)$. Since $t - s \le s$ for any $s \in [t/2, t]$, we get

$$\begin{aligned}
\big|V_{[\alpha_1]} u(t,x)\big| &\le C_1(p) \Big[1 + t^{(1-\|\alpha_1\|)/2} \mathbb{E}\big[\big| \nabla h(X_t^x) \big|^p \big]^{1/p} \\
&\qquad + \int_{t/2}^t \sum_{\beta_1 \in \mathcal{A}_0(m)} (t-s)^{[(\|\beta_1\| - \|\alpha_1\|)_+ - 1]/2} \mathcal{Q}_{\beta_1}^1(s,t,x) ds \Big], \\
t^{1/2} \big|V_{[\alpha_1]} V_i u(t,x)\big| &\le C_1(p) \Big[1 + t^{(1-\|\alpha_1\|)/2} \mathbb{E}\big[\big| \nabla h(X_t^x) \big|^p \big]^{1/p} \\
&\qquad + \int_{t/2}^t \sum_{\beta_1 \in \mathcal{A}_0(m)} (t-s)^{[(\|\beta_1\| - \|\alpha_1\|)_+ - 1]/2} \mathcal{Q}_{\beta_1}^1(s,t,x) ds \Big],
\end{aligned}$$

where $i = 1, \dots, N$. By the bound $s \geq t/2$ again, both inequalities can be incorporated into:

$$t^{(\|\alpha_1\|-1)/2}\Big[\big|V_{[\alpha_1]}u(t,x)\big| + t^{1/2}\sum_{i=1}^{N}\big|V_{[\alpha_1]}V_iu(t,x)\big|\Big]$$
$$\leq C_1(p)\Big[1+\mathbb{E}\big[\big|\nabla h(X_t^x)\big|^p\big]^{1/p} + \sum_{\beta_1\in\mathcal{A}_0(m)}\int_{t/2}^{t}(t-s)^{-1/2}s^{(\|\beta_1\|-1)/2}\mathcal{Q}^1_{\beta_1}(s,t,x)\Big]ds,$$

the constant $C_1(p)$ possibly varying from line to line hereafter.

Choosing x of the form X^x_{r-t}, with $r \geq t$, taking the L^p moment, applying Minkowski's integral inequality, and then summing over $\alpha_1 \in \mathcal{A}_0(m)$ and $i \in \{1, \dots, N\}$, we eventually obtain

$$\sum_{\alpha_1\in\mathcal{A}_0(m)} t^{(\|\alpha_1\|-1)/2}\Big[\mathbb{E}\big[\big|\big(V_{[\alpha_1]}u\big)(t,X^x_{r-t})\big|^p\big]^{1/p} + \sum_{i=1}^{N}t^{1/2}\mathbb{E}\big[\big|\big(V_{[\alpha_1]}V_iu\big)(t,X^x_{r-t})\big|^p\big]^{1/p}\Big]$$
$$\leq C_1(p)\Big[1+\mathbb{E}\big[\big|\nabla h(X_r^x)\big|^p\big]^{1/p} + \sum_{\beta_1\in\mathcal{A}_0(m)}\int_{t/2}^{t}(t-s)^{-1/2}s^{(\|\beta_1\|-1)/2}\mathcal{Q}^1_{\beta_1}(s,r,x)ds\Big].$$

We emphasize that the left-hand side is nothing but $\sum_{\alpha_1\in\mathcal{A}_0(m)} t^{(\|\alpha_1\|-1)/2}\mathcal{Q}^1_{\beta_1}(t,r,x)$. By Lemma 2.13 (applied in the forward sense), we complete the proof of $\mathcal{P}_1$.

We turn to the proof of the induction property. Assume that $\mathcal{P}_k$ holds for every $1 \leq k \leq n-1$, for some rank $2 \leq n \leq K-m-1$. We make use of Proposition 3.5 at rank n. We have two cases: $i_k = n$ and $i_k < n$. When $i_k = n$, the sum over $\boldsymbol{\beta}$ actually reduces to a sum over $\boldsymbol{\beta} = (\beta_1, \dots, \beta_n) \in [\mathcal{A}_0(m)]^n$ and the product of the V's reduces to a single term of the form $V_{[\beta_1]}\dots V_{[\beta_n]}v$, v running over the set $\{u(s,\cdot), V_\ell u(s,\cdot), 1 \leq \ell \leq N\}$. In this case, we do not use the induction property. When $i_k < n$, all the possible i_j's, $1 \leq j \leq k$, are also (strictly) less than n. That is, the terms of the form $V_{[\beta_{1,j}]}\dots V_{[\beta_{i_j,j}]}v_j$ fulfill the induction property, i.e., for any $1 \leq j \leq k$,

$$\big|\big(V_{[\beta_{1,j}]}\dots V_{[\beta_{i_j,j}]}v_j\big)(s,X^x_{t-s})\big| \leq C_n(p)s^{\delta/2-(\sum_{\ell=1}^{i_j}\|\beta_{\ell,j}\|)/2}\big[1+\mathbb{E}\big[\big|\nabla h(X_t^x)\big|^{i_jp}|\mathcal{F}_{t-s}\big]^{1/p}\big],$$

with δ being equal to 1 when $v_j(s,\cdot)$ matches $u(s,\cdot)$ and being equal to 0 when $v_j(s,\cdot)$ matches some $V_iu(s,\cdot)$, $1 \leq i \leq N$. Clearly, the worst rates hold for the term

$$\big|\big(V_{[\beta_{1,j}]}\dots V_{[\beta_{i_j,j}]}v_j\big)(s,X^x_{t-s})\big| \leq C_n(p)s^{-(\sum_{\ell=1}^{i_j}\|\beta_{\ell,j}\|)/2}\big[1+\mathbb{E}\big[\big|\nabla h(X_t^x)\big|^{i_jp}|\mathcal{F}_{t-s}\big]^{1/p}\big]. \tag{27}$$

We then obtain

$$\begin{aligned}
&\prod_{j=1}^{k}\Big[\mathbb{E}\big[\big|\big(V_{[\beta_{1,j}]}\dots V_{[\beta_{i_j,j}]}v_j\big)(s,X^x_{t-s})\big|^{np/i_j}\big]\Big]^{i_j/(np)}\\
&\leq C_n(p)s^{-(\sum_{j=1}^{k}\sum_{\ell=1}^{i_j}\|\beta_{\ell,j}\|)/2}\prod_{j=1}^{k}\big[1+\mathbb{E}\big[\big|\nabla h(X_t^x)\big|^{np}\big]\big]^{i_j/(np)}\\
&\leq C_n(p)s^{-\|\boldsymbol{\beta}\|/2}\mathbb{E}\big[\big(1+\big|\nabla h(X_t^x)\big|\big)^{np}\big]^{1/p},
\end{aligned} \tag{28}$$

where $\boldsymbol{\beta}$ stands for the k-tuple of multi-indices $((\beta_{\ell,j})_{1\leq\ell\leq i_j})_{1\leq j\leq k}$.

Plugging these bounds into Proposition 3.5, we obtain (up to a modification of the constant $C_n(p)$)

$$
\begin{aligned}
\big|V_{[\alpha_1]} \dots V_{[\alpha_n]} u(t,x)\big| &\le C_n(p) \Big[1 + t^{(1-\|\boldsymbol{\alpha}\|)/2} \mathbb{E}\big[\big| \nabla h(X_t^x) \big|^p \big]^{1/p} \Big] \\
&+ C_n(p) \sum_{\boldsymbol{\beta}=(\beta_1,\dots,\beta_k) \in [\mathcal{A}_0(m)]^k, k<n} \int_{t/2}^t s^{-\|\boldsymbol{\beta}\|/2} (t-s)^{(\|\boldsymbol{\beta}\|-\|\boldsymbol{\alpha}\|)^+/2} \mathcal{R}(t,x) ds \\
&+ C_n(p) \sum_{\beta_1,\dots,\beta_n \in \mathcal{A}_0(m)} \int_{t/2}^t (t-s)^{[(\|\boldsymbol{\beta}\|-\|\boldsymbol{\alpha}\|)^+ - 1]/2} \mathcal{Q}^n_{\beta_1,\dots,\beta_n}(s,t,x) ds \\
&= T_1(t,x) + T_2(t,x) + T_3(t,x),
\end{aligned}
\tag{29}
$$

with

$$
\begin{aligned}
\mathcal{R}(t,x) &= \mathbb{E}\big[\big(1 + \big| \nabla h(X_t^x) \big| \big)^{np} \big]^{1/p} \\
\mathcal{Q}^n_{\beta_1,\dots,\beta_n}(s,t,x) &= \mathbb{E}\big[\big| \big(V_{[\beta_1]} \dots V_{[\beta_n]} u \big)(s, X_{t-s}^x) \big|^p \big]^{1/p} \\
&\qquad + s^{1/2} \sum_{i=1}^N \mathbb{E}\big[\big| \big(V_{[\beta_1]} \dots V_{[\beta_n]} V_i u \big)(s, X_{t-s}^x) \big|^p \big]^{1/p}.
\end{aligned}
$$

By replacing x with X^x_{r-t}, $r \ge t$, taking the L^p moment and using Minkowski's integral inequality we get

$$
t^{(\|\boldsymbol{\alpha}\|-1)/2} \mathbb{E}\big[\big| T_2(t, X^x_{r-t}) \big|^p \big]^{1/p} \le C_n(p) t^{1/2} \mathbb{E}\big[\big(1 + \big| \nabla h(X_r^x) \big| \big)^{np} \big]^{1/p}. \tag{30}
$$

Similarly,

$$
\begin{aligned}
&t^{(\|\boldsymbol{\alpha}\|-1)/2} \mathbb{E}\big[\big| T_3(t, X^x_{r-t}) \big|^p \big]^{1/p} \\
&\qquad \le C_n(p) \sum_{\beta_1,\dots,\beta_n \in \mathcal{A}_0(m)} \int_{t/2}^t (t-s)^{-1/2} s^{(\|\boldsymbol{\beta}\|-1)/2} \mathcal{Q}^n_{\beta_1,\dots,\beta_n}(s,r,x) ds.
\end{aligned}
\tag{31}
$$

By (29), (30) and (31), we deduce

$$
\begin{aligned}
t^{(\|\boldsymbol{\alpha}\|-1)/2} \mathbb{E}\big[\big| \big(V_{[\alpha_1]} \dots V_{[\alpha_n]} u \big)(t, X^x_{r-t}) \big|^p \big]^{1/p} &\le C_n(p) \mathbb{E}\big[\big(1 + \big| \nabla h(X_r^x) \big| \big)^{np} \big]^{1/p} \\
&+ C_n(p) \sum_{\beta_1,\dots,\beta_n \in \mathcal{A}_0(m)} \int_{t/2}^t (t-s)^{-1/2} s^{(\|\boldsymbol{\beta}\|-1)/2} \mathcal{Q}^n_{\beta_1,\dots,\beta_n}(s,r,x) ds.
\end{aligned}
\tag{32}
$$

By a similar argument,

$$
\begin{aligned}
t^{\|\boldsymbol{\alpha}\|} \sum_{i=1}^N \mathbb{E}\big[\big| \big(V_{[\alpha_1]} \dots V_{[\alpha_n]} V_i u \big)(t, X^x_{r-t}) \big|^p \big]^{1/p} &\le C_n(p) \mathbb{E}\big[\big(1 + \big| \nabla h(X_r^x) \big| \big)^{np} \big]^{1/p} \\
&+ C_n(p) \sum_{\beta_1,\dots,\beta_n \in \mathcal{A}_0(m)} \int_{t/2}^t (t-s)^{-1/2} s^{(\|\boldsymbol{\beta}\|-1)/2} \mathcal{Q}^n_{\beta_1,\dots,\beta_n}(s,r,x) ds.
\end{aligned}
\tag{33}
$$

Summing (32) and (33) over $(\alpha_1,\dots,\alpha_n)\in[\mathcal{A}_0(m)]^n$, we obtain

$$\begin{aligned}\sum_{\alpha_1,\dots,\alpha_n\in\mathcal{A}_0(m)} t^{(\|\boldsymbol{\alpha}\|-1)/2}\mathcal{Q}^n_{\alpha_1,\dots,\alpha_n}(t,r,x) &\le C_n(p)\mathbb{E}\big[\big(1+\big|\nabla h(X_r^x)\big|\big)^{np}\big]^{1/p}\\ &+C_n(p)\sum_{\beta_1,\dots,\beta_n\in\mathcal{A}_0(m)}\int_{t/2}^t (t-s)^{-1/2}s^{(\|\boldsymbol{\beta}\|-1)/2}\mathcal{Q}^n_{\beta_1,\dots,\beta_n}(s,x)ds.\end{aligned}\tag{34}$$

By Lemma 2.13 (applied in the forward sense), we complete the induction proof. □

3.4. **Proofs of Lemma 3.4 and Proposition 3.5.** The proofs rely on the technical lemma:

Lemma 3.6. *Consider three random jointly measurable functions $\Psi:(\omega,t,x)\in\Omega\times[0,T]\times\mathbb{R}^d\mapsto\Psi(\omega,t,x)\in\mathbb{R}^{d_1}$, $\Phi:(\omega,t,s,x)\in\Omega\times[0,T]^2\times\mathbb{R}^d\mapsto\Phi(\omega,t,x)\in\mathbb{R}_+$ and $F:(\omega,t,s,x,\zeta)\in\Omega\times[0,T]^2\times\mathbb{R}^d\times\mathbb{R}^{d_1}\mapsto F(\omega,t,s,x,\zeta)\in\mathbb{R}^{d_1}$ such that, a.s., for any $t\in[0,T]$, the mappings $x\mapsto\Psi(\omega,t,x)$ and $(x,\zeta)\mapsto F(\omega,t,x,\zeta)$ are continuously differentiable. Assume in addition that $(\Phi(\omega,s,t,x))_{0\le s<t\le T,x\in\mathbb{R}^d}$ is in $L^p(\Omega)$, uniformly in $0\le s<t\le T$ and $x\in\mathbb{R}^d$, for any $p\ge1$. Assume finally that*

$$\begin{aligned}&|\Psi(\omega,t,0)|\le\Phi(\omega,0,t,0),\quad |F(\omega,s,t,x,0)|\le\Phi(\omega,s,t,x),\\ &|\nabla_x\Psi(\omega,t,x)|\le\Phi(\omega,0,t,x),\quad |\nabla_xF(\omega,s,t,x,\zeta)|\le\Phi(\omega,s,t,x)(1+|\zeta|),\\ &|\nabla_\zeta F(\omega,s,t,x,\zeta)|\le\Phi(\omega,s,t,x).\end{aligned}\tag{35}$$

If $\bar v:[0,T]\times\mathbb{R}^d\to\mathbb{R}^{d_1}$ is a function in $L^\infty([0,T],\mathcal{C}_b(\mathbb{R}^d))$ that satisfies

$$\bar v(t,x)=\mathbb{E}\bigg[\Psi(t,x)+\int_t^T(s-t)^{-1/2}F\big(\omega,t,s,x,\bar v(s,X_s^{t,x})\big)ds\bigg],\tag{36}$$

then $\bar v(t,\cdot)$ is Lipschitz continuous, uniformly in t. Moreover, if, a.s., for any $s\in[0,T]$, $x\in\mathbb{R}^d$, $\zeta\in\mathbb{R}^{d_1}$, the functions $t\in[0,T]\mapsto\Psi(t,x)$ and $t\in(0,s)\mapsto F(\omega,t,s,x,\zeta)$ are continuous, then $\bar v$ is continuous on $[0,T]\times\mathbb{R}^d$.

Proof. We introduce the following mapping

$$\begin{aligned}&\Phi: L^\infty\big([0,T],\mathcal{C}_b(\mathbb{R}^d)\big)\to L^\infty\big([0,T],\mathcal{C}_b(\mathbb{R}^d)\big)\\ &\quad v\mapsto\bigg(w:(t,x)\in[0,T]\times\mathbb{R}^d\mapsto\mathbb{E}\bigg[\Psi(t,x)+\int_t^T\frac{F\big(\omega,t,s,x,v(s,X_s^{t,x})\big)}{(s-t)^{1/2}}ds\bigg]\bigg).\end{aligned}$$

There exists a constant C (whose value may vary below) such that, for any $t\in[0,T]$ and $x\in\mathbb{R}^d$,

$$\|(w_1-w_2)(t,\cdot)\|_\infty\le C\int_t^T\frac{\|(v_1-v_2)(s,\cdot)\|_\infty}{(s-t)^{1/2}}ds,\tag{37}$$

with $w_1=\Phi(v_1)$ and $w_2=\Phi(v_2)$. By Lemma 2.13,

$$\int_0^T\exp(\lambda t)\|(w_1-w_2)(t,\cdot)\|_\infty dt\le\frac12\int_0^T\exp(\lambda s)\|(v_1-v_2)(s,\cdot)\|_\infty ds,$$

for some $\lambda>0$. Thus, the mapping Φ is a contraction on $L^\infty([0,T],\mathcal{C}_b(\mathbb{R}^d))$ endowed with the semi-norm $v\mapsto\int_0^T\exp(\lambda t)\|v(t,\cdot)\|_\infty dt$. In particular, if $\bar v$ satisfies (36) and $\tilde v$ is a fixed point of Φ,

then, for a.e. $t \in [0,T]$, $\tilde{v}(t,\cdot) = \bar{v}(t,\cdot)$. By (37), $\tilde{v}(t,\cdot) = \bar{v}(t,\cdot)$ for any $t \in [0,T]$. Similarly, for a recursive sequence $(v_{n+1} = \Phi(v_n))_{n\geq 0}$, $v_0 = 0$, we get

$$\lim_{n\to+\infty} \int_0^T \exp(\lambda t) \|(v_n - \bar{v})(t,\cdot)\|_\infty dt = 0.$$

By (37) and Lemma 2.13 again,

$$\sup_{0\leq t\leq T} \|(v_{n+1} - \bar{v})(t,\cdot)\|_\infty \leq \frac{1}{2} \sup_{0\leq s\leq T} \|(v_n - \bar{v})(s,\cdot)\|_\infty + C \int_0^T \|(v_n - \bar{v})(s,\cdot)\|_\infty ds. \tag{38}$$

We deduce that $\sup_{0\leq t\leq T} \|v_n(t,\cdot) - \bar{v}(t,\cdot)\|_\infty$ converges towards 0. Therefore, if the functions $((v_n(t,\cdot))_{t\in[0,T]})_{n\geq 1}$ are Lipschitz continuous, uniformly in t and in n, $\bar{v}(t,\cdot)$ is Lipschitz continuous as well, uniformly in $t \in [0,T]$. By induction, it is clear that all the $v_n(t,\cdot)$ are continuously differentiable. By (35),

$$\|\nabla_x v_{n+1}(t,\cdot)\|_\infty \leq C + C \int_t^T \frac{\|\nabla_x v_n(s,\cdot)\|_\infty}{(t-s)^{1/2}} ds, \tag{39}$$

since the functions $(v_n)_{n\geq 1}$ are bounded, uniformly in n. (The value of C may vary below.) We use Lemma 2.13 again. For a possibly new value of λ,

$$\int_0^T \exp(\lambda t) \|\nabla_x v_{n+1}(t,\cdot)\|_\infty dt \leq C + \frac{1}{2} \int_0^T \exp(\lambda t) \|\nabla_x v_n(t,\cdot)\|_\infty dt.$$

Iterating the bound, we get $\int_0^T \exp(\lambda t) \|\nabla_x v_n(t,\cdot)\|_\infty dt \leq C$. In particular, by (39) and Lemma 2.13

$$\|\nabla_x v_{n+1}(t,\cdot)\|_\infty \leq C + \frac{1}{2} \sup_{0\leq s\leq T} \|\nabla_x v_n(s,\cdot)\|_\infty. \tag{40}$$

Iterating, we obtain that $\sup_{n\geq 1} \sup_{0\leq t\leq T} \|\nabla_x v_n(t,\cdot)\|_\infty < +\infty$.

When the random functions Ψ and F satisfy the prescribed continuity conditions w.r.t. the time parameter, all the functions $(v_n)_{n\geq 1}$ are continuous on $[0,T]\times\mathbb{R}^d$; by local uniform convergence of the sequence $(v_n)_{n\geq 1}$ towards $\bar{v}$, $\bar{v}$ is continuous. □

Proof (Lemma 3.4). The first-order continuous differentiability of $u(t,\cdot)$ is a straightforward consequence of Pardoux and Peng [27]. Moreover, for any initial condition $(t,x) \in [0,T]\times\mathbb{R}^d$, the solution $(\nabla_x Y_s^{t,x}, \nabla_x Z_s^{t,x})_{t\leq s\leq T}$ to the derivative BSDE

$$\begin{aligned} \nabla_x Y_s^{t,x} &= \nabla h(X_T^{t,x}) \nabla_x X_T^{t,x} - \int_s^T dB_r^\top \nabla_x Z_r^{t,x} \\ &+ \int_s^T \big[\nabla_x f\big(\Theta_r^{t,x}\big) \nabla_x X_r^{t,x} + \nabla_y f\big(\Theta_r^{t,x}\big) \nabla_x Y_r^{t,x} + \nabla_z f\big(\Theta_r^{t,x}\big) \nabla_x Z_r^{t,x}\big] dr, \end{aligned} \tag{41}$$

with $\Theta_r^{t,x} = (T-r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})$, satisfies

$$\lim_{h\to 0} \mathbb{E}\big[\sup_{t\le s\le T} \big|\frac{Y_s^{t,x+h}-Y_s^{t,x}}{h} - \nabla_x Y_s^{t,x}\big|^2 + \sup_{t\le s\le T} \big|\nabla_x Y_s^{t,x+h} - \nabla_x Y_s^{t,x}\big|^2\big] = 0, \quad x\in\mathbb{R}^d, \tag{42}$$

$$\lim_{h\to 0} \mathbb{E}\Big[\int_t^T \big|\frac{Z_s^{t,x+h}-Z_s^{t,x}}{h} - \nabla_x Z_s^{t,x}\big|^2 ds + \int_t^T \big|\nabla_x Z_s^{t,x+h} - \nabla_x Z_s^{t,x}\big|^2 ds\Big] = 0, \quad x\in\mathbb{R}^d. \tag{43}$$

Clearly, (41) yields $\sup_{0\le t\le T}\|\nabla_x u(t,\cdot)\|_\infty < +\infty$, since $\nabla_x f$, $\nabla_y f$ and $\nabla_z f$ are bounded, that is

$$\sup_{0\le t\le T} \big\|\nabla_x u(t,\cdot)\big\|_\infty \le C(\Lambda_1, T), \tag{44}$$

where $C(\Lambda_1, T)$ depends on Λ_1, T and the bounds of the derivatives of the vector fields $V_0,\dots,V_N$ only. Precisely, by Proposition 3.2 in Briand et al. [2], we have that for any $p>1$

$$\forall (t,x)\in[0,T)\times\mathbb{R}^d, \quad \big|\nabla_x u(t,x)\big| \le C(\Lambda_1, p, T)\big[1+\mathbb{E}\big[\big|\nabla h(X_t^x)\big|^p\big]^{1/p}\big], \tag{45}$$

for some constant $C(\Lambda_1, p, T)$ depending on Λ_1, p, T and the bounds of the derivatives of the vector fields $V_0,\dots,V_N$ only.

We now go back to the backward formulation of $u(t,\cdot)$:

$$u(T-t,x) = \mathbb{E}\big[h(X_T^{t,x})\big] + \int_t^T \mathbb{E}\big[f\big(T-s, X_s^{t,x}, u(T-s,X_s^{t,x}), (Vu)^\top(T-s,X_s^{t,x})\big)\big]ds.$$

By the example in Subsection 2.3 and by Lebesgue dominated theorem, we know that the right-hand side is in $\mathcal{D}_V^{1/2}(\mathbb{R}^d)$ and that for any $1\le i\le N$,

$$\begin{aligned} V_i u(T-t,x) &= \mathbb{E}\big[\nabla h(X_T^{t,x}) V_i X_T^{t,x}\big] \\ &\quad + \int_t^T (s-t)^{-1/2}\mathbb{E}\big[f\big(T-s, X_s^{t,x}, u(T-s,X_s^{t,x}), (Vu)^\top(T-s,X_s^{t,x})\big)\theta_t^*(\psi_i)(s-t,x)\big]ds, \end{aligned}$$

where $V_i X_T^{t,x}$ being understood as $\nabla_x X_T^{t,x} V_i(x)$. Above, ψ_i stands for a Kusuoka-Stroock function in $\mathcal{K}_0^T(K-m-1)$ and $\theta_t^*(\psi_i)$ indicates that the randomness is evaluated after shifting. (See Subsection 2.2.) Clearly, we can rewrite the above expression as

$$V_i u(T-t,x) = \mathbb{E}\big[\nabla h(X_T^{t,x}) V_i X_T^{t,x}\big] + \int_t^T \frac{\mathbb{E}\big[f\big(T-s, X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x}\big)\theta_t^*(\psi_i)(s-t,x)\big]}{(s-t)^{1/2}}ds. \tag{46}$$

We need to apply (42) and (43) to differentiate the right-hand side under the integral. However $(\nabla_x Z_s^{t,x})_{t\le s\le T}$ is in $L^2([t,T]\times\Omega)$ only so that the convergence of the integral of $(s-t)^{-1/2}|\nabla_x Z_s^{t,x}|$ is not guaranteed.

We now make use of Lemma 3.6. Since $\theta_t^*(\psi_i)(s-t,x)$ is centered, we can replace $f(T-s, X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})$ by $f(T-s, X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x}) - f(T-s,x,u(s,x),0)$ in (46) and then apply Lemma 3.6 with $\Psi(t,x) = \nabla h(X_T^{t,x}) V_i X_T^{t,x}$, $F(t,s,x,\zeta) = [f(T-s, X_s^{t,x}, u(s,X_s^{t,x}),\zeta) - f(T-s,x,u(s,x),0)]\theta_t^*(\psi)(s-t,x)$, and obviously, $\bar v(t,x) = (Vu)^\top(t,x)$. We then deduce that $Vu(t,\cdot)$ is Lipschitz continuous, uniformly in t. Writing $\mathbb{E}[f(T-s, X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})\theta_t^*(\psi_i)(s-t,x)]$ as $\mathbb{E}[f(T-s, X_{s-t}^x, u(s, X_{s-t}^x), (Vu)^\top(s, X_{s-t}^x))\psi_i(s-t,x)]$ and then taking advantage of the time-continuity of X and ψ_i, we also deduce from Lemma 3.6 that Vu is continuous on $(0,T]\times\mathbb{R}^d$. In

particular, for any $0 \le t \le s \le T$, the mapping $x \mapsto Z_s^{t,x} = (Vu)^\top(s, X_s^{t,x})$ is locally Lipschitz continuous, i.e. for any $x \in \mathbb{R}^d$,

$$\sup_{y,y' \in \mathbb{B}(x,1)} |Z_s^{t,y'} - Z_s^{t,y}| \le \vartheta(x)|y' - y|,$$

where ϑ is a random variable in any L^p, uniformly in x and s. In particular, by (43), we can choose a version of $\nabla_x Z_s^{t,x}$ that is in any $L^p(\Omega)$, uniformly in s and x.

We now go back to (46). By (43), we know that the term inside the integral is continuously differentiable for any $s > t$. Since $\nabla_x Z_s^{t,x}$ is in any $L^p(\Omega)$, uniformly in s and x, we deduce that $Vu(t,\cdot)$ is continuously differentiable as well and that $\nabla_x Vu(t,\cdot)$ is bounded uniformly in t.

The proof is completed by an induction step. We now assume that, for a given $1 \le n \le K-m-2$, $u(t,\cdot)$ and $Vu(t,\cdot)$ are n-times continuously differentiable in all the directions of the space, with bounded derivatives, uniformly in t. We also assume that, for any $0 \le k \le n-1$, the functions $\nabla_x^k u$ and $\nabla_x^k Vu$ are continuous on $(0,T] \times \mathbb{R}^d$.

By Lemma 2.2, we can differentiate the pair $(Y_s^{t,x}, Z_s^{t,x})_{t \le s \le T}$ pathwise n times. The dynamics of the derivative process $(\nabla_x^n Y_s^{t,x}, \nabla_x^n Z_s^{t,x})_{t \le s \le T}$ may be summarized as follows:

$$\begin{aligned} \nabla_x^n Y_s^{t,x} &= H^n(t,x) + \int_s^T \big[F^n(t,s,x) + \nabla_y f(\Theta_r^{t,x}) \nabla_x^n Y_r^{t,x} + \nabla_z f(\Theta_r^{t,x}) \nabla_x^n Z_r^{t,x}\big] dr \\ &\quad - \int_s^T dB_r^\top \nabla_x^n Z_r^{t,x}, \end{aligned} \tag{47}$$

where $H^n(t,x)$ is an $\mathcal{F}_T$-measurable r.v., bounded in any $L^p(\Omega)$, $p \ge 1$, uniformly in (t,x), and $(F^n(t,s,x))_{t<s\le T}$ is a progressively-measurable process (w.r.t. s), bounded in any $L^p(\Omega)$, $p \ge 1$, uniformly in $0 \le t < s \le T$ and in x. Obviously, $H^n(t,x)$ is given by the differentiation of the boundary condition and $F^n(t,s,x)$ by the differentiation of the driver of the BSDE: $F^n(t,s,x)$ contains all the derivatives of X up to order n and all the derivatives of (Y,Z) up to order $n-1$. In particular, $F^n(t,s,x)$ is a.s. continuously differentiable w.r.t. x, with bounded derivatives in any $L^p(\Omega)$, $p \ge 1$, uniformly in $0 \le t < s \le T$ and in x (by the induction assumption).

Following the strategy developed in (46) and differentiating n times therein, we obtain as generic equation for $\nabla_x^n Vu(t,\cdot)$:

$$\begin{aligned} \nabla_x^n V_i u(T-t,x) &= \mathbb{E}\big[H^{n+1/2}(t,x)\big] \\ &\quad + \int_t^T \frac{\mathbb{E}\big[G^n(t,s,x) + \big(\nabla_y f(\Theta_r^{t,x}) \nabla_x^n Y_r^{t,x} + \nabla_z f(\Theta_r^{t,x}) \nabla_x^n Z_r^{t,x}\big) \theta_t^*(\psi_i)(s-t,x)\big]}{(s-t)^{1/2}} dr, \end{aligned} \tag{48}$$

$1 \le i \le N$, for some $\psi_i \in \mathcal{K}_0^T(K-m-1)$. Above, G^n is obtained by differentiating both the driver of the BSDE and the Kusuoka-Stroock function in (46). In particular, by centering f as in (46), we can assume that G^n satisfies the same properties as F^n. Moreover, $H^{n+1/2}(t,x)$ is a.s. continuously differentiable, with derivatives in any $L^p(\Omega)$, for any $p \ge 1$. (Basically, $H^{n+1/2}(t,x)$ is obtained by differentiating $(n+1)$-times the boundary condition. Since $n+2 \le K$ and $(t,x) \mapsto X_{T-t}^x$ is in $\mathcal{K}_T^0(K)$, $H^{n+1/2}$ is continuously differentiable w.r.t. x.)

Making use of (47) and (48) and applying the time-space continuity argument in Lemma 3.6 to the pair $(\nabla_x^n u, \nabla_x^n Vu)$, we deduce that $(\nabla_x^n u, \nabla_x^n Vu)$ is continuous on $(0,T] \times \mathbb{R}^d$. By the same strategy as in (41), we also deduce that the pair $(\nabla_x^{n+1} Y_s^{t,x}, \nabla_x^{n+1} Z_s^{t,x})_{t \le s \le T}$ exists as in (42) and (43). (See

also Footnote[5].) Clearly, $\nabla_x^{n+1}u(t,\cdot)$ is well-defined and continuous, and it is bounded, uniformly in t. To establish the continuous differentiability of $\nabla_x^n Vu(T-t,\cdot)$, we use the same strategy as in the case $n=1$ by applying first Lemma 3.6 to (48). This proves that $\nabla_x^{n+1}Vu(T-t,\cdot)$ is a continuous function and that it is bounded, uniformly in t. Writing the dynamics for $\nabla_x^{n+1}u(T-t,\cdot)$ and $\nabla_x^{n+1}Vu(T-t,\cdot)$ and applying the time-space continuity argument in Lemma 3.6, we finally establish that $\nabla_x^{n+1}u(T-\cdot,\cdot)$ and $\nabla_x^{n+1}Vu(T-\cdot,\cdot)$ are continuous on $[0,T)\times\mathbb{R}^d$. □

At last we are in a position to give the proof of Proposition 3.5. In the following we estimate the higher order-derivatives of u along the vector fields. We write, for all $t>0$ and $x\in\mathbb{R}^d$,

$$u(t,x)=P_{t/2}\big[u\big(\frac{t}{2},\cdot\big)\big](x)+\int_{t/2}^t P_{t-s}\big[f\big(s,\cdot,u(s,\cdot),(Vu)^\top(s,\cdot)\big)\big](x)ds. \tag{49}$$

For n given multi-indices $\alpha_1,\dots,\alpha_n$ in $\mathcal{A}_0(m)$,

$$\begin{aligned} V_{[\alpha_1]}\dots V_{[\alpha_n]}u(t,x) &= V_{[\alpha_1]}\dots V_{[\alpha_n]}P_{t/2}\big[u\big(\frac{t}{2},\cdot\big)\big](x)\\ &\quad+\int_{t/2}^t V_{[\alpha_1]}\dots V_{[\alpha_n]}P_{t-s}\big[f\big(s,\cdot,u(s,\cdot),(Vu)^\top(s,\cdot)\big)\big](x)ds\\ &= T_1(t,x)+T_2(t,x). \end{aligned} \tag{50}$$

By Theorem 2.5 (see also Corollaries 3.10 and 3.14 in [21]), we can find a family of Kusuoka-Stroock functions $(\phi^j_{\alpha_1,\dots,\alpha_n})_{1\le j\le N}$ in $\mathcal{K}_0^T(K-m-n+1)$ such that

$$\begin{aligned} T_1(t,x) &= V_{[\alpha_1]}\dots V_{[\alpha_{n-1}]}\mathbb{E}\big[V_{[\alpha_n]}\big(u\big(\frac{t}{2},X^x_{t/2}\big)\big)\big]\\ &= \sum_{j=1}^d V_{[\alpha_1]}\dots V_{[\alpha_{n-1}]}\mathbb{E}\big[\big(J_{t/2,x}V_{[\alpha_n]}(x)\big)_j\frac{\partial u}{\partial x_j}\big(\frac{t}{2},X^x_{t/2}\big)\big)\big]\\ &= t^{-(1/2)\sum_{i=1}^{n-1}\|\alpha_i\|}\sum_{j=1}^d\mathbb{E}\big[\phi^j_{\alpha_1,\dots,\alpha_n}\big(\frac{t}{2},x\big)\frac{\partial u}{\partial x_j}\big(\frac{t}{2},X^x_{t/2}\big)\big]. \end{aligned} \tag{51}$$

Therefore, for any $p>1$, we can find a constant $C_n(p)$, depending on T and the bounds for the higher-order derivatives of the vector fields only and possibly varying from line to line, such that

$$\begin{aligned} |T_1(t,x)| &\le C_n(p)t^{1/2-(1/2)\|\boldsymbol{\alpha}\|}\mathbb{E}\big[\big|\nabla_x u\big(\frac{t}{2},X^x_{t/2}\big)\big|^p\big]^{1/p}\\ &\le C_n(p)t^{1/2-(1/2)\|\boldsymbol{\alpha}\|}\mathbb{E}\big[1+\big|\nabla h\big(X^x_t\big)\big|^p\big]^{1/p}, \end{aligned} \tag{52}$$

the last line following from (45). We emphasize that the exponent in t is $1/2(1-\|\boldsymbol{\alpha}\|)$, where $\|\boldsymbol{\alpha}\|=|\alpha_1|+\dots+|\alpha_n|$. Compared with (51), the additional $1/2$ follows from the term $|\alpha_n|$, which is not taken into account in (51). We here see that the smoothing decay of the boundary condition behaves as in the linear case exactly. The major hurdle is to handle the nonlinear term.

[5]We note that (42) and (43) stand for continuous differentiability in L^2-mean. Although, this is weaker than pathwise continuous differentiability, it is sufficient in our setting. To establish differentiability in L^2-mean, there is no need to apply Kolmogorov continuity theorem and thus no need to assume Hölder continuity of the derivatives of the coefficients.

By Corollary 3.3 with (φ,θ) therein possibly depending on s, that is with φ of the form $u(s,\cdot)$ and $\Theta(X^x_{t-s})$ of the form $\Theta(s,X^x_{t-s})=\Theta(s,X^x_{t-s},u(s,X^x_{t-s}),(V_iu(s,X^x_{t-s}))_{1\le i\le N})$, we write

$$T_2(t,x)=\int_{t/2}^t\sum_{k,\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}\mathbb{E}\Big[\Big(\prod_{j=1}^k\big(V_{[\beta_{1,j}]}\dots V_{[\beta_{i_j,j}]}v_j\big)(s,X^x_{t-s})\Big)\times\phi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}(t-s,x)\psi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}\big(\Theta(s,X^x_{t-s})\big)\Big]ds, \tag{53}$$

where the shorten notation $(k,\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta})$ is as in Corollary 3.3: it stands for $k\in\{0,\dots,n\}$, $\boldsymbol{i}\in\mathcal{I}_k(n)$, $\boldsymbol{v}\in\mathcal{U}_k(u(s,\cdot))$ and $\boldsymbol{\beta}=(\beta_{1,\ell},\dots,\beta_{i_\ell,\ell})_{1\le\ell\le k}\in\prod_{\ell=1}^k[\mathcal{A}_0(m)]^{i_\ell}$. Keeping in mind that $\phi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}\in\mathcal{K}_{(\|\boldsymbol{\beta}\|-\|\boldsymbol{\alpha}\|)^+}(K-m-n)$ and that $\psi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}$ is bounded, we deduce that, for any $p>1$,

$$\begin{aligned}&|T_2(t,x)|\\&\le C_n(p)\sum_{k,\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}\int_{t/2}^t(t-s)^{(\|\boldsymbol{\beta}\|-\|\boldsymbol{\alpha}\|)^+/2}\mathbb{E}\Big[\prod_{j=1}^k\big|V_{[\beta_{1,j}]}\dots V_{[\beta_{i_j,j}]}v_j(s,X^x_{t-s})\big|^p\Big]^{1/p}ds\\&\le C_n(p)\sum_{k,\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}\int_{t/2}^t(t-s)^{(\|\boldsymbol{\beta}\|-\|\boldsymbol{\alpha}\|)^+/2}\prod_{j=1}^k\mathbb{E}\Big[\big|V_{[\beta_{1,j}]}\dots V_{[\beta_{i_j,j}]}v_j(s,X^x_{t-s})\big|^{np/i_j}\Big]^{i_j/(np)}ds,\end{aligned} \tag{54}$$

the constant $C_n(p)$ possibly depending on Λ_n as well. Similarly, we can compute, for any index $i\in\{1,\dots,N\}$, $V_{[\alpha_1]}\dots V_{[\alpha_n]}V_iu(t,x)$

$$\begin{aligned}V_{[\alpha_1]}\dots V_{[\alpha_n]}V_iu(t,x)&=V_{[\alpha_1]}\dots V_{[\alpha_n]}V_iP_{t/2}\big[u\big(\frac{t}{2},\cdot\big)\big](x)\\&\quad+\int_{t/2}^tV_{[\alpha_1]}\dots V_{[\alpha_n]}V_iP_{t-s}\big[f\big(s,\cdot,u(s,\cdot),(Vu)^\top(s,\cdot)\big)\big](x)ds\\&=S_1(t,x)+S_2(t,x).\end{aligned} \tag{55}$$

Following the proof of (52), we obtain

$$|S_1(t,x)|\le C_n(p)t^{-\|\boldsymbol{\alpha}\|/2}\mathbb{E}\big[\big|\nabla_xu\big(\frac{t}{2},X^x_{t/2}\big)\big|^p\big]^{1/p}\le C_n(p)t^{-\|\boldsymbol{\alpha}\|/2}\big[1+\mathbb{E}\big[\big|\nabla h\big(X^x_t\big)\big|^p\big]^{1/p}\big].$$

We now turn to S_2. By Integration by Parts (see Corollary 3.12 in [21]), we emphasize that

$$\begin{aligned}V_iP_{t-s}\big[f\big(s,\cdot,u(s,\cdot),(Vu)^\top(s,\cdot)\big)\big](x)&=V_i\mathbb{E}\big[f\big(\Theta(s,X^x_{t-s})\big)\big]\\&=(t-s)^{-1/2}\mathbb{E}\big[f\big(\Theta(s,X^x_{t-s})\big)\phi^0_i(t-s,x)\big],\end{aligned}$$

for some Kusuoka-Stroock function $\phi^0_i\in\mathcal{K}^T_0(K-m-1)$. Therefore,

$$\begin{aligned}&V_{[\alpha_1]}\dots V_{[\alpha_n]}V_iP_{t-s}\big[f\big(s,\cdot,u(s,\cdot),(Vu)^\top(s,\cdot)\big)\big](x)\\&=(t-s)^{-(1/2)}V_{[\alpha_1]}\dots V_{[\alpha_n]}\mathbb{E}\big[f\big(\Theta(s,X^x_{t-s})\big)\phi^0_i(t-s,x)\big].\end{aligned}$$

Differentiating the product, we obtain

$$
\begin{aligned}
&V_{[\alpha_1]}\dots V_{[\alpha_n]}V_iP_{t-s}\big[f\big(s,\cdot,u(s,\cdot),(Vu)^\top(s,\cdot)\big)\big](x)\\
&=(t-s)^{-(1/2)}\sum_{k=1}^{n}\sum_{1\le\ell_1<\dots<\ell_k\le n}\mathbb{E}\big[V_{[\alpha_{\ell_1}]}\dots V_{[\alpha_{\ell_k}]}\{f\big(\Theta(s,X_{t-s}^x)\big)\}\phi_i^{\ell_1,\dots,\ell_k}(t-s,x)\big]\\
&\quad+(t-s)^{-(1/2)}\mathbb{E}\big[f\big(\Theta(s,X_{t-s}^x)\big)\phi_i^n(t-s,x)\big]\\
&=T_3(s,t,x)+T_4(s,t,x)
\end{aligned}
\tag{56}
$$

for new Kusuoka-Stroock functions $\phi_i^{\ell_1,\dots,\ell_k},\phi_i^n\in\mathcal{K}_0^T(K-m-n-1)$.

To bound $T_4(s,t,x)$, we observe that $\phi_i^n(t-s,x)$ is centered, so that

$$
|T_4(s,t,x)|=(t-s)^{-1/2}\big|\mathbb{E}\big[\big\{f(\Theta(s,X_{t-s}^x))-\mathbb{E}\big[f(\tilde\Theta(s,x))\big]\big\}\phi_i^n(t-s,x)\big]\big|,
\tag{57}
$$

with $\tilde\Theta(s,x)=(s,x,u(s,X_{t-s}^x),\mathbb{E}[(Vu)^\top(s,X_{t-s}^x)])$. By the Lipschitz property of f, we deduce

$$
\begin{aligned}
|T_4(s,t,x)|&\le C(t-s)^{-1/2}\big(1+\mathbb{E}\big[|Vu(s,X_{t-s}^x)|^p\big]^{1/p}\big)\\
&\quad+(t-s)^{-1/2}\big|\mathbb{E}\big[\big\{f(\tilde\Theta(s,x))-\mathbb{E}\big[f(\tilde\Theta(s,x))\big]\big\}\phi_i^n(t-s,x)\big]\big|.
\end{aligned}
$$

By Clark-Ocone formula and then by integration by parts formula,

$$
\begin{aligned}
|T_4(s,t,x)|&\le C(t-s)^{-1/2}\big(1+\mathbb{E}\big[|Vu(s,X_{t-s}^x)|^p\big]^{1/p}\big)\\
&\quad+(t-s)^{-1/2}\left|\mathbb{E}\int_0^{t-s}\langle\mathbb{E}[\partial_y f(\tilde\Theta(s,x))D_r[u(s,X_{t-s}^x)]|\mathcal{F}_r],D_r\phi_i^n(t-s,x)\rangle dr\right|.
\end{aligned}
$$

By definition of a Kusuoka-Stroock function, the process $(D_r\phi_i^n(r,x))_{0\le r\le t-s}$ belongs to the space $L^q(\Omega,d\mathbb{P};L^2([0,t-s],dr))$, for any $q\ge1$, so that, for any $\varepsilon>0$,

$$
\begin{aligned}
|T_4(s,t,x)|&\le C(p,\varepsilon)(t-s)^{-1/2}\Bigg(1+\mathbb{E}\big[|Vu(s,X_{t-s}^x)|^p\big]^{1/p}\\
&\qquad+\mathbb{E}\Bigg[\bigg(\int_0^{t-s}\mathbb{E}[|D_r[u(s,X_{t-s}^x)]|\,|\mathcal{F}_r]^2dr\bigg)^{(1+\varepsilon)/2}\Bigg]^{1/(1+\varepsilon)}\Bigg).
\end{aligned}
\tag{58}
$$

By the well-known relationship $D_r^iX_{t-s}^x=J_{t-s}^x(J_r^x)^{-1}V_i(X_r^x)$ and by Lemma 2.3, we claim

$$
\begin{aligned}
D_r^iu(s,X_{t-s}^x)&=\nabla_xu(s,X_{t-s}^x)J_{t-s}^x(J_r^x)^{-1}V_i(X_r^x)\\
&=\sum_{\gamma_1\in\mathcal{A}_0(m)}a_{i,\gamma_1}(r,x)\nabla_xu(s,X_{t-s}^x)J_{t-s}^xV_{[\gamma_1]}(x)\\
&=\sum_{\gamma_1,\gamma_2\in\mathcal{A}_0(m)}a_{i,\gamma_1}(r,x)b_{\gamma_1,\gamma_2}(t-s,x)\nabla_xu(s,X_{t-s}^x)V_{[\gamma_2]}(X_{t-s}^x).
\end{aligned}
\tag{59}
$$

Since a_{i,γ_1} is time-progressively measurable and belongs to $\mathcal{K}^T_{(\|\gamma_1\|-1)^+}(K-m)$ and b_{γ_1,γ_2} belongs to $\mathcal{K}^T_{(\|\gamma_2\|-\|\gamma_1\|)^+}(K-m)$, we deduce, for the specific choice $1+3\varepsilon=p$,

$$
\begin{aligned}
&\mathbb{E}\Big[\Big(\int_0^{t-s}\mathbb{E}\big[|D_r\big(u(s,X^x_{t-s})\big)|\,|\mathcal{F}_r\big]^2dr\Big)^{(1+\varepsilon)/2}\Big]^{1/(1+\varepsilon)}\\
&\le\sum_{\gamma_1,\gamma_2\in\mathcal{A}_0(m)}\mathbb{E}\Big[\Big(\int_0^{t-s}a^2_{i,\gamma_2}(r,x)\mathbb{E}\big[|b_{\gamma_1,\gamma_2}(t-s,x)V_{[\gamma_2]}u(s,X^x_{t-s})|\,|\mathcal{F}_r\big]^2dr\Big)^{(1+\varepsilon)/2}\Big]^{1/(1+\varepsilon)}\\
&\le\sum_{\gamma_1,\gamma_2\in\mathcal{A}_0(m)}\mathbb{E}\big[\sup_{0\le r\le t-s}\mathbb{E}\big[|b_{\gamma_1,\gamma_2}(t-s,x)V_{[\gamma_2]}u(s,X^x_{t-s})|\,|\mathcal{F}_r\big]^{(1+2\varepsilon)}\big]^{1/(1+2\varepsilon)}\\
&\qquad\qquad\times\mathbb{E}\Big[\Big(\int_0^{t-s}a^2_{i,\gamma_2}(r,x)dr\Big)^{(1+\varepsilon)(1+2\varepsilon)/(2\varepsilon)}\Big]^{\varepsilon/[(1+\varepsilon)(1+2\varepsilon)]}\\
&\le C(p)\big(1+\sum_{\gamma\in\mathcal{A}_0(m)}(t-s)^{(\|\gamma\|-1)/2}\mathbb{E}\big[|V_{[\gamma]}u(s,X^x_{t-s})|^p\big]^{1/p}\big),
\end{aligned}
$$

the last line following from Doob's inequality for martingales. By (58)

$$
|T_4(s,t,x)|\le C(t-s)^{-1/2}\big(1+\sum_{\gamma\in\mathcal{A}_0(m)}(t-s)^{(\|\gamma\|-1)/2}\mathbb{E}\big[|V_{[\gamma]}u(s,X^x_{t-s})|^p\big]^{1/p}\big). \tag{60}
$$

To handle $T_3(s,t,x)$, we apply Corollary 3.3 again. For any $1\le\ell_1<\cdots<\ell_k\le n$, we can write

$$
\begin{aligned}
&V_{[\alpha_{\ell_1}]}\ldots V_{[\alpha_{\ell_k}]}\{f\big(\Theta(s,X^x_{t-s})\big)\}\\
&=\sum_{k'=0}^k\sum_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}\Big[\prod_{j=1}^{k'}\big(V_{[\beta_{1,j}]}\ldots V_{[\beta_{i_j,j}]}v_j(s,X^x_{t-s})\big)\phi^{\ell_1,\ldots,\ell_k}_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}(t-s,x)\psi^{\ell_1,\ldots,\ell_k}_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}\big(\Theta(s,X^x_{t-s})\big)\Big],
\end{aligned}
\tag{61}
$$

where the notation $(\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta})$ stands for $\boldsymbol{i}\in\mathcal{I}_{k'}(k)$, $\boldsymbol{v}\in\mathcal{U}_{k'}(u(s,\cdot))$ and $\boldsymbol{\beta}=(\beta_{1,j},\ldots,\beta_{i_j,j})_{1\le j\le k'}\in\prod_{j=1}^{k'}[\mathcal{A}_0(m)]^{i_j}$, $\phi^{\ell_1,\ldots,\ell_k}_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}$ stands for a Kusuoka-Stroock function belonging to $\mathcal{K}^T_{(\|\boldsymbol{\beta}\|-\sum_{p=1}^k\|\alpha_{\ell_p}\|)^+}(K-m-k)$ and $\psi^{\ell_1,\ldots,\ell_k}_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}$ stands for a bounded function.

Therefore, denoting by $\boldsymbol{\ell}$ the increasing sequence $1\le\ell_1<\cdots<\ell_k\le n$ and gathering (56), (60) and (61)

$$
\begin{aligned}
&V_{[\alpha_1]}\ldots V_{[\alpha_n]}V_iP_{t-s}\big[f\big(s,\cdot,u(s,\cdot),(Vu)^\top(s,\cdot)\big)\big](x)\\
&\le C_n(p)(t-s)^{-1/2}\\
&\quad\times\sum_{k=0}^n\sum_{\boldsymbol{\ell}}\sum_{k'=0}^k\sum_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}(t-s)^{(\|\boldsymbol{\beta}\|-\|\boldsymbol{\alpha}\|)^+/2}\prod_{j=1}^{k'}\mathbb{E}\Big[\big|V_{[\beta_{1,j}]}\ldots V_{[\beta_{i_j,j}]}v_j(s,X^x_{t-s})\big|^{np/i_j}\Big]^{i_j/(np)}\\
&\le C_n(p)\sum_{k=0}^n\sum_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}(t-s)^{(\|\boldsymbol{\beta}\|-\|\boldsymbol{\alpha}\|)^+/2-1/2}\prod_{j=1}^{k}\mathbb{E}\Big[\big|V_{[\beta_{1,j}]}\ldots V_{[\beta_{i_j,j}]}v_j(s,X^x_{t-s})\big|^{np/i_j}\Big]^{i_j/(np)},
\end{aligned}
$$

where the shorter notation in the last line above stands for $\boldsymbol{i}\in\mathcal{I}_k(n)$, $\boldsymbol{v}\in\mathcal{U}_k(u(s,\cdot))$ and $\boldsymbol{\beta}=(\beta_{1,j},\ldots,\beta_{i_j,j})_{1\le j\le k}\in\prod_{j=1}^k[\mathcal{A}_0(m)]^{i_j}$. We emphasize that the case $k=0$ is the constant case: the

product is understood as being equal to 1; we also notice that the case $k = 1$ contains inequality (60): choose $i_1 = 1$, $\beta_{1,1} = \gamma$ and $v_1(s,\cdot) = u(s,\cdot)$. On the right-hand sides of the two estimates in the statement of Proposition 3.5, the sum over k starts from $k = 1$: the case when $k = 0$ is contained in the additional 1 in the boundary term. □

3.5. **Proof of Theorem 3.1 in the general case.** The first step is to obtain the representation formula (26) in the smooth setting. For a given $s \in [t, S)$, it follows from (50), (53), (55), (56) and (61) replacing therein the initial point (t, x) of the diffusion process by its current position $(s, X_s^{t,x})$ and noting that the random variable $\mathbb{E}[\nabla_x u(T-S, X_{S-s}^x)\phi_{\boldsymbol{\alpha}}(S-s,x)]_{|x=X_s^{t,x}}$ is a version of $\mathbb{E}[\nabla_x u(T-S, X_S^{t,x})\theta_s^*[\phi_{\boldsymbol{\alpha}}](S-s, X_s^{t,x})|\mathcal{F}_s]$. To prove that, almost-surely, (26) holds for any $s \in [t, S)$, some continuity argument is necessary. By Lemma 3.4, $(Y_s^{\boldsymbol{\alpha}})_{t\le s<S}$ and $(Z_s^{\boldsymbol{\alpha}})_{t\le s<S}$ are continuous w.r.t. s. Clearly, the conditional expectations of the integrals from s to S of $\boldsymbol{F}_{\boldsymbol{\alpha}}$ and $\boldsymbol{G}_{\boldsymbol{\alpha}}$ are continuous as well. Finally, $\mathbb{E}[\nabla_x u(T-S, X_{S-s}^y)\phi_{\boldsymbol{\alpha}}(S-s,y)]_{|y=X_s^{t,x}}$ is continuous with respect to s since $\nabla_x u$ is time-continuous.

3.5.1. *Mollification of the Boundary Condition.* When the boundary condition h is Lipschitz continuous only, we denote by $(h_\ell)_{\ell\ge1}$ a sequence of mollifications of h converging towards h uniformly on compact sets and we denote by $(u_\ell)_{\ell\ge1}$ the associated family of solutions. Using the stability property (see for example [26] and [27]) of the BSDE (11), the sequence of corresponding solutions $(u_\ell)_{\ell\ge1}$ converges towards u uniformly on compact subsets of $[0,T]\times\mathbb{R}^d$. By the standard maximum principle, there exists a constant C, independent of ℓ, such that

$$\forall (t,x) \in [0,T]\times\mathbb{R}^d, \quad |u_\ell(t,x)| \le C(1+|x|). \tag{62}$$

By (44) (for a possibly new value of C),

$$\forall (t,x) \in [0,T]\times\mathbb{R}^d,\ \forall 1\le i\le N, \quad |V_i u_\ell(t,x)| \le C. \tag{63}$$

3.5.2. *Representation Formula for the Mollified Solutions.* To get the convergence of the derivatives of u_ℓ, we notice that the terminal condition in (26) may be written in terms of $u(T-S, X_S^{t,x})$ itself instead of $\nabla_x u(T-S, X_S^{t,x})$. Specifically, for any $1 \le n \le K-m-1$, $\ell \ge 1$ and $x \in \mathbb{R}^d$, the family of derivative pair processes

$$\Big(\big(Y_s^{\ell,\boldsymbol{\alpha}} = (V_{[\alpha_1]}\dots V_{[\alpha_n]}u_\ell)(T-s, X_s^{t,x}),\ Z_s^{\ell,\boldsymbol{\alpha}} = ((V_{[\alpha_1]}\dots V_{[\alpha_n]}V_i u_\ell)(T-s, X_s^{t,x}))_{1\le i\le N}\big)_{t\le s<T}\Big)$$

indexed by sequences of multi-indices $\boldsymbol{\alpha} = (\alpha_1,\dots,\alpha_n) \in [\mathcal{A}_0(m)]^n$, satisfies

$$\begin{aligned}
Y_s^{\ell,\boldsymbol{\alpha}} &= (S-s)^{-\|\boldsymbol{\alpha}\|/2}\mathbb{E}\big[u_\ell(T-S, X_S^{t,x})\theta_s^*[\phi_{\boldsymbol{\alpha}}](S-s, X_s^{t,x})|\mathcal{F}_s\big] \\
&\quad + \mathbb{E}\bigg[\int_s^S \boldsymbol{F}_{\boldsymbol{\alpha}}\big(\omega,s,r,x,Y_r^{\ell,t,x}, Z_r^{\ell,t,x}, (Y_r^{\ell,\boldsymbol{\beta}})_{\sharp(\boldsymbol{\beta})\le n}, (Z_r^{\ell,\boldsymbol{\beta}})_{\sharp(\boldsymbol{\beta})\le n}\big)dr\big|\mathcal{F}_s\bigg], \\
(Z_s^{\ell,\boldsymbol{\alpha}})_i &= (S-s)^{-[1+\|\boldsymbol{\alpha}\|]/2}\mathbb{E}\big[u_\ell(T-S, X_S^{t,x})\theta_s^*[\psi_{\boldsymbol{\alpha}}](S-s, X_s^{t,x})|\mathcal{F}_s\big] \\
&\quad + \mathbb{E}\bigg[\int_s^S (r-s)^{-1/2}\boldsymbol{G}_{\boldsymbol{\alpha}}^i\big(\omega,s,r,x,Y_r^{\ell,t,x}, Z_r^{\ell,t,x}, (Y_r^{\ell,\boldsymbol{\beta}})_{\sharp(\boldsymbol{\beta})\le n}, (Z_r^{\ell,\boldsymbol{\beta}})_{\sharp(\boldsymbol{\beta})\le n}\big)dr\big|\mathcal{F}_s\bigg],
\end{aligned} \tag{64}$$

with $1 \le i \le N$, where $Y_r^{\ell,t,x} = u_\ell(T-r, X_r^{t,x})$ and $Z_r^{\ell,t,x} = (V_i u_\ell(T-r, X_r^{t,x}))_{1\le i\le N}$ and where the functions $\phi_{\boldsymbol{\alpha}}$ and $\psi_{\boldsymbol{\alpha}}$ will differ from the original ones in (26). (Here they are $\mathbb{R}$-valued.)

3.5.3. *Convergence of the Sequence* $(Z^{\ell,t,x})_{\ell\geq 1}$. We emphasize that the second line in (64) makes sense when $\boldsymbol{\alpha}=\emptyset$. It provides a representation formula for $(Z_s^{\ell,t,x})_{t\leq s<T}$ of the form

$$\begin{aligned} Z_s^{\ell,t,x} &= (S-s)^{-1/2}\mathbb{E}\big[u_\ell(T-S,X_S^{t,x})\theta_s^*[\psi_\emptyset](S-s,X_s^{t,x})|\mathcal{F}_s\big] \\ &\quad + \mathbb{E}\bigg[\int_s^S (r-s)^{-1/2}\boldsymbol{G}_\emptyset(\omega,s,r,x)f(r,X_r^{t,x},Y_r^{\ell,t,x},Z_r^{\ell,t,x})\big|\mathcal{F}_s\bigg], \end{aligned} \tag{65}$$

$t\leq s<S$, $\boldsymbol{G}_\emptyset(\omega,s,r,x)$ being a random functional with values in $\mathbb{R}^N$ such that, for any $p>1$, $\mathbb{E}[|\boldsymbol{G}_\emptyset(\omega,s,r,x)|^p|\mathcal{F}_s]$ is uniformly bounded in randomness, in $x\in\mathbb{R}^d$ and in $t\leq s<r<S$. By Cauchy-Schwarz inequality, we can find a constant C (independent of ℓ_1, ℓ_2, t and x) such that

$$\begin{aligned} |Z_s^{\ell_1,t,x}-Z_s^{\ell_2,t,x}|^2 &\leq C(S-s)^{-1}\mathbb{E}\big[|u_{\ell_1}(T-S,X_S^{t,x})-u_{\ell_2}(T-S,X_S^{t,x})|^2|\mathcal{F}_s\big] \\ &\quad + C\int_s^S (r-s)^{-1/2}\big(\mathbb{E}\big[|Z_r^{\ell_1,t,x}-Z_r^{\ell_2,t,x}|^2|\mathcal{F}_s\big]+\mathbb{E}\big[|Y_r^{\ell_1,t,x}-Y_r^{\ell_2,t,x}|^2|\mathcal{F}_s\big]\big)dr. \end{aligned} \tag{66}$$

Taking the expectation and then the supremum over $\ell_1,\ell_2\geq\ell$, we get by (63), that for any $S'\in(t,S)$ and $s\in[t,S']$, we have

$$\begin{aligned} &\sup_{\ell_1,\ell_2\geq\ell}\mathbb{E}\big[|Z_s^{\ell_1,t,x}-Z_s^{\ell_2,t,x}|^2\big] \\ &\leq C(S-S')^{-1}\sup_{\ell_1,\ell_2\geq\ell}\mathbb{E}\big[|u_{\ell_1}(T-S,X_S^{t,x})-u_{\ell_2}(T-S,X_S^{t,x})|^2\big]+C(S-S')^{1/2} \\ &\quad +C\sup_{\ell_1,\ell_2\geq\ell}\sup_{t\leq r\leq S}\mathbb{E}\big[|Y_r^{\ell_1,t,x}-Y_r^{\ell_2,t,x}|^2\big]+C\int_s^{S'}(r-s)^{-1/2}\sup_{\ell_1,\ell_2\geq\ell}\mathbb{E}\big[|Z_r^{\ell_1,t,x}-Z_r^{\ell_2,t,x}|^2\big]dr. \end{aligned}$$

By Lemma 2.13, for any $t\leq s\leq S'$

$$\begin{aligned} &\sup_{\ell_1,\ell_2\geq\ell}\mathbb{E}\big[|Z_s^{\ell_1,t,x}-Z_s^{\ell_2,t,x}|^2\big] \\ &\leq C(S-S')^{-1}\sup_{\ell_1,\ell_2\geq\ell}\mathbb{E}\big[|u_{\ell_1}(T-S,X_S^{t,x})-u_{\ell_2}(T-S,X_S^{t,x})|^2\big]+C(S-S')^{1/2} \\ &\quad +C\sup_{\ell_1,\ell_2\geq\ell}\sup_{t\leq r\leq S}\mathbb{E}\big[|u_{\ell_1}(T-r,X_r^{t,x})-u_{\ell_2}(T-r,X_r^{t,x})|^2\big]. \end{aligned} \tag{67}$$

Taking the supremum w.r.t. $x\in\mathcal{K}$ in (67), $\mathcal{K}$ standing for a compact subset of $\mathbb{R}^d$, we deduce that

$$\lim_{\ell\to+\infty}\sup_{x\in\mathcal{K}}\sup_{\ell_1,\ell_2\geq\ell}\sup_{t\leq s\leq S'}\mathbb{E}\big[|Z_s^{\ell_1,t,x}-Z_s^{\ell_2,t,x}|^2\big]=0. \tag{68}$$

Below, we show that the supremum over $s\in[t,S']$ can be put inside the expectation. Going back to (66), taking the supremum therein w.r.t. $s\in[t,S']$, applying Doob's inequality for martingales to the first term in the right-hand side and applying Hölder's inequality with exponents $(4/3,4)$ to the

second term in the right-hand side, we obtain

$$\begin{aligned}
&\mathbb{E}\big[\sup_{t\le s\le S'}|Z_s^{\ell_1,t,x}-Z_s^{\ell_2,t,x}|^2\big]\\
&\le C(S-S')^{-1}\mathbb{E}\big[|u_{\ell_1}(T-S,X_S^{t,x})-u_{\ell_2}(T-S,X_S^{t,x})|^2\big]\\
&\quad+C\mathbb{E}\bigg[\sup_{t\le s\le S'}\int_s^S(r-s)^{-1/2}\big(\mathbb{E}\big[|Z_r^{\ell_1,t,x}-Z_r^{\ell_2,t,x}|^2+|Y_r^{\ell_1,t,x}-Y_r^{\ell_2,t,x}|^2|\mathcal{F}_s\big]\big)dr\bigg]\\
&\le C(S-S')^{-1}\mathbb{E}\big[|u_{\ell_1}(T-S,X_S^{t,x})-u_{\ell_2}(T-S,X_S^{t,x})|^2\big]\\
&\quad+C\sup_{t\le s\le S'}\bigg(\int_s^S(r-s)^{-2/3}dr\bigg)^{3/4}\\
&\qquad\times\bigg(\int_t^S\mathbb{E}\big[\sup_{t\le s\le S'}\big(\mathbb{E}\big[|Z_r^{\ell_1,t,x}-Z_r^{\ell_2,t,x}|^8|\mathcal{F}_s\big]+\mathbb{E}\big[|Y_{r+s}^{\ell_1,t,x}-Y_{r+s}^{\ell_2,t,x}|^8|\mathcal{F}_s\big]\big)\big]dr\bigg)^{1/4}.
\end{aligned}$$

By Doob's inequality again, we deduce

$$\begin{aligned}
\mathbb{E}\big[\sup_{t\le s\le S'}|Z_s^{\ell_1,t,x}-Z_s^{\ell_2,t,x}|^2\big]&\le C(S-S')^{-1}\mathbb{E}\big[|u_{\ell_1}(T-S,X_S^{t,x})-u_{\ell_2}(T-S,X_S^{t,x})|^2\big]\\
&\quad+C\bigg(\int_t^S\mathbb{E}\big[|Z_r^{\ell_1,t,x}-Z_r^{\ell_2,t,x}|^{16}+|Y_r^{\ell_1,t,x}-Y_r^{\ell_2,t,x}|^{16}\big]dr\bigg)^{1/8}.
\end{aligned}$$

By the bounds (62) and (63) and by (68), we finally deduce that, for any $t\le S'<S$,

$$\lim_{\ell\to+\infty}\sup_{x\in\mathcal{K}}\sup_{\ell_1,\ell_2\ge\ell}\mathbb{E}\big[\sup_{t\le s\le S'}|Z_s^{\ell_1,t,x}-Z_s^{\ell_2,t,x}|^2\big]=0. \tag{69}$$

We deduce that, for any $t\le S<T$, the processes $((Z_s^{\ell,t,x})_{t\le s\le S})_{\ell\ge1}$ are convergent w.r.t. the norm $\mathbb{E}[\sup_{t\le s\le S}|\cdot_s|^2]^{1/2}$, uniformly with respect to x taking value in compact subsets of $\mathbb{R}^d$.

3.5.4. *Proof that $u(t,\cdot)$ Belongs to $\mathcal{D}_V^{1/2,\infty}(\mathbb{R}^d)$.* Taking $s=t$ in (69), we deduce that $(Vu_\ell(t,x))_{\ell\ge1}$ is uniformly convergent w.r.t. x in compact subsets of $\mathbb{R}^d$. This shows that $u(t,\cdot)\in\mathcal{D}_V^{1/2,\infty}(\mathbb{R}^d)$ for any $t>0$.

3.5.5. *Existence of Higher-Order Derivatives.* From the preliminary result (69) and from the bounds we have for $(Y^{\ell,\boldsymbol{\alpha}},Z^{\ell,\boldsymbol{\alpha}})$ (see Theorem 3.1 in the mollified setting), we know that, for any $p\ge1$,

$$\begin{aligned}
\delta_\ell^{(p)}(S)=\sup_{t\le s<r\le S}\sup_{\ell_1,\ell_2\ge\ell}\mathbb{E}\big[\big|&\big(\boldsymbol{F}_{\boldsymbol{\alpha}},(\boldsymbol{G}_{\boldsymbol{\alpha}}^i)_i\big)\big(\omega,s,r,x,Y_r^{\ell_1,t,x},Z_r^{\ell_1,t,x},(Y_r^{\ell_1,\boldsymbol{\beta}})_{\sharp(\boldsymbol{\beta})\le n},(Z_r^{\ell_1,\boldsymbol{\beta}})_{\sharp(\boldsymbol{\beta})\le n}\big)\\
&-\big(\boldsymbol{F}_{\boldsymbol{\alpha}},(\boldsymbol{G}_{\boldsymbol{\alpha}}^i)_i\big)\big(\omega,s,r,x,Y_r^{\ell_2,t,x},Z_r^{\ell_2,t,x},(Y_r^{\ell_1,\boldsymbol{\beta}})_{\sharp(\boldsymbol{\beta})\le n},(Z_r^{\ell_1,\boldsymbol{\beta}})_{\sharp(\boldsymbol{\beta})\le n}\big)\big|^p\big]
\end{aligned}$$

in (64) converges towards 0 as ℓ tends to $+\infty$, uniformly in x in compact sets. We can follow (67) to derive from (64)

$$\begin{aligned}
\sup_{\sharp\boldsymbol{\alpha}\le n}\sup_{\ell_1,\ell_2\ge\ell}\mathbb{E}\big[|Y_s^{\ell_1,\boldsymbol{\alpha}}-Y_s^{\ell_2,\boldsymbol{\alpha}}|^2+|Z_s^{\ell_1,\boldsymbol{\alpha}}-Z_s^{\ell_2,\boldsymbol{\alpha}}|^2\big]&\le C(S-S')^{1/2}+C\delta_\ell^{(2)}(S)\\
&\quad+C(S-S')^{-(n+1)/2}\sup_{\ell_1,\ell_2\ge\ell}\mathbb{E}\big[|u_{\ell_1}(T-S,X_S^{t,x})-u_{\ell_2}(T-S,X_S^{t,x})|^2\big].
\end{aligned}$$

Note that C depends on S. Following the proof of (69), we can also prove that, for any $t \le S < T$,

$$\lim_{\ell\to+\infty} \sup_{\sharp\boldsymbol{\alpha}\le n} \sup_{x\in\mathcal{K}} \sup_{\ell_1,\ell_2\ge\ell} \mathbb{E}\Big[\sup_{t\le s\le S} |Y_s^{\ell_1,\boldsymbol{\alpha}} - Y_s^{\ell_2,\boldsymbol{\alpha}}|^2 + \sup_{t\le s\le S} |Z_s^{\ell_1,\boldsymbol{\alpha}} - Z_s^{\ell_2,\boldsymbol{\alpha}}|^2 \Big] = 0, \tag{70}$$

so that, for any $t \le S < T$, the sequence $((Y_s^{\ell,\boldsymbol{\alpha}}, Z_s^{\ell,\boldsymbol{\alpha}})_{t\le s\le S})_{\ell\ge1}$ is Cauchy with respect to the norm $\mathbb{E}[\sup_{t\le s\le S} |\cdot_s|^2]^{1/2}$. In particular, it converges towards some $(Y_s^{\boldsymbol{\alpha}}, Z_s^{\boldsymbol{\alpha}})_{t\le s\le S}$ for the same norm.

Taking in particular $s = t$ in (70), we deduce that the sequences $(V_{[\alpha_1]} \dots V_{[\alpha_n]} u_\ell(t,x))_{\ell\ge1}$ and $((V_{[\alpha_1]} \dots V_{[\alpha_n]} V_i u_\ell(T-t,x))_{1\le i\le N})_{\ell\ge1}$ are convergent, uniformly with respect to x in an arbitrary compact subset of $\mathbb{R}^d$. This shows that $u(t,\cdot)$ belongs to $\mathcal{D}_V^{K-m-1/2,\infty}(\mathbb{R}^d)$ for any $t > 0$.

We use now (26) but in the mollified setting. (That is replacing u by u_ℓ and $(Y^{\boldsymbol{\alpha}}, Z^{\boldsymbol{\alpha}})$ by $(Y^{\ell,\boldsymbol{\alpha}}, Z^{\ell,\boldsymbol{\alpha}})$ therein.) We know that the sequence $(\nabla_x u_\ell(T-S, X_S^{t,x}))_{\ell\ge1}$ is bounded. We can denote by $\nabla_x u(T-S, X_S^{t,x})$ a possible weak limit in $L^2(\Omega)$. (We will show below that $\nabla_x u$ exists as a true function when h is continuously differentiable.) Multiplying the dynamics of $Y^{\ell,\boldsymbol{\alpha}}$ in (26) by a test random variable ξ_s that is square integrable and $\mathcal{F}_s$-measurable and then letting ℓ tend to $+\infty$, we deduce that, for any $s \in [t, S)$, (26) holds true almost-surely in the limit setting. To prove that, almost-surely, it holds true for any $s \in [t, S)$, we apply a continuity argument. By (70), we know that the limit processes $(Y_s^{\boldsymbol{\alpha}})_{t\le s<T}$ and $(Z_s^{\boldsymbol{\alpha}})_{t\le s<T}$ are almost-surely continuous. In particular, the left-hand sides in (26) are continuous. By the martingale representation theorem, the conditional expectations of the integrals involving $\boldsymbol{F}_{\boldsymbol{\alpha}}$ and $\boldsymbol{G}_{\boldsymbol{\alpha}}$ are continuous as well. This says that there exists a continuous modification of the conditional expectation $(\mathbb{E}[\nabla_x u(T-S, X_S^{t,x})\theta_s^*[\phi_{\boldsymbol{\alpha}}](S-s, X_s^{t,x})|\mathcal{F}_s])_{t\le s<S}$. Choosing this modification of the conditional expectation, we deduce that the formula holds true almost-surely for any $t \le s < S$.

3.5.6. *Continuously Differentiable Case.* If h is continuously differentiable, then ∇h exists as a continuous function. In this case we apply (41) with h_ℓ instead of h. Using standard stability ([26], [27]) results for BSDEs and taking the expectation in (41), we deduce the equicontinuity property for the family of functions $(\nabla_x u_\ell)_{\ell\ge1}$ over compact subsets of $[0,T]\times\mathbb{R}^d$. Letting ℓ tend to $+\infty$, we deduce that $\nabla_x u$ exists as a continuous function over the whole space. By the convergence of $(\nabla_x u_\ell)_{\ell\ge1}$ towards $\nabla_x u$ on compact subsets (up to a subsequence), this shows that $\nabla_x u(T-S, X_S^{t,x})$ in (26) is understood as the *true* gradient of u: in particular, we check that the conditional expectation $\mathbb{E}[\nabla_x u(T-S, X_S^{t,x})\theta_s^*[\phi_{\boldsymbol{\alpha}}](S-s, X_s^{t,x})|\mathcal{F}_s]$ also reads $\mathbb{E}[\nabla_x u(T-S, X_{S-s}^{y})\phi_{\boldsymbol{\alpha}}(S-s,y)]_{|y=X_s^{t,x}}$, which is a continuous process, as expected.

3.5.7. *Bounds in the Lipschitz Setting.* The bounds in Theorem 3.1 are obtained by passing to the limit along the bounds obtained in the mollified setting. When ∇h exists as a continuous function, it is immediate to pass to the limit in the right-hand side in (25). When, h is not continuously differentiable, it is possible to bound the limit quantity in the right-hand side in terms of the limit $\limsup_{\varepsilon\to0} |\varepsilon|^{-d} \int_{\{|y|\le\varepsilon\}} |\nabla h(X_t^x + y)| dy$, as specified in the statement. □

4. Measurable Boundary Condition

In this section we dispense with the Lipschitz condition and assume that the boundary condition h is of polynomial growth and possibly discontinuous. The driver f satisfies the same assumption as in

Section 3 together with the stronger growth condition: $|f(t,x,y,z)| \le \Lambda(1+|y|+|z|)$. Basically, this growth condition ensures that, for any $T>0$ and $p>1$, there exists a constant $C_p>0$ such that

$$|u(t,x)| \le C_p\big(1+\mathbb{E}\big[|h(X_t^x)|^p\big]^{1/p}\big). \tag{71}$$

Eq. (71) must be seen as the counterpart of (45). It follows from Briand et al. [2] as well.

As already stated (see Theorem 2.5) when $f=0$ and h is bounded and smooth, it is known that, for any $T>0$, $p>1$, $n\ge 1$, $(\alpha_1,\dots,\alpha_n)\in[\mathcal{A}_0(m)]^n$ and $(t,x)\in(0,T]\times\mathbb{R}^d$,

$$\big|V_{[\alpha_1]}\dots V_{[\alpha_n]}u(t,x)\big| \le C_n(p)t^{-\|\boldsymbol{\alpha}\|/2}\mathbb{E}\big[\big|h(X_t^x)\big|^p\big]^{1/p}. \tag{72}$$

for some constant $C_n(p)$, independent of h. The main result of this section is

Theorem 4.1. *Let $(V_i)_{0\le i\le N}$ be $N+1$ vector fields satisfying Definition 1.1. Then, for any $t>0$, $u(t,\cdot)$ belongs to $\cap_{p\ge1}\mathcal{D}_V^{K-m-1/2,p}(\mathbb{R}^d)$. Moreover, for any $T>0$, $p>1$, $n=1,2$ and $\alpha_1,\alpha_n\in\mathcal{A}_0(m)$, there exists a constant $C_n(p)$, depending on Λ_n, n, p, T and the vector fields $V_0,\dots,V_N$ only, such that for all $t\in[0,T)$ and almost every $x\in\mathbb{R}^d$,*

$$\big|V_{[\alpha_1]}V_{[\alpha_n]}u(t,x)\big| \le C_n(p)t^{-\|\boldsymbol{\alpha}\|/2}\big[1+\mathbb{E}\big[|h|^{np}(X_t^x)\big]^{1/p}\big], \tag{73}$$

and for any $\delta>0$, $3\le n\le K-m-1$ and $\alpha_1,\dots,\alpha_n\in\mathcal{A}_0(m)$, there exists a constant $C_n(p,\delta)$, depending on δ, Λ_n, n, p, T and the vector fields $V_0,\dots,V_N$ only, such that for all $t\in(0,T]$ and almost every $x\in\mathbb{R}^d$,

$$\begin{aligned}&\big|V_{[\alpha_1]}\dots V_{[\alpha_n]}u(t,x)\big|\\ &\quad\le C_n(p,\delta)t^{-\|\boldsymbol{\alpha}\|/2}\big[1+t^{-n/2+1+\min(1/\|\alpha_{(1)}\|,1/2+1/(2\|\alpha_{(2)}\|))-\delta}\big]\big[1+\mathbb{E}\big[|h|^{np}(X_t^x)\big]^{1/p}\big],\end{aligned} \tag{74}$$

with $1\le i\le N$, where $\alpha_{(1)}$ and $\alpha_{(2)}$ stand for multi-indices in the family $\alpha_1,\dots,\alpha_n$ such that $\|\alpha_{(1)}\|\le\|\alpha_{(2)}\|$ are the two smallest elements in the family $\|\alpha_1\|,\dots,\|\alpha_n\|$. In particular, when $n=3$ and $\|\alpha_{(1)}\|=1$, Eq. (73) holds as well.

Finally, given $0\le t<S<T$, for any bounded $\mathcal{F}_t$-measurable random variable ξ with an absolutely continuous distribution on $\mathbb{R}^d$ (see[6]), the derivative pair processes

$$\Big(\big(Y_s^{\boldsymbol{\alpha}}=(V_{[\alpha_1]}\dots V_{[\alpha_n]}u)(T-s,X_s^{t,\xi}),Z_s^{\boldsymbol{\alpha}}=((V_{[\alpha_1]}\dots V_{[\alpha_n]}V_iu)(T-s,X_s^{t,\xi}))_{1\le i\le N}\big)_{t\le s<T}\Big)_{\boldsymbol{\alpha}}$$

indexed by the n-tuples of multi-indices $\boldsymbol{\alpha}=(\alpha_1,\dots,\alpha_n)\in[\mathcal{A}_0(m)]^n$ are continuous and satisfy the generalized BSDE

$$\begin{aligned}Y_s^{\boldsymbol{\alpha}}&=(S-s)^{-\|\boldsymbol{\alpha}\|/2}\mathbb{E}\big[u(T-S,X_S^{t,\xi})\theta_s^*[\phi_{\boldsymbol{\alpha}}](S-s,X_s^{t,\xi})|\mathcal{F}_s\big]\\&\quad+\mathbb{E}\bigg[\int_s^S\boldsymbol{F}_{\boldsymbol{\alpha}}\big(s,r,x,Y_r^{t,\xi},Z_r^{t,\xi},(Y_r^{\boldsymbol{\beta}})_{\sharp(\boldsymbol{\beta})\le n},(Z_r^{\boldsymbol{\beta}})_{\sharp(\boldsymbol{\beta})\le n}\big)dr\big|\mathcal{F}_s\bigg],\\(Z_s^{\boldsymbol{\alpha}})_i&=(S-s)^{-(1+\|\boldsymbol{\alpha}\|)/2}\mathbb{E}\big[u(T-S,X_S^{t,\xi})\theta_s^*[\psi_\alpha^i](S-s,X_s^{t,\xi})|\mathcal{F}_s\big]\\&\quad+\mathbb{E}\bigg[\int_s^S(r-s)^{-1/2}\boldsymbol{G}_{\boldsymbol{\alpha}}^i\big(s,r,x,Y_r^{t,x},Z_r^{t,x},(Y_r^{\boldsymbol{\beta}})_{\sharp(\boldsymbol{\beta})\le n},(Z_r^{\boldsymbol{\beta}})_{\sharp(\boldsymbol{\beta})\le n}\big)dr\big|\mathcal{F}_s\bigg],\end{aligned} \tag{75}$$

[6]Here, the probability space Ω must be enlarged to define random variables that are independent of the Wiener process. A standard way consists in considering the tensorial product of $\mathbb{R}^d$ and of the canonical Wiener space. This construction preserves the shift operator as defined in Subsubsection 2.2.1.

with $t \le s < S$, the coefficients satisfying the same properties as in Theorem 3.1. (Here, ϕ_α and ψ^i_α, $1 \le i \le N$, are $\mathbb{R}$-valued.)

When h is continuous, $u(t,\cdot)$ belongs to $\mathcal{D}_V^{K-m-1/2,\infty}(\mathbb{R}^d)$ for any $t > 0$, and (73) *and* (74) *hold for any $x \in \mathbb{R}^d$. Moreover,* (75) *hold for $\xi = x$, i.e. ξ deterministic.*

We observe that $n = 3$ is the threshold after which the small time behaviour of the solution to the nonlinear equation is worse than in the linear case. In the following section we give an example of a simple degenerate semilinear PDE for which the small time asymptotic behaviour is indeed worse than in the linear case beyond $n \ge 3$. In the uniformly elliptic setting, all the α's in $\mathcal{A}_0(m)$ have length 1, so that $-n/2 + 1 + \min(1/\|\alpha_{(1)}\|, 1/2 + 1/(2\|\alpha_{(2)}\|)) = -n/2 + 2 = -(n-4)/2$: the threshold is $n = 4$ or even $n = 5$ if the additional δ in the bound for the fourth-order derivatives is forgotten. In what follows we also give an example of a nondegenerate semilinear PDE for which the small time asymptotic behaviour is indeed worse than in the linear case beyond $n \ge 5$. In the uniformly elliptic setting, it is not clear whether the additional δ when $n = 4$ is sharp or not.

From a technical point of view, the threshold occurs because of the product

$$G_n(s,t;k,\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}) = \prod_{j=1}^{k} \mathbb{E}\Big[\big| \big(V_{[\beta_{1,j}]} \dots V_{[\beta_{i_j,j}]} v_j\big)(s, X^x_{t-s}) \big|^{np/i_j} \Big]^{i_j/(np)}, \tag{76}$$

that appears in Proposition 3.5 (and which will be used in this case). For $k = n$ (i.e. when all the β's in (76) are of length 1), this product is of order s^{-n} whereas it was of order $s^{-n/2}$ under the assumption of Theorem 3.1. Clearly, this is much more than the gap between the rates in the L^∞ and $W^{1,\infty}$ cases for a linear equation: in the linear setting, the gap is constant, equal to $1/2$.

Nevertheless, the gap in the product is not felt for low values of n since the nonlinear term f is integrated over the interval $[0,t]$: for n small, this additional integration permits to balance the gap between the L^∞ and $W^{1,\infty}$ cases. Obviously, the effect of the integration is limited: beyond some rank, the gap in the term $G_n(s,t;k,\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta})$ affects the small time asymptotic behaviour of the derivatives.

4.1. **Keystone in the Smooth Setting.** Again, we investigate first the case of a smooth boundary condition: below we will assume that h is bounded, infinitely differentiable with bounded derivatives of any order. The precise mollifying procedure is discussed in Subsection 4.5 following the model of Subsection 3.5. The keystone for the estimate is the following analogue of Proposition 3.5:

Proposition 4.2. *For any $p > 1$ and $T > 0$, there exists a constant $C_n(p)$, depending on Λ_n, p and T only, such that, for any $(\alpha_1,\dots,\alpha_n) \in (\mathcal{A}_0(m))^n$ and any $(t,x) \in (0,T] \times \mathbb{R}^d$,*

$$\begin{aligned} \big|V_{[\alpha_1]} \dots V_{[\alpha_n]} u(t,x)\big| &\le C_n(p) \Big[1 + t^{-\|\boldsymbol{\alpha}\|/2} \mathbb{E}\big[\big| h(X^x_t) \big|^p \big]^{1/p} \Big] \\ &+ \int_{t/2}^{t} \sum_{k=1}^{n} \sum_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}} (t-s)^{(\|\boldsymbol{\beta}\| - \|\boldsymbol{\alpha}\|)^+/2} \prod_{j=1}^{k} \mathbb{E}\Big[\big| \big(V_{[\beta_{1,j}]} \dots V_{[\beta_{i_j,j}]} v_j\big)(s, X^x_{t-s}) \big|^{np/i_j} \Big]^{i_j/(np)} ds, \end{aligned}$$

and

$$\big|V_{[\alpha_1]}\dots V_{[\alpha_n]}V_iu(t,x)\big| \le C_n(p)\bigg[1+t^{-(1+\|\boldsymbol{\alpha}\|)/2}\mathbb{E}\big[\big|h(X_t^x)\big|^p\big]^{1/p}\bigg] + \int_{t/2}^t \sum_{k=1}^n \sum_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}} (t-s)^{[(\|\boldsymbol{\beta}\|-\|\boldsymbol{\alpha}\|)^+-1]/2} \prod_{j=1}^k \mathbb{E}\bigg[\big|\big(V_{[\beta_{1,j}]}\dots V_{[\beta_{i_j,j}]}v_j\big)(s,X_{t-s}^x)\big|^{np/i_j}\bigg]^{i_j/(np)} ds\bigg].$$

Above, both sums run over the indices $\boldsymbol{i}=(i_1,\dots,i_k)\in\mathcal{I}_k(n)$, $\boldsymbol{v}=(v_1,\dots,v_k)\in\mathcal{U}_k(u(s,\cdot))$ *and* $\boldsymbol{\beta}=((\beta_{1,\ell},\dots,\beta_{i_\ell,\ell})_{1\le\ell\le k})\in\prod_{\ell=1}^k[\mathcal{A}_0(m)]^{i_\ell}$.

The proof is identical to that of Proposition 3.5 up to the additional estimate (71) in place of (45). Clearly, the price to pay in comparison with Proposition 3.5 is the additional exponent $-1/2$ in the boundary terms of both upper bounds. As announced above, this correction doesn't propagate linearly to the estimates of the higher order derivatives: because of the non-linearity, a break occurs beyond which the small time asymptotic behaviour of the derivatives is higher than in the analogue linear case.

4.2. **Proof of the Estimates for the first and second order derivatives in the Smooth Setting.** We start by proving the announced estimates when $n=1,2$.

For $n=1$, the proof is similar to that of Theorem 3.1. The only difference comes from the linear bounds of the first and second order derivatives (put it differently, it comes from the boundary terms in Proposition 4.2). At this stage of the proof, the nonlinearity doesn't affect the small time asymptotic behaviour: the product in (76) always reduces to a single term since k matches 1, that is everything works as in a linear setting with a non-zero source term.

Actually, one can deduce better estimate than the announced bound for $n=1$. As in the proof of Theorem 3.1, we also obtain a bound for $|V_{[\alpha_1]}V_iu(t,x)|$, $i\in\{1,\dots,N\}$. Clearly, we get the same bound as for $|V_{[\alpha_1]}u(t,x)|$, but the exponent of the explosion rate is augmented by 1/2, i.e.

$$\big|V_{[\alpha_1]}V_iu(t,x)\big| \le C_1(p)t^{-1/2-\|\alpha_1\|/2}\big(1+\mathbb{E}\big[|h(X_t^x)|^p\big]^{1/p}\big), \tag{77}$$

for some constant $C_1(p)$, depending on Λ_1, p, T and the vector fields only. (Eq. (77) is a little bit better than the announced estimate since the exponent in the power of $|h|$ is p and not $2p$ as it would be by applying (73) directly.)

For $n=2$, the method consists in examining the factors in (76) carefully. Since $k\le 2$ therein, we notice that the factors in the product (76) are of three possible forms: $\mathbb{E}[|(V_{[\beta_1]}v_1)(s,X_{t-s}^x)|^{2p}]^{1/(2p)}$, $\mathbb{E}[|(V_{[\beta_1]}v_1)(s,X_{t-s}^x)|^{2p}]^{1/(2p)}\times\mathbb{E}[|(V_{[\beta_2]}v_2)(s,X_{t-s}^x)|^{2p}]^{1/(2p)}$ and $\mathbb{E}[|(V_{[\beta_1]}V_{[\beta_2]}v_1)(s,X_{s-t}^x)|^p]^{1/p}$, with $\beta_1,\beta_2\in\mathcal{A}_0(m)$ and $v_1,v_2\in\{u,V_iu\}$, $1\le i\le N$. Using the bounds for $n=1$, we can follow the proof of Theorem 3.1 (see (29)) and then deduce that there exists a constant $C_2(p)$, depending on

Λ_2, p, and T only (the value of $C_2(p)$ possibly varying from line to line), such that

$$
\begin{aligned}
t^{(\|\alpha_1\|+\|\alpha_2\|)/2}\big|V_{[\alpha_1]}V_{[\alpha_2]}u(t,x)\big| &\le C_2(p)\big(1+\mathbb{E}\big[|h|^p(X_t^x)\big]^{1/p}\big)\\
&+ C_2(p)\int_{t/2}^t\Big[s^{-1/2}\big(1+\mathbb{E}\big[|h|^{2p}(X_t^x)\big]^{1/(2p)}\big)+s^{-1}\big(1+\mathbb{E}\big[|h|^{2p}(X_t^x)\big]^{1/p}\big)\Big]ds\\
&+ C_2(p)\sum_{\beta_1,\beta_2}\Bigg[\int_{t/2}^t s^{(\|\beta_1\|+\|\beta_2\|)/2}\mathbb{E}\big[\big|\big(V_{[\beta_1]}V_{[\beta_2]}u\big)(s,X_{t-s}^x)\big|^p\big]^{1/p}ds\\
&+\sum_{j=1\ldots N}\int_{t/2}^t(t-s)^{-1/2}s^{(1+\|\beta_1\|+\|\beta_2\|)/2}\mathbb{E}\big[\big|\big(V_{[\beta_1]}V_{[\beta_2]}V_ju\big)(s,X_{t-s}^x)\big|^p\big]^{1/p}ds\Bigg],
\end{aligned}
\tag{78}
$$

the sum running over $\beta_1,\beta_2\in\mathcal{A}_0(m)$.

Similarly, for any $1\le i\le N$,

$$
\begin{aligned}
t^{(1+\|\alpha_1\|+\|\alpha_2\|)/2}\big|V_{[\alpha_1]}V_{[\alpha_2]}V_iu(t,x)\big| &\le C_2(p)\big(1+\mathbb{E}\big[|h|^{2p}(X_t^x)\big]^{1/p}\big)\\
&+ C_2(p)\sum_{\beta_1,\beta_2}\Bigg[\int_{t/2}^t(t-s)^{-1/2}s^{(\|\beta_1\|+\|\beta_2\|)/2}\mathbb{E}\big[\big|\big(V_{[\beta_1]}V_{[\beta_2]}u\big)(s,X_{t-s}^x)\big|^p\big]^{1/p}ds\\
&+\sum_{j=1\ldots N}\int_{t/2}^t(t-s)^{-1/2}s^{(1+\|\beta_1\|+\|\beta_2\|)/2}\mathbb{E}\big[\big|\big(V_{[\beta_1]}V_{[\beta_2]}V_ju\big)(s,X_{t-s}^x)\big|^p\big]^{1/p}ds\Bigg].
\end{aligned}
\tag{79}
$$

By (78) and (79),

$$
\begin{aligned}
&t^{(\|\alpha_1\|+\|\alpha_2\|)/2}\big|V_{[\alpha_1]}V_{[\alpha_2]}u(t,x)\big|+t^{(1+\|\alpha_1\|+\|\alpha_2\|)/2}\big|V_{[\alpha_1]}V_{[\alpha_2]}V_iu(t,x)\big|\\
&\le C_2(p)\big(1+\mathbb{E}\big[|h|^{2p}(X_t^x)\big]^{1/p}\big)\\
&+ C_2(p)\sum_{\beta_1,\beta_2}\Bigg[\int_{t/2}^t(t-s)^{-1/2}s^{(\|\beta_1\|+\|\beta_2\|)/2}\mathbb{E}\big[\big|\big(V_{[\beta_1]}V_{[\beta_2]}u\big)(s,X_{t-s}^x)\big|^p\big]^{1/p}ds\\
&+\sum_{j=1\ldots N}\int_{t/2}^t(t-s)^{-1/2}s^{(1+\|\beta_1\|+\|\beta_2\|)/2}\mathbb{E}\big[\big|\big(V_{[\beta_1]}V_{[\beta_2]}V_ju\big)(s,X_{t-s}^x)\big|^p\big]^{1/p}ds\Bigg].
\end{aligned}
$$

Summing over $\alpha_1,\alpha_2\in\mathcal{A}_0(m)$ and $i\in\{1,\ldots,N\}$, choosing x of the form X_{r-t}^x, $r\ge t>0$, as in (34), taking the L^p moment and applying Lemma 2.13, we complete the proof when $n=2$. □

4.3. **Crude Estimates for $n\ge 3$.** When n is larger than 3, we first prove the following crude estimates:

Proposition 4.3. *For any $T>0$, $p>1$ and $1\le n\le K-m-1$, there exists a constant $C_n(p)$, depending on Λ_n, n, p, T and the vector fields only, such that, for any $(\alpha_1,\ldots,\alpha_n)\in(\mathcal{A}_0(m))^n$ and $(t,x)\in(0,T]\times\mathbb{R}^d$,*

$$
\begin{aligned}
&\big|V_{[\alpha_1]}\ldots V_{[\alpha_n]}u(t,x)\big|\le C_n(p)t^{-\|\boldsymbol{\alpha}\|/2-(n-2)^+/2}\big[1+\mathbb{E}\big[\big|h(X_t^x)\big|^{np}\big]^{1/p}\big],\\
&\big|V_{[\alpha_1]}\ldots V_{[\alpha_n]}V_iu(t,x)\big|\le C_n(p)t^{-\|\boldsymbol{\alpha}\|/2-(n-2)^+/2-1/2}\big[1+\mathbb{E}\big[\big|h(X_t^x)\big|^{np}\big]^{1/p}\big],\quad 1\le i\le N.
\end{aligned}
$$

Proof. We proceed by induction. By Subsection 4.2, the estimates hold true when $n = 1, 2$. Assume next that they hold true up to $n-1$, where n is such that $2 \leq n \leq K - m - 1$. We then establish the announced bounds for rank n.

The strategy is the same as for the Lipschitz case. It relies on Proposition 4.2, applied at rank n. We thus consider $\alpha_1, \ldots, \alpha_n \in (\mathcal{A}_0(m))^n$. With the same notation as in (76), we are to analyze $G_n(s,t;k,\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta})$.

When all the $(i_j)_{1\leq j\leq k}$ in $G_n(s,t;k,\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta})$ are less than or equal to $n-1$, we make use of the induction property to bound $G_n(s,t;k,\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta})$. Following (27) and (28), we obtain

$$\begin{aligned}\big|G_n(s,t;k,\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta})\big| &\leq C_n(p)\prod_{j=1}^{k}\Big[s^{-\sum_{\ell=1}^{i_j}\|\beta_{\ell,j}\|/2-(i_j-2)^+/2-\mathbf{1}_{\{i_j\geq 1\}}/2}\big(1+\mathbb{E}\big[\big|h(X_t^x)\big|^{np}\big]^{i_j/(np)}\big)\Big]\\ &\leq C_n(p)s^{-\|\boldsymbol{\beta}\|/2-\sum_{j=1}^{k}[(i_j-2)^+ +\mathbf{1}_{\{i_j\geq 1\}}]/2}\big(1+\mathbb{E}\big[\big|h(X_t^x)\big|^{np}\big]^{1/p}\big).\end{aligned}$$

Since $\sum_{j=1}^{k}[(i_j-2)^+ + \mathbf{1}_{\{i_j\geq 1\}}] = \sum_{j=1}^{k} i_j + \sum_{j=1}^{k}(\mathbf{1}_{\{i_j=1\}} - 1) \leq n$, we deduce that

$$\big|G_n(s,t;k,\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta})\big| \leq Cs^{-\|\boldsymbol{\beta}\|/2-n/2}\big(1+\mathbb{E}\big[\big|h(X_t^x)\big|^{np}\big]^{1/p}\big), \tag{80}$$

when all the $(i_j)_{1\leq j\leq k}$ in $G_n(s,t;k,\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta})$ are less than or equal to $n-1$. Plugging (80) into Proposition 4.2 and following (29) and (33), we deduce that

$$\begin{aligned}\big|V_{[\alpha_1]}\ldots V_{[\alpha_n]}u(t,x)\big| &\leq C_n(p)\Big[\big(1+t^{-\|\boldsymbol{\alpha}\|/2-(n-2)/2}\big)\mathbb{E}\big[\big|h(X_t^x)\big|^{np}\big]^{1/p}\\ &\quad+\sum_{\beta_1,\ldots,\beta_n}\int_{t/2}^{t}(t-s)^{(\|\boldsymbol{\beta}\|-\|\boldsymbol{\alpha}\|)^+/2}s^{-1/2}\mathcal{Q}^n_{\beta_1,\ldots,\beta_n}(s,t,x)ds\Big],\end{aligned} \tag{81}$$

and,

$$\begin{aligned}\big|V_{[\alpha_1]}\ldots V_{[\alpha_n]}V_iu(t,x)\big| &\leq C_n(p)\Big[\big(1+t^{-\|\boldsymbol{\alpha}\|/2-(n-1)/2}\big)\mathbb{E}\big[\big|h(X_t^x)\big|^{p}\big]^{1/p}\\ &\quad+\sum_{\beta_1,\ldots,\beta_n}\int_{t/2}^{t}(t-s)^{-1/2+(\|\boldsymbol{\beta}\|-\|\boldsymbol{\alpha}\|)^+/2}s^{-1/2}\mathcal{Q}^n_{\beta_1,\ldots,\beta_n}(s,t,x)ds\Big],\end{aligned} \tag{82}$$

where

$$\begin{aligned}\mathcal{Q}^n_{\beta_1,\ldots,\beta_n}(s,t,x) &= \mathbb{E}\big[\big|\big(V_{[\beta_1]}\ldots V_{[\beta_n]}u\big)(s,X_{t-s}^x)\big|^p\big]^{1/p}\\ &\quad+s^{1/2}\sum_{i=1}^{N}\mathbb{E}\big[\big|\big(V_{[\beta_1]}\ldots V_{[\beta_n]}V_iu\big)(s,X_{t-s}^x)\big|^p\big]^{1/p}.\end{aligned}$$

Choosing (t,x) of the form (t,X_{r-t}^x) in (81) and (82) for some $r \geq t$, taking the L^p-norm and applying Minkowski's integral inequality, we deduce

$$\begin{aligned}\sum_{\alpha_1,\ldots,\alpha_n} t^{\|\boldsymbol{\alpha}\|/2+(n-2)/2}\mathcal{Q}^n_{\alpha_1,\ldots,\alpha_n}(t,r,x) &\leq C_n(p)\Big[1+\mathbb{E}\big[\big|h(X_r^x)\big|^{np}\big]^{1/p}\\ &\quad+\sum_{\beta_1,\ldots,\beta_n}\int_{t/2}^{t}(t-s)^{-1/2}s^{\|\boldsymbol{\beta}\|/2+(n-2)/2}\mathcal{Q}^n_{\beta_1,\ldots,\beta_n}(s,r,x)ds\Big],\end{aligned} \tag{83}$$

for any $0 < t \leq r$. By Lemma 2.13, the proof is easily completed. □

4.4. **Proof of Theorem 4.1 in the Smooth Setting.** The proof of Theorem 4.1 relies on a suitable version of Corollary 3.3. Recall that $\mathcal{I}_k(n)$ is the set of non-decreasing sequences of (possibly zero) integers $i_1, \ldots, i_k$ such that $i_1 + \cdots + i_k \leq n$. Also for any $k \in \{0, \ldots, n\}$, let $\mathcal{U}_k(\varphi)$ stands for the set of k-tuples of functions of the form $(v_1, \ldots, v_k)$, with v_i being equal either to φ or $V_\ell\varphi$, $1 \leq \ell \leq N$ (When $k = 0$, $\mathcal{U}_k(\varphi) = \emptyset$). We claim the following:

Corollary 4.4. *Let F be a $(K-m-3)$-times differentiable function from $\mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^N$ into $\mathbb{R}$ with bounded derivatives of any order up to $K-m-3$ and φ be in $\mathcal{D}_V^{n+1/2}(\mathbb{R}^d)$, $3 \leq n \leq K-m-1$. Then, for any n-tuple of indices $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in [\mathcal{A}_0(m)]^n$*

$$
\begin{aligned}
V_{[\alpha_n]} \ldots V_{[\alpha_1]} \mathbb{E}\big[F\big(\Theta(X_s^x)\big)\big] \quad &= \quad s^{-(\|\alpha_{(1)}\|+\|\alpha_{(2)}\|)/2} \Bigg(\mathbb{E}\big[F(\Theta(X_s^x))\phi_0(s,x)\big] \\
&+ \sum_{k=0}^{n-2} \sum_{i \in \mathcal{I}_k(n-2)} \sum_{\boldsymbol{v} \in \mathcal{U}_k(\varphi)} \sum_{\boldsymbol{\beta} = (\beta_{\ell,j})_{1 \leq \ell \leq i_j, 1 \leq j \leq k} \in \prod_{1 \leq j \leq k} [\mathcal{A}_0(m)]^{i_j}} \\
&\qquad \mathbb{E}\Bigg[\prod_{j=1}^{k} \big(V_{[\beta_{1,j}]} \ldots V_{[\beta_{i_j,j}]} v_j \big)(X_s^x) \phi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}(s,x) \psi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}\big(\Theta(X_s^x)\big) \Bigg] \Bigg),
\end{aligned}
\tag{84}
$$

where $\|\alpha_{(1)}\| \leq \|\alpha_{(2)}\|$ stand for the two smallest lengthes among the family $(\|\alpha_i\|)_{1 \leq i \leq n}$, where $\phi_0 \in \mathcal{K}_0^T(K-m-n)$ and $\phi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}} \in \mathcal{K}^T_{(\|\boldsymbol{\beta}\|-\|\boldsymbol{\alpha}\|+\|\alpha_{(1)}\|+\|\alpha_{(2)}\|)^+}(K-m-n)$, with $\|\boldsymbol{\alpha}\| = \sum_{i=1}^n \|\alpha_i\|$ and $\|\boldsymbol{\beta}\| = \sum_{j=1}^k \sum_{i=1}^{i_j} \|\beta_{i,j}\|$, and where $\psi_{\boldsymbol{i},\boldsymbol{v},\boldsymbol{\beta}}$ is bounded.

A similar version holds with $\|\alpha_{(1)}\|$ only. In this case, F is assumed to be $(K-m-2)$-time differentiable and k runs over $\{0, \cdots, n-1\}$.

Proof. The proof is quite straightforward. Assume that the smallest indices at which $\alpha_{(1)}$ and $\alpha_{(2)}$ appear in the sequence $\alpha_1, \ldots, \alpha_n$ are p_1 and p_2 (not necessarily in a respective way), with $p_1 < p_2$. Apply then Corollary 3.3 to $V_{[\alpha_{p_1-1}]} \ldots V_{[\alpha_1]}[F(\Theta(X_s))]$ and then take the expectation to get a representation of $V_{[\alpha_{p_1-1}]} \ldots V_{[\alpha_1]}\mathbb{E}[F(\Theta(X_s))]$. Apply an integration by parts to compute $V_{[\alpha_{p_1}]} V_{[\alpha_{p_1-1}]} \ldots V_{[\alpha_1]}\mathbb{E}[F(\Theta(X_s))]$ without differentiating the function of X involved in the representation of $V_{[\alpha_{p_1-1}]} \ldots V_{[\alpha_1]}\mathbb{E}[F(\Theta(X_s))]$. (See, for example, Corollary 3.12 in [21].) Next apply Corollary 3.3 again to write $V_{[\alpha_{p_2-1}]} \ldots V_{[\alpha_1]}\mathbb{E}[F(\Theta(X_s))]$ and, then, a new integration by parts again, and finally Corollary 3.3 again.

The first term in the right-hand side in (84) appears when $p_1 = 1$: in such a case, we first perform an integration by parts; the resulting Kusuoka-Stroock function is then differentiated $n-1$ times. □

We are now in position to complete the proof of Theorem 4.1 when the boundary condition is smooth. We go back to (49) and (50). Clearly, we can bound $T_1(t,x)$ therein by (compare with (52))

$$
|T_1(t,x)| \leq C_n(p) t^{-\|\boldsymbol{\alpha}\|/2} \big(1 + \mathbb{E}\big[|h(X_t)|^p\big]^{1/p}\big).
$$

To bound $T_2(t,x)$ in (50), we use an interpolation argument. For $\varepsilon \in [0,1]$, we have the trivial inequality

$$|T_2(t,x)| \le \int_{t/2}^t \big|V_{[\alpha_1]} \dots V_{[\alpha_n]} P_{t-s}\big[f\big(s,\cdot,u(s,\cdot),(Vu)^\top(s,\cdot)\big)\big](x)\big|^{1-\varepsilon} \\ \times \big|V_{[\alpha_1]} \dots V_{[\alpha_n]} P_{t-s}\big[f\big(s,\cdot,u(s,\cdot),(Vu)^\top(s,\cdot)\big)\big](x)\big|^{\varepsilon} ds.$$

To bound the first factor $|P_{s-t}[f(s,\cdot,u(s,\cdot),Vu(s,\cdot))](x)|^{1-\varepsilon}$ in the integral above, we follow (53) and (54). Using Proposition 4.3, we deduce that, for any $p>1$,

$$|T_2(t,x)| \le C_n(p) \sum_{k=0}^n \sum_{\boldsymbol{i}} \int_{t/2}^t \big\{ s^{-\|\boldsymbol{\alpha}\|/2 - \sum_{j=1}^k [(i_j-2)^+ + \mathbf{1}_{\{i_j \ge 1\}}]/2} \big[1 + \mathbb{E}\big[|h|^{np}(X_t^x)\big]^{1/p}\big]\big\}^{1-\varepsilon} \\ \times \big|V_{[\alpha_1]} \dots V_{[\alpha_n]} P_{t-s}\big[f\big(s,\cdot,u(s,\cdot),(Vu)^\top(s,\cdot)\big)\big](x)\big|^{\varepsilon} ds,$$

$\boldsymbol{i}$ running over the indices $(i_1,\dots,i_k)$ such that $\sum_{j=1}^k i_j \le n$. Following the proof of Proposition 4.3, $\sum_{j=1}^k [(i_j-2)^+ + \mathbf{1}_{\{i_j\ge 1\}}]/2 \le n/2$, so that

$$|T_2(t,x)| \le C_n(p) \int_{t/2}^t \big\{ s^{-\|\boldsymbol{\alpha}\|/2 - n/2} \big[1 + \mathbb{E}\big[|h|^{np}(X_t^x)\big]^{1/p}\big]\big\}^{1-\varepsilon} \\ \times \big|V_{[\alpha_1]} \dots V_{[\alpha_n]} P_{t-s}\big[f\big(s,\cdot,u(s,\cdot),(Vu)^\top(s,\cdot)\big)\big](x)\big|^{\varepsilon} ds.$$

To bound the second factor in the above integral, we apply Corollary 4.4 together with Proposition 4.3. Basically, it permits to reduce n into $n-1$ or $n-2$. We then obtain

$$|T_2(t,x)| \le C_n(p)\big[1 + \mathbb{E}\big[|h|^{np}(X_t^x)\big]^{1/p}\big] \int_{t/2}^t s^{-\|\boldsymbol{\alpha}\|/2 - n/2} \big\{ (t-s)^{-\|\alpha_{(1)}\|/2} s^{\|\alpha_{(1)}\|/2+1/2}\big\}^{\varepsilon_1} \\ \times \big\{ (t-s)^{-(\|\alpha_{(1)}\|+\|\alpha_{(2)}\|)/2} s^{(\|\alpha_{(1)}\|+\|\alpha_{(2)}\|)/2+1}\big\}^{\varepsilon_2} ds,$$

with $\varepsilon_1+\varepsilon_2=\varepsilon$, $0\le\varepsilon_1,\varepsilon_2\le 1$. (The first term in (84) is handled as in (57) and (60).) The critical values for $(\varepsilon_1,\varepsilon_2)$ to ensure integrability satisfy: $\varepsilon_1\|\alpha_1\|/2 + \varepsilon_2(\|\alpha_{(1)}\|+\|\alpha_{(2)}\|)/2 = 1$. Forgetting for a while the divergence of the integral of $(t-s)^{-1}$, we then understand that the critical bound for $|T_2(t,x)|$ is $C_n(p)[1+\mathbb{E}[|h|^{np}(X_t^x)]^{1/p}] t^{-\|\boldsymbol{\alpha}\|/2 - n/2 + 1 + \varepsilon_1/2 + \varepsilon_2}$. Therefore, the point is to maximize $\varepsilon_1/2+\varepsilon_2$ under the constraints $\varepsilon_1,\varepsilon_2\ge 0$, $\varepsilon_1+\varepsilon_2\le 1$ and $\varepsilon_1\|\alpha_1\|/2 + \varepsilon_2(\|\alpha_{(1)}\|+\|\alpha_{(2)}\|)/2 = 1$. It is plain to see that it is the same as maximizing $2/(\|\alpha_{(1)}\|+\|\alpha_{(2)}\|) + (\|\alpha_{(2)}\|-\|\alpha_{(1)}\|)/[2(\|\alpha_{(1)}\|+\|\alpha_{(2)}\|)]\varepsilon_1$ under the constraints $0\le\varepsilon_1\le \min(1, 2/\|\alpha_{(1)}\|, (\|\alpha_{(1)}\|+\|\alpha_{(2)}\|-2)/\|\alpha_{(2)}\|)$. The optimum is given by $\varepsilon_1 = \min(1, 2/\|\alpha_{(1)}\|, (\|\alpha_{(1)}\|+\|\alpha_{(2)}\|-2)/\|\alpha_{(2)}\|)$ since $\|\alpha_{(2)}\|\ge\|\alpha_{(1)}\|$. Therefore, the critical values are

$$\begin{cases} \varepsilon_1 = (\|\alpha_{(2)}\|-1)/\|\alpha_{(2)}\|, & \varepsilon_2 = 1/(\|\alpha_{(2)}\|), & \text{if } \|\alpha_{(1)}\|=1, \\ \varepsilon_1 = 2/\|\alpha_{(1)}\|, & \varepsilon_2 = 0, & \text{if } \|\alpha_{(1)}\|\ge 2. \end{cases}$$

(In short, the above result says that we try to saturate the integral with a first-order derivative. When the first order derivative doesn't saturate the integral, we saturate it with a second-order derivative. In this way, the integral is always saturated and there is no need to look at higher-order derivatives.) To

take into account the divergence of the intregal of $(t-s)^{-1}$, we must subtract some small $\delta > 0$ to ε_1. We finally obtain, for any $\delta > 0$,

$$|T_2(t,x)| \le C_n(p,\delta)\big[1+\mathbb{E}\big[|h|^{np}(X_t^x)\big]^{1/p}\big] t^{-\|\boldsymbol{\alpha}\|/2-n/2+1+\min(1/\|\alpha_{(1)}\|,1/2+1/(2\|\alpha_{(2)}\|))-\delta}. \quad \square$$

4.5. **Proof of Theorem 4.1 in the General Setting.** We follow here the same strategy as in Subsection 3.5.

4.5.1. *Mollification of the Boundary Condition.* If h is continuous, it can be mollified as in Subsection 3.5. If it is measurable only, the sequence of mollified coefficients $(h_\ell)_{\ell\ge1}$ converges towards h, in $L^p_{\rm loc}(\mathbb{R}^d)$ only, for any $p \ge 1$. In any case, the sequence of solutions $(u_\ell)_{\ell\ge1}$ is at most of linear growth on the whole $[0,T]\times\mathbb{R}^d$, uniformly in ℓ. (See (71).)

Following Subsection 7.3, for any $t > 0$, $u_\ell(t,\cdot) \to u(t,\cdot)$ as $\ell \to +\infty$ in any $L^p_{\rm loc}(\mathbb{R}^d)$, for any $p \ge 1$. If h is continuous, the convergence holds in supremum norm on compact sets, as in subsection 3.5. Following Subsubsection 3.5.2, (64) holds here as well.

4.5.2. *Convergence of the Sequence* $(Z^{\ell,t,\xi})_{\ell\ge1}$. Eq. (67) holds true, but we cannot pass to the limit on it since the convergence of the sequence $(u^\ell)_{\ell\ge1}$ holds in $\bigcap_{p\ge1} L^p_{\rm loc}(\mathbb{R}^d)$ only. To overcome this difficulty, we choose as initial condition for X at time t a random variable ξ, bounded and $\mathcal{F}_t$-measurable, with an absolutely continuous distribution μ over $\mathbb{R}^d$. (See Footnote[6].) There is no difficulty to replace (t,x) by (t,ξ) in (67). By Lemma 2.10, $\lim_{\ell\to+\infty}\sup_{\ell_1,\ell_2\ge\ell}\mathbb{E}[|u_{\ell_1}(S,X_S^{t,\xi}) - u_{\ell_2}(S,X_S^{t,\xi})|^2] = 0$, so that (68) and (69) holds with (t,x) replaced by (t,ξ). (And forgetting the sup with respect to x therein.) By the new version of (69), $\lim_{\ell\to+\infty}\sup_{\ell_1,\ell_2\ge\ell}\mathbb{E}[|Z_t^{\ell_1,t,\xi} - Z_t^{\ell_2,t,\xi}|^2] = 0$, for any $t\in[0,T)$, that is $\lim_{\ell\to+\infty}\sup_{\ell_1,\ell_2\ge\ell}\int_{\mathbb{R}^d}|Vu_{\ell_1}(t,x) - Vu_{\ell_2}(t,x)|^2 d\mu(x) = 0$.

By the a priori bounds we have on $(Vu_\ell(t,\cdot))_{\ell\ge1}$ (see Theorem 4.1), we deduce that, for any $t\in[0,T)$, $(Vu_\ell(t,\cdot))_{\ell\ge1}$ converges towards $Vu(t,\cdot)$ in any $L^p_{\rm loc}(\mathbb{R}^d)$, $p>1$.

4.5.3. *Completion of the proof.* The end of the proof is then similar to Subsection 3.5. (Using in particular the bounds for $(V_{[\alpha_1]}\dots V_{[\alpha_n]}V_i u(t,x))_{1\le i\le N}$ in Proposition 4.3 when $n = K-m-1$, since nothing is said about it in Theorem 4.1.)

When h is continuous, there is no need to introduce ξ, since the convergence of $(u_\ell)_{\ell\ge1}$ towards u is uniform on compact subsets. The whole argument is then similar to Subsection 3.5. Moreover, by standard stability properties on BSDEs, u is continuous on the whole $[0,T]\times\mathbb{R}^d$. $\square$

5. Counter-Examples

In this section we give two counter-examples:

(1) In the first example, the second order differential operator is the one-dimensional Laplace operator and the boundary condition is bounded but not Lipschitz (it is, in fact, discontinuous). Since the operator is uniformly elliptic, Theorem 4.1 says that the exponent of the explosion rate of the derivatives of order less than 3 is the same as in the linear case and that the exponent of the explosion rate of the derivatives of order 4 is almost the same as in the linear case, up to a small correction of the exponent. On the opposite, Theorem 4.1 suggests that the exponent of the explosion rate of the derivatives of order greater than 5 might be higher. For a specific choice of the boundary condition and of the nonlinear term, we show that the

exponent of the derivatives of order greater than 5 is indeed worse than the corresponding exponent in the linear setting. This confirms that, as suggested by Theorem 4.1, order 5 appears as a threshold above which the small time behaviour of the derivatives deteriorates because of the nonlinearity.

(2) In the second example, we investigate a nonlinear equation driven by a weak Hörmander operator of dimension 2, close to the hypoelliptic Kolmogorov operator. Basically, the operator is driven by two vector fields V_0 and V_1 satisfying UFG condition with $m = 3$ and weak Hörmander condition as well. Theorem 4.1 says that the bound for the derivatives of order less than 2 is the same as in the linear case but suggests that a threshold might exist at order 3. For a suitable boundary condition and a suitable nonlinear term, we show that bound for the derivatives of order 3 is indeed worse than in the linear case. In other words, the simultaneity of the nonlinearity and of the degeneracy here modifies the threshold above which the small time behaviour of the derivatives deteriorates.

In both cases, we show that the right exponent for the rate of the derivatives exactly fits the exponent suggested by Theorem 4.1, up to the additional correction δ therein. This may be seen as a justification of the title of the paper: "sharp estimates".

5.1. **Counter-Example in the Uniformly Elliptic Setting.** In the whole subsection, we assume that $d = N = 1$ and we choose a smooth function f from $\mathbb{R}$ to $[-1, 1]$. By Theorem 4.1, we know that the solution u to the nonlinear equation

$$\partial_t u(t,x) = \frac{1}{2}\partial^2_{x,x} u(t,x) + f\big(\partial_x u(t,x)\big), \quad t \in (0,1],\ x \in \mathbb{R}, \tag{85}$$

with $u(0,x) = \mathbf{1}_{\{x>0\}}$ as boundary condition satisfies $|\partial^n_{x,\dots,x} u(t,x)| \le C_n t^{-n/2}$, $t \in (0,1]$, $x \in \mathbb{R}$, n=1,2,3, where C_n is some nonnegative constant. Moreover, for any $\delta > 0$ and any $n \ge 4$, there exists a constant $C_n(\delta)$ such that $|\partial^n_{x,\dots,x} u(t,x)| \le C_n(\delta) t^{2-n-\delta}$, $t \in (0,1]$, $x \in \mathbb{R}$.

5.1.1. *Diffusive Scaling.* Having in mind to take advantage of the diffusive scaling, we then set, for any integer $p \in \mathbb{N}^*$, $u_p(t,x) = u(p^{-2}t, p^{-1}x)$, so that, for any $t \in (0,1]$, $x \in \mathbb{R}$,

$$\begin{aligned} &|\partial^n_{x,\dots,x} u_p(t,x)| \le C_n t^{-n/2},\ n = 1,2,3, \\ &|\partial^n_{x,\dots,x} u_p(t,x)| \le C_n(\delta) p^{2\delta+n-4} t^{2-n-\delta},\ \delta > 0,\ n \ge 4, \end{aligned} \tag{86}$$

and

$$\partial_t u_p(t,x) = \frac{1}{2}\partial^2_{x,x} u_p(t,x) + p^{-2} f\big(p\partial_x u_p(t,x)\big), \quad t \in (0,1],\ x \in \mathbb{R}. \tag{87}$$

In particular, the functions $(\partial_t u_p)_{p\ge1}$ are uniformly bounded in compact subsets of $(0,1] \times \mathbb{R}$, so that the functions $(u_p)_{p\ge1}$ are uniformly convergent on compact subsets of $(0,1] \times \mathbb{R}$ towards the solution of the linear equation

$$\partial_t u_0(t,x) = \frac{1}{2}\partial^2_{x,x} u_0(t,x), \quad t \in (0,1],\ x \in \mathbb{R},$$

with $u(0,x) = \mathbf{1}_{\{x>0\}}$ as boundary condition. That is, $u_0(t,x) = (2\pi t)^{-1/2} \int_{-x}^{+\infty} \exp[-y^2/(2t)] dy$.

We first identify the rate of convergence:

Lemma 5.1. *For any $(t,x) \in (0,1] \times \mathbb{R}$ and any $p \geq 1$,*

$$|u_p(t,x) - u_0(t,x)| \leq p^{-2}, \ |\partial_x u_p(t,x) - \partial_x u_0(t,x)| \leq C p^{-2},$$

for some universal constant $C \geq 0$.

Proof. It is clear that

$$u_p(t,x) = u_0(t,x) + p^{-2} \int_0^t \int_{\mathbb{R}} f\big(p \partial_x u_p(t-s,y)\big) g(s,x-y) ds dy,$$

where g is the standard Gaussian kernel, hence the first inequality. To get the second inequality, we differentiate the above formula to obtain

$$|\partial_x u_p(t,x) - \partial_x u_0(t,x)| \leq p^{-2} \int_0^t s^{-1} \int_{\mathbb{R}} |f|\big(p \partial_x u_p(t-s,y)\big) |x-y| g(s,x-y) ds dy. \qquad \square$$

The rate of convergence of the second-order derivative is slightly different:

Lemma 5.2. *There exists a constant $C \geq 0$, such that for any $(t,x) \in (0,1] \times \mathbb{R}$,*

$$|\partial^2_{x,x}(u_p - u_0)(t,x)| \leq C p^{-1} t^{-1/2}.$$

Proof. We write

$$\begin{aligned}(u_p - u_0)(t,x) = &\int_{\mathbb{R}} (u_p - u_0)(t/2, x-y) g(t/2,y) dy \\ &+ p^{-2} \int_0^{t/2} \int_{\mathbb{R}} f\big(p \partial_x u_p(t-s,x-y)\big) g(s,y) ds dy,\end{aligned}$$

so that, after differentiating once, making a change of variable and differentiating once again, we get

$$\begin{aligned}&\partial^2_{x,x}(u_p - u_0)(t,x) \\ &= -2t^{-1} \int_{\mathbb{R}} \partial_x (u_p - u_0)(t/2,y)(x-y) g(t/2, x-y) dy \\ &\quad - p^{-1} \int_0^{t/2} s^{-1} \int_{\mathbb{R}} f'\big(p \partial_x u_p(t-s,y)\big) \partial^2_{x,x} u_p(t-s,y)(x-y) g(s,x-y) ds dy.\end{aligned}$$

Therefore, by (86) and by Lemma 5.1, we can find a constant C, such that

$$|\partial^2_{x,x}(u_p - u_0)(t,x)| \leq C t^{-1/2} p^{-2} + C p^{-1} t^{-1/2}. \qquad \square$$

5.1.2. *Sharpness of the Bounds of the Derivatives.* We are now ready to complete the analysis of the first counter-example. By differentiating the PDE (85) n times and by applying the chain rule formula (or the so-called Faà di Bruno's formula),

$$\begin{aligned}\partial_t \partial_{x,\dots,x} u_p^n(t,x) = &\frac{1}{2} \partial_{x,\dots,x} u_p^{n+2}(t,x) \\ &+ p^{-2} \sum \beta_{n,m_1,\dots,m_n} p^{m_1+\cdots+m_n} f^{(m_1+\cdots+m_n)}\big(p \partial_x u_p(t,x)\big) \prod_{j=1}^n \big(\partial^{j+1}_{x,\dots,x} u_p(t,x)\big)^{m_j},\end{aligned}$$

for some weights $(\beta_{n,m_1,\dots,m_n})_{n,m_1,\dots,m_n}$, the sum running over n-tuples $(m_j)_{1 \leq j \leq n}$ such that $m_1 + 2m_2 + \cdots + n m_n = n$.

By Itô's formula, we deduce for a given stopping time τ less than some prescribed real $\theta < 1/2$,

$$\begin{aligned}\partial_{x,\dots,x} u_p^n(1,-1) \\ = \mathbb{E}\big[\partial_{x,\dots,x} u_p^n(1-\tau,-1+B_\tau)\big] + \sum \beta_{n,m_1,\dots,m_n} p^{m_1+\cdots+m_n-2} T^{(p)}_{n,m_1,\dots,m_n},\end{aligned} \tag{88}$$

where

$$T^{(p)}_{n,m_1,\dots,m_n} = \mathbb{E}\int_0^\tau f^{(m_1+\cdots+m_n)}\big(p\partial_x u_p(1-s,-1+B_s)\big)\prod_{j=1}^n \big(\partial^{j+1}_{x,\dots,x} u_p(1-s,-1+B_s)\big)^{m_j} ds$$

and $(B_t)_{t\geq 0}$ stands for a one-dimensional Brownian motion.

Below, we choose τ as the first exit time $\tau = \inf\{t \geq 0 : |B_t| \geq \theta p^{-1}\} \wedge (\theta^2 p^{-2})$, so that τ has the same law as $\theta^2 p^{-2}(\rho \wedge 1)$, where ρ stands for the first exit time of a Brownian motion from $(-1,1)$. We deduce that $\theta^2 p^{-2}\mathbb{P}\{\rho \geq 1\} \leq \mathbb{E}(\tau) \leq \theta^2 p^{-2}\mathbb{E}(\rho)$.

By (86), for every $\delta > 0$, we can find a constant C_δ such that

$$\begin{aligned} p^{m_1+\cdots+m_n-2}|T^{(p)}_{n,m_1,\dots,m_n}| &\leq C_\delta \theta^2 p^{\delta-4} p^{\sum_{j=1}^n m_j}\prod_{j=1}^n p^{(j-3)^+ m_j} \\ &\leq C_\delta \theta^2 p^{\delta-4} p^{\sum_{j=1}^n m_j} p^{\sum_{j=1}^n (j-3)m_j + \sum_{j=1}^2 (3-j)m_j} \\ &= C_\delta \theta^2 p^{n+\delta-4} p^{-2\sum_{j=3}^n m_j - m_2}. \end{aligned} \tag{89}$$

(Keep in mind that $\sum_{j=1}^n j m_j = n$.) Therefore, when $m_1 < n$ (i.e. $m_i \geq 1$ for some $i \in \{2,\dots,n\}$),

$$\limsup_{p\to+\infty} p^{4-n} p^{m_1+\cdots+m_n-2}|T^{(p)}_{n,m_1,\dots,m_n}| = 0. \tag{90}$$

Now, when $m_1 = n$,

$$p^{n-2}T^{(p)}_{n,n,0,\dots,0} = p^{n-2}\mathbb{E}\int_0^\tau f^{(n)}\big(p\partial_x u_p(1-s,-1+B_s)\big)\big(\partial^2_{x,x} u_p(1-s,-1+B_s)\big)^n ds.$$

By Lemmas 5.1 and 5.2 and by Taylor's formula, we can find a constant $C \geq 1$ such that

$$\begin{aligned} &p^{n-2}T^{(p)}_{n,n,0,\dots,0} \\ &= p^{n-2}\mathbb{E}\int_0^\tau f^{(n)}\big(p\partial_x u_0(1-s,-1+B_s)\big)\big(\partial^2_{x,x} u_0(1-s,-1+B_s)\big)^n ds + O_p(p^{n-3})\mathbb{E}(\tau) \\ &\geq p^{n-2}\mathbb{E}(\tau)\inf_{|x|\leq C\theta}\big[f^{(n)}\big(p\partial_x u_0(1,-1)+x\big)\big]\inf_{|x|\leq C\theta}\big[\partial^2_{x,x} u_0(1,-1)+x\big]^n + O_p(p^{n-3})\mathbb{E}(\tau) \\ &\geq C\theta^2 p^{n-4}\inf_{|x|\leq C\theta}\big[f^{(n)}\big(p\partial_x u_0(1,-1)+x\big)\big]\inf_{|x|\leq C\theta}\big[\partial^2_{x,x} u_0(1,-1)+x\big]^n + O_p(p^{n-5}), \end{aligned}$$

where $O_p(\cdot)$ stands for the Landau notation (as p tends to $+\infty$). We now compute

$$\partial_x u_0(t,x) = (2\pi t)^{-1/2}\exp[-x^2/(2t)], \quad \partial^2_{x,x} u_0(t,x) = -(2\pi)^{-1/2}t^{-3/2}x\exp[-x^2/(2t)],$$

so that $\partial_x u_0(1,-1) = c_1 > 0$, $\partial^2_{x,x} u_0(1,-1) = c_2 > 0$. Choose now $f(z) = \cos[(2\pi/c_1)z - n(\pi/2)]$. Then, $f^{(n)}(z) = (2\pi/c_1)^n\cos[(2\pi/c_1)z]$, so that

$$f^{(n)}\big(p\partial_x u_0(1,-1)+x\big) = (2\pi/c_1)^n\cos[(2\pi/c_1)x] \geq (2\pi/c_1)^n/2,$$

for $(2\pi/c_1)|x| \leq \pi/4$.

Therefore, for θ small enough, $p^{n-2}T^{(p)}_{n,n,0,\dots,0} \geq c_3 p^{n-4} + O_p(p^{n-5})$, with $c_3 > 0$. Finally,

$$\liminf_{p\to+\infty} \big[p^{4-n}(p^{n-2}T^{(p)}_{n,n,0,\dots,0})\big] > 0. \tag{91}$$

5.1.3. *Conclusion.* Assume now that, for some $\delta > 0$ and $n \geq 5$, the bound

$$|\partial^n_{x,\dots,x} u(t,x)| \leq C_n t^{-n+2+\delta}, \quad t \in (0,1],\ x \in \mathbb{R}, \tag{92}$$

holds. By scaling,

$$|\partial^n_{x,\dots,x} u_p(t,x)| \leq C_n p^{n-4-2\delta} t^{-n+2+\delta}, \quad t \in (0,1],\ x \in \mathbb{R}.$$

Plugging the above inequality in (88) and multiplying (88) by p^{4-n}, we understand from (90) that all the terms but $p^{4-n}(p^{n-2}T^{(p)}_{n,n,0,\dots,0})$ vanish as p tends to $+\infty$. By (91), there is a contradiction hence the bound (92) cannot hold. □

5.2. **Counter-Example in the Degenerate Setting.** Consider now the following family of PDEs:

$$\partial_t u_p(t,x,y) = \frac{1}{2}\partial^2_{x,x} u_p(t,x,y) + \varphi(x)\partial_y u_p(t,x,y) + f\big(\partial_x u_p(t,x,y)\big), \quad t > 0,\ (x,y) \in \mathbb{R}^2, \tag{93}$$

with $u_p(0,x,y) = -\mathrm{sign}(x)\mathrm{sign}(y) + \lambda \mathrm{sign}(x + 1/p)$ as boundary condition, the function $|f|$ being bounded by 1 and the parameter λ being real. Both f and λ will be chosen later on.

In Eq. (93) above, φ stands for the function

$$\varphi(x) = \int_0^x \exp[-\phi(u)]du,$$

where ϕ is a nonnegative smooth function with bounded derivatives of any order satisfying:

$$\phi(u) = u^2, \quad |u| \leq 1\,; \quad \phi(u) = |u|, \quad |u| \geq 2\,; \quad \phi(u) \leq \min(u^2, 2|u|), \quad u \in \mathbb{R}.$$

In particular φ is smooth and has bounded derivatives of any order. Moreover, $\varphi(0) = 0$ and $\varphi'(0) = 1$. Eq. (93) is degenerate but satisfies the weak Hörmander condition since $[\partial_x, \varphi(x)\partial_y] = \exp[-\phi(x)]\partial_y$, that is $\mathcal{A}_0(3) = \{\partial_x, \exp[-\phi(x)]\partial_y\}$ spans $\mathbb{R}^2$ at any point $(x,y) \in \mathbb{R}^2$. Similarly, $[\partial_x, \exp[-\phi(x)]\partial_y] = -\phi'(x)\exp[-\phi(x)]\partial_y$ so that $\mathcal{A}_0(4)$ may be expressed as $\mathcal{A}_0(4) = \{\partial_x, \exp[-\phi(x)]\partial_y, -\phi'(x)\exp[-\phi(x)]\partial_y\}$. Since ϕ' is smooth and bounded, we deduce that all the elements of $\mathcal{A}_0(4)$ can be expressed as a smooth and bounded combination of the elements of $\mathcal{A}_0(3)$. In other words, the UFG property is checked with $m = 3$ and $K = +\infty$ (see Definition 1.1).

Equation (93) may be seen as a nonlinear generalization of the so-called Kolmogorov hypoelliptic example: in the earlier paper [14], Kolmogorov noticed that the operator driving the nonlinear equation above admitted a smooth density of Gaussian type when $\varphi(x) = x$, despite the degeneracy of the diffusion matrix. (Below, the operator $(1/2)\partial^2_{x,x} + x\partial_y$ will be referred to as Kolmogorov operator.)

5.2.1. *Gaussian Fundamental Solution when $\varphi(x) = x$.* We notice that

$$|\varphi(x) - x| \leq \int_0^{|x|} \phi(u)du \leq \int_0^{|x|} u^2 du = |x|^3/3, \quad x \in \mathbb{R}, \tag{94}$$

that is $\varphi(x)$ is very close to x in the neighborhood of zero. In particular, the derivatives of the solution u to (93) are expected to be close to the derivatives of the solution to (93) but driven by the Kolmogorov operator. (Obviously, we cannot choose $\varphi(x) = x$, $x \in \mathbb{R}$, since it is not bounded.)

The Kolmogorov operator is of great interest since its fundamental solution is explicitely known. It is given by the Gaussian density associated with the covariance matrix of the two-dimensional Gaussian process $G_t = (B_t, \int_0^t B_s ds)_{t\geq 0}$, $(B_t)_{t\geq 0}$ here standing for a one-dimensional Brownian motion. The covariance matrix of G_t, at a given time $t > 0$, reads

$$K_t = \begin{pmatrix} t & t^2/2 \\ t^2/2 & t^3/3 \end{pmatrix}.$$

Therefore, the kernel of Eq. (93) when $\varphi(x) = x$, may be expressed as $\mathbb{P}\{G_t \in dx'dy' | G_0 = (x,y)\} = g(t, x'-x, y'-(y+tx))$ with

$$g(t,x,y) = \frac{3^{1/2}}{\pi t^2}\exp\Big(-\frac{|K_t^{-1/2}(x,y)^*|^2}{2}\Big) = \frac{3^{1/2}}{\pi t^2}\exp\Big(-2\frac{x^2}{t} - 6\frac{y^2}{t^3} + 6\frac{xy}{t^2}\Big). \tag{95}$$

That is, u_p has the form

$$\begin{aligned} u_p(t,x,y) = &\int_{\mathbb{R}^2} u_p(0,x',y')g(t,x-x',y+tx-y')dx'dy' \\ &+ \int_{\mathbb{R}^2}\int_0^t f\big(\partial_x u_p(t-s,x',y')\big)g(s,x-x',y+sx-y')dx'dy', \quad t>0,\ x,y\in\mathbb{R}^2, \end{aligned}$$

when $\varphi(x) = x$.

We observe that the covariance matrix has two scales: $1/2$ stands for the exponent of the fluctuations of the coordinate x and $3/2$ for the exponent of the fluctuations of the coordinate y; $1/2$ may also be understood as the half-length of the vector field $V_1(x) = 1$ and $3/2$ as the half-length of the vector field $[V_1, V_0]$, with $V_0 = x\partial_y$.

5.2.2. *Rescaling Argument.* Following the previous subsection, we consider a rescaled version of u_p according to the scaling exponents $(1/2, 3/2)$. We set:

$$\hat{u}_p(t,x,y) = u_p\big(p^{-2}t, p^{-1}x, p^{-3}y\big), \quad t>0,\ x,y\in\mathbb{R},$$

for any $p \geq 1$. By Theorem 4.1 (and by maximum principle to bound $\hat{u}_p$ itself), we have

Lemma 5.3. *There exists a constant C, independent of p, such that $|\hat{u}_p(t,x,y)| \leq C$ and*

$$\begin{aligned} &|\partial_x\hat{u}_p(t,x,y)| \leq Ct^{-1/2},\ |\partial_y\hat{u}_p(t,x,y)| \leq C\exp[\phi(x/p)]t^{-3/2},\ |\partial^2_{x,x}\hat{u}_p(t,x,y)| \leq Ct^{-1}, \\ &|\partial^2_{x,y}\hat{u}_p(t,x,y)| \leq C\exp[\phi(x/p)]t^{-2},\ |\partial^2_{y,y}\hat{u}_p(t,x,y)| \leq C\exp[2\phi(x/p)]t^{-3}, \\ &|\partial^3_{x,x,y}\hat{u}_p(t,x,y)| \leq C\exp[\phi(x/p)]t^{-5/2},\ |\partial^3_{x,y,y}\hat{u}_p(t,x,y)| \leq C\exp[2\phi(x/p)]t^{-7/2}, \end{aligned}$$

$x, y \in \mathbb{R}$ and $t \in (0,1]$. Moreover, for any $\delta > 0$ and any $n \geq 3$, there exists a constant $C_n(\delta)$, independent of p, such that

$$\begin{aligned} &|\partial^n_{y,\dots,y}\hat{u}_p(t,x,y)| \leq C_n(\delta)\exp[n\phi(x/p)]p^{n-8/3+2\delta}t^{-2n+4/3-\delta}, \\ &|\partial^{n+1}_{x,y,\dots,y}\hat{u}_p(t,x,y)| \leq C_n(\delta)\exp[n\phi(x/p)]p^{n-7/3+2\delta}t^{-2n+2/3-\delta}, \\ &|\partial^{n+2}_{x,x,y,\dots,y}\hat{u}_p(t,x,y)| \leq C_n(\delta)\exp[n\phi(x/p)]p^{n-2+2\delta}t^{-2n-\delta}, \end{aligned}$$

$x, y \in \mathbb{R}$ and $t \in (0,1]$. The last inequality above is also true when $n = 2$.

We now investigate the limit behaviour of $\hat{u}_p$, as p tends to $+\infty$. The equation for $\hat{u}_p$ has the form

$$\partial_t \hat{u}_p(t,x,y) = \frac{1}{2}\partial^2_{x,x}\hat{u}_p(t,x,y) + p\varphi(x/p)\partial_y \hat{u}_p(t,x,y) + p^{-2} f\big(p\partial_x \hat{u}_p(t,x,y)\big), \tag{96}$$

$t>0$, $x,y\in\mathbb{R}$, with $\hat{u}_p(0,x,y) = \mathrm{sign}(x)\mathrm{sign}(y) + \lambda \mathrm{sign}(x+1)$ as boundary condition. Below, we set $\hat{u}(0,x,y) = \mathrm{sign}(x)\mathrm{sign}(y) + \lambda\mathrm{sign}(x+1)$. (That is, we get rid of the index p in $\hat{u}_p(0,\cdot,\cdot)$ since it is independent of p.) Since $\varphi(0)=0$ and $\varphi'(0)=1$, the limit is expected to be $\hat{u}_0$, solution to the PDE

$$\partial_t \hat{u}_0(t,x,y) = \frac{1}{2}\partial^2_{x,x}\hat{u}_0(t,x,y) + x\partial_y \hat{u}_0(t,x,y), \quad t>0,\ x,y\in\mathbb{R}, \tag{97}$$

with $\hat{u}_0(0,\cdot,\cdot) = \hat{u}(0,\cdot,\cdot)$ as boundary condition. It is immediate to see that Eq. (97) is well-posed and that the solution $\hat{u}_0$ is given by

$$\hat{u}_0(t,x,y) = \int_{\mathbb{R}^2} \hat{u}(0,x',y') g(t,x-x',y+tx-y')dx'dy',$$

with g as in (95). As a corollary, we deduce

Lemma 5.4. *We can find a constant C such that*

$$|\hat{u}_p(t,x,y) - \hat{u}_0(t,x,y)| \le C(1+|x|^3)\exp(2|x|)p^{-2}t^{-1/2}, \quad t\in(0,1],\ x,y\in\mathbb{R}.$$

Proof. We write $\hat{u}_p$ as the solution of the PDE

$$\begin{aligned}\partial_t \hat{u}_p(t,x,y) &= \frac{1}{2}\partial^2_{x,x}\hat{u}_p(t,x,y) + x\partial_y\hat{u}_p(t,x,y)\\ &\quad + \big(p\varphi(x/p)-x\big)\partial_y\hat{u}_p(t,x,y) + p^{-2}f\big(p\partial_x\hat{u}_p(t,x,y)\big), \quad t\in(0,1],\ x,y\in\mathbb{R},\end{aligned}$$

so that

$$\begin{aligned}\hat{u}_p(t,x,y) &= \hat{u}_0(t,x,y) + R_p^{(1)}(t,x,y) + R_p^{(2)}(t,x,y),\\ R_p^{(1)}(t,x,y) &= \int_0^t\int_{\mathbb{R}^2}\big(p\varphi(x'/p)-x'\big)\partial_y\hat{u}_p(t-s,x',y')g(s,x-x',y+sx-y')dx'dy'ds,\\ R_p^{(2)}(t,x,y) &= p^{-2}\int_0^t\int_{\mathbb{R}^2} f\big(p\partial_x\hat{u}(t-s,x',y')\big)g(s,x-x',y+sx-y')dx'dy'ds.\end{aligned} \tag{98}$$

By boundedness of f, we can find a constant C, independent of p, such that $|R_p^{(2)}(t,x,y)|\le Cp^{-2}$, $t\in(0,1]$, $x,y\in\mathbb{R}$. (C may vary below.) We turn now to $R_p^{(1)}(t,x,y)$. By integration by parts,

$$\begin{aligned}R_p^{(1)}(t,x,y) &\le \int_0^t\Bigg\{\left|\int_{\mathbb{R}^2}\big(p\varphi(x'/p)-x'\big)\partial_y\hat{u}_p(t-s,x',y')g(s,x-x',y+sx-y')dx'dy'\right|^{1/2}\\ &\qquad\times\left|\int_{\mathbb{R}^2}\big(p\varphi(x'/p)-x'\big)\hat{u}_p(t-s,x',y')\partial_y g(s,x-x',y+sx-y')dx'dy'\right|^{1/2}\Bigg\}ds\\ &= \int_0^t\Bigg\{\left|\int_{\mathbb{R}^2}R_p^{(1,1)}(t-s,x',y')g(s,x-x',y+sx-y')dx'dy'\right|^{1/2}\\ &\qquad\times\left|\int_{\mathbb{R}^2}R_p^{(1,2)}(t-s,s,x',y')g(s,x-x',y+sx-y')dx'dy'\right|^{1/2}\Bigg\}ds.\end{aligned}$$

By (94) and Lemma 5.3, we deduce that $|R_p^{(1,1)}(t-s,x',y')| \le C(t-s)^{-3/2}p^{-2}|x'|^3\exp(2|x'|)$, $0 \le s < t \le 1$, $x', y' \in \mathbb{R}$, for some possibly new value of C. Similarly, by (95), $|R_p^{(1,2)}(t-s,x',y')| \le Cs^{-3/2}p^{-2}|x'|^3(s^{-1/2}|x'-x| + s^{-3/2}|y+sx-y'|)$, $0 \le s < t \le 1$, $x', y' \in \mathbb{R}$. Performing a change of variable in the integrals above, we obtain

$$|R_p^{(1)}(t,x,y)| \le C(1+|x|^3)\exp(2|x|)p^{-2}\int_0^t s^{-3/4}(t-s)^{-3/4}ds \le C(1+|x|^3)\exp(2|x|)p^{-2}t^{-1/2}.$$

This completes the proof. □

As a corollary, we deduce

Lemma 5.5. *We can find a constant C such that, for any $t \in (0,1]$, $x, y \in \mathbb{R}$,*

$$|\partial_x\hat{u}_p(t,x,y) - \partial_x\hat{u}_0(t,x,y)| \le C(1+|x|^3)\exp(2|x|)p^{-2}t^{-1},$$
$$|\partial_{x,y}\hat{u}_p(t,x,y) - \partial_{x,y}\hat{u}_0(t,x,y)| \le C(1+|x|^3)\exp(4|x|)p^{-1}t^{-5/2}.$$

Proof. We consider a variation of (98).

$$\begin{aligned}
&\hat{u}_p(t,x,y) = \hat{u}_0(t,x,y) + S_p^{(1)}(t,x,y) + S_p^{(2)}(t,x,y) + S_p^{(3)}(t,x,y),\\
&S_p^{(1)}(t,x,y) = \int_{\mathbb{R}^2}\big[\hat{u}_p(t/2,x',y') - \hat{u}_0(t/2,x',y')\big]g(t/2,x-x',y+(t/2)x-y')dx'dy',\\
&S_p^{(2)}(t,x,y) = \int_0^{t/2}\int_{\mathbb{R}^2}\big[p\varphi\big(\frac{x'}{p}\big) - x'\big]\partial_y\hat{u}_p(t-s,x',y')g(s,x-x',y+sx-y')dx'dy'ds,\\
&S_p^{(3)}(t,x,y) = p^{-2}\int_0^{t/2}\int_{\mathbb{R}^2} f\big(p\partial_x\hat{u}(t-s,x',y')\big)g(s,x-x',y+sx-y')dx'dy'ds.
\end{aligned} \tag{99}$$

Convergence of $\partial_x\hat{u}_p$. We start with $\partial_x S_p^{(1)}$. By Lemma 5.4,

$$\partial_x S_p^{(1)}(t,x,y) = \int_{\mathbb{R}^2}\big[\hat{u}_p(t/2,x',y') - \hat{u}_0(t/2,x',y')\big]\partial_x\big[g\big(t/2,x-x',y+(t/2)x-y'\big)\big]dx'dy',$$

so that

$$\begin{aligned}
\big|\partial_x S_p^{(1)}(t,x,y)\big| &\le Cp^{-2}t^{-1/2}\int_{\mathbb{R}^2}\big\{(1+|x'|^3)\exp(2|x'|)\big(t^{-1}|x-x'| + t^{-2}|y+\frac{t}{2}x-y'|\big)\\
&\qquad\qquad \times g(\frac{t}{2},x-x',y+\frac{t}{2}x-y')\big\}dx'dy'\\
&\le C(1+|x|^3)\exp(2|x|)p^{-2}t^{-1}.
\end{aligned} \tag{100}$$

By a similar argument and by Lemma 5.3,

$$\begin{aligned}
\big|\partial_x S_p^{(2)}(t,x,y)\big| &\le Cp^{-2}t^{-3/2}\int_0^{t/2}\int_{\mathbb{R}}|x'|^3\exp(2|x'|)\big(s^{-1}|x-x'| + s^{-2}|y+sx-y'|\big)\\
&\qquad\qquad \times g(s,x-x',y+sx-y')dx'dy'ds\\
&\le C(1+|x|^3)\exp(2|x|)p^{-2}t^{-1}.
\end{aligned} \tag{101}$$

By the same method, it is plain to check that $|\partial_x S_p^{(3)}(t,x,y)| \le Cp^{-2}$. Together with (100) and (101), we complete the proof of the convergence of $\partial_x\hat{u}_p$.

Convergence of $\partial^2_{x,y}\hat{u}_p$. We start with $\partial^2_{x,y}S^{(1)}_p$. Following (100),

$$
\begin{aligned}
&\big|\partial_{x,y}S^{(1)}_p(t,x,y)\big| \\
&\le Cp^{-2}t^{-1/2}\int_{\mathbb{R}^2}\big\{(1+|x'|^3)\exp(2|x'|)\big[t^{-1}\big(t^{-1}|x-x'|+t^{-2}|y+sx-y'|\big)^2+t^{-2}\big] \\
&\qquad\qquad\times g(s,x-x',y+sx-y')\big\}dx'dy' \\
&\le C(1+|x|^3)\exp(2|x|)p^{-2}t^{-5/2}.
\end{aligned}
\tag{102}
$$

To deal with $\partial_{x,y}S^{(2)}_p(t,x,y)$, we perform a change of variable:

$$
\begin{aligned}
\partial_{x,y}S^{(2)}_p(t,x,y)=\int_0^{t/2}\int_{\mathbb{R}^2}\big(p\varphi(x'/p)-x'\big)\partial^2_{y,y}\hat{u}_p(t-s,x',y+sx-y') \\
\times\big[\partial_x g(s,x-x',y')+s\partial_y g(s,x-x',y')\big]dx'dy'ds,
\end{aligned}
$$

so that, by (94) and Lemma 5.3,

$$
\begin{aligned}
&\big|\partial_{x,y}S^{(2)}_p(t,x,y)\big| \\
&\le Ct^{-3}p^{-2}\int_0^{t/2}\int_{\mathbb{R}^2}|x'|^3\exp(4|x'|)(s^{-1}|x-x'|+s^{-2}|y'|)g(s,x-x',y')dx'dy'ds \\
&\le C(1+|x|^3)\exp(4|x|)t^{-5/2}p^{-2}.
\end{aligned}
\tag{103}
$$

By a similar argument,

$$
\begin{aligned}
\partial_{x,y}S^{(3)}_p(t,x,y)=p^{-1}\int_{\mathbb{R}^2}\int_0^{t/2}\big\{f'\big(p\partial_x\hat{u}_p(t-s,x',y+sx-y')\big) \\
\times\partial^2_{x,y}\hat{u}_p(t-s,x',y+sx-y')\big[\partial_x g(s,x-x',y')+s\partial_y g(s,x-x',y)\big]\big\}dx'dy',
\end{aligned}
$$

so that, by Lemma 5.3,

$$
\begin{aligned}
&\big|\partial_{x,y}S^{(3)}_p(t,x,y)\big| \\
&\le Cp^{-1}t^{-2}\int_0^{t/2}\int_{\mathbb{R}^2}\exp(2|x'|)(s^{-1}|x-x'|+s^{-2}|y'|)g(s,x-x',y')dx'dy'ds \\
&\le C\exp(2|x|)p^{-1}t^{-3/2}.
\end{aligned}
\tag{104}
$$

By (102), (103) and (104), the proof is over. □

5.2.3. *Criticality of order 3 in Theorem 4.1.* We investigate $\partial^3_{y,y,y}\hat{u}_p$. Specifically, we assume that it satisfies the bound $|\partial^3_{y,y,y}\hat{u}_p(t,x,y)|\le C(\delta)p^{1/3-2\delta}t^{-9/2-1/6+\delta}$ for any $t\in(0,1]$, $|x|\le 1$, $y\in\mathbb{R}$ and some $\delta>0$. (Compare with Lemma 5.3.) We will establish below a contradiction showing that the order 3 in y is critical.

In what follows, we denote by $(X^{1,p}_t,X^{2,p}_t)_{t\ge0}$ the two-dimensional process associated with the operator $(1/2)\partial^2_{x,x}+p\varphi(x/p)\partial_y$. Differentiating three times equation (96) w.r.t. y, we apply Itô's formula to $(\partial^3_{y,y,y}\hat{u}_p(t-s,X^{1,p}_s,X^{2,p}_s))_{0\le s<t}$, $t>0$ being given. (With $X^{1,p}_0=x$ and $X^{2,p}_0=y$.)

For a stopping time τ less than θ, for θ small (in particular, $\theta < t/2 \le 1/2$), we have

$$\begin{aligned}\partial^3_{y,y,y}\hat{u}_p(t,x,y) &= \mathbb{E}\big[\partial^3_{y,y,y}\hat{u}_p(t-\tau, X^{1,p}_\tau, X^{2,p}_\tau)\big]\\ &\quad + p\mathbb{E}\int_0^\tau f^{(3)}\big(p\partial_x\hat{u}_p(t-s,X^{1,p}_s,X^{2,p}_s)\big)\big(\partial^2_{x,y}\hat{u}_p(t-s,X^{1,p}_s,X^{2,p}_s)\big)^3 ds\\ &\quad + 3\mathbb{E}\int_0^\tau f^{(2)}\big(p\partial_x\hat{u}_p(t-s,X^{1,p}_s,X^{2,p}_s)\big)\partial^2_{x,y}\hat{u}_p(t-s,X^{1,p}_s,X^{2,p}_s)\\ &\qquad\qquad \times \partial^3_{x,y,y}\hat{u}_p(t-s,X^{1,p}_s,X^{2,p}_s)ds\\ &\quad + p^{-1}\mathbb{E}\int_0^\tau f'\big(p\partial_x\hat{u}_p(t-s,X^{1,p}_s,X^{2,p}_s)\big)\partial^4_{x,y,y,y}\hat{u}_p(t-s,X^{1,p}_s,X^{2,p}_s)ds\\ &= T^{(1)}_p(t,x,y)+T^{(2)}_p(t,x,y)+T^{(3)}_p(t,x,y)+T^{(4)}_p(t,x,y).\end{aligned} \tag{105}$$

By Lemma 5.3, for any $\delta>0$, $p^{-2/3-\delta}\partial^4_{x,y,y,y}\hat{u}_p$ is bounded on every compact subset of $(0,1]\times\mathbb{R}^2$, uniformly in p. Similarly, $\partial^3_{x,y,y}\hat{u}_p$ is bounded on every compact subset of $(0,1]\times\mathbb{R}^2$, uniformly in p. When τ is the first exit time of a compact subset of $(0,1]\times[-1,1]\times\mathbb{R}$, $T^{(3)}_p(t,x,y)$ and $T^{(4)}_p(t,x,y)$ are bounded, uniformly in p.

By Lemma 5.5, the asymptotic behavior of $T^{(2)}_p(t,x,y)$ is given by

$$\begin{aligned}&T^{(2)}_p(t,x,y)\\ &\quad = p\mathbb{E}\int_0^\tau f^{(3)}\big(p\partial_x\hat{u}_0(t-s,X^{1,p}_s,X^{2,p}_s)\big)\big(\partial^2_{x,y}\hat{u}_0(t-s,X^{1,p}_s,X^{2,p}_s)\big)^3 ds + O_p(1),\end{aligned} \tag{106}$$

where $O_p(1)$ stands for the Landau symbol and denotes a bounded sequence in p. (Again, τ is the first exit time from a compact subset of $(0,1]\times[-1,1]\times\mathbb{R}$.)

Assume now that we can find $t\in(0,1]$ such that $\partial_y\hat{u}_0(t,0,0)=\partial^2_{x,x}\hat{u}_0(t,0,0)=\partial^3_{x,x,x}\hat{u}_0(t,0,0)=0$ (see Subsubsection 5.2.5). Choose then $X^{1,p}_0=X^{2,p}_0=0$ and τ as the first exit time $\tau=\inf\{t\ge 0 : |X^{1,p}_t|\ge\theta p^{-1/3},\ |X^{2,p}_t|\ge\theta^3p^{-1}\}\wedge\theta^2p^{-2/3}$. Differentiating PDE (97) w.r.t. x, we also have $\partial^2_{t,x}\hat{u}_0(t,0,0)=0$. Performing a Taylor expansion in (106), we obtain

$$\begin{aligned}&T^{(2)}_p(t,0,0)\\ &\quad = p\mathbb{E}\int_0^\tau f^{(3)}\big(p\partial_x\hat{u}_0(t,0,0)+\theta O_p(1)\big)\big(\partial^2_{x,y}\hat{u}_0(t,0,0)+\theta O_p(p^{-1/3})\big)^3 ds + O_p(1).\end{aligned} \tag{107}$$

In particular, there exists a constant $\gamma\ge 0$, such that, for any power $\delta>0$,

$$\begin{aligned}\liminf_{p\to+\infty} p^{-\delta}T^{(2)}_p(t,0,0) \ge \liminf_{p\to+\infty}\big\{&p^{1-\delta}\mathbb{E}[\tau]\inf_{|x|\le\gamma\theta}\big[f^{(3)}\big(p\partial_x\hat{u}_0(t,0,0)+x\big)\big]\\ &\times\inf_{|x|\le\gamma\theta}\big[\big(\partial^2_{x,y}\hat{u}_0(t,0,0)+x\big)^3\big]\big\}.\end{aligned} \tag{108}$$

Let us return to (105). We claim that the bound $|\partial^3_{y,y,y}\hat{u}_p(s,x,y)|\le Cp^\eta s^{-9/2-\eta/2}$, $s\in[t/2,t]$, $|x|\le 1$, $y\in\mathbb{R}$, cannot be true if the limit below is infinite:

$$\liminf_{p\to+\infty}\big\{p^{1-\eta}\mathbb{E}[\tau]\inf_{|x|\le\gamma\theta}\big[f^{(3)}\big(p\partial_x\hat{u}_0(t,0,0)+x\big)\big]\inf_{|x|\le\gamma\theta}\big[\big(\partial^2_{x,y}\hat{u}_0(t,0,0)+x\big)^3\big]\big\}=+\infty. \tag{109}$$

Indeed, by (108), (109) implies $\liminf_{p\to+\infty} p^{-\eta} T_p^{(2)}(t,x,y) = +\infty$. Multiplying (105) by $p^{-\eta}$, we then obtain a contradiction.

In particular, the bound $|\partial^3_{y,y,y} u_p(t,x,y)| \le C t^{-9/2-\eta/2}$, $t \in (0,1]$, $|x| \le 1$, $y \in \mathbb{R}$, cannot be true if (109) holds true. Indeed, if $|\partial^3_{y,y,y} u_p(t,x,y)| \le C t^{-9/2-\eta/2}$, then, for $t \in (0,1]$, $|x| \le 1$, $y \in \mathbb{R}$,

$$|\partial^3_{y,y,y} \hat{u}_p(t,x,y)| = p^{-9} \big|\partial^3_{y,y,y} u_p\big(p^{-2}t, p^{-1}x, p^{-3}y\big)\big| \le C p^{\eta} t^{-9/2-\eta/2}.$$

5.2.4. *Lower Bound for* $\mathbb{E}[\tau]$. It now remains to bound $\mathbb{E}(\tau)$ from below. Define $\tau' = \inf\{t \ge 0 : |X_t^{1,p}| \ge \theta p^{-1/3}\}$. Since

$$\big|X_t^{2,p}\big| = \left| p \int_0^t \varphi\big(X_s^{1,p}/p\big) ds \right| \le \int_0^t \big|X_s^{1,p}\big| ds, \quad t \ge 0,$$

we obtain that $|X_t^{2,p}| < \theta t p^{-1/3}$, $t < \tau'$. In particular, $|X_t^{2,p}| < \theta^3 p^{-1}$, $t < \tau'$ and $t \le \theta^2 p^{-2/3}$. Therefore, $\mathbb{E}[\tau] \ge \theta^2 \mathbb{P}\{\tau' \ge \theta^2 p^{-2/3}\} p^{-2/3}$. Since $\tau' \sim \theta^2 p^{-2/3} \rho$, where ρ is the first exit time of a Brownian motion from $(-1,1)$, we deduce that

$$\mathbb{E}[\tau] \ge \theta^2 \mathbb{P}\{\rho \ge 1\} p^{-2/3}. \tag{110}$$

Therefore, (109) holds for $\eta < 1/3$, provided

$$\liminf_{p\to+\infty} \Big\{ \inf_{|x|\le\gamma\theta} \big[f^{(3)}\big(p\partial_x \hat{u}_0(t,0,0) + x\big)\big] \inf_{|x|\le\gamma\theta} \big[\big(\partial^2_{x,y} \hat{u}_0(t,0,0) + x\big)^3\big] \Big\} > 0. \tag{111}$$

That is, the bound $|\partial^3_{y,y,y} u_p(t,x,y)| \le C t^{-9/2-\eta/2}$, $t > 0$, $x, y \in \mathbb{R}$, cannot be true for $\eta < 1/3$. This exactly fits the threshold in Theorem 4.1 and Lemma 5.3.

5.2.5. *Computation of the Derivatives.* It now remains to find $t \in (0,1]$ such that $\partial_y \hat{u}_0(t,0,0) = \partial^2_{x,x} \hat{u}_0(t,0,0) = \partial^3_{x,x,x} \hat{u}_0(t,0,0) = 0$ and to check (111).

We first notice that $\hat{u}_0$ can be split into terms $\hat{u}_0 = \hat{u}_0^{(1)} + \lambda \hat{u}_0^{(2)}$, $\hat{u}_0^{(1)}$ and $\hat{u}_0^{(2)}$ both satisfying Equation (97) but with different boundary conditions:

$$\hat{u}_0^{(1)}(0,x,y) = -\mathrm{sign}(x)\mathrm{sign}(y), \quad \hat{u}_0^{(2)}(0,x) = \mathrm{sign}(x+1).$$

We emphasize that

$$\hat{u}_0^{(1)}(t,x,y) = \int_{\mathbb{R}^2} \hat{u}_0^{(1)}(0,x',y') g(t, x-x', y+tx-y') dx' dy'.$$

Since $\hat{u}_0^{(1)}(0,-x',-y') = \hat{u}_0^{(1)}(0,x',y')$, it is immediate to see, by a change of variable, that

$$\hat{u}_0^{(1)}(t,-x,-y) = \hat{u}_0^{(1)}(t,x,y), \quad t > 0,\ x, y \in \mathbb{R}.$$

By differentiation, we deduce that $\partial_y \hat{u}_0^{(1)}(t,0,0) = \partial^3_{x,x,x} \hat{u}_0^{(1)}(t,0,0) = 0$.

We now compute

$$\partial_x \hat{u}_0^{(1)}(t,x,y) = -2 \int_{\mathbb{R}} \mathrm{sign}(y+tx-y') g(t,x,y') dy' - 2t \int_{\mathbb{R}} \mathrm{sign}(x-x') g(t,x',y+tx) dx'$$

$$\partial^2_{x,y} \hat{u}_0^{(1)}(t,x,y) = -4 g(t,x,y+tx) - 2t \int_{\mathbb{R}} \mathrm{sign}(x-x') \big(-12 \frac{y+tx}{t^3} + 6 \frac{x'}{t^2}\big) g(t,x',y+tx) dx'.$$

In particular,

$$\partial^2_{x,y}\hat{u}_0^{(1)}(t,0,0) = -4g(t,0,0) - 12t^{-1}\int_{\mathbb{R}} \operatorname{sign}(-x')x'g(t,x',0)dx' = c_1 t^{-2},$$

with $c_1 = 2\sqrt{3}/\pi > 0$.

We now investigate $\hat{u}_0^{(2)}(t,x)$. It is given by

$$\hat{u}_0^{(2)}(t,x) = (2\pi)^{-1/2}\int_{\mathbb{R}} \operatorname{sign}(x - t^{1/2}x' + 1)\exp\big(-\frac{(x')^2}{2}\big)dx'.$$

Therefore,

$$\begin{aligned}
&\partial_x\hat{u}_0^{(2)}(t,x) = 2(2\pi)^{-1/2}t^{-1/2}\exp\big(-\frac{(x+1)^2}{2t}\big),\\
&\partial^2_{x,x}\hat{u}_0^{(2)}(t,x) = -2(2\pi)^{-1/2}t^{-3/2}(x+1)\exp\big(-\frac{(x+1)^2}{2t}\big),\\
&\partial^3_{x,x,x}\hat{u}_0^{(2)}(t,x) = 2(2\pi)^{-1/2}\big(t^{-5/2}(x+1)^2 - t^{-3/2}\big)\exp\big(-\frac{(x+1)^2}{2t}\big).
\end{aligned}$$

In particular, $\partial^2_{x,x}\hat{u}_0^{(2)}(1,0) = -c_2 < 0$ and $\partial^3_{x,x,x}\hat{u}_0^{(2)}(1,0) = 0$. Finally,

$$\begin{aligned}
&\partial^2_{x,x}\hat{u}_0(1,0,0) = \partial^2_{x,x}\hat{u}_0^{(1)}(1,0,0) + \lambda\partial^2_{x,x}\hat{u}_0^{(2)}(1,0) = \partial^2_{x,x}\hat{u}_0^{(1)}(1,0,0) - \lambda c_2,\\
&\partial^3_{x,x,x}\hat{u}_0(1,0,0) = \partial^3_{x,x,x}\hat{u}_0^{(1)}(1,0,0) + \lambda\partial^3_{x,x,x}\hat{u}_0^{(2)}(1,0) = 0,\\
&\partial^2_{x,y}\hat{u}_0(1,0,0) = \partial^2_{x,y}\hat{u}_0^{(1)}(1,0,0) = c_1 > 0.
\end{aligned} \tag{112}$$

Choose now λ so that $\partial^2_{x,x}\hat{u}_0^{(1)}(1,0,0) - \lambda c_2 = 0$. (This is possible since $c_2 > 0$.) For this choice, the required conditions $\partial_y\hat{u}_0(1,0,0) = \partial^2_{x,x}\hat{u}_0(1,0,0) = \partial^3_{x,x,x}\hat{u}_0(1,0,0) = 0$ are satisfied.

5.2.6. *Conclusion.* We now choose f:

$$\begin{aligned}
&f(z) = -\sin\big(2\pi z/|\partial_x\hat{u}_0(1,0,0)|\big), \quad z\in\mathbb{R}, \quad \text{if } \partial_x\hat{u}_0(1,0,0)\neq 0,\\
&f(z) = -\sin\big(z\big), \quad z\in\mathbb{R}, \quad \text{if } \partial_x\hat{u}_0(1,0,0) = 0.
\end{aligned} \tag{113}$$

In particular, there are two cases in (111). If $\partial_x\hat{u}_0(1,0,0)\neq 0$,

$$\begin{aligned}
\inf_{|x|\le\gamma\theta}\big[f^{(3)}\big(p\partial_x\hat{u}_0(1,0,0)+x\big)\big] &\ge \big(2\pi/|\partial_x\hat{u}_0(1,0,0)|\big)^3\inf_{|x|\le\gamma\theta}\big[\cos\big(\pm 2\pi p + 2\pi x/|\partial_x\hat{u}_0(1,0,0)|\big)\big]\\
&= \big(2\pi/|\partial_x\hat{u}_0(1,0,0)|\big)^3\inf_{|x|\le\gamma\theta}\big[\cos\big(2\pi x/|\partial_x\hat{u}_0(1,0,0)|\big)\big].
\end{aligned}$$

Choosing $\gamma\theta < |\partial_x\hat{u}_0(1,0,0)|/8$, we then obtain

$$\inf_{|x|\le\gamma\theta}\big[f^{(3)}\big(p\partial_x\hat{u}_0(1,0,0)+x\big)\big] \ge 2^{-1/2}\big(2\pi/|\partial_x\hat{u}_0(1,0,0)|\big)^3. \tag{114}$$

If $\partial_x\hat{u}_0(1,0,0) = 0$,

$$\inf_{|x|\le\gamma\theta}\big[f^{(3)}\big(p\partial_x\hat{u}_0(1,0,0)+x\big)\big] = \inf_{|x|\le\gamma\theta}\big[\cos(x)\big].$$

Choosing $\gamma\theta < \pi/4$, we then obtain

$$\inf_{|x|\le\gamma\theta}\big[f^{(3)}\big(p\partial_x\hat{u}_0(1,0,0)+x\big)\big] \ge 2^{-1/2}. \tag{115}$$

Let us examine now the second term in (109). For $\gamma\theta < c_1/2$,

$$\inf_{|x|\le\gamma\theta}\big[\big(\partial^2_{x,y}\hat{u}_0(t,0,0)+x\big)^3\big] \ge (c_1/2)^3. \tag{116}$$

From (110), (114), (115) and (116), we deduce that (109) holds true with $\eta < 1/3$. This shows criticallity at order 3.

5.2.7. *Generalization at any Order $n \ge 3$.* Following Subsubsection 5.1.2, we can generalize the result to any order $n \ge 3$. The point is to differentiate (96) n times w.r.t. y and to apply Itô's formula as in (105). We then obtain

$$\begin{aligned}
&\partial^n_{y,\dots,y}\hat{u}_p(t,x,y)\\
&= \mathbb{E}\big[\partial^n_{y,\dots,y}\hat{u}_p(t-\tau,X^{1,p}_\tau,X^{2,p}_\tau)\big]\\
&\quad + p^{-2}\sum \beta_{n,m_1,\dots,m_n}p^{m_1+\cdots+m_n}\mathbb{E}\int_0^\tau\Big[f^{(m_1+\cdots+m_n)}\big(p\partial_x\hat{u}_p(t-s,X^{1,p}_s,X^{2,p}_s)\big)\\
&\qquad\qquad \times\prod_{j=1}^n\big(\partial^{j+1}_{x,y,\dots,y}\hat{u}_p(t-s,X^{1,p}_s,X^{2,p}_s)\big)^{m_j}\Big]ds\\
&= \mathbb{E}\big[\partial^n_{y,\dots,y}\hat{u}_p(t-\tau,X^{1,p}_\tau,X^{2,p}_\tau)\big] + \sum\beta_{n,m_1,\dots,m_n}p^{m_1+\cdots+m_n-2}T^{(p)}_{n,m_1,\dots,m_n}.
\end{aligned} \tag{117}$$

(The sum running over $m_1,\dots,m_n$ such that $\sum_{j=1}^n jm_j = n$.) Following (89) and applying Lemma 5.3, for any $\delta > 0$, we can find a constant $C_\delta > 0$ such that

$$\begin{aligned}
p^{m_1+\cdots+m_n-2}|T^{(p)}_{n,m_1,\dots,m_n}| &\le C_\delta\mathbb{E}(\tau)p^{\delta-2}p^{\sum_{j=1}^n m_j}\prod_{j=1}^n p^{(j-7/3)^+m_j}\\
&\le C_\delta\mathbb{E}(\tau)p^{\delta-2}p^{\sum_{j=1}^n m_j}p^{\sum_{j=1}^n(j-7/3)m_j+m_2/3+4m_1/3}\\
&= C_\delta\mathbb{E}(\tau)p^{n+\delta-2}p^{-(4/3)\sum_{j=3}^n m_j-m_2}.
\end{aligned}$$

Keeping in mind that $\tau \le p^{-2/3}$, we deduce that $\lim_{p\to+\infty}p^{-n+8/3}p^{m_1+\cdots+m_n-2}|T^{(p)}_{n,m_1,\dots,m_n}| = 0$ when $m_1 < n$.

When $m_1 = n$, we can follow (107), (108) and (110). We deduce $\liminf_{p\to+\infty}p^{-n+8/3}T^{(p)}_{n,1,0,\dots,0} > 0$, provided

$$\liminf_{p\to+\infty}\inf_{|x|\le\gamma\theta}\big[f^{(n)}\big(p\partial_x\hat{u}_0(t,0,0)+x\big)\big]\inf_{|x|\le\gamma\theta}\big[\big(\partial_{x,x}\hat{u}_0(t,0,0)+x\big)^n\big] > 0. \tag{118}$$

Following (113), (118) holds true for

$$\begin{aligned}
&f(z) = \cos\big(2\pi z/|\partial_x\hat{u}_0(1,0,0)| - n(\pi/2)\big), \quad z\in\mathbb{R}, \quad \text{if } \partial_x\hat{u}_0(1,0,0)\neq 0,\\
&f(z) = \cos\big(z - n(\pi/2)\big), \quad z\in\mathbb{R}, \quad \text{if } \partial_x\hat{u}_0(1,0,0) = 0.
\end{aligned}$$

Going back to (117), we deduce that the bound $|\partial^n_{y,\dots,y}\hat{u}_p(t,x,y)| \le Cp^{n-8/3-2\delta}t^{-2n+4/3+\delta}$, $t\in(0,1]$, $|x|\le 1$, $y\in\mathbb{R}$, cannot be true for some $\delta > 0$. By scaling, we deduce that the bound

$|\partial^n_{y,\dots,y} u_p(t,x,y)| \leq Ct^{-2n+4/3+\delta}$, $t \in (0,1]$, $|x| \leq 1$, $y \in \mathbb{R}$, cannot be true. This shows sharpness of the bound in Theorem 4.1 for the current example. □

6. Quadratic Case

Semilinear PDEs with quadratic nonlinearities appear in solving certain optimization problems encountered in mathematical finance (see [11, 28]). Their corresponding BSDE (11) is said to be quadratic if the growth of the driver f with respect to z is quadratic. Here, we will assume $|f(t,x,y,z)| \leq \Lambda_1(1+|y|+|z|^2)$, for some constant Λ_1 (independent of t). The exponent 2 is the critical one for the growth of the nonlinear term with respect to the spatial derivatives: it is known that existence and uniqueness may fail for higher exponents.

Following Dos Reis [9] (see Assumptions (HY1) and (HY1+) in Theorems 3.1.9 and 3.1.11 therein), we here investigate the case when the source term in (5) is $K-m-1$ times continuously differentiable w.r.t. x, y and z, $K \geq m+3$, with bounded derivatives of order greater than or equal to 2, and with first order derivatives of the following growth:

$$|\nabla_x f(t,x,y,z)| \leq \Lambda_1(1+|z|^2),\ |\nabla_y f(t,x,y,z)| \leq \Lambda_1, \quad |\nabla_z f(t,x,y,z)| \leq \Lambda_1(1+|z|).$$

(Below, Λ_n denotes a bound for the derivatives of order k between 2 and n, with $2 \leq n \leq K-m-1$.)

In this framework, BSDE (11) is well-posed provided the boundary condition h is bounded: we refer the reader to the original paper by Kobylanski [13]. Basically, the boundedness property ensures that the martingale driving the BSDE (11) is BMO. The BMO property plays a crucial role: under the BMO condition of the martingale part, one can apply Girsanov transformation to get rid of the quadratic part of the equation. We refer to Hu, Imkeller and Müller [11], Ankirchner, Imkeller and Dos Reis [1] and Dos Reis [9] for a review of this strategy. For this reason, the most natural approach is to estimate the first-order derivatives in terms of the L^∞ norm of h (and not in terms of L^p norms of h as in Theorem 4.1). We remind the reader of the following (see e.g. Lemma 1.2.13 in Dos Reis [9]):

Proposition 6.1. *Choose the driver f in* (11) *as above, then* (11) *is uniquely solvable for any starting point (t,x) of X. Moreover, the BMO-norm of the martingale part*

$$\left\| \int_t^\cdot \langle Z_s, dB_s \rangle \right\|_{\mathrm{BMO}} = \sup_{\text{Stopping Times } t \leq \tau \leq T} \mathbb{E}\left[\int_\tau^T Z_s^2 ds \middle| \mathcal{F}_\tau \right]^{1/2}$$

is finite and bounded by a constant C, depending on Λ_1, T and $\|h\|_\infty$ only.

As announced, Girsanov assumption holds under BMO property (see Theorem 3.1 in Kazamaki [12]):

Proposition 6.2. *For any progressively-measurable process $(\mu_t)_{0\leq t\leq T}$ with values in $\mathbb{R}^N$ such that $(M_t = \int_0^t \langle \mu_s, dB_s \rangle)_{0\leq t\leq T}$ has a finite BMO-norm, there exists an exponent $q^* > 1$, depending on the BMO-norm of $(M_t)_{0\leq t\leq T}$ only, such that the $L^{q^*}(\mathbb{P})$-norm of the exponential martingale of $(M_t)_{0\leq t\leq T}$ is finite and bounded by a constant, depending on the BMO-norm of $(M_t)_{0\leq t\leq T}$ only.*

We have the following:

Theorem 6.3. *Let $(V_i)_{0\leq i\leq N}$ be $N+1$ vector fields satisfying Definition 1.1. Assume that the source term in* (5) *is as in Proposition 6.1 and that h is a bounded Lipschitz function. Then, for any $t > 0$,*

$u(t,\cdot)$ belongs to $\mathcal{D}_V^{K-m-1/2,\infty}(\mathbb{R}^d)$. Moreover, for any $T>0$, $n \le K-m-1$ and $\alpha_1,\dots,\alpha_n \in \mathcal{A}_0(m)$, there exists a constant C_n, depending on Λ_1, Λ_n, n, T, the L^∞-bound of h, the Lipschitz constant of h and the vector fields $V_0,\dots,V_N$ only, such that for all $(t,x) \in (0,T]\times\mathbb{R}^d$,

$$\big|V_{[\alpha_1]}\dots V_{[\alpha_n]}u(t,x)\big| \le C_n t^{(1-\|\boldsymbol{\alpha}\|)/2},$$
$$\big|V_{[\alpha_1]}\dots V_{[\alpha_n]}V_i u(t,x)\big| \le C_n t^{-\|\boldsymbol{\alpha}\|/2}, \quad 1\le i\le N.$$

Proof. The proof is identical with the case when f is assumed to be Lipschitz. The reason is quite simple: when h is smooth, the gradient is known to exist and to be bounded in any directions of the space in terms of the Lipschitz constant of h. This is proved by Dos Reis [9], see Lemma 3.1.4 and Theorem 3.1.11 therein. As a consequence, quadratic growth does not affect the small time asymptotic behaviour of the higher order derivatives, but only the dependence of the constant C_n on the L^∞-bound and Lipschitz constant of h. Using a mollification argument as in the proof of Theorem 3.1, we complete the proof. □

The non-Lipschitz case is much more involved. Here we no longer have available the result of Dos Reis [9] for the control of the first order derivatives. The first step is to obtain a bound for the first order derivatives. Once obtained, the analysis is handled as in the non-quadratic case.

Lemma 6.4. *Let $(V_i)_{0\le i\le N}$ be $N+1$ vector fields satisfying Definition 1.1. Assume that the source term in* (5) *has the same structure as in Proposition 6.1 and that h is a bounded continuous function*[7]*. Then, for any $t>0$, $u(t,\cdot)$ belongs to $\mathcal{D}_V^{3/2,\infty}(\mathbb{R}^d)$ and, for any $T>0$, there exists a constant C, depending on Λ_1, T, $\|h\|_\infty$ and the vector fields only, such that, for any $\alpha\in\mathcal{A}_0(m)$ and $(t,x)\in(0,T]\times\mathbb{R}^d$, $|V_{[\alpha]}u(t,x)| \le Ct^{-\|\alpha\|/2}$.*

Proof. As above, we first mollify the boundary condition. We then need to prove (in the mollified setting) the announced estimates in terms of the parameters Λ_1, T and $\|h\|_\infty$ only.

By Kobylanski [13], we know that u is bounded in terms of Λ_1 and T only. This point is crucial in what follows. Let $(X_s^{t,x},Y_s^{t,x},Z_s^{t,x})_{t\le s\le T}$ be the solution of the equation (11), with $X_t^x = x\in\mathbb{R}^d$ as initial condition. By Lemma 1.2.13 in Dos Reis [9], for any $p\ge 1$, there exists a constant C_p, depending on Λ_1, T and $\|h\|_\infty$ only, such that

$$\mathbb{E}\left[\left(\int_t^T |Z_r^{t,x}|^2 dr\right)^p\right] \le C_p. \tag{119}$$

By Theorem 3.1.9 in Dos Reis [9], we can differentiate $(X_s^{t,x},Y_s^{t,x},Z_s^{t,x})_{t\le s\le T}$ with respect to x as a function from $\mathbb{R}^d$ into the space of $\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^N$-valued processes $(\xi_s,\Upsilon_s,\zeta_s)_{t\le s\le T}$ endowed with the norm $\mathbb{E}[\sup_{t\le s\le T}(|\xi_s|^2+|\Upsilon_s|^2)+\int_t^T|\zeta_s|^2ds]^{1/2}$. The derivative process satisfies

$$\begin{aligned} d\big[V_{[\alpha]}(x)Y_s^{t,x}\big] &= -\nabla_x f(\Theta_s)V_{[\alpha]}(x)X_s^{t,x}ds - \nabla_y f(\Theta_s)V_{[\alpha]}(x)Y_s^{t,x}ds \\ &\quad - \nabla_z f(\Theta_s)V_{[\alpha]}(x)Z_s^{t,x}ds + dB_s^\top V_{[\alpha]}(x)Z_s^{t,x}, \end{aligned}$$

where $\Theta_s = (s,X_s^{t,x},Y_s^{t,x},Z_s^{t,x})$. By Theorem 3.1.11 in [9], the process $(Y_s^{t,x})_{t\le s\le T}$ is pathwise continuously differentiable w.r.t. x. In particular, for any $t>0$, $u(t,\cdot)$ is continuously differentiable and $V_{[\alpha]}(x)[Y_s^{t,x}] = \nabla_x u(T-s,X_s^{t,x})\nabla_x X_s^{t,x}V_{[\alpha]}(x)$.

[7] For the sake of clarity, we only give the statement for continuous boundary condition. The statement for the discontinuous case follows the model of Theorem 4.1.

First Step. Girsanov Transformation. Owing to Propositions 6.1 and 6.2 (or taking advantage of the mollified setting), we know that the exponential martingale

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\biggl(\int_t^T \nabla_z f(\Theta_r) dB_r - \frac{1}{2}\int_t^T \bigl|\nabla_z f(\Theta_r)\bigr|^2 dr\biggr)$$

defines a new probability measure $\mathbb{Q}$ under which the process $(\bar{B}_s = B_s - \int_t^s (\nabla_z f)^\top(\Theta_r) dr)_{t\le r\le s}$ is a Brownian motion.

In particular, under $\mathbb{Q}$, the process $(V_{[\alpha]}(x)Y_s^{t,x})_{t\le s\le T}$ admits the following semi-martingale decomposition:

$$\begin{aligned} &d\bigl[V_{[\alpha]}(x)Y_s^{t,x}\bigr] \\ &= -\nabla_x f(\Theta_s) V_{[\alpha]}(x) X_s^{t,x} ds - \nabla_y f(\Theta_s) V_{[\alpha]}(x) Y_s^{t,x} ds + (d\bar{B}_s)^\top V_{[\alpha]}(x) Z_s^{t,x}. \end{aligned} \tag{120}$$

By standard BSDE results (see, for example, [2]), for any $p \ge 1$, we can find a constant C_p' (whose value may vary from line to line), depending on Λ_1, p, T and $\|h\|_\infty$ only, such that

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}\biggl[\biggl(\int_t^{(T+t)/2} |V_{[\alpha]}(x) Z_s^{t,x}|^2 ds\biggr)^p\biggr] &\le C_p' \sup_{t\le r\le (T+t)/2} \mathbb{E}^{\mathbb{Q}}\bigl[\bigl|V_{[\alpha]}(x) Y_r^{t,x}\bigr|^{2p}\bigr] \\ &+ C_p' \mathbb{E}^{\mathbb{Q}}\biggl[\sup_{t\le s\le (T+t)/2} |\nabla_x X_s^{t,x}|^{2p} \cdot \biggl|\int_t^{(T+t)/2} \bigl(1 + |Z_r^{t,x}|^2\bigr) dr\biggr|^{2p}\biggr]. \end{aligned}$$

By the BMO condition (see Proposition 6.2), we know that the density $d\mathbb{Q}/d\mathbb{P}$ belongs to the space $L^{q^*}(\mathbb{P})$, for some $q^* > 1$, the $L^{q^*}(\mathbb{P})$-norm being bounded in terms of known parameters. By (119), we deduce that

$$\mathbb{E}^{\mathbb{Q}}\biggl[\biggl(\int_t^{(T+t)/2} |V_{[\alpha]}(x) Z_s^{t,x}|^2 ds\biggr)^p\biggr] \le C_p'\bigl(1 + \sup_{t\le r\le (T+t)/2} \mathbb{E}^{\mathbb{Q}}\bigl[\bigl|V_{[\alpha]}(x) Y_r^{t,x}\bigr|^{2p}\bigr]\bigr). \tag{121}$$

By Lemma 2.3, we have

$$\begin{aligned} V_{[\alpha]}(x) Y_r^{t,x} &= \nabla_x u\bigl(T-r, X_r^{t,x}\bigr) \nabla_x X_r^{t,x} V_{[\alpha]}(x) \\ &= \sum_{\beta\in\mathcal{A}_0(m)} \theta_t^*(b_{\alpha,\beta})\bigl(r-t, x\bigr)\bigl(V_{[\beta]} u\bigr)\bigl(T-r, X_r^{t,x}\bigr). \end{aligned} \tag{122}$$

Using again the bound for $d\mathbb{Q}/d\mathbb{P}$ in $L^{q^*}(\mathbb{P})$, we deduce that

$$\mathbb{E}^{\mathbb{Q}}\bigl[\bigl|V_{[\alpha]}(x) Y_r^{t,x}\bigr|^{2p}\bigr] \le C_p' \sum_{\beta\in\mathcal{A}_0(m)} (r-t)^{p(\|\beta\|-\|\alpha\|)^+} \sup_{y\in\mathbb{R}^d} \bigl|\bigl(V_{[\beta]} u\bigr)\bigl(T-r, y\bigr)\bigr|^{2p}. \tag{123}$$

Finally, we emphasize from Definition 2.2 in Kazamaki [12] that $d\mathbb{P}/d\mathbb{Q}$ is in $L^{r^*}(\mathbb{P})$ for some $r^* > 0$, that is $d\mathbb{P}/d\mathbb{Q}$ is in $L^{1+r^*}(\mathbb{Q})$. (The norms in $L^{r^*}(\mathbb{P})$ and $L^{1+r^*}(\mathbb{Q})$ being controlled in terms of known parameters, see Theorem 2.4 in [12].) Therefore,

$$\mathbb{E}\biggl[\biggl(\int_t^{(T+t)/2} |V_{[\alpha]}(x) Z_s^{t,x}|^2 ds\biggr)^p\biggr] \le C\mathbb{E}^{\mathbb{Q}}\biggl[\biggl(\int_t^{(T+t)/2} |V_{[\alpha]}(x) Z_s^{t,x}|^2 ds\biggr)^{p(1+r^*)/r^*}\biggr]^{r^*/(1+r^*)}.$$

Finally, (121) and (123) yield

$$\mathbb{E}\Big[\Big(\int_t^{(T+t)/2} |V_{[\alpha]}(x) Z_s^{t,x}|^2 ds\Big)^p\Big] \leq C_p' \sum_{\beta \in \mathcal{A}_0(m)} \sup_{t \leq r \leq (T+t)/2} \big[(T-r)^{p(\|\beta\|-\|\alpha\|)^+} \big\| V_{[\beta]} u(r,\cdot)\big\|_\infty^{2p}\big]. \tag{124}$$

Second Step. Integration by Parts. By (119) and the trivial inequality

$$C_{2p} \geq \mathbb{E}\Big[\Big(\int_{t+(T-t)/2}^{t+3(T-t)/4} |Z_r^{t,x}|^2 dr\Big)^{2p}\Big] \geq \sum_{\ell=1}^L \mathbb{E}\Big[\Big(\int_{t_{\ell-1}}^{t_\ell} |Z_r^{t,x}|^2 dr\Big)^{2p}\Big]$$

that holds for any mesh $t + (T-t)/2 = t_0 \leq t_1 \leq \cdots \leq t_L = t + 3(T-t)/4$, we deduce, by choosing $t_\ell = t + [1/2 + \ell/(4L)](T-t)$, that, for a given value of p (that will be chosen later on) and for any large enough integer L, there exists a certain $s \in [t + (T-t)/2, t + [3/4 - 1/(4L)](T-t)]$ such that

$$\mathbb{E}\Big[\Big(\int_s^{s+(T-t)/(4L)} |Z_r^{t,x}|^2 dr\Big)^{2p}\Big] \leq C_{2p}/L. \tag{125}$$

We now come back to (11). By integration by parts (see Theorem 2.5), we know that

$$\begin{aligned}
&(V_{[\alpha]} u)(T-t,x) \\
&= [(T-t)/(4L)]^{-\|\alpha\|/2} \mathbb{E}\big[u\big([1-1/(4L)](T-t), X_{t+(T-t)/(4L)}^{t,x}\big) \theta_t^*[\phi_\alpha]\big((T-t)/(4L), x\big)\big] \\
&\quad + \mathbb{E}\int_t^{t+(T-t)/(4L)} \big[\nabla_x f(\Theta_r) V_{[\alpha]}(x) X_r^{t,x} + \nabla_y f(\Theta_r) V_{[\alpha]}(x) Y_r^{t,x} + \nabla_z f(\Theta_r) V_{[\alpha]}(x) Z_r^{t,x}\big] dr.
\end{aligned}$$

Taking the power $2p$ and using the boundedness of u, we obtain

$$\begin{aligned}
\big|V_{[\alpha]} u(T-t,x)\big|^{2p} &\leq C_p' \big[1 + L^{p\|\alpha\|}(T-t)^{-p\|\alpha\|}\big] \\
&\quad + C_p' [(T-t)/L]^{2p} \sup_{t \leq r \leq t+(T-t)/(4L)} \mathbb{E}\big[\big|V_{[\alpha]}(x) Y_r^{t,x}\big|^{2p}\big] \\
&\quad + C_p' \mathbb{E}\Big[\Big(\int_t^{t+(T-t)/(4L)} (1 + |Z_r^{t,x}|^2) dr\Big)^{2p}\Big]^{1/2} \mathbb{E}\Big[\Big(\int_t^{t+(T-t)/(4L)} \big|V_{[\alpha]}(x) Z_r^{t,x}\big|^2 dr\Big)^{2p}\Big]^{1/2}.
\end{aligned}$$

Applying Lemma 2.3 to expand $V_{[\alpha]}(x) Y_r^{t,x}$ as in (122) and using (124) to bound the $L^{2p}(\mathbb{P})$-moment of $\int_t^{t+(T-t)/(4L)} \big|V_{[\alpha]}(x) Z_r^{t,x}\big|^2 dr$,

$$\begin{aligned}
\big|V_{[\alpha]} u(T-t,x)\big|^{2p} &\leq C_p' \big(1 + L^{p\|\alpha\|}(T-t)^{-p\|\alpha\|}\big) \\
&\quad + C_p' \Big[[(T-t)/L]^p + \mathbb{E}\Big[\Big(\int_t^{t+(T-t)/(4L)} |Z_r^{t,x}|^2 dr\Big)^{2p}\Big]^{1/2}\Big] \\
&\quad \times \sum_{\beta \in \mathcal{A}_0^m} \sup_{t \leq r \leq (T+t)/2} \big[(T-r)^{p(\|\beta\|-\|\alpha\|)^+} \|(V_{[\beta]} u)(T-r,\cdot)\big\|_\infty^{2p}\big].
\end{aligned}$$

Clearly, we can replace t by s and then x by $X_s^{t,x}$ in the above inequality, with s as in (125). Taking the expectation and using the Markov property, we obtain

$$\begin{aligned}&\mathbb{E}\big[\big|V_{[\alpha]}u(T-s,X_s^{t,x})\big|^{2p}\big] \le C_p'\big(1+L^{p\|\alpha\|}(T-s)^{-p\|\alpha\|}\big)\\&\quad+(C_p'/L^{1/2})\bigg[1+\sum_{\beta\in\mathcal{A}_0(m)}\sup_{s\le r\le(T+s)/2}\big[(T-r)^{p(\|\beta\|-\|\alpha\|)^+}\big\|(V_{[\beta]}u)(T-r,\cdot)\big\|_\infty^{2p}\big]\bigg].\end{aligned}$$

Since $s\le t+3(T-t)/4$, we can replace $T-s$ in the second term above by $T-t$ by modifying C_p'. Moreover, $s\le t+3(T-t)/4$ implies $(T+s)/2\le(7T+t)/8$. We deduce

$$\begin{aligned}&\mathbb{E}\big[\big|V_{[\alpha]}u(T-s,X_s^{t,x})\big|^{2p}\big] \le C_p'\big(1+L^{p\|\alpha\|}(T-t)^{-p\|\alpha\|}\big)\\&\quad+(C_p'/L^{1/2})\bigg[1+\sum_{\beta\in\mathcal{A}_0(m)}\sup_{t\le r\le(7T+t)/8}\big[(T-r)^{p(\|\beta\|-\|\alpha\|)^+}\big\|(V_{[\beta]}u)(T-r,\cdot)\big\|_\infty^{2p}\big]\bigg].\end{aligned}\tag{126}$$

Third Step. Girsanov Transformation again. By (120), keep in mind that (with the same s as above)

$$\begin{aligned}&V_{[\alpha]}(x)u(T-t,x)\\&=\mathbb{E}^{\mathbb{Q}}\bigg[V_{[\alpha]}u(T-s,X_s^{t,x})+\int_t^s\big[\nabla_x f(\Theta_r)V_{[\alpha]}(x)X_r^{t,x}+\nabla_y f(\Theta_r)V_{[\alpha]}(x)Y_r^{t,x}\big]dr\bigg].\end{aligned}$$

Recall that the density $d\mathbb{Q}/d\mathbb{P}$ belongs to $L^{q^*}(\mathbb{P})$, with a well-controlled norm. (See Theorem 2.4 in [12].) Choosing $2p$ greater than the conjugate exponent of p^* (since s depends on p, this says that s is now fixed), we deduce from Hölder's inequality and from (122) that

$$\begin{aligned}&\big|V_{[\alpha]}(x)u(T-t,x)\big|^{2p}\\&\le C_p'\mathbb{E}\big[\big|V_{[\alpha]}(x)u(T-s,X_s^{t,x})\big|^{2p}\big]\\&\quad+C_p'\mathbb{E}\bigg[\bigg|\int_t^s\big[\nabla_x f(\Theta_r)V_{[\alpha]}(x)X_r^{t,x}+\nabla_y f(\Theta_r)V_{[\alpha]}(x)Y_r^{t,x}\big]dr\bigg|^{2p}\bigg]\\&\le C_p'\mathbb{E}\big[\big|V_{[\alpha]}(x)u(T-s,X_s^{t,x})\big|^{2p}\big]\\&\quad+C_p'(T-t)^{2p-1}\bigg[1+\sum_{\beta\in\mathcal{A}_0(m)}\sup_{t\le r\le(3T+t)/4}\big[(T-r)^{p(\|\beta\|-\|\alpha\|)^+}\big\|V_{[\beta]}u(T-r,\cdot)\big\|_\infty^{2p}\big]dr\bigg].\end{aligned}$$

By (126),

$$\begin{aligned}&\big|V_{[\alpha]}u(T-t,x)\big|^{2p}\le C_p'\big(1+L^{p\|\alpha\|}(T-t)^{-p\|\alpha\|}\big)\\&\quad+C_p'\big(T-t+1/L^{1/2}\big)\sum_{\beta\in\mathcal{A}_0(m)}\sup_{t\le r\le(7T+t)/8}\big[(T-r)^{p(\|\beta\|-\|\alpha\|)^+}\big\|(V_{[\beta]}u)(T-r,\cdot)\big\|_\infty^{2p}\big].\end{aligned}$$

Multiplying by $(T-t)^{p\|\alpha\|}$, using the bound $(T-t)^{p\|\alpha\|}(T-r)^{p(\|\beta\|-\|\alpha\|)^+}\le C(T-r)^{p\|\beta\|}$ for $t\le r\le(7T+t)/8$, taking the supremum over $x\in\mathbb{R}^d$ and then choosing L large enough and $T-t$ small enough, we complete the proof. (Clearly, the bound is proven on some small interval of the form $[T-\delta,T)$, $\delta>0$. By a similar argument, the bound holds on any $[t-\delta/2,t+\delta/2)$, $\delta/2\le t\le T-\delta/2$. That is, $V_{[\alpha]}u(t,\cdot)$ is uniformly bounded for $0\le t\le T-\delta/2$.) $\square$

The lemma gives us the gradient bounds for the higher order derivatives as in the case when f is Lipschitz.

Theorem 6.5. *Let $(V_i)_{0\le i\le N}$ be $N+1$ vector fields satisfying Definition 1.1, let f be as in Proposition 6.1, and let h be a bounded continuous function (see Footnote[7]). Then, for any $t>0$, $u(t,\cdot)$ belongs to $\mathcal{D}_V^{K-m-1/2,\infty}(\mathbb{R}^d)$.*

Moreover, for any $T>0$ and $\alpha_1,\alpha_2\in\mathcal{A}_0(m)$, there exists a constant C_2, depending on Λ_1, Λ_2, T, $\|h\|_\infty$ and the vector fields $V_0,\dots,V_N$ only, such that for all $t\in[0,T)$ and $x\in\mathbb{R}^d$,

$$\big|V_{[\alpha_1]}V_{[\alpha_2]}u(t,x)\big|\le C_2 t^{-(\|\alpha_1\|+\|\alpha_2\|)/2}, \tag{127}$$

and for any $\delta>0$, $3\le n\le K-m-1$ and $\alpha_1,\dots,\alpha_n\in\mathcal{A}_0(m)$, there exists a constant $C_n(\delta)$, depending on δ, Λ_1, Λ_n, n, T, $\|h\|_\infty$ and the vector fields $V_0,\dots,V_N$ only, such that for all $t\in(0,T]$ and $x\in\mathbb{R}^d$,

$$\big|V_{[\alpha_1]}\dots V_{[\alpha_n]}u(t,x)\big|\le C_n(\delta)t^{-\|\boldsymbol{\alpha}\|/2}\big[1+t^{-n/2+1+\min(1/\|\alpha_{(1)}\|,1/2+1/(2\|\alpha_{(2)}\|))-\delta}\big]. \tag{128}$$

with $1\le i\le N$, where $\alpha_{(1)}$ and $\alpha_{(2)}$ stand for multi-indices in the family $\alpha_1,\dots,\alpha_n$ such that $\|\alpha_{(1)}\|\le\|\alpha_{(2)}\|$ are the two smallest elements in the family $\|\alpha_1\|,\dots,\|\alpha_n\|$.

Proof. There are not so many differences with the case when f is at most of linear growth: most of the work has been done in Lemma 6.4. Comparing with the proof of Theorem 4.1, we understand that we first have to check the validity of Propositions 4.2 and 4.3 and of Corollary 4.4.

Extension of Proposition 4.2 to the quadratic case. We first notice that Lemma 3.4 holds in the quadratic but smooth framework: following the proof of Theorem 6.3 (or applying Theorem 6.3 directly), we know that Vu exists and is bounded when the boundary condition is Lipschitz continuous, that is the driver f may be assumed to be bounded when the boundary condition is smooth, so that Lemma 3.4 applies in the smooth framework. The first line in Proposition 4.2 is then proven by differentiating the representation formula for $u(t,x)$ n times. Since the derivatives of f of order greater than 2 are here bounded, most of the terms in Proposition 4.2 remain unchanged in the quadratic case. Basically, we must pay attention to the boundary condition, which is now estimated in L^∞ through a non-explicit constant as in Lemma 6.4. We must also pay attention to the terms involving the first derivatives of f w.r.t. x or z, i.e. to the term $\nabla_x f(\Theta(s,X^x_{t-s}))V_{[\alpha_1]}\dots V_{[\alpha_n]}[X^x_{t-s}]$ and to the term $\nabla_z f(\Theta(s,X^x_{t-s}))V_{[\alpha_1]}\dots V_{[\alpha_n]}[(Vu)^\top(s,X^x_{t-s})]$ in the proof of Corollary 3.3, Corollary 3.3 being the keystone of the proof of Proposition 4.2. Here, $\Theta(s,X^x_{t-s})$ stands for the 4-tuple $(s,X^x_{t-s},u(s,X^x_{t-s}),(Vu)^\top(s,X^x_{t-s}))$. Clearly, $\nabla_x f(\Theta(s,X^x_{t-s}))V_{[\alpha_1]}\dots V_{[\alpha_n]}[X^x_{t-s}]$ is of order s^{-1} by Lemma 6.4. Since it is integrated over an interval of length $t/2$, it doesn't affect the decay of the boundary condition. The term $\nabla_z f(\Theta(s,X^x_{t-s}))V_{[\alpha_1]}\dots V_{[\alpha_n]}[(Vu)^\top(s,X^{t,x}_s)]$ is more difficult to handle. By the linear growth of $\nabla_z f$ in z, it is of order $s^{-1/2}|V_{[\alpha_1]}\dots V_{[\alpha_n]}[(Vu)(s,X^x_{t-s})]|$. Following the proof of Corollary 3.3, we are to evaluate the $(V_{[\alpha_j]})_{1\le j\le n}$ at X^x_{t-s}. Using Lemma 2.3, in the first line in Proposition 4.2, we get new terms of the form

$$\sum_{k=1}^n\sum_{\boldsymbol{\beta}}\int_{t/2}^t s^{-1/2}(t-s)^{(\|\boldsymbol{\beta}\|-\|\boldsymbol{\alpha}\|)^+/2}\mathbb{E}\big[\big|\big(V_{[\beta_1]}\dots V_{[\beta_k]}Vu\big)(s,X^x_{t-s})\big|^p\big]^{1/p}ds, \tag{129}$$

$\boldsymbol{\beta}$ running over the k-tuples of multi-indices $(\beta_1,\dots,\beta_k) \in [\mathcal{A}_0(m)]^k$. Below, the terms in the integral in (129) will be referred to as "non-product terms" since the iterated derivatives are not multiplied between them (compare with Proposition 4.2).

Now, we must do the same job for the second line, that is for the terms deriving from the integration by parts used to obtain the second line. Clearly, the terms $\nabla_x f(s,\Theta_s)V_{[\alpha_1]}\dots V_{[\alpha_n]}[X^x_{t-s}]$ and $\nabla_z f(s,\Theta_s)V_{[\alpha_1]}\dots V_{[\alpha_n]}[(Vu)^\top(s,X^x_{t-s})]$ modify the second inequality as they modify the first one: the term $\nabla_x f(\Theta(s,X^x_{t-s}))V_{[\alpha_1]}\dots V_{[\alpha_n]}[X^x_{t-s}]$ doesn't change anything to the final rate; and the term $\nabla_z f(\Theta(s,X^x_{t-s}))V_{[\alpha_1]}\dots V_{[\alpha_n]}[(Vu)^\top(s,X^x_{t-s})]$ generates a new $s^{-1/2}$ in the integrals of the non-product terms. Anyhow, we must also pay attention to $T_4(s,t,x)$ in (56). Since we do not take into account the dependence of the final constants upon $\|h\|_\infty$, it is here enough to bound $|f(\Theta(s,X^x_{t-s}))|$ by $C(1+s^{-1})$. Obviously, this doesn't affect the resulting control of the boundary condition in the second line in Proposition 4.2.

Extension of Proposition 4.3 to the quadratic case. As for Proposition 4.2, the dependence upon $\|h\|_\infty$ cannot be made explicit in the new version of Proposition 4.3. Up to this restriction, Proposition 4.3 holds true for $n=1$: this is Lemma 6.4.

To see how the property propagates with n, we are to analyse how the new version of Proposition 4.2 affects the induction. Assuming that Proposition 4.3 holds true up to $n-1\ge 1$ in the quadratic case (up to the shape of the dependence upon $\|h\|_\infty$), we then plug (129) in the induction property: for $k=1,\dots,n-1$, the worst contribution in the first line of Proposition 4.3 is of order $t^{-\|\boldsymbol{\alpha}\|/2}t^{-(n-3)/2}$; in the second line of Proposition 4.3, the worst contribution is of order $t^{-\|\boldsymbol{\alpha}\|/2}t^{-(n-2)/2}$; in the end, the final bound is not affected. (Actually, this is well-expected: nonlinearity affects the final bound through product terms only.) The difficult point is in (81) and (82): there is an additional $s^{-1/2}$ in the second lines because of the additional $s^{-1/2}$ in (129). As a consequence, (83) reads

$$\sum_{\alpha_1,\dots,\alpha_n} t^{\|\boldsymbol{\alpha}\|/2+(n-2)/2}\mathcal{Q}^n_{\alpha_1,\dots,\alpha_n}(t,r,x) \le C_n(p)\Big[1+ \\ + \sum_{\beta_1,\dots,\beta_n}\int_{t/2}^t (t-s)^{-1/2}s^{-1/2}s^{\|\boldsymbol{\beta}\|/2+(n-2)/2}\mathcal{Q}^n_{\beta_1,\dots,\beta_n}(s,r,x)ds\Big],$$

To make it tractable, we proceed as follows. Following the proof of Lemma 6.4, the idea is to replace the lower bound $t/2$ in the integral by $[(L-1)/L]t$ for L large. This makes very short the length of the interval over which the integration is performed. Basically, this just deteriorates the constant of the integration by parts in the new version of Proposition 4.2, that is the bound therein reads as $L^{\|\boldsymbol{\alpha}\|/2}t^{-\|\boldsymbol{\alpha}\|/2}$. Therefore, we get

$$\sum_{\alpha_1,\dots,\alpha_n} t^{\|\boldsymbol{\alpha}\|/2+(n-2)/2}\mathcal{Q}^n_{\alpha_1,\dots,\alpha_n}(t,r,x) \le C_n(p)\Big[1+L^{\|\boldsymbol{\alpha}\|/2} \\ + \sum_{\beta_1,\dots,\beta_n}\sup_{0<s\le r}\big[s^{\|\boldsymbol{\beta}\|/2+(n-2)/2}\mathcal{Q}^n_{\beta_1,\dots,\beta_n}(s,r,x)\big]\int_{[(L-1)/L]t}^t (t-s)^{-1/2}s^{-1/2}ds\Big], \tag{130}$$

Now, notice that

$$\int_{[(L-1)/L]t}^{t} (t-s)^{-1/2} s^{-1/2} ds \le L^{1/2}(L-1)^{-1/2} t^{-1/2} \times L^{-1/2} t^{1/2} = (L-1)^{-1/2}.$$

Therefore, choosing L large enough and taking the supremum w.r.t. $t \in (0, r]$ in (130), we can complete the proof of the new version of Proposition 4.3.

Extension of Corollary 4.4 to the quadratic case. The new version of Corollary 4.4, that is when F therein satisfies the same growth properties as f, is obtained as the new version of Proposition 4.2: the terms for which $k = 1$ in (84) are affected by an additional $s^{-1/2}$ following from the growth of $\nabla_z f$, on the same model as in (129); the terms for which $k = 0$ are affected by an additional s^{-1} following from the growth of $\nabla_x f$; and F itself, in the product $F(\Theta(X_s^x))\phi_0(s,x)$, increases as s^{-1} for s small.

Completion of the Proof in the Smooth Setting. For $n = 2$, the extension of Proposition 4.2 already applies. For $n \ge 3$, we follow the end of the proof of Theorem 4.1. We must check that the additional terms in the new versions of Proposition 4.2 and Corollary 4.4 do not affect the final estimate. In the original proof of Theorem 4.1, the worst possible bound is $s^{-\|\boldsymbol{\alpha}\|/2-n/2}$ when differentiating n times $f(\Theta(s, X_s^x))$, $s^{-\|\boldsymbol{\alpha}\|/2-n/2} s^{(\|\alpha_{(1)}\|+1)/2}$ when differentiating it $(n-1)$ times and $s^{-\|\boldsymbol{\alpha}\|/2-n/2} s^{(\|\alpha_{(1)}\|+\|\alpha_{(2)}\|)/2+1}$ when differentiating it $(n-2)$ times. We now compare this bound with the bound of the so-called "non-product terms", that is the terms affected by the additional $s^{-1/2}$, as in (129). All these terms count a single factor of the form $V_{[\beta_1]} \dots V_{[\beta_k]} Vu$: using Proposition 4.3, the worst bound for all of them is $s^{-1/2} s^{-\|\boldsymbol{\alpha}\|/2-(n-2)/2-1/2}$, i.e. $s^{-\|\boldsymbol{\alpha}\|/2-n/2}$ exactly! Obviously, the same holds when differentiating $(n-1)$ or $(n-2)$ times only. It then remains to see how the terms affected by the additional s^{-1} behave: keep in mind that all these ones are free of any terms of the form $V_{[\beta_1]} \dots V_{[\beta_k]} Vu$. The worst bound for all these terms is s^{-1}, which is less than $s^{-\|\boldsymbol{\alpha}\|/2-(n-2)/2}$.

The general case. Generally speaking, the proof is the same as in the case when f is at most of linear growth w.r.t. z. Basically, only the starting point is different: we here use stability results for quadratic BSDEs to derive the convergence of the mollified sequence $(u_\ell)_{\ell \ge 1}$ towards u, with the same notation as in Subsubsection 4.5.1. Stability results for quadratic BSDEs may be found in Lemma 2.1.2 in Dos Reis [9]. The end of the proof is completely similar: away from the boundary, Lemma 6.4 applies and the driver f is bounded. □

7. Connection with PDEs

We prove here Propositions 2.8 and 2.12.

7.1. **Proof of Proposition 2.8.**

The proof relies on the following version of Itô's formula :

Proposition 7.1. *Let v satisfy part (1) in Definition 2.6 and be at most of polynomial growth as in* (21). *Then, for any $T > 0$ and $x \in \mathbb{R}^d$, a.s., for any $t \le s < T$,*

$$v(T-s, X_s^{t,x}) = v(T,x) + \int_t^s \left[-\mathcal{V}_0 v + \frac{1}{2} \sum_{i=1}^N V_i^2 v \right](T-r, X_r^{t,x}) dr + \int_t^s V v(T-r, X_r^{t,x}) dB_r.$$

We first assume that Proposition 7.1 holds true and prove first that the unique solvability of the PDE (5) holds.

7.1.1. *Solvability.* We first check that u satisfies (1) and (3) in Definition 2.6. To do so, we consider an approximating sequence $(h_\ell)_{\ell\geq 1}$ of h as in Subsection 3.5 or as in Subsection 4.5 for the continuous case and we denote by u_ℓ the associated solutions to the PDE (5). Since h is continuous, $(h_\ell)_{\ell\geq 1}$ here converges towards h uniformly on compact sets. Following Subsection 3.5, we know that $(u_\ell)_{\ell\geq 1}$ converges towards u uniformly on compact subsets of $[0,T]\times\mathbb{R}^d$. In particular, u is continuous up to the boundary. Taking the supremum over (t,x) in a compact subset of $(0,T]\times\mathbb{R}^d$ in (67), we deduce that $(Vu_\ell)_{\ell\geq 1}$ converges towards Vu uniformly on compact subsets of $(0,T]\times\mathbb{R}^d$. By the same argument, for any $\alpha_1,\alpha_2\in\mathcal{A}_0(m)$, $(V_{[\alpha_1]}u_\ell)_{\ell\geq 1}$ and $(V_{[\alpha_1]}V_{[\alpha_2]}u_\ell)_{\ell\geq 1}$ converge towards $V_{[\alpha_1]}u$ and $V_{[\alpha_1]}V_{[\alpha_2]}u$ uniformly on compact subsets of $(0,T]\times\mathbb{R}^d$. This proves that $V_{[\alpha_1]}u$ and $V_{[\alpha_1]}V_{[\alpha_2]}u$ are continuous on $(0,T]\times\mathbb{R}^d$. In the smooth setting, we know from Pardoux and Peng [27] that u_ℓ satisfies PDE (5) in the classical sense. Therefore, $(\mathcal{V}_0u_\ell)_{\ell\geq 1}$ is uniformly convergent on compact subsets of $(0,T]\times\mathbb{R}^d$: this shows that u belongs to $\mathcal{D}^{1,\infty}_{\mathcal{V}_0}((0,+\infty)\times\mathbb{R}^d)$. Passing to the limit in PDE in (5), we deduce that u satisfies (2).

7.1.2. *Uniqueness.* Uniqueness also follows from Proposition 7.1. Note first that the martingale term in Proposition 7.1 is local only. However, we can prove it to be a true martingale under the standing assumption (see Subsection 2.1). Indeed, by the PDE structure, for any starting point $(t,x)\in[0,T)\times\mathbb{R}^d$, the pair $(v(T-s,X_s^{t,x}),Vv(T-s,X_s^{t,x}))_{t\leq s<T}$ satisfies the BSDE (12) on $[t,T)$. By standard Young's inequality, it is then possible to prove that

$$\mathbb{E}\int_t^T|Vv(T-s,X_s^{t,x})|^2ds\leq C\sup_{t\leq s\leq T}\mathbb{E}\big[|v(T-s,X_s^{t,x})|^2\big],$$

for a constant C possibly depending on T. By the growth property of v, this proves that the martingale term is square integrable. Moreover, by the continuity of v up to the boundary, Eq. (12) is shown to hold up to time T. The initial condition of the diffusion being given, uniqueness of the classical solution easily follows by uniqueness of the solution to the BSDE (12).

7.2. **Proof of Proposition 7.1.** Clearly, Proposition 7.1 is true when v is smooth. When v is not smooth, the point is to approximate it by a sequence of smooth functions $(v_p)_{p\geq 1}$ such that

$$\text{(131)}\quad\forall r\geq 1,\quad\lim_{p\to+\infty}\sup_{1/r\leq t\leq T}\|v_p(t,\cdot)-v(t,\cdot)\|^{V,2}_{\mathbb{B}(0,r),\infty}=0,\ \lim_{p\to+\infty}\|v_p-v\|^{\mathcal{V}_0,1}_{[1/r,T]\times\mathbb{B}(0,r),\infty}=0.$$

Indeed, introducing the stopping times $(\tau_q=\inf\{s\geq t:|X_s^{t,x}|\geq q\})_{q\geq 1}$ ($\inf\emptyset=+\infty$), we can apply Itô's formula to $(v_p(T-s,X_s^x))_{0\leq s\leq\tau_q\wedge(T-\varepsilon)}$, ε standing for a small positive real, and then let p tend to $+\infty$. Property (131) then implies Itô's formula for $(v(T-s,X_s^x))_{0\leq s\leq\tau_q\wedge(T-\varepsilon)}$ until time $\tau_q\wedge(T-\varepsilon)$. Letting q tend to $+\infty$, this completes the proof.

It thus remains to prove (131). It is a consequence of the following convolution argument, the proof of which is left to the reader. □

Lemma 7.2. *For two smooth densities ρ_1 and ρ_d over $\mathbb{R}$ and $\mathbb{R}^d$, both with compact support, and for a solution v to the PDE as in Definition 2.6, define for all $\varepsilon > 0$*

$$v^\varepsilon(t,x) = \int_{\mathbb{R}^{d+1}} v(t-\varepsilon s, x-\varepsilon y)\mathbf{1}_{\{t-\varepsilon s>0\}}\rho_1(s)\rho_d(y)dsdy.$$

Then,

$$\forall r \geq 1, \quad \lim_{\varepsilon\to 0} \sup_{1/r\leq t\leq r} \|v^\varepsilon(t,\cdot) - v(t,\cdot)\|^{V,2}_{\mathbb{B}(0,r),\infty} = 0, \ \lim_{\varepsilon\to 0} \|v^\varepsilon - v\|^{\mathcal{V}_{0,1}}_{[1/r,r]\times\mathbb{B}(0,r),\infty} = 0.$$

7.3. **Proof of Proposition 2.12.** The proof of the proposition is based on a suitable version of Itô's formula. Because of the L^p setting, it cannot be true for any given starting point. We prove the following:

Proposition 7.3. *Let v satisfy part (1) in Definition 2.11 and be at most of polynomial growth as in* (21). *Then, for any $T > 0$ and any bounded $\mathcal{F}_t$-measurable (see Footnote[6]) and $\mathbb{R}^d$-valued random vector ξ, $0 \leq t < T$, with an absolutely continuous distribution w.r.t. the Lebesgue measure on $\mathbb{R}^d$, Itô's formula holds on the same model as in Proposition 7.1, but replacing $X_s^{t,x}$ by $X_s^{t,\xi}$ therein.*

In particular, the process $(v(T-s, X_s^{t,\xi}))_{t\leq s\leq T}$ admits a continuous version.

We emphasize that, in Itô's formula, all the terms are uniquely defined even if the derivatives of v are defined up to sets of zero Lebesgue measure. This a consequence of Lemma 2.10.

We first assume that Proposition 7.3 holds true and then prove that the unique solvability of the PDE (5) holds as well.

7.3.1. *Solvability.* We first check that u satisfies (1) in Definition 2.11. To do so, we consider an approximating sequence $(h_\ell)_{\ell\geq 1}$ of h as in Subsection 4.5 and we denote by $(u_\ell)_{\ell\geq 1}$ the associated solutions to the PDE (5). By (71), all the $(u_\ell)_{\ell\geq 1}$ are at most of polynomial growth on $[0,T]\times\mathbb{R}^d$, uniformly in ℓ. For a real $t \in [0,T)$ and an $\mathcal{F}_t$-measurable bounded random variable ξ with an absolutely continuous distribution, we deduce from standard stability results on BSDEs:

$$\begin{aligned}\sup_{t\leq s\leq T} \mathbb{E}\big[|(u-u_\ell)(T-s,X_s^{t,\xi})|^2\big] &= \sup_{t\leq s\leq T}\int_{\mathbb{R}^d} \mathbb{E}\big[|(u-u_\ell)(T-s,X_s^{t,x})|^2\big]\mu(x)dx \\ &\leq C\int_{\mathbb{R}^d} \mathbb{E}\big[|(h-h_\ell)(X_T^{t,x})|^2\big]\mu(x)dx,\end{aligned} \tag{132}$$

where μ stands for the density of the distribution of ξ. By Lemma 2.10, the above right-hand side converges to 0 as ℓ tends to $+\infty$, uniformly w.r.t. t in $[0,T]$. By polynomial growth of $(u_\ell)_{\ell\geq 1}$, the sequence $(u_\ell(t,\cdot))_{\ell\geq 1}$ converges towards $u(t,\cdot)$ in $\cap_{p\geq 1}L^p_{\rm loc}(\mathbb{R}^d)$, uniformly in $t\in[0,T]$. Applying (67) with $s=t$, $S=T$, $S'=T-\delta$, for δ small, and x replaced by ξ therein and then taking the supremum w.r.t. t in $[0,T-\delta]$, we deduce that $(Vu_\ell(t,\cdot))_{\ell\geq 1}$ converges towards $Vu(t,\cdot)$ in $L^2_{\rm loc}(\mathbb{R}^d)$, uniformly in t in compact subsets of $(0,T]$. By the bounds in Theorem 4.1, the convergence holds in any $L^p_{\rm loc}(\mathbb{R}^d)$, $p\geq 1$, uniformly in t in compact subsets of $(0,T]$. By the same argument, for any $\alpha_1,\alpha_2\in\mathcal{A}_0(m)$, $(V_{[\alpha_1]}u_\ell)_{\ell\geq 1}$ and $(V_{[\alpha_1]}V_{[\alpha_2]}u_\ell)_{\ell\geq 1}$ converge towards $V_{[\alpha_1]}u$ and $V_{[\alpha_1]}V_{[\alpha_2]}u$ in $\cap_{p\geq 1}L^p_{\rm loc}(\mathbb{R}^d)$, uniformly in t in compact subsets of $(0,T]$. This proves that $V_{[\alpha_1]}u$ and $V_{[\alpha_1]}V_{[\alpha_2]}u$ are measurable on $(0,T]\times\mathbb{R}^d$. (For any $t\in(0,T]$, $V_{[\alpha_1]}u(t,x)$ is the almost-everywhere limit of $\varepsilon^{-d}\int_{|r|\leq\varepsilon} V_{[\alpha_1]}u(t,x+r)dr$, which is time-space measurable. The same for

$V_{[\alpha_1]}V_{[\alpha_2]}u(t,x)$.) By PDE (5) (which holds in the classical sense in the smooth setting), $(\mathcal{V}_0 u_\ell)_{\ell\geq 1}$ converges in $\cap_{p\geq 1}L^p_{\rm loc}(\mathbb{R}^d)$, uniformly in t in compact subsets of $(0,T]$: this shows that u belongs to $\cap_{p\geq 1}\mathcal{D}^{1,p}_{\mathcal{V}_0}((0,+\infty)\times\mathbb{R}^d)$. Passing to the limit in (5), this proves (2) in Definition 2.11.

It finally remains to check that u satisfies the boundary condition (3) in Definition 2.11. By (71), the solution u is at most of polynomial growth. Taking the expectation in (12) and using the *a priori* estimates in Theorem 4.1, we then write $\mathbb{E}[Y_t^{t,x}]$ as $\mathbb{E}[h(X_T^{t,x})]+O((T-t)^{1/2})$, the Landau notation $O(\cdot)$ being uniform w.r.t. x on compact subsets. Therefore, with μ as above, $\lim_{t\to T}\int_{\mathbb{R}^d}|u(T-t,x)-\mathbb{E}[h(X_T^{t,x})]|\mu(x)dx=0$. We deduce that

$$\lim_{t\to T}\int_{\mathbb{R}^d}\big|u(T-t,x)-h(x)\big|\mu(x)dx=0, \tag{133}$$

provided

$$\lim_{t\to T}\int_{\mathbb{R}^d}\big|\mathbb{E}[h(X_T^{t,x})]-h(x)\big|\mu(x)dx=0. \tag{134}$$

Eq. (134) holds true when h is continuous. When h is not continuous, we can approximate it by a smooth function in $L^1_{\rm loc}(\mathbb{R}^d)$ and then apply Lemma 2.10. This implies (3) in Definition 2.11. □

7.3.2. *Connection with BSDE* (12). We emphasize here that, for an initial condition ξ as in Proposition 7.3, $(Y_s^{t,\xi})_{t\leq s\leq T}$ is a continuous version of $(u(T-s,X_s^{t,\xi}))_{t\leq s\leq T}$. When h is smooth, it holds true since $((Y_s^{t,x})_{t\leq s\leq T})_{t\in[0,T),x\in\mathbb{R}^d}$ defines a continuous flow (w.r.t the initial condition x): see Pardoux and Peng [27]. In the case when h is measurable only, things are less obvious since u might be discontinuous. Nevertheless, it can be proven that $(Y_s^{t,\xi})_{t\leq s\leq T}$ and $(u(T-s,X_s^{t,\xi}))_{t\leq s\leq T}$ coincide by approximating the terminal condition: we can approximate h by a sequence of bounded smooth functions $(h_\ell)_{\ell\geq 1}$, uniformly of a polynomial growth and converging towards h almost everywhere (for the Lebesgue measure). Then, by standard stability results on BSDEs, it is known that

$$\mathbb{E}\big[\sup_{t\leq s\leq T}|Y_s^{t,\xi}-u_\ell(T-s,X_s^{t,\xi})|^2\big]\leq C\mathbb{E}\big[|h(X_T^{t,\xi})-h_\ell(X_T^{t,\xi})|^2\big], \tag{135}$$

where u_ℓ is associated with the boundary condition h_ℓ by (13). Above, the right-hand side tends to 0 since the law of $X_T^{t,\xi}$ is absolutely continuous w.r.t. the Lebesgue measure (apply Lemma 2.10). By (132), we deduce that $(Y_s^{t,\xi})_{t\leq s\leq T}$ is a continuous version of $(u(T-s,X_s^{t,\xi}))_{t\leq s\leq T}$. (Put it differently, $(Y_s^{t,\xi})_{t\leq s\leq T}$ coincides with the continuous version of $(u(T-s,X_s^{t,\xi}))_{t\leq s\leq T}$ given by Proposition 7.3.)

7.3.3. *Uniqueness.* Given a solution v to the PDE with polynomial growth, the point is to prove that $(v(T-s,X_s^{t,\xi}))_{t\leq s\leq T}$ satisfies the BSDE (12) (for the same ξ as above). Basically, this follows from Itô's formula. As in the continuous case, the polynomial growth property together with the standing assumption on f imply the martingale part in the BSDE to be square integrable on $[t,T]$, that is $\mathbb{E}\int_t^T|Vv(T-s,X_s^{t,\xi})|^2ds<+\infty$. As a consequence, the martingale part $(\int_t^s Vv(T-s,X_s^{t,\xi})dB_s)_{t\leq s<T}$ has an a.s. limit as s tends to T, as the limit of an L^2-martingale. Similarly, by the Cauchy criterion,

$$\left(\int_t^s f\big(T-r,X_r^{t,\xi},v(T-r,X_r^{t,\xi}),(Vv)^\top(T-r,X_r^{t,\xi})\big)dr\right)_{t\leq s<T}$$

has an a.s. limit as well. Therefore, $(v(T-s,X_s^{t,\xi}))_{t\le s<T}$ has also an a.s. limit as s tends to T. We can identify it as an L^1 limit:

$$\mathbb{E}\big[|v(T-s,X_s^{t,\xi})-h(X_T^{t,\xi})|\big]\le \mathbb{E}\big[|v(T-s,X_s^{t,\xi})-h(X_s^{t,\xi})|\big]+\mathbb{E}\big[|h(X_T^{t,\xi})-h(X_s^{t,\xi})|\big].$$

By Lemma 2.10 and by (3) in Definition 2.11, the first term in the right-hand side tends to 0 as s tends to T. The second one also tends to 0 when h is continuous: approximating h in $L^1_{\rm loc}(\mathbb{R}^d)$ by a continuous function and applying Lemma 2.10 again, it tends to 0 as well when h is measurable only.

Finally, there is a version of $(v(T-s,X_s^{t,\xi}))_{t\le s\le T}$ that satisfies (12) with $h(X_T^{t,\xi})$ as boundary condition. By uniqueness of the solution to the BSDE, we deduce that $(Y_s^{t,\xi})_{t\le s\le T}$ and the continuous version of $(v(T-s,X_s^{t,\xi}))_{t\le s\le T}$ coincide, that is $(v(T-s,X_s^{t,\xi}))_{t\le s\le T}$ and $(u(T-s,X_s^{t,\xi}))_{t\le s\le T}$ have the same continuous version. Here, we emphasize that we cannot choose $s=t$ directly because of the possible discontinuities of v and u. Anyhow, we can always claim that

$$\forall t\in[0,T),\ \forall t\le s<T,\quad \mathbb{E}\int_t^s\big|v(T-r,X_r^{t,\xi})-u(T-r,X_r^{t,\xi})\big|dr=0.$$

By Lemma 7.4 below, we deduce that u and v match almost everywhere. □

Lemma 7.4. *Let $\psi:[0,T]\times\mathbb{R}^d\to\mathbb{R}$ be a function such that, for any $t\in[0,T]$ and $x\in\mathbb{R}^d$, $|\psi(t,x)|\le C(1+|x|^r)$ for some $r\ge 0$, and, for any $t\in[0,T)$ and $s\in[t,T)$, $\mathbb{E}\int_t^s\psi(r,X_r^{t,\xi})dr=0$. Then, ψ is zero almost-everywhere for the Lebesgue measure.*

Proof (Lemma 7.4). For any $t\in[0,T)$, there exists a Borel subset $\mathcal{N}_t\subset[t,T]$, of zero Lebesgue measure, such that, for all $s\in\mathcal{N}_t^{\complement}\cap[t,T)$, the integral $\int_{\mathbb{R}^d}\psi(s,y)d\mathbb{P}_{X_s^{t,\xi}}(y)$ is zero. Setting $\mathcal{N}=\cup_{t\in\mathbb{Q}\cap[0,T)}\mathcal{N}_t$, we deduce, that for all $s\in\mathcal{N}^{\complement}\cap[0,T)$, for all $t\in[0,s)\cap\mathbb{Q}$, the integral is zero. In particular, we can let t tend to s: as t tends to s, $X_s^{t,\xi}$ tends in law towards ξ. Since ξ has a density, there is no need of continuity on ψ to pass to the limit in the above expression. (That is, by Lemma 2.10, we can approximate ψ by a continuous function in $L^1_{\rm loc}([0,T]\times\mathbb{R}^d)$.) We deduce that, for all $s\in\mathcal{N}^{\complement}\cap[0,T)$, $\int_{\mathbb{R}^d}\psi(s,y)\mu(y)dy=0$. Choosing μ in a countable total subset of densities with compact support, we deduce that ψ is zero almost-everywhere. □

7.4. **Proof of Proposition 7.3.** Again, the proof follows via a mollification argument. We need to find a sequence $(v_\ell)_{\ell\ge1}$ of smooth functions such that, for all $p\ge1$,

$$\forall r\ge1,\quad \lim_{\ell\to+\infty}\sup_{1/r\le t\le T}\|v_\ell(t,\cdot)-v(t,\cdot)\|^{V,2}_{\mathbb{B}(0,r),p}=0,\ \lim_{\ell\to+\infty}\|v_\ell-v\|^{\mathcal{V}_{0,1}}_{[1/r,T]\times\mathbb{B}(0,r),p}=0. \tag{136}$$

Indeed, introducing the stopping times $(\tau_q=\inf\{s\ge t:|X_s^{t,x}|\ge q\})_{q\ge1}$ ($\inf\emptyset=+\infty$), we can apply Itô's formula to $(v_\ell(T-s,X_s^x))_{0\le s\le\tau_q\wedge(T-\varepsilon)}$, for some small positive real ε.

Therefore, for any $\ell\ge1$ and any $t\le s<T$, we have $v_\ell(T-s,X_s^{t,\xi})-v_\ell(T-t,\xi)=\mathcal{I}_\ell(s)$, with

$$\mathcal{I}_\ell(s)=\int_t^s\left[-\mathcal{V}_0v_\ell+\frac12\sum_{i=1}^N V_i^2v_\ell\right](T-r,X_r^{t,\xi})dr\int_t^s Vv_\ell(T-r,X_r^{t,\xi})dB_r.$$

By Lemma 2.10, the following quantity makes sense:

$$\mathcal{I}(s) = \int_t^s \left[-\mathcal{V}_0 v + \frac{1}{2}\sum_{i=1}^N V_i^2 v \right](T-r, X_r^{t,\xi})dr + \int_t^s V v(T-r, X_r^{t,\xi}) dB_r.$$

By Lemma 2.10 again, $\lim_{\ell\to+\infty} \mathbb{E}\big[\sup_{t\le s\le \tau_q\wedge(T-\varepsilon)} |\mathcal{I}(s) - \mathcal{I}_\ell(s)|\big] = 0$. Therefore,

$$\lim_{\ell\to+\infty} \sup_{k\ge 0} \mathbb{E}\Big[\sup_{t\le s\le \tau_q\wedge(T-\varepsilon)} |v_{\ell+k}(T-s, X_s^{t,\xi}) - v_\ell(T-s, X_s^{t,\xi})|\Big] = 0.$$

We deduce that we can find a continuous adapted process $(\Xi_s)_{t\le s<T}$ such that

$$\lim_{\ell\to+\infty} \mathbb{E}\Big[\sup_{t\le s\le \tau_q\wedge(T-\varepsilon)} |\Xi_s - v_\ell(T-s, X_s^{t,\xi})|\Big] = 0. \tag{137}$$

The point is now to identify $(\Xi_s)_{t\le s<T}$ as a version of $(v(T-s, X_s^{t,\xi}))_{t\le s<T}$. By Lemma 2.10,

$$\lim_{\ell\to+\infty} \mathbb{E}\big[|v(T-s, X_s^{t,\xi}) - v_\ell(T-s, X_s^{t,\xi})|\big] = 0. \tag{138}$$

By (137) and (138), we deduce that, for any $s\in[t,T)$, $\mathbb{P}\{\Xi_s \neq v(T-s, X_s^{t,\xi}), \sup_{t\le s\le T}|X_s^{t,\xi}| \le q\} = 0$. Letting q tend to $+\infty$, this completes the proof. □

Now, (136) follows again from a convolution argument, the proof of which is left to the reader. □

Lemma 7.5. *For two smooth densities ρ_1 and ρ_d over $\mathbb{R}$ and $\mathbb{R}^d$, both with compact support, and for a solution v to the PDE as in Definition 2.11, define for all $\varepsilon>0$*

$$v^\varepsilon(t,x) = \int_{\mathbb{R}^{d+1}} v(t-\varepsilon s, x-\varepsilon y)\mathbf{1}_{\{t-\varepsilon s>0\}}\rho_1(s)\rho_d(y)dsdy.$$

Then, for all $p\ge 1$,

$$\forall r\ge 1, \quad \lim_{\varepsilon\to 0}\sup_{1/r\le t\le r} \|v^\varepsilon(t,\cdot) - v(t,\cdot)\|^{V,2}_{\mathbb{B}(0,r),p} = 0, \ \lim_{\varepsilon\to 0}\|v^\varepsilon - v\|^{\mathcal{V}_0,1}_{[1/r,r]\times\mathbb{B}(0,r),p} = 0.$$

ACKNOWLEDGMENT

The work of D. Crisan was partially supported by the EPSRC Grant No: EP/H0005500/1.

The work of F. Delarue was partially supported by a British Council–Ministères français des Affaires étrangères et européennes et de l'Enseignement supérieur et de la Recherche/Alliance Programme grant.

REFERENCES

[1] S. Ankirchner, P. Imkeller and G. Dos Reis. Classical and variational differentiability of BSDEs with quadratic growth *Electronic Journal of Probability, 12 (2007), 1418–1453*

[2] P. Briand, P. Delyon, Y. Hu, É. Pardoux and L. Stoïca. L^p solutions of backward stochastic differential equations *Stochastic Processes and Applications 108 (2003), 109–129.*

[3] P. Briand and Y. Hu. BSDE with quadratic growth and unbounded terminal value. *Probab. Theory Related Fields 136 (2006), 604–618.*

[4] P. Briand and F. Confortola. BSDEs with stochastic Lipschitz condition and quadratic PDEs in Hilbert spaces. *Stochatic Processes and Applications 118 (2008), 818–838.*

[5] D. Crisan and K. Manolarakis. Probabilistic methods for semilinear partial differential equations. Applications to finance. *M2AN Math. Model. Numer. Anal. 44 (2010), 1107–1133.*

[6] D. Crisan and K. Manolarakis. Solving Backward Stochastic Differential Equations using the Cubature Method, SIAM J. Finan. Math., 3(1), 534-571, 2012.
[7] D. Crisan and S. Ghazali. On the convergence rates of a general class of weak approximations of SDEs. *Stochastic differential equations: theory and applications*, 221–248, Interdiscip. Math. Sci., 2, World Sci. Publ., Hackensack, NJ, 2007.
[8] F. Delarue and G. Guatteri. Weak Existence and Uniqueness for FBSDEs. *Stochastic Processes and Applications 116 (2006), 1712–1742.*
[9] G. Dos Reis. *Some Advances on Quadratic BSDE: Theory–Numerics–Applications*. LAP LAMBERT Academic Publishing, 2011.
[10] N. El Karoui, S. Peng and M.-C. Quenez. Backward stochastic differential equations in finance. *Math. Finance 7 (1997), 1–71.*
[11] Y. Hu, P. Imkeller and M. Müller. Utility maximization in incomplete markets. *The Annals of Applied Probability 15 (2005), 1691–1712.*
[12] N. Kazamaki. *Continuous Exponential Martingales and BMO*. Lecture Notes in Mathematics, 1579. Springer-Verlag Berlin and Heidelberg GmbH & Co. K, 1994.
[13] M. Kobylanski. Backward stochastic differential equations and partial differential equations with quadratic growth. *Ann. Probab. 28 (2000), 558–602.*
[14] A.N. Kolmogorov. Zufllige Bewegungen (zur Theorie der Brownschen Bewegung). *Ann. of Math. 35 (1934), 116–117*.
[15] S. Kusuoka. Approximation of expectation of diffusion processes and mathematical finance. *Advanced Studies in Pure Mathematics* **10** (2001), 261–277.
[16] S. Kusuoka and D.W. Stroock. Applications of the Malliavin Calculus - I. *Stochastic analysis (Katata/Kyoto, 1982) (1982), 271–306.*
[17] S. Kusuoka and D.W. Stroock. Applications of the Malliavin Calculus - II. *Journal of the Faculty of Science, Univ. of Tokyo 1 (1985) 1–76.*
[18] S. Kusuoka, D.W. Stroock. Applications of the Malliavin Calculus - III. *Journal of the Faculty of Science, Univ. of Tokyo 2 (1987), 391–442.*
[19] S. Kusuoka. Malliavin calculus revisited. *J. Math. Sci. Univ. Tokyo, 10 (2003), 261–277.*
[20] T. Lyons and N. Victoir. Cubature on Wiener Space. *Proc. R. Soc. Lond. A (2004), 169–198.*
[21] C. Nee. *Sharp Gradient Bounds for the Diffusion Semigroup*. PhD Dissertation, Imperial College London, 2011.
[22] S. Ninomiya. A New Simulation Scheme of Diffusion Processes: Application of the Kusuoka Approximation to Finance Problems. *3rd IMACS Seminar on Monte Carlo Methods—MCM 2001 (Salzburg). Math. Comput. Simulation 62 (2003), 479–486.*
[23] S. Ninomiya. A partial sampling method applied to the Kusuoka approximation. *Monte Carlo Methods Appl. 9 (2003), 27–38.*
[24] S. Ninomiya and N. Victoir. Weak Approximations of Stochastic Differential Equations and Application to Derivative Pricing. *Preprint.*
[25] D. Nualart. *The Malliavin Calculus and Related Topics*. Probability and its Applications. Springer Verlag, 1995.
[26] É. Pardoux and S.G. Peng. Adapted solution of a backward stochastic differential equation. *Systems Control Lett. 14 (1990), 55–61.*
[27] É. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. In Stochastic partial differential equations and their applications (Charlotte, NC, 1991), 200217, volume 176 of *Lecture Notes in Control and Inform. Sci., Springer, Berlin, 1992.*
[28] R. Rouge and N. El Karoui. Pricing via utility maximization and entropy. *INFORMS Applied Probability Conference (Ulm, 1999). Math. Finance 10 (2000), no. 2, 259-276.*

DEPARTMENT OF MATHEMATICS, HUXLEY BUILDING, IMPERIAL COLLEGE LONDON, 180 QUEENS GATE, LONDON SW7 2BZ, UNITED KINGDOM
E-mail address: d.crisan@imperial.ac.uk

LABORATOIRE J.-A. DIEUDONNÉ, UNIVERSITÉ DE NICE SOPHIA-ANTIPOLIS, PARC VALROSE, 06108 CEDEX 02, NICE, FRANCE (Corresponding Author)
E-mail address: delarue@unice.fr